We Trip The Light Fantastic
**Forensic Science International** is an international journal publishing original contributions in the many different scientific disciplines pertaining to the forensic sciences. Fields include forensic pathology and histochemistry, chemistry, biochemistry and toxicology (including drugs, alcohol, etc.), biology (including the identification of hairs and fibres), serology, odontology, psychiatry, anthropology, the physical sciences, firearms, and document examination, as well as investigations of value to public health in its broadest sense, and the important marginal area where science and medicine interact with the law.

Review Articles and Preliminary Communications (where brief accounts of important new work may be announced with less delay than is inevitable with major papers) may be accepted after correspondence with the appropriate Editor. Case Reports will be accepted only if they contain some important new information for the readers.

**Submission of Articles:** Manuscripts prepared in accordance with Instructions to Authors should be sent to the Editor-in-Chief or one of the Associate Editors for: Forensic Science International.

**Editor-in-Chief**

P. Saukko – (for: Experimental Forensic Pathology, Traffic Medicine and subjects not listed elsewhere) Department of Forensic Medicine, University of Turku, SF-20520 Turku, Finland

Tel.: (+358) 2 3337543; Fax: (+358) 2 3337600; E-mail: psaukko@utu.fi

**Secretary to the Editor-in-Chief**

A. Saarenpää – Address as for P. Saukko.

Tel: (+358) 2 3337438; Fax: (+358) 2 3337600; E-mail: ansaare@utu.fi

**Associate Editors**

A. Carracedo – (for: Forensic Genetics)

Instituto de Medicina Legal,

Facultad de Medicina,

15705 Santiago de Compostela, Galicia, Spain

Tel.: (+34) 981 580336

O.H. Drummer – (for: Toxicology)

Department of Forensic Medicine,

Victorian Institute of Forensic Medicine

57-83 Kavanagh Street,

Southbank 3006, Victoria, Australia

Tel.: (+61)-3-9684-4334;

Fax: (+61)-3-9682-7353

Cristina Cattaneo – (Anthropology and Osteology)

Instituto de Medicina Legal, Università degli Studi

Via Mangiagalli 37, 20133 Milano, Italy

Tel: 0039 2 5031 5724;

Fax: 0039 2 5031 5724

**Editorial Board**

A. Amorim (Porto, Portugal)

J. Buckleton (Auckland, New Zealand)

B. Budowle (Quintico, VA, USA)

J. Clement (Melbourne, Australia)

S.D. Cohle (Grand Rapids, MI, USA)

S. Cordner (South Melbourne, Australia)

P. Dickens (Buxton, UK)

M. Durigon (Garches, France)

A. Eriksson (Umeå, Sweden)

J.A.J. Ferris (Auckland, New Zealand)

M.C. Fishbein (Encino, USA)

P. Gill (Solihull, UK)

C. Henssge (Essen, Germany)

M.A. Huestis (Baltimore, MD, USA)

A.W. Jones (Linköping, Sweden)

H. Kalimo (Helsinki, Finland)

B. Kneubuehl (Thun, Switzerland)

S. Leadbeatter (Cardiff, UK)

P.J. Lincoln (Surrey, UK)

A. Luna Maldonado (Espinardo (Murcia), Spain)

B. Madea (Bonn, Germany)

N. Morling (Copenhagen, Denmark)

B. Olaisen (Oslo, Norway)

V. Pascali (Rome, Italy)

S. Pollak (Freiburg i. Br., Germany)

M.S. Pollanen (Toronto, Canada)

D. Pounder (Dundee, UK)

O. Prokop (Berlin, Germany)

K. Püschel (Hamburg, Germany)

G. Quatrehomme (Nice, France)

J. Robertson (Canberra, Australia)

P.M. Schneider (Cologne, Germany)

S. Seto (Tokyo, Japan)

J. Simonsen (Copenhagen, Denmark)

P. Sótonyi (Budapest, Hungary)

M. Steyn (Pretoria, South Africa)

F. Tagliaro (Verona, Italy)

T. Takatori (Chiba, Japan)

S. Tsunenari (Kumamoto, Japan)

D.N. Vieira (Coimbra, Portugal)

X. Xu (Shantou, People’s Republic of China)

J. Zhu (Guangzhou, People’s Republic of China)
Introduction

On behalf of the scientific committee of EAFS2006 welcome to this special edition of Forensic Science International which is specially produced to commemorate the Conference held in Helsinki, 13–16 June, 2006.

I think it is important to outline the review process for this special edition as it differs from the normal system.

All authors from the Conference were invited to submit their work for consideration to the Journal. All members of the scientific committee were given access to all submissions. During the Conference the committee met and reviewed their decisions as a group. Any member who had contact with a particular institute abstained from the discussion when papers from that institute was being discussed.

Any paper where two or more of the committee rejected the submission were eliminated from discussion and the remaining papers were read again and included for publication. The process was of necessity quite harsh because papers were submitted directly for publication and were not altered. I believe this process favoured authors who were more familiar with publication requirements. A number of the papers reviewed contained good quality material and their authors were encouraged to rewrite or carry out some additional work and resubmit to the editor for consideration at another time.

The selection of papers in this special edition may not be fully representative of the range of material presented at the Conference. There was much soul searching in relation to which papers were suitable for publication and the main criteria was to have a high quality special edition.

Members of the Scientific Committee are gratefully acknowledged for their efforts and contribution to this special issue: Dr. Bramley (UK); Dr. Broeders (The Netherlands), Dr. Houck (USA), Prof. Kopp (Sweden), Prof. Piekoszewski (Poland), Dr. Sippola (Chair and Organiser of EAFS2006; Finland), Dr. Willis (Chair of the Scientific Committee) and Prof. Pierre Margot.

Sheila Willis, Chair,
EAFS Standing Committee,
EAFS2006 Scientific Committee

Special Issue Editor,
Associate Editor
Pierre Margot*

*Corresponding author. Tel.: +41 21 692 4600; fax: +41 21 692 4605
E-mail address: Pierre.Margot@unil.ch

Available online 7 September 2006
The difference between drug money and a “lifetime’s savings”

Karl A. Ebejer a, Jane Winn a, James F. Carter a,*, Richard Sleeman a, Jill Parker b, Fritjof Körber b

a Mass Spec Analytical Ltd., Building 20F, Golf Course Lane, P.O. Box 77, Filton, Bristol BS34 7QS, UK
b Faculty of Applied Science, University of the West of England, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, UK

Received 8 June 2006; accepted 14 June 2006
Available online 4 August 2006

Abstract

In many countries, monies suspected of being associated with drug trafficking can be seized by the authorities. One of the ways of investigating this association is through the analysis of seized banknotes for traces of controlled drugs.

We report three studies which may assist the expert in assessing whether banknotes contaminated with diamorphine are part of the general population of notes in circulation or whether they show unusual contamination patterns which require explanation.

Study 1 is based on three plausible contamination scenarios as they may occur during the various stages of an illicit drug transaction and seizure. It shows that notes which have been in direct contact with visible traces of diamorphine show significantly higher contamination to those in more indirect contact with the drug.

Study 2 investigates the transfer of diamorphine from one highly contaminated note to other notes in a bundle over a period of 10 weeks with and without agitation. It was found that the total amount of drug transferred was smaller than 6% and no more than 4 out of a bundle of 10 previously clean notes became lightly contaminated.

Based on extensive background data, study 3 proposes a probabilistic model to assess whether an observed proportion of diamorphine bearing banknotes is likely to have been contaminated by chance. The model predicts that there is only a 0.3% chance that a bundle of 100 notes from the general banknote population contains more than six contaminated specimens.

Jointly, the three studies give useful indications for the spread of contamination throughout a sample and the amounts of heroin which may be expected given plausible contamination scenarios.

© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Diamorphine; Banknotes; Transfer

1. Introduction

The Financial Action Task Force (FATF) lists 40 recommendations, which member countries are committed to implement [1]. The first of these recommendations is to criminalise money laundering on the basis of the United Nations Convention against Illicit Traffic in Narcotic Drugs and Psychotropic Substances, 1988 (the Vienna Convention) [1]. The United Kingdom is one of the 33 members of FATF and, in the UK, the Proceeds of Crime Act allows for the seizure of cash reasonably suspected to be gained unlawfully [2]. Where illegal drug activity is suspected, the results of the analysis of the banknotes for traces of illicit substances can be used as circumstantial evidence. However, this requires that the banknotes be distinguishable from those typically in general circulation [3].

Banknote analysis by gas chromatography mass spectrometry (GC/MS) has been performed successfully by a number of workers [3–7]. A major drawback of any method based on GC/MS is the long preparation and analysis time required, in comparison with thermal desorption atmospheric pressure chemical ionisation tandem mass spectrometry (APCI MS/MS) [8–10]. The latter is currently the method of choice for law enforcement agencies in the UK. The speed of this technique has allowed a huge number of banknotes from general circulation, so called “background” banknotes, to be analysed individually. This background database of illicit drug traces detected on banknotes from the UK is maintained and routinely used for comparison with drug traces detected on banknotes seized from suspects thought to be involved in illegal drug activities.
Since the background database contains data from individual banknotes, the contamination of bundles of banknotes can be described in terms of both the quantity of drug contamination [11] and the proportion of banknotes in a bundle with detectable traces of a drug [12]. The advantage of having two complementary methods for data interpretation is that either or both can be employed as appropriate to the drug of interest. Cocaine, for instance, occurs on virtually all Bank of England banknotes in general circulation [10]. Therefore, any attempt to differentiate on the basis of contaminated proportion is futile. Such an approach, however, is valid for diamorphine contamination, where the drug occurs on much smaller numbers of banknotes in the background.

The proportion of banknotes with diamorphine contamination has been used as the discriminating factor in a large number of legal proceedings, but this forms only one aspect of the evidence. Defendants commonly claim that large quantities of banknotes in their possession represent a lifetime’s savings, a gift, an inheritance, the profits from trading in a grey economy, etc. They also maintain that the diamorphine contamination came to be on the banknotes by mechanisms unknown and unrelated to them. It is, therefore, useful to consider not only whether a seizure is contaminated to an unusual degree, but also to assess the pattern of contamination on the seized banknotes and the likelihood that they became contaminated through means other than illegal drug activity.

In this paper, we report on the contamination patterns of banknotes subjected to three plausible scenarios associated with activities related to drug dealing. A further study investigated the transfer of diamorphine from a highly contaminated note to adjacent notes over a period of 10 weeks. Lastly, a probabilistic model was fitted to the diamorphine contamination background database to permit a more reliable assessment of which contamination proportions may be classed as “unusual”.

2. Materials, methods, techniques

All banknote analyses were performed using triple-quadrupole mass spectrometers (MDS Sciex, Concorde, ON., Canada) with custom-built thermal desorption inlets consisting of two metal plates heated to 285 °C. Insertion of a banknote or paper swab, for approximately 1 s, between the metal plates caused volatilisation of debris including diamorphine, from the sample. The vapours became entrained in ambient air flowing into the APCI source of the instrument [12]. The mass spectrometer was programmed to monitor two product ions from the protonated molecule of diamorphine (m/z 370). The product ions of interest for diamorphine were m/z 328 and m/z 268. Peak areas were obtained from the extracted ion chromatograms (Analyst 1.4, MDS Sciex, Concorde, ON., Canada) using software written in MATLAB (release 12, The Mathworks, Natick, MA, USA) for the identification and integration of peaks [11].

Instrument performance was monitored by injection of a solution containing 2 ng of diamorphine in methanol. Swabs from the analyst’s gloves and worktop were analysed, and shown to be free from diamorphine, prior to making contact with the banknotes.

The advantage of the described analytical technique is that a single banknote can be analysed twice—once for each end. The results from two ends of an individual banknote may vary, due to localised deposits. However, average results from the two ends are found to converge when large numbers of banknotes are analysed (unpublished results). In forensic casework, exhibits that have been analysed once can be reanalysed if required.

2.1. Study 1—primary, secondary and tertiary transfer to banknotes

British banknotes, being composed of a mixture of cotton and linen, were shown to be simulated best by cotton paper, rather than wood pulp based products as previously used [13]. A study of diamorphine transfer showed cotton paper to have similar properties to banknotes (average correlation coefficient 0.84) in contrast to a wood pulp based copy paper (average correlation coefficient 0.79) (unpublished results).

Contaminated banknotes were modelled using Crane’s Crest cotton paper (100% rag, 90GSM, Crane & Co. Inc. Dalton, MA). The paper was cut to the approximate size of a £10 sterling banknote. This substitution removed the need to correct for the natural background levels of drugs on banknotes and also eliminated differences between the quality of banknotes (e.g. worn, new, dirty, clean). Indeed, it proved difficult to obtain a sufficiently large number of banknotes in a similar physical condition.

Three stages were used to simulate different steps in a drug-dealing scenario. Each stage was physically isolated on a 30 cm × 30 cm sheet of laminated chipboard of a type commonly used in furniture manufacture. The intention was that the amounts of diamorphine present on each of the three surfaces were representative of those transferred by primary, secondary and tertiary contacts with the drug.

In the first stage, approximately 10 mg portions of street heroin (Avon & Somerset Constabulary, Scientific Investigations, approximately 10% diamorphine) were weighed into cigarette papers to make 10 typical street-sized dealer’s “wraps”. Samples were weighed using Diamond brand “professional mini” electronic scales, typical of the variety seized from suspected drugs dealers. Ten cotton paper notes were counted onto this surface into a single pile, which was then inverted and the notes recounted. Subsequently, the notes were analysed.

In the second stage the wraps, prepared above, were placed on a surface previously shown to be clean. Ten cotton paper notes were then counted, as before, onto this surface and then again onto a third, clean surface prior to analysis.

In the third stage, 10 cotton paper notes were counted onto the third surface and analysed.

Gloves were worn during the weighing and wrapping of heroin, but cotton paper notes were counted using bare hands because this appeared more realistic. Hands were washed with copious amounts of warm water between counting stages. The
entire sequence was repeated using clean surfaces and cling film to manufacture the wraps.

2.2. Study 2—transfer between banknotes

To simulate the contaminants as realistically as possible, particulate street heroin and household dust were used to contaminate clean paper. Household dust is likely to contain skin oils as found on banknotes [10] and may be expected to simulate the daily environment to which banknotes are exposed. A quantity of street heroin (as above) was mixed in a plastic receptacle with a dust taken from a domestic vacuum cleaner and passed though a 1.7 mm sieve followed by a 0.5 mm sieve.

A number of cotton paper notes were agitated manually for a period of several minutes with the spiked dust. The notes were withdrawn, shaken to remove loose deposits and tested and the response to diamorphine was measured. The amount of heroin in the dust was adjusted until the response was approximately at the mid-point of the dynamic range of the MS–MS instrument. This mixture was then used to spike further cotton paper notes in the same manner. A few of the spiked notes were analysed to check that sufficient diamorphine had been transferred to mimic the contamination levels typically detected on contaminated banknotes observed in case work.

A sample consisting of 120 banknotes (£10 denomination) was drawn over the counter from a bank in Bristol, UK and one end of each note analysed for the presence of diamorphine. A record was kept of the order in which the notes were analysed, and this order was maintained while dividing the banknotes into 12 groups of 10 banknotes each. A spiked cotton note was then inserted exactly in the middle of each bundle (Fig. 3), and the bundles were sealed inside close fitting zip-lock bags. All bags were then sealed inside a tamper-evident bag and stored in the dark.

Over the course of 10 weeks, each bundle, except two control bundles, was counted on a weekly basis. The zip-lock bags were distributed to different co-workers who opened the bag and counted out the banknotes, including the spiked note, at least once, in any manner they pleased, but without altering the sequence of notes within the bundle. After a few minutes, the banknotes were returned to their respective zip-lock bags. In weeks 2, 4, 6, 8 and 10, two bundles were analysed and removed from the pool. The analysis was performed on both ends of each note. The remaining samples were resealed inside a tamper-evident bag and returned to dark storage.

2.3. Study 3—comparing contamination with the background database

One hundred and eighty-six bundles of banknotes, each comprising 100 banknotes or more, were selected from the background database. Diamorphine contamination on both ends of the banknotes was measured using the procedure described. The mean number of banknotes in each sample that gave a true positive response for diamorphine was determined [12]. This number was then normalised to the total number of banknotes in the sample (true positives per 100 banknotes) to correct for sample size variation, and rounded to the nearest integer.

The arithmetic mean of the number of contaminated banknotes per 100 ($\mu$) was also determined from the same data set. A Poisson distribution, using $\mu$ as the predicted outcome ($\lambda$) was then generated and compared to the background database. Data were processed using SYSTAT 11 for Windows (SYSTAT software Inc., Richmond, CA).

3. Results

3.1. Study 1—primary, secondary and tertiary transfer to banknotes

Fig. 1(a and b) show the intensity of the diamorphine product ions, as a function of time, detected on ten cotton paper notes from each of the three stages simulated. The trends observed for cigarette paper wraps and cling film wraps can better be visualised from a histogram representing the relative, average diamorphine transfer for each category (Fig. 2). The number of contaminated notes decreases with the level of contact in both the cigarette paper and cling film wrap scenarios. Secondary...
transfer from cigarette paper wraps produced about half the number of contaminated notes compared to primary transfer. However, for secondary transfer from cling film wraps almost all the notes in the bundle were contaminated, making it difficult to distinguish primary from secondary transfer on the basis of contamination proportion. Tertiary transfer resulted in a single note becoming contaminated from both types of wrapping material.

The relative intensity plot (Fig. 2) shows that the quantities transferred in the various stages differ considerably and an analysis of variance was performed to establish whether the apparent differences are statistically significant. For a given experimental situation, i.e. for each bundle, the distribution of contamination (peak areas) was found to be approximately log-normal. Therefore a two factor analysis with replication (Microsoft Excel 97) and a Tukey test [14] were performed on the logarithms of the peak areas for the m/z 268 and the m/z 328 product ions. The tests established which parameters influence the mean quantity of drug detected on each analysed bundle and whether one can distinguish the various stages of a drugs deal by the quantity of drug transferred.

In the two factor analysis the main factor was the type of transfer scenario (primary, secondary or tertiary transfer). This variable was responsible for approximately 65% of the total variation in the sample set, as measured by the total sum of squares. Its contribution was more than twice as large as that from random fluctuations, showing that this factor is highly significant \( p < 10^{-14} \). The second factor analysed was the type of material used for the wraps. This factor was responsible for less than 5% to the variation but was still a statistically significant source of difference \( p < 0.01 \). As might be expected, the interaction between the two factors is small and at any reasonably chosen level of significance the two factors are not correlated.

The Tukey test identified which of the bundles were significantly different with regards to the amount of drug transferred. The test was performed independently for each ion on all six bundles (three transfer types and two wrap materials) at a confidence level of 95%. The test reliably differentiated between primary transfer scenarios in one group and secondary and tertiary transfer scenarios in a second group, i.e. the notes in direct contact with spilled heroin on the cutting table were significantly more contaminated than those from indirectly contaminated bundles. Despite the apparent trend towards less contamination for the tertiary stage it was not possible to differentiate between secondary and tertiary transfer, based solely on the quantity of material transferred, at the 95% confidence level. By reference to Fig. 2 it is, however, immediately clear that secondary and tertiary transfer can be distinguished on the basis of the proportion of contaminated notes.

3.2. Study 2—transfer between banknotes

A signal to noise ratio of 4.5 was used as a threshold for peak detection. Peak areas for the two ions were summed. The real peak data for each end of a given banknote or spiked note were also summed to give a better representation of the total diamorphine contamination detected on the note. Every two weeks average of the data from two bundles, was calculated to arrive at a single value for the response of the MS/MS instrument to a banknote in a given environment.

The diamorphine contamination detected on banknotes at a given sequential distance (Fig. 3) from the spiked note was
calculated as the total contamination on the two contributing notes. The degree to which diamorphine was transferred or retained was expressed as a proportion of the total amount of diamorphine detected (Fig. 4).

In addition, the number of contaminated banknotes in a bundle was plotted as a function of time (bar chart, Fig. 5). The amount of diamorphine detected on the spiked paper, corrected for instrument variation, was also plotted (line graph, Fig. 5).

3.3. Study 3—comparing contamination with the background database

A total of 186 bundles of banknotes in the background database were found to have an average of 1.85 true positives for diamorphine per 100 banknotes. The number of contaminated banknotes in a bundle was shown to follow a Poisson distribution ($\chi^2 = 2.30, p = 0.81$) (Fig. 6). Using this Poisson distribution it is possible to estimate the chance that a sample of banknotes with a high degree of diamorphine contamination has been drawn from the background by chance. The highest proportions of contaminated banknotes observed in the 186 backgrounds were two incidents of 6 contaminated banknotes per 100, and a single incident of 8 notes per 100. These upper values are consistent with those predicted by a Poisson distribution based on a predicted outcome ($\lambda$) of 1.85 contaminated notes per 100 (Table 1).

The chance of a seizure having 10 contaminated banknotes per 100, or higher, is very slight if the banknotes had a similar history to that experienced by banknotes in the background.

Fig. 4. Proportion of diamorphine detected on the spike and adjacent banknotes.
4. Discussion

Although quantitative amounts of diamorphine could easily have been applied to the simulated notes as solutions, it was considered that the use of impure, solid ‘street heroin’ samples would provide a more realistic model for ‘real-life’ heroin contamination in several ways. First, solutions are likely to ‘wick’ into the fibres of the paper, rather than remaining on the paper surface, or being trapped in the interstices between fibres as would be expected of particles. In turn, this may render the diamorphine less amenable to liberation by thermal desorption. Secondly, the evaporation of the solvent would lead to the formation of smaller crystal sizes than would typically be present in ‘street heroin’, and most certainly would not represent a true reflection of the range of particle sizes found in the real world [15]. It is reasonable to speculate that larger crystals may transfer more easily between banknotes as they become trapped less readily between the fibres. Since most of the mass transferred will be in the form of larger crystals, applying diamorphine in solution is likely to represent a poor model for real world contamination, even though the spiking would be more quantitative and less variable. Further, application of a solution could be achieved by dipping (leading to homogeneous distribution of the material) or via a syringe (resulting in contamination concentrated in certain areas), neither of which accurately reflects the likely distribution of particulate material. Hence, although the method used to spike cotton notes lacks both repeatability and reproducibility it was considered the best model available.

In study 1 it was, perhaps, unsurprising that cotton notes counted on a surface with visible deposits of heroin retained significant traces of diamorphine which could subsequently be liberated by thermal desorption. The intensity of the responses observed from these notes was similar to those observed when analysing banknotes from suspected drugs dealers. Wraps prepared on the same surface would be exposed to similar amounts of heroin and can reasonably be expected to retain external deposits. Some proportion of these deposits may then fall onto the stage 2 surface, becoming available for retention by the cotton notes counted on this surface.

From Figs. 1 and 2, it is apparent that some secondary transfer to the cotton notes has indeed occurred. Cotton notes counted after cigarette paper wraps had been in contact with the surface, however, retained much smaller traces of diamorphine than those counted after cling film wraps had been in contact with the surface.

Different mechanisms are undoubtedly responsible for the retention of particles by paper and by cling film. In the latter case the mechanism is almost certainly electrostatic, possibly leading to larger initial retention and a larger resulting deposit. In addition, the retention mechanism of the cling film may be weaker than the paper, resulting in a deposit that is more readily lost. In contrast, particles retained by paper surfaces are likely to be retained as inclusions in surface features [10] or by dissolution in surface dirt (finger grease, etc.) or water associated with the cellulose. These different mechanisms may well exhibit different affinities for different crystal sizes

### Table 1

<table>
<thead>
<tr>
<th>Number of contaminated banknotes in a sample of 100</th>
<th>Approximate odds of drawing these by chance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>6:1</td>
</tr>
<tr>
<td>1</td>
<td>3:1</td>
</tr>
<tr>
<td>2</td>
<td>4:1</td>
</tr>
<tr>
<td>3</td>
<td>6:1</td>
</tr>
<tr>
<td>4</td>
<td>12:1</td>
</tr>
<tr>
<td>5</td>
<td>33:1</td>
</tr>
<tr>
<td>6</td>
<td>100:1</td>
</tr>
<tr>
<td>8</td>
<td>2000:1</td>
</tr>
<tr>
<td>10</td>
<td>50,000:1</td>
</tr>
</tbody>
</table>

Fig. 5. Total number of contaminated banknotes (excluding the spiked notes) detected from two bundles of ten banknotes after introducing a spiked note to each bundle and counting on a weekly basis.

Fig. 6. Data observed (red) and predicted by a Poisson distribution (blue) for diamorphine contamination of banknotes taken from general circulation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)
resulting in preferential retention of different materials. It is possible that deposits retained by both the cotton and cigarette papers are strongly bound, not readily lost to the stage 2 and 3 surfaces and, therefore, not available for subsequent retention by clean cotton notes. This conjecture is supported by the small deposits retained by cotton notes counted on the third surface.

The statistical analysis shows that the quantities of drugs transferred to the notes directly from the cutting table are significantly different from those transferred at the secondary and tertiary stages. This information would be of limited value unless the study was performed at similar primary surface contamination levels as observed in real life scenarios. Whilst this cannot be asserted, the aforementioned similarity in intensity observed in seized samples gives an indication that this may indeed be the case.

In study 2 the source of contamination (the spike) was easily identifiable at every stage of counting. The amount of diamorphine remaining on the spiked note through the study varied from 94 to 99% of the total amount detected on all banknotes in a bundle. Transfer occurred primarily from the spiked note to one or both banknotes immediately adjacent (Adj1) to it in a bundle. Where the spiked note lost as much as 6% of its diamorphine to other banknotes in the bundle, the amount detected on a single adjacent banknote was less than 6%. This is in line with previous findings [10]. After 4 weeks and after 8 weeks, diamorphine was also detected on a banknote two steps (Adj2) and three steps (Adj3) away from the spiked note, respectively. On banknotes between these and the spiked note, no contamination was detected, suggesting that transfer in these instances may have occurred via the inner surface of the zip-lock bag in which a bundle was stored, or via the counting surface.

Since a single source of contamination was, in general, found to contaminate up to two banknotes with diamorphine, it might be argued that two-thirds of the contamination detected in an exhibit could have arisen by this process. This would assume minimal bunching of contaminated notes in a bundle. To have an appreciable effect, banknotes with “primary” contamination would need to be dispersed throughout a bundle of otherwise uncontaminated banknotes.

In the first 8 weeks of the study, the number of contaminated banknotes in every bundle of 20 notes increased with the number of counting events. This was, however, concomitant with a general increase in the diamorphine contamination of the spiked paper in these bundles. One of the drawbacks of using a solid material to create the spiked notes was that there was no way to precisely regulate the amount of drug present on them. The observed trend in the spike note intensity reflects the random variations in this process.

A combination of the amount of drug initially present on the spiked note, and the number of mixing events, is probably closer to the truth. In the 10th week, a sudden drop in the number of contaminated banknotes in the bundles that had been subjected to weekly counting corresponded with a fall in the amount of contamination on the spiked paper. However, the spike sizes were similar in the counted and uncounted control samples, but the number of contaminated banknotes in that sample was higher. One explanation may be that the majority of transfer occurs in the first contact event, experienced by all bundles when the spiked paper was introduced. Subsequently, transfer is minimal, and each counting occasion results in losses from the bundles and more exposure to light and moist air that could increase the natural degradation of the diamorphine [16–18].

In study 3, the database against which comparisons were made comprises mainly samples drawn from over the counter at banks throughout the UK. A smaller component of the database consists of samples from small businesses such as pubs and shops. These are all considered to be representative of banknotes in general circulation. There has been extensive discussion on the degree to which banknotes mix in general circulation [12]. If the small proportion of banknotes contaminated with diamorphine in general circulation is indeed randomly distributed, then the probability of drawing a sample bearing a given number of contaminated banknotes can be described by a Poisson distribution [19]. This is a good approximation for a binomial probability for large numbers where the event of interest occurs with a low probability. There is no need to specify sample size as long as an expected frequency can be calculated for the number of events observed [19].

The study has illustrated that the chance of detecting highly contaminated bundles of banknotes in samples drawn at random from general circulation, is slight. Indeed, the detection of high frequencies of diamorphine contamination implies that the banknotes may have originated from an isolated population, such as drug dealers or drug users. It is possible that contaminated monies could, for example, be paid into a bank and subsequently withdrawn by an innocent individual prior to mixing. Extensive mixing of such banknotes with banknotes in general circulation would account for the low frequency observed in the database. Finding a highly contaminated bundle therefore cries out for an explanation.

5. Conclusions

The limited studies discussed in this paper have produced some interesting findings, although further work will be necessary to strengthen the conclusions that can be drawn. Cotton notes in direct contact with surfaces bearing visible deposits of heroin were found to retain significant amounts diamorphine. Such deposits did not transfer readily between banknotes when mixed and counted in bundles (less than 6% transfer). A single highly contaminated note is unlikely to contaminate more than the two adjacent banknotes. Since only a small proportion of the initial deposit is transferred the primary and secondary deposits are readily distinguishable. The extent of transfer between banknotes appears to be dependent upon the amount of material on the initial spiked note but may also be dependent upon the physical handling of the notes, which can increase or decrease the size of secondary deposits.

Other materials, e.g. cling film and cigarette papers can transfer deposits to a secondary surface where they are available for retention by cotton notes. These secondary deposits do not then transfer readily to other surfaces to become available for retention by other notes. Hence, counting contaminated notes on a surface along with clean notes does not appear to spread
contamination. These data suggest that two retention mechanisms are involved, one by which material is strongly retained and one by which material is weakly retained and readily lost; the latter providing a means of secondary transfer.

As a result of poor transfer, only a very limited number of banknotes in general circulation (approximately 2%) carry traces of diamorphine. The mixing and circulation of banknotes provides a mechanism by which these contaminated notes are randomly dispersed in the general population producing a homogeneous background. Knowing the average number of contaminated banknotes in the background it is, therefore, possible to estimate the chance of drawing a given number of contaminated banknotes using a Poisson distribution. Using this predictor, the chance of drawing a large proportion is very slight. Even if the contamination on these banknotes had spread to adjacent notes, the overall conclusion of this study is that banknotes known to have been in direct contact with deposits of heroin can be distinguished from those contaminated by secondary or tertiary processes.

Acknowledgements

The authors wish to thank the following people for their help. Lauren Baba of the Centre for Chemometrics, School of Chemistry, University of Bristol for her work in comparing the distribution of contamination on banknotes. Paula Hodge and Martin Scrase of Lloyds TSB for information about the circulation of UK banknotes. MSA staff for populating the background database.

References

Latent fingermark visualisation using a scanning Kelvin probe

Geraint Williams*, Neil McMurray

Materials Research Centre, School of Engineering, University of Wales Swansea, Singleton Park, Swansea SA2 8PP, UK

Received 8 June 2006; received in revised form 10 August 2006; accepted 25 August 2006
Available online 4 October 2006

Abstract

The current state of the art in fingermark visualisation on metallic surfaces by a scanning Kelvin probe (SKP) technique is described. Latent eccrine fingermarks deposited on a range of polished and roughened metallic surfaces can be effectively imaged. Results are presented which show that the SKP technique is able to visualise fingermarks obscured beneath optically opaque soot films and retrieve ridge detail in instances where fingermarks have been physically removed (e.g. by rubbing with a tissue) from a metal surface. SKP Volta potential mapping of small, severely non-planar metal objects such as fired brass cartridge cases is demonstrated.

# 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Fingermark; Scanning Kelvin probe; Visualisation; Latent; Metal surface

1. Introduction

Over a century after fingerprint evidence was first used to obtain a criminal conviction, the detection and identification of fingerprints remains at the forefront of modern crime investigation [1,2]. When a fingertip is brought into contact with a surface, pores on the papillary ridges deposit a residue of perspiration, consisting of 99% water with the remainder made up of inorganic salts such as NaCl and organic substances such as urea. In addition, fingertip contact with the face or scalp leads to the presence of sebaceous materials, such as fatty acids, within the fingermark residue. The complex patterns formed by the ridge contours are unique to a given individual, a premise which underpins the use of fingermarks as a means of identification admissible in a court of law. Invisible or latent fingermarks are detected by the use of a “developer” which produces a high degree of visual contrast between ridge deposit patterns and the surface on which the fingermark is deposited.

In cases where fingermarks are deposited on metallic surfaces, development techniques range from powder dusting [3], to various chemical treatments such as cyanoacrylate fuming [4] or exposure to ruthenium tetroxide [5] and detect either the moisture present within the fingermark residue and/or one or more of the organic components associated with the deposit. Other water-based development techniques, such as selenous acid [6], ammoniacal silver nitrate [7], palladium salts [8] and gunblueing mixtures [7,9] rely on electrochemical interactions with exposed metal between fingermark ridge deposits to enhance contrast. Such reagents generally work best when the fingermarks are substantially sebaceous in nature and are ineffective in visualising water-soluble eccrine deposits. In a recent paper we showed how fingermarks deposited on metallic surfaces could be visualised directly, without development or any other form of perturbation, by Volta potential mapping using a scanning Kelvin probe (SKP) [10]. Uniquely, since the technique is based on purely potentiometric measurements, fingermark patterns could be detected beneath intact, substantially insulating films such as clear and pigmented lacquers. Furthermore, because the Volta potential is primarily influenced by non-volatile inorganic salts, fingermarks could also be visualised even after the metal surface had been heated to temperatures sufficient to volatilise the organic components of the deposit. In the work presented here, further developments in the SKP technique as a potential forensic tool for fingermark visualisation are described. Improvements in instrument performance by modifying probe tip design and scan parameters are outlined along with modifications to the existing design to enable height profiling of highly non-planar objects such as coins or fired cartridge cases.

We have also sought to identify circumstances where SKP analysis may be more effective than established methods at visualising fingermarks on metal surfaces. The ability of SKP to effectively visualise latent eccrine fingerprint patterns is

* Corresponding author. Tel.: +44 1792 295666; fax: +44 1792 295244. E-mail address: Geraint.Williams@swan.ac.uk (G. Williams).
demonstrated on both iron and brass, for both polished and heavily gritted (roughened) surfaces. In addition it will also be shown that discernible ridge detail can be retrieved by scanning a fingerprinted metal surface after the fingerprint deposit had been removed by rubbing with a tissue. The effect of smoke contamination on fingerprint visualisation by SKP is also presented showing that although eccrine print detail becomes indistinct, sebaceous patterns remain clear even when covered with thick soot layers. Finally, examples of SKP visualisation of latent eccrine fingermarks, deposited post-firing on brass cartridge cases will be given. Although not exploiting the previously observed ability of SKP to detect fingermarks after exposure to high temperatures [10], such examples serve to demonstrate the capability of the technique to deal with “real”, highly non-planar surfaces.

2. Experimental details

2.1. Materials

Samples of 1 mm thick iron, copper, nickel, titanium and aluminium (99.9% purity) were obtained from Goodfellow Metals Ltd. 1 mm thick brass foil, 37% zinc 63% copper, was obtained from Advent Research Materials Ltd. Spent brass cartridge cases of 0.45 in. calibre were provided by the UK Forensic Science Service. Prior to fingerprint deposition, all metal samples were abrasively cleaned and polished using aqueous slurries of 5 µm polishing alumina. Cleaned samples were rinsed with distilled water, followed by acetone, and allowed to dry in air. “Natural” fingermarks from a single donor were deposited using constant pressure from the left forefinger after freshly washing the hands with soap and water. Sebum-rich fingermarks were deposited as above but after rubbing the side of the nose with the fingertip. Eccrine deposits (i.e. inorganic-salt-rich) were produced by rinsing the left forefinger with ethanol followed by 30 min enclosure in an airtight plastic bag. Unless otherwise specified, SKP mapping of fingermarked samples was typically carried out within 6 h of fingermark deposition. The effect of smoke contamination on SKP fingerprint visualisation was studied by passing an inverted, fingermarked soot film at a wavelength of 500 nm.

2.2. Methods

The scanning Kelvin probe apparatus used in this work is described in detail elsewhere [10,11]. The vibrating reference probe assembly was mounted in a fixed position above a moving test sample. The reference probe itself consisted of a gold wire, which was vibrated along the vertical axis using a moving coil electromechanical actuator. Scan parameter optimisation was carried out using a 50 µm diameter wire, while subsequent, high-definition scans of fingerprinted metal surfaces were performed using a more robust 100 µm diameter wire, mechanically profiled to a 20 µm tip. The probe vibration frequency was 280 Hz and the vibration amplitude was 40 µm peak-to-peak. Reference probe vibration amplitudes were checked using stroboscopic observation in conjunction with a travelling microscope. The experimental arrangement was such that the tip of the vibrating reference probe was held at earth potential and positioned inside a stainless steel environment chamber, which was also at earth potential. The electromechanical actuator and vibrating electronics were positioned outside the environment chamber in order to ensure effective electrostatic and magnetic shielding of both the reference probe and sample. Vibration was conducted to the probe tip via a 80 mm long glass push rod.

Positioning and scanning of the test sample was carried out using a micromanipulation stage consisting of three orthogonally arranged (x, y, z), stepper motor driven, linear bearings (Time and Precision Ltd.). The probe was held at a constant height above the sample surface and scanned in a raster of parallel lines to generate a regular array of values which could be mapped using commercially available cartography software (Surfer®, Golden software). For the majority of the work presented here, data points were acquired at 0.05 mm intervals and at a mean probe-to-sample height of 50 µm. To perform each measurement the ac current, \( i(t) \), generated in the external circuit connecting the sample and vibrating probe, was amplified and converted into an ac voltage signal, \( V(t) \), using a dc biased transconductance amplifier circuit. The \( V(t) \) signal was detected using a lock-in amplifier (EG&G model 7265). The dc output of the lock-in amplifier, \( V_{dc} \), was transmitted to a feedback system based on an integrator circuit which controlled the dc bias, \( E \), applied to the sample via the transconductance amplifier so as to automatically null \( i(t) \). The magnitude of the dc bias (equivalent to the Kelvin potential \( -E_{kp} \) or the Volta potential difference between the probe and sample) applied via the integrator, was digitised and logged. Probe scanning and data logging were all carried out automatically under computer control. Unless specified, all scans were performed in ambient air (nominal temperature 22 °C, RH 50%). A photographic image showing the major components of the SKP instrument used in this work is given in Fig. 1.

![Fig. 1. Photographic image of the SKP instrument.](image-url)
Surface analysis of fingerprint residue by scanning SIMS was carried out using a Millbrook Instruments Ltd. MC300 (Mk II) mini-SIMS. The instrument has a lateral resolution of 10 μm, and a mass resolution of ±0.1 Da. Samples for analysis were prepared by freshly depositing a “natural” fingerprint on to a fragment of clean room grade silicon wafer. A Topometrix Explorer atomic force microscope (AFM) was used to study the topography of mechanically ground iron surfaces.

3. Results and discussion

It is believed that inorganic salts present in papillary ridge sweat deposits are principally responsible for producing the observed fingerprint-induced Volta potential depressions which can be mapped by SKP. To confirm the presence of inorganic salts, the elemental composition of fingerprint residue, freshly deposited on a fragment of clean-room grade silicon wafer was determined by imaging SIMS. Broad area positive and negative ion spectra taken of a 3 mm × 3 mm section of fingerprint deposit, revealed the presence of sodium, potassium and chloride ions, along with characteristic peaks of a complex mixture of organic components. Elemental maps, recorded at fixed atomic masses showing areas of high abundance as bright regions of grey scale images, are given in Fig. 2. Sodium (ii), and chlorine (iii) occurrence coincides with the position of fingerprint ridge deposits as shown in the secondary electron (topographical) image given in (i). A mass image obtained for potassium (m/e = +39) was identical to that shown for Na⁺. Mass imaging of silicon (iv) generates the opposite situation, where areas of high abundance concur with regions of substrate that are untouched by the residue.

By scanning an area of metallic surface encompassing a fingerprint deposit, the SKP can be used to generate a map showing spatially resolved Volta potential (E_{kp}) patterns. Fingermarks deposited on a range of planar metal substrates such as iron, aluminium, nickel, titanium, copper, zinc and brass were shown to locally depassivate the metal surface and yield well-defined E_{kp} patterns which coincided with visible ridge deposit detail. Iron and copper surfaces gave the greatest Volta potential differences, typically up to 200 mV, between ridge deposits and the background. Other metal surfaces, including Al and Ti which develop highly passive surface oxide films, exhibit lower E_{kp} variations of typically 50–100 mV, measured over residue-covered and untouched metal regions.

Optimisation experiments, carried out in order to increase SKP resolution and improve the clarity of the fingerprint Volta potential maps, were performed out on a “natural” fingerprint deposited on a brass surface. Previously published work on the effect of probe diameter (D) and probe-to-sample distance (d) on the lateral resolution of a SKP[12] predicts that a reduction in both D and d should improve the sharpness of SKP-derived Volta potential maps of fingermarked metal samples. Two separate sets of experiments were carried out on the same fingermarked brass sample. In the first, the effect of probe-to-sample distance (d) on the resolution of the Volta potential maps of the same fingermarked region was investigated. For each scan, a constant data point density of 20 pts/mm and a probe diameter (D) of 50 μm was employed. Results showed that using d < D had little effect on the resolution of the greyscale Volta potential maps obtained. However when d was increased, the Volta potential maps obtained became progressively more blurred and much fine detail was lost. In a second set of experiments, the effect of data point density on greyscale Volta potential map clarity was studied at a fixed d of 50 μm. Increasing the data point density in both x and y planes gave a significant improvement in the clarity of the Volta potential

Fig. 2. Elemental maps of a 4.5 mm × 4.5 mm area of fingermark-deposited silicon wafer obtained using imaging SIMS. The secondary electron image of the area analysed is shown in (i), while the areas abundant in (ii) Na⁺, (iii) Cl⁻ and (iv) Si⁺ are represented by the bright areas of the grey scale maps.
maps obtained. Unsurprisingly, using a low data point density (e.g. 5 pts/mm) gave rise to a highly pixellated map, although employing such scan parameters over a large sample area may be useful as a pre-scan to pinpoint regions which may bear a fingermark deposit. The highest image definition was obtained at a point density of 40 pts/mm where image quality was sufficient to reveal a high degree of ridge deposit pore structure (see Fig. 3). Using the present configuration, high resolution SKP scans, requiring the acquisition of large numbers of data points take a considerable time to complete (e.g. a 30 h scan time was typically required to generate the $2.4 \times 10^5$ point data grid from which Fig. 3 was derived). However, in practice a data point density of 20 pts/mm is adequate to effectively visualise a complete fingermark pattern in sufficient detail for identification purposes and allow Volta potential maps to be acquired in a significantly shorter time period (typically 6–8 h). These optimisation studies confirm that significant improvement in SKP spatial resolution and thus fingermark ridge definition can be obtained by (i) decreasing the probe tip diameter, (ii) increasing the data point density and (iii) decreasing the probe-to-sample height to a point where $d \approx D$.

3.1. Effect of fingermark composition

Sebaceous, lipid-rich fingermarks deposited on polished metal surfaces are usually visible to the naked eye and may not require any development to enhance image quality [13]. In contrast, eccrine fingermarks composed mainly of an aqueous solution of inorganic salts and non-lipid organic compounds are typically latent, i.e. invisible to the naked eye. As mentioned previously, aqueous methods for developing fingerprints on metal surfaces are ineffective in such instances because of the high solubility of eccrine deposits. In addition, physical vapour deposition of thin multi-metal layers are also problematic because eccrine residue is less efficient than sebum-rich deposits in inhibiting metal condensation [13]. To assess the efficiency of SKP in visualising different types of fingermark, both sebaceous and eccrine fingermarks were deposited on planar polished iron surfaces and scanned using a 100 μm wire profiled probe and employing a probe-to-sample height of 50 μm and a point density of 20 pts/mm. Interpolated SKP-derived greyscale Volta potential maps of the fingermarked regions are shown in Fig. 4. Both the highly visible sebaceous print and the latent eccrine fingermark patterns are clearly visualised by SKP Volta potential mapping. In the case of the sebaceous print, the fingermarked region shows a broad, depassivated (by ca. 80 mV compared to the surrounding metal) area encompassing the fingermark ridge pattern. Fine ridge detail is discernible within this region where a ca. 100 mV $E_{kp}$ contrast between ridge peak and furrow areas is observed. The eccrine pattern shows a greater contrast of ca. 200 mV between the exposed substrate and the areas in contact with the fingermark residue. The SKP-derived map obtained for a sebaceous fingermark confirms the observations of Thomas and Reynolds [13], who report that regions between sebaceous ridge deposits can be connected by a continuous film of thickness up to 10 μm. It is likely therefore that the darkest (i.e.
most negative $E_{kp}$ fine ridge structure is principally due to the interaction of inorganic salts with the underlying metal while the surrounding “shadow” is caused by the presence of this interconnecting surface film. The observation that SKP is highly effective at visualising latent eccrine deposits confirms previous assertions that Volta potential depression appears to be produced primarily by the non-volatile inorganic salt component of sweat deposits [10].

Eccrine fingermarks were also deposited on iron substrates which had previously been subjected to mechanical grinding using a range of different grade silicon carbide papers to study the influence of surface roughness on SKP-derived Volta potential images. For all surface finishes ranging from 1200 (smoothest) to 60 grit (roughest), the fingermark ridge detail could be clearly visualised. A typical example is shown in Fig. 5, where a Volta potential map of a region of 60 grit treated iron surface bearing a latent eccrine fingermark has been generated by SKP analysis. An atomic force microscope investigation of the topography of the 60 grit surface showed up to 40 $\mu$m undulations in the surface. The SKP-derived Volta potential map shows that although the majority of the fingermark ridge pattern can be clearly visualised, some detail becomes indistinct at areas where capillary action has caused eccrine deposit to be drawn along grooves in ground surface.

3.2. Visualisation of “removed” fingermarks

A serendipitous finding of our research is that the physical removal of a fingermark from a metal surface, e.g. by rubbing with a tissue, may not necessarily mean that all the fingermark pattern information is lost. The Volta potential depressions which form the basis of SKP fingermark visualisation are caused primarily by interaction of inorganic salts in fingermark residue with the underlying metal. Since these interactions appear to commence the instant that the fingermark residue contacts the metal surface, it may be assumed that aggressive ions present in the residue, such as chloride ($\text{Cl}^-$), which break down passive films and activate metal surfaces through the formation of soluble metal chloride complexes, can rapidly be incorporated into the protective oxide film at the metal/fingermark interface. Physical removal of the fingermark residue by rubbing with a cloth or tissue may not necessarily remove species which have already chemically interacted with the metal surface.

This hypothesis was tested by depositing ideal “natural” fingermarks on polished iron surfaces and allowing varying time periods to elapse before removing the residue by rubbing...
vigorously with a paper tissue. Areas of the iron surfaces which had borne fingermarks prior to their removal were then scanned by SKP. Fig. 6 shows a series of SKP-derived Volta potential maps obtained for fingermark/metal contact times of (i) 5 min, (ii) 1 h and (iii) 48 h prior to removal. Even a short contact time of 5 min appears sufficient to allow some permanent interaction of fingermark residue components with the metal surface. The areas of higher potential than the background metal probably derive from organic adsorbates while regions of depressed potential are caused by chloride-induced activation of the iron surface. After a 48 h contact period, much of the ridge detail remains after fingermark removal and SKP visualisation is achieved with an excellent Volta potential contrast (ca. 100 mV) between ridge deposits and the background metal surface. Similar results were obtained with a brass surface, although a fingermark/metal holding time in excess of 1 h was required to yield discernible ridge pattern detail when subjected by SKP visualisation. These preliminary findings suggest that the SKP technique may be worthy of further investigation as a forensic tool to visualise fingermarks on small metal objects which investigators suspect may have been handled by a criminal and subsequently “wiped clean” at a later time.

3.3. Effect of smoke contamination

The effectiveness of SKP in visualising fingermarks on metallic samples which had been contaminated by a soot layer was investigated. Studies were carried out on planar iron samples onto which fresh sebaceous or eccrine fingermarks were deposited. SKP scanning of the soot affected fingermarks showed that sebaceous deposits could be clearly visualised, even in the presence of a dense soot layer. A typical Volta potential map of a soot-contaminated fingermark is shown in Fig. 7, where a thick soot layer of 5% optical transmittance (measured at 500 nm wavelength) is obscuring the fingermarked region. Eccrine deposits could be imaged through light soot contamination, but characteristic Volta potential patterns became indistinguishable from the background in the presence of heavy contamination. The influence of soot layer thickness, quantified by optical transmittance measured at 500 nm, on the contrast of SKP derived Volta potential fingermark images is shown in Fig. 8. The contrast was calculated by taking an $E_{kn}$–distance profile through a central portion of the fingermark pattern (indicated by the dashed line in Fig. 7) and calculating a mean value for $E_{kn}$ differences between 10 adjacent ridge deposit—background metal regions. Again the plots highlight the difference in fingermark composition-dependent behaviour, where Volta potential contrast decreases markedly for eccrine deposits (Fig. 8, curve (ii)), even in the presence of a thin soot covering, while becoming enhanced for sebaceous deposits at similar contaminating layer thicknesses (Fig. 8, curve (i)). The likelihood that thick, sebum-rich deposits can block electron transfer between the underlying metal and conductive carbon particles in the soot layer probably accounts for the observed difference in the efficiency of visualisation.

3.4. Volta potential mapping of non-planar objects

Although all the experiments outlined in the previous sections were performed on planar samples, non-planar samples can also be scanned provided the surface profile is known. Surface profiling may be carried out using conventional mechanical or optical methods. However, recent work [14] has demonstrated that surface profiling is possible using the SKP itself by exploiting the fact that harmonic distortion in the Kelvin probe current ($i(t)$) is dependent on probe to sample distance. This methodology has now been adapted in the SKP instrumentation described in Section 2, resulting in the capability to map the Volta potential patterns of small three-dimensional metallic items such as coins and cartridge cases. The procedure for maintaining a constant probe-to-sample distance over a non-planar surface involves a carrying out preliminary height scan of the area of interest. In height...
profiling mode, a high dc bias is applied between the probe and the sample (typically +2 V) and the Kelvin probe currents at the fundamental ($i(t_1)$) and second harmonic frequencies ($i(t_2)$) are measured. After adjusting the probe-to-sample distance to the desired value (typically 50 $\mu$m), the resulting $i(t_2)/i(t_1)$ ratio is logged. The probe is then scanned over the test surface under active $z$-axis control to maintain a constant $i(t_2)/i(t_1)$ ratio at all times. In this way, a grid of laterally resolved $z$-axis movements is generated, which is subsequently re-traced when a Volta potential scan is performed.

The use of same Au wire probe to carry out both height profiling and Volta potential mapping eliminates many of the alignment problems associated with other surface profiling methods. Preliminary results on samples of significant forensic interest, such as brass cartridge cases, show that Volta potential mapping can reveal characteristic fingermark ridge detail, even on examples where the fingermark is invisible to the naked eye, i.e. latent. Fig. 9a shows a 0.45 in. calibre brass cartridge case, polished to a 5 $\mu$m finish using an aqueous Al$_2$O$_3$ slurry and mounted on a cylindrical SKP sample holder. The area indicated by the white dashed rectangle bears a latent eccrine fingermark, deposited by the method described in Section 2 post-firing. The height profile of this region, scanned at a point density of 20 pts/mm and a mean probe-to-sample distance of 50 $\mu$m, is given in Fig. 9b, along with its corresponding Volta potential map. Fingermark ridge deposits are clearly revealed as regions where $E_{kp}$ is typically ca. 100 mV lower than the remainder of the brass surface.

4. Conclusions

Certain circumstances have been identified where SKP imaging may be more effective than established development techniques for visualising fingermarks on metal surfaces. It has been demonstrated that the SKP can effectively visualise late eccrine fingermark patterns on both iron and brass, for both polished and heavily roughened surfaces. It has also been demonstrated that discernible ridge detail can be retrieved by scanning a metal surface from which fingermark deposits had been physically removed by rubbing with a tissue. The extent of discernible detail was shown to be dependent upon the time of fingermark-metal contact prior to removal. The effect of smoke contamination on fingermark visualisation by SKP has also been studied. Eccrine fingermark detail becomes indistinct under smoke film contamination, while sebaceous patterns remain clear even when covered with thick, optically opaque smoke layers. SKP visualisation of latent eccrine fingermarks, deposited post-firing on brass cartridge cases has been achieved, demonstrating the capability of the technique to deal with “real”, highly non-planar surfaces.

The limitation of SKP use to conductive surfaces, relatively small scan areas, protracted scan times required to visualize fingermarks and lack of commercially available instrumentation make it unlikely that the SKP technique will be immediately suitable for routine work as carried out in police forensic facilities. However, it is much more likely that SKP may provide a significant increase in the power available to visualise latent fingermarks on metallic items and fragments associated with serious crimes (murder and attempted murder) and cases of terrorist attack.

Acknowledgements

This research was funded by the UK Engineering and Physical Sciences Research Council (EPSRC) via its ‘Technologies for Crime Prevention and Detection’ programme. The authors would like to thank the UK Forensic Science Service for supplying test samples.

References

Clothing damage analysis in alleged sexual assaults—The need for a systematic approach

C.A. Boland *, S.D. McDermott, J. Ryan

Forensic Science Laboratory, Department of Justice Equality and Law Reform, Garda Headquarters, Phoenix Park, Dublin 8, Ireland

Received 9 June 2006; accepted 14 June 2006
Available online 4 August 2006

Abstract

Clothing damage analysis is an integral part of the examinations carried out in sexual assault type cases. This analysis can be used to corroborate different versions of events and is at its most powerful in elucidating false allegation cases and consent cases. The purpose of this study was to determine to what extent people with varying levels of forensic awareness, experience and training could correctly carry out damage analysis.

Two participant groups were asked to take part in this study. Group A ('forensic group') comprised participants at a forensic science conference, and Group B ('student group') comprised students undertaking a degree course in Forensic Science. Each group was given a practical workshop consisting of a lecture outlining common fabric types and general features observed in different damage types. Each participant was subsequently shown 25 pieces of ‘damage’ and asked to identify both the type of fabric construction (knit or weave) and the type of damage (cut, tear, rip, wear and tear). The ability to identify fabric construction and damage types varied within the two groups studied and across the groups. The forensic group performed better both in fabric and damage assessment than the student group.

This paper suggests a systematic approach to clothing damage analysis to maximise the benefits that can be obtained from this area of forensic science and to minimise the subjectivity within the field.

Keywords: Forensic science; Sexual assault; Damage; Clothing; Training

1. Introduction

The fundamental role of a forensic scientist is to help those who address the burdensome issue of guilt or innocence in a court of law. Did he rape her? Did she murder him? Were they supplying drugs? A large proportion of our work is quantifiable: a DNA profile can be reported along with a match probability, narcotics can be identified using longstanding techniques such as HPLC and mass spectroscopy. However, an equally large proportion of our work is subjective. Histological semen identification, footmark identification and even fingerprint identification are techniques, which are at least in part, subjective. They are robust subjective techniques, which we all accept as accurate and discriminating. Damage interpretation certainly does not enjoy the same accolades.

As forensic scientists, we often find ourselves in situations, where the most pertinent information can only be gleaned using subjective testing. A large percentage of crimes against the person, dealt with by forensic science laboratories, are crimes of sexual assault. In Ireland, this constitutes approximately 56% of the cases received in the Biology section (approx. 450 cases per year). The majority of these cases either begin with or acquire a consent defence by the time they reach our courts. In these cases, finding semen and in fact getting a matching DNA profile, may offer no additional evidential value to the case. Other examinations, such as damage interpretation, possibly indicating a struggle or that force was used, may be critical. This analysis may be used to corroborate or refute a particular scenario and indeed, in a small, but significant number of cases, damage interpretation may be critical in preventing false allegations proceeding to prosecution [1]. We believe that the potential usefulness of clothing damage analysis requires us at least attempt to measure our ability to correctly assess damage and propose a systematic approach to damage analysis to minimise the subjectivity within the field.

Previous work has shown the benefits of capturing the features specific to cuts and tears in different fabrics [2]. This information can be used to help ascertain if the damage was
recent and if one can tell what type of implement may have caused the damage [3,4]. These theories have been applied to casework studies [5,6], which have highlighted the huge potential of damage analysis and interpretation in crime investigation.

The aim of our study was to address the issue of whether or not trained professionals had greater competency in clothing damage analysis, and further, to devise guidelines, which might aid less subjective interpretation.

2. Materials and methods

Two participant groups were asked to identify damage types, which included cuts, tears, rips and wear and tear in 25 damage test areas on different items of clothing. In addition, the groups were asked to designate the damaged test fabrics as either knit or weave for each sample fabric.

2.1. Participants

2.1.1. Forensic group (N = 46)

This group comprised attendees at a forensic science conference. The group included a range of professionals including forensic scientists, pathologists and police officers. About 41% of this group indicated they assessed clothing damage, while 24% indicated they reported it. Only 37% of the group indicated that they had any specific training. The range of professional experience for the group was from none (26%) to over 10 years (20%).

2.1.2. Participant group B (N = 35)

This group comprised students in their second, third and fourth year of a forensic science degree course. None of the students had previous exposure or training in damage analysis. The general age profile of the group was 18–22 years.

2.2. Workshop

Each participant group was given identical workshops in terms of content and time allocation. Workshops were two and a half hours in total. These consisted of a lecture detailing the background and significance of damage analysis. Different types of fabric construction were introduced before outlining the main types of damage encountered in casework, i.e., cuts, tears, rips and wear and tear. For each type of damage, the specific features for the different damage types were explained. This took approximately 1 h in total. The remainder of time was allocated to participants to assess test damage presented. The definitions given to the participants were as follows; cut: a severance with neat edges caused by a sharp implement; tear: a severance in the fabric caused by pulling with some force leaving ragged or irregular edges; rip: broken or unravelling sewing thread stitches (e.g., at hems and seams); wear and tear: the general damage seen on clothing from day to day wear and use.

2.3. Types of test damage

A total of 25 test pieces of damage were shown to each participant. Of these, 14 were made with knit fabrics, and eleven were made of weave construction. Within these test pieces, there were a total of five cuts, seven tears, six rips and six areas of wear and tear. Some test pieces had more than one type of damage present and the participants were not told which ones (e.g., a cut and a tear along one severance). Each area of test damage was marked and given an identifying number.

2.4. Survey

Each person completed a questionnaire to capture information about previous experience and training with special reference to damage training. In order to keep the damage identification uniform and to make interpretation of the survey simpler, the participants were asked to fill in a check box form containing all the options required per item of damage (see Table 1).

3. Results

3.1. Fabric identification

The total number of test pieces was 25 and of these 14 were constructed by a knit method and 11 by a weave method. Each score was included if the correct identification was made but was not subtracted for selecting the incorrect fabric. The mean correct over all the fabric pieces is shown in Table 2 and Fig. 1 for the two groups.

![Fabric Construction Identification](image)

![Fabric Construction Identification](image)

Fig. 1. Graph of the correct identification of fabric construction as a knit or a weave.
3.2. Damage identification

The scoring for the identification of damage was quite stringent. The participant was marked correct if all types of damage were identified in each test piece. Participants were not negatively scored if they chose an incorrect type of damage but did not gain a mark if a type of damage was missed or if the test piece was not answered at all.

The number of test damage types is shown in Table 3 with the figures showing the percentage of each group that got the correct answer.

Fig. 2a–d shows the breakdown of the proportion of correct identifications made by the two groups for the four individual damage types and Fig. 3 shows the mean identification for both groups for all four damage types.

### Table 3

Represents the percentage of each group to correctly identify varying numbers of the different types of damage

<table>
<thead>
<tr>
<th>Damage</th>
<th>Group (%)</th>
<th>No. of damage test pieces</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>Cut</td>
<td>Forensic</td>
<td>0 0 13 61 26 0</td>
</tr>
<tr>
<td></td>
<td>Student</td>
<td>3 6 11 23 40 17</td>
</tr>
<tr>
<td>Tear</td>
<td>Forensic</td>
<td>2 11 9 30 26 13 2</td>
</tr>
<tr>
<td></td>
<td>Student</td>
<td>3 6 34 37 11 6 3</td>
</tr>
<tr>
<td>Rip</td>
<td>Forensic</td>
<td>4 26 4 28 22 12 4</td>
</tr>
<tr>
<td></td>
<td>Student</td>
<td>9 14 37 20 14 6 0</td>
</tr>
<tr>
<td>Wear/tear</td>
<td>Forensic</td>
<td>2 7 8 20 28 22 13</td>
</tr>
<tr>
<td></td>
<td>Student</td>
<td>6 6 23 23 11 8</td>
</tr>
</tbody>
</table>

Fig. 2. Ability of both groups to identify the cuts (a), tears (b), rips (c) and wear and tear (d) test pieces.
3.3. Damage assessment accuracy

In order to measure the accuracy of individual’s damage assessment ability, we looked at the number of correct responses as a proportion of the total responses. In other words, participants were scored only on the damage they were prepared to assess. In addition, there were some test pieces that individuals did not have a chance or time to complete and thus were not answered at all and these items were discounted. Table 4 gives the figures of the accuracy of the answers completed. This is expressed as a percentage of the group that got the assessment correct.

When represented graphically (Fig. 4), it can be seen that in the answers given, there is a much higher success rate for accuracy than for total identification (see Fig. 3).

3.4. Specific examples

The authors considered that some of the damage test pieces contained very obvious damage types, which should have been easily identified by people with only the basic awareness attained from the workshop. Other test pieces were considered to be more difficult and it was felt that these should be correctly identified by trained personnel, but that untrained people would find them more difficult. Table 5 lists five examples of specific damage test pieces with the proportion in each group that correctly identified each piece. The test pieces considered to be obvious types of damage and easily identifiable were a scissor cut (6B), a rip of approximately forty centimetres long along the side seam of a shirt (4B) and a tear at the pocket of a pair of trousers (9A). Items 3A (button torn off a shirt) and 11A (tears at a cuff from wear and tear) were considered by the authors to be more difficult to identify, but should have been correctly identified by trained assessors. Correct assessment of damage excludes participants who had not answered.

4. Discussion

The capturing of information via a survey has been successful in highlighting the need for training in other areas of forensic science [7].

The initial aim of the study was to help address the issue of whether or not trained professionals were competent in clothing damage analysis. With this in mind, two participant groups were chosen and tested. There were issues that should be kept in mind when considering the findings: 26% of the forensic group (the professionals) actually had no experience, being newly recruited. Lighting and microscopic equipment normally

<table>
<thead>
<tr>
<th>Table 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy of damage assessment</td>
</tr>
<tr>
<td>Mean correct (%)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>0–24</td>
</tr>
<tr>
<td>Forensic group (%)</td>
</tr>
<tr>
<td>Student group (%)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Excludes damage types omitted or not answered.

<table>
<thead>
<tr>
<th>Table 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison of both groups’ ability to identify obvious damage (6B, 4B, 9A) or more difficult damage (3A, 11A)</td>
</tr>
<tr>
<td>Group</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Forensic (%)</td>
</tr>
<tr>
<td>Student (%)</td>
</tr>
</tbody>
</table>
available in laboratory conditions were not available. Magnifying glasses (3 ×) were provided. Participants were time restricted. Reconstruction of damage was not possible. Taking the above factors into account, the authors felt that there should still be a discernable difference between the participant groups, if damage assessment was a ‘trainable’ competency.

The forensic group consistently performed better than the student group in fabric construction and damage type identification. As the students had no experience or training, while some of the forensic group had, this may account for the higher scores by the professionals. However, there may be other factors involved such as life experiences. The age profile of the student group was generally 18–21 years but some people in the forensic group had been working for 10–20 years and therefore were considerably older. In fact, one of the best results obtained in the entire study was from an individual with no specific damage training, with 25 years experience in a police force.

Overall, there was a wide distribution of ability within the groups for identifying the different fabric constructions. The majority of the forensic group could differentiate between knits and weaves in more than half of the items while the students had more difficulty, with the highest proportion only identifying between a quarter and half of the items correctly (Table 2, Fig. 1). As the fundamental step in damage analysis, fabric construction identification is extremely important. The poor identification emphasizes the need for better awareness and training in fabric construction for forensic scientists.

The damage types of tears,rips and wear and tear, were better identified by the forensic group. A higher proportion of the forensic group identified more of these damage types and at the same time a lower proportion of the forensic group scored poorly (none or one piece identified per type) (Fig. 2b–d). However, the student group scored better for the cut identification (Fig. 2a). It was felt that perhaps the confidence of the student in selecting the cuts might also be coupled with the ‘guess’ scenario. Although they had the option to fill in ‘Don’t know’, very few did. Although 57% of the student group identified four or five of the cuts, 9% of the group also identified none or only one of the cuts. None of the forensic group identified so few of the cuts.

The mean damage identification for the four types of damage shows that 74% of the forensic group correctly identified more than half of the test pieces compared to 46% of the student group (Fig. 3). This figure is not unexpected but still shows room for improvement.

While total damage identification may be very stringent, it does not give the full picture as to the assessment accuracy of the groups. When the correct identification was expressed as a percentage of the answered items, both groups performed much better (Fig. 4). The forensic group made accurate choices in their assessment (65% of the group getting more than 75% of the damage correct) and made much fewer errors in the assessments than in the student group (none of the forensic group got below 50% of the damage items correct).

In an attempt to isolate why certain types and pieces of damage posed more problems than others, five specific test pieces were addressed. The scissor cut (6B) was easily identified by both the forensic and student groups (Table 5). Both groups also performed well in identifying the rip (4B) and the tear (9A), however, there was a wider difference in ability between the two groups. The two test pieces of damage that were deemed to be more difficult were a button torn from a shirt (3A) and tears arising from wear and tear (11A). The forensic group showed a much better ability than the student group in identifying 11A, which in our view demonstrates that this is a skill that can be learnt. There were enough feature characteristics to unambiguously identify the damage in 3A, but even the majority of the forensic group did not correctly identify it. It was felt that in the forensic group among forensic scientists, the current training methods are not performing or correctly addressing the problems with subjectivity in damage analysis.

For forensic practitioners we propose that with specific training using a systematic approach to damage assessment, identification and interpretation, more conformity can be reached to reduce some of the subjectivity.

The approach to training in assessment of damage interpretation we propose, is a four-tier system of (1) gross morphology; (2) microscopic examination; (3) reconstruction; and (4) assessment and interpretation.

4.1. Gross morphology

There should be a systematic approach to the examination of items for the presence of damage. This includes detailed notes, diagrams, and/or photographs of the garment and damage, outlining the fabric type as this influences the specific features to look for in different damage types. The damage should be noted also in relation to other items of clothing, i.e., multi-layered damage. For each piece of damage, measurements should be taken for future comparison with possible weapons, and pathology reports.

4.2. Microscopic examination

For each area of damage, low power microscopic examination should be used to identify features, e.g., the presence of snippets in a knit fabric but not in a weave [5]. Under magnification, planar array can be confirmed, pulled thread ends can be visualised, etc. Microscopy can be very powerful in finding associated evidence, e.g., blood on the cut thread ends of a stab cut indicating that the blood got there from a wet bloodied blade. How recently the damage was caused may also be elucidated with the help of magnification if clean areas are exposed in a ripped seam (recent) or different fibres are entangled in a severed edge of the damage (not recent).

4.3. Reconstruction

The use of reconstructions for recreating scenarios is invaluable before coming to a conclusion or interpretation as to the likely cause of the damage. This can never reproduce the exact circumstances of the alleged incident due to unknown factors such as force, position, sequence of events, stretching over the body, etc. but may give information about the degree of
force, behaviour of fabric, etc. When carrying out reconstructions a similar undamaged area of the garment should be used if possible, or the same type of garment. Detailed notes should be kept on the way the scenario was simulated in order that it may be reproduced if required. The subsequent experimental damage should be compared both macroscopically and microscopically to the ‘evidence’ damage.

4.4. Assessment and interpretation

There may be limitations with interpretation and defining an exact cause of the damage. Therefore, assessment of the damage should be made on scientific findings if the indicator features are present. It may be assessed as normal wear and tear and no further interpretation is required. However, for other damage types, the interpretation must take into account the circumstances of the case. The circumstances of the case that must be available to the forensic scientist include, statements from the injured party and suspect and medical notes detailing any injuries that may correspond to damage, e.g., stab wounds. The strength of the interpretation in the statement of evidence ranges from very obvious damage (e.g., “the damage was caused by an implement such as a scissors”) to less obvious damage (“cannot eliminate normal wear and tear”).

This four-tier system can be applied to the examination of clothing using standard operating procedures for uniformity within a laboratory and across laboratories.

5. Conclusion

The ability to identify fabric construction and damage types varied within the two groups studied and across the groups. The forensic group performed better both in fabric and damage assessment than the student group.

With the introduction of laboratory standard operating procedures outlining a systematic approach to damage assessment, the subjectivity in the area can be minimised, ensuring that the maximum benefit can be obtained from this very powerful piece of evidence.

Acknowledgements

The authors would like to thank Dr. John Fox, School of Chemical and Pharmaceutical Sciences, Dublin Institute of Technology, Kevin St., Dublin 8, Ireland for facilitating the student workshop and the Forensic Science Society for facilitating the conference workshop.

References

The potential (negative) influence of observational biases at the analysis stage of fingermark individualisation

Beatrice Schiffer*, Christophe Champod

Université de Lausanne, Ecole des Sciences Criminelles, Institut de Police Scientifique, Quartier UNIL-Sorge, Bâtiment Batochime, Lausanne CH-1015, Switzerland

Received 9 June 2006; accepted 14 June 2006
Available online 28 July 2006

Abstract

Recent cases of erroneous identification have strengthened critical comments on the reliability of fingerprint identification. This goes hand in hand with recent publications regarding the lack of scientific foundation of the discipline. Combined with “legislative” needs, such as for instance the admissibility criteria under Daubert, or experimental studies revealing potential bias, the call for research on the identification process has become more urgent.

That background set the basis of this research project financed by the Swiss National Science Foundation (SNSF) which includes, among other parts, experimental tests to study potential observational biases in the analysis stage of fingerprint individualisation. These tests have been submitted to several groups of forensic science students at the University of Lausanne.

The aim is to study factors potentially influencing the analysis of fingermarks, more specifically the influence of training/education (test I) as well as the potential impact of case contextual information or known print availability (test II). For all tests students were given 11 or 12 fingermarks of a medium to difficult quality, with a range of 8–15 minutiae. For all tests the task was always the same for the participants but carried out in different contexts: to analyse the marks, to annotate the minutiae observed, to designate them and to decide on the status of the mark in two categories, exploitable and identifiable. The aim was to see how the fingermarks were annotated by different individuals so as to have an idea of the variation in annotation and counting in the analysis stage only.

For test I, students were submitted the same 12 fingerprints before and after having followed specific training in fingermark individualisation. The aim was to see how training/education impacts the analysis of fingermarks. For test II, were participants given eleven fingermarks so as to study whether the presence of a comparison print changes the amount of minutiae found and whether low/high-profile background information influences the analysis stage.

Results show that for test I the effect of training can be observed, among other, in an increase of minutiae annotated and a higher consensus between participants. For test II no effect of the stimuli used to induce observational biases has been observed by all of the factors studied.

Keywords: Fingermarks; Observational biases; Errors; Misidentification; Experimental study

1. Introduction

Traditional forensic identification evidence – especially fingermark individualisation – though accepted in court for over 100 years, has been challenged lately. Reasons are, among others, the highly publicised erroneous identification by the FBI in the Mayfield affair [1] or legislative needs such as renewed attention on admissibility criteria [2]. The misleading influence of observational biases in forensic science [3] is mentioned as a possible explanation for errors in the Mayfield misidentification [1]. It has been theorised that the lower the quality of the fingermark, the more demanding and subjective the analysis process will be, and accordingly the more vulnerable to stimuli potentially inducing observational biases [4]. Those stimuli can be “circular reasoning” (looking for features found in the comparison print on the fingermark), disregarding of the “one discrepancy rule” (explaining away discrepancies) and incorrect verification of the results when questioned [1]. All these problems do have their origin in the inaccurate following of the fingermark individualisation process Analysis, Comparison,
Evaluation and Verification (ACE-V) [5], especially jeopardising the independence of the four stages. Within this process basically three elements exist. They are:

1. the fingermark: the object to be observed;
2. tools and methods used: the way the observation is made;
3. the fingermark examiner: the observer.

Observational biases can have their influence on all three elements and might happen all along the ACE-V process. We will focus on the analysis stage, as Langenburg [6] showed that significant differences exist already in the beginning of the process. Indeed, the amount of minutiae found by fingermark examiners in contrast to lay persons increased. Enlargement of the fingermark was also observed to modify, that is to increase the amount of minutiae observed. Experimental tests on the influence of bottom-up and bottom-down factors on fingermark individualisation have shown that the quality of the fingermark (contrast, potential distortion and apposition) is an important factor in creating a difficult decisional situation (for lay-persons) and that this might be a critical factor in the individualisation process [4]. Furthermore, a study with five experienced fingermark examiners showed that, when submitted twice the same fingermark in a highly different (and emotionally charged) context three of them changed their initial opinion [7]. Kerstholt et al. studied observational biases in shoe print comparison [8]. They evaluated how expectations (background of the case) and complexity of task (difficulty of the comparison) as well as experience influence the evaluation of simulated case work of 12 shoe print examiners. Contrary to their expectations they did not find any of these influences having a bearing on the examiners results. Only experience changed the way decisions were justified, but not the results as such.

The examiner will be conditioned by all these previous factors, namely the fingermark and the protocol to be followed for observation. However, in addition to them various other external and internal elements will influence him. Internal factors will be understood as those characteristics which are part of the individual. This comprises for instance his visual faculties, his training and his experience. They will not vary considerably from one fingermark comparison to another, though they might evolve. External factors will be understood as elements that will influence one particular case because of exceptional circumstances. These might comprise special media coverage due to the profile of the case or working under special pressures due to time or other constraints.

All these factors might be leading to irreproducible analysis, especially if added to two peculiarities of forensic science work. First, there is the environment and the aim of forensic work that can be tense due to time and heavy context associated with the case. Second, known comparison material may be available upfront, and originate mostly from a “relevant” source from the investigation perspective. Expectation would be that the chance to find a match between the mark and the known comparison material might be a high a priori for the examiner. This phenomenon might even be reinforced by experiencing this happening repeatedly.

To summarise, it can be stated that the lower the quality of the fingermark, the more demanding and subjective the analysis process will be. According to Tversky and Kahneman [9] the tendency to rely on additional, though not necessarily relevant information increases when the data present does not offer enough information for a clear decision. It is then that case relevant and not mark relevant information will be used to reach an opinion. In short, fingermark individualisation is more vulnerable to expectations and biases if the difficulty of the task is increased.

The aim of the research was thus to study the potential (negative) influence of observational biases at the analysis stage of fingermark individualisation by an experimental approach using low quality marks. Such research should help assess whether or not observational biases can develop into errors compromising the reliability of fingermark individualisation as advanced by some authors [3,10].

Several factors were studied. The first was the effect of training on the analysis stage of fingermarks (test I). Forensic science students were tested before having acquired thorough knowledge in fingermark individualisation and after having followed a course on forensic identification and having carried out practical examination. It was predicted that the number of minutiae annotated (per fingermark and per individual) would increase with training and that the overall variation between individuals for the minutiae counted for a given fingerprint would diminish. It was also predicted that the amount of fingermarks considered exploitable (useful for comparison purposes) and identifiable would increase as well with training. Langenburg [6] showed already the higher efficiency of trained professionals compared to lay person. We felt important to attempt to show that such a trend (if any) can be related to specialised training.

Secondly, the potential impact of observational effects has been explored (test II). Stimuli in form of the presence of a matching and a non-matching comparison fingerprint was expected to vary the amount of minutiae found by the group submitted to condition A (group A) versus condition B (group B). The same phenomenon was expected for the disclosure of a high-profile (terrorism) versus low-profile (attempted petty burglary) background scenario.

2. Method

2.1. Participants

Participants were forensic science students of the School of Criminal Sciences of the University of Lausanne. Depending on the test, different classes were sampled. However, all forensic science students have followed the same theoretical and practical lessons in fingermark individualisation.

2.2. Materials and design

Fingermarks from practical student case work were given for analysis to three experienced fingermark examiners in order to determine among others the number of minutiae found. Their findings were used as criterion for selecting fingermarks for the tests. A total of 16 fingermarks relatively close to the Swiss threshold of 12 minutiae and of varying nature were included in both tests (12 in test I and 11 in test II).
For test I (experience) a within group design over a period of time was used. Experience was manipulated by testing the same population (39 students) before and after (29 students) having acquired specific theoretical and practical knowledge, that is having followed a full course in forensic identification. For test II (observational biases) a between-group (between subjects) design was chosen. Observational biases were manipulated by subjecting group A to one condition and group B to the second condition for the same fingermark (Table 1). The sample included 20 master and 28 bachelor students, all having followed the same forensic identification course. To study the influence of the presence of a known comparison print at the analysis stage a matching and a non-matching case were given to each group. The same design was used for the high/low-profile manipulation (attempted petty burglary versus terrorism case). One reference fingermark was presented to both groups without specific stimuli, so as to have a direct between test condition comparison. The test sheets for groups A and B were randomly distributed within the class.

Materials comprised the four times enhanced fingermarks all printed with a photo-quality Fuji Pictrostrat printer for both tests I and II. The results to be observed were the same for both tests I and II. For each fingermark, the response consisted in the total number of minutiae annotated, the respective type of minutiae – ridge ending, bifurcation, point or unknown – as well as the classification of the fingermark into exploitable or identifiable.

2.3. Procedure

For both tests I and II, participants carried out the task during a normal class. They first read a written instruction sheet explaining the task to them and the way to proceed. Explanations as to the aim of the research were minimal and kept very general. Then, the task was described, namely, to analyse each given fingermark by annotating each minutiae found, to determine the type of minutiae found – ridge ending, bifurcation, point or unknown – as well as the classification of the fingermark into exploitable or identifiable.

Participants were encouraged to differentiate the minutiae marked either by letter (A, B, P and U) and/or by colours. Participants were expected to use the standard procedure of analysing a fingermark learned during their theoretical and practical lessons based on Ashbaugh [5].

3. Results

3.1. Test I: training

The aim was to observe possible differences in the annotation of minutiae before and after training of forensic science students. In a first time the mean number of minutiae found by participants over all fingermarks was compared. If they did find an average of 8.1 minutiae when novice, they found 11.5 when having been trained, that is an average of 3.4 minutiae more when more experienced. Increase was highest for fingermarks with background noise and/or bad contrast. Then, for each fingermark the means and the difference of minutiae observed were compared. The minimal and maximal number of minutiae found per fingermark and the range of minutiae found (difference between minimum and maximum) were noted as well.

For each of the 12 fingermarks the mean of minutiae found does increase considerably from the novice to the experienced condition, with a minimum of two minutiae more observed, up to a maximum of nearly seven minutiae (Graph 1). All of these values are highly significant by the statistical t-test. The minimum amount of minutiae found per fingermark does increase as well for all fingermarks. The same is true for the maximum, excluding however 3 out of 12 fingermarks. In general, the range of minutiae found by all participants for the same fingermark does decrease after having been trained in fingermark examination (two out of three fingermarks). As to the types of minutiae found, it has been observed that over all minutiae marked the total number of ridge endings passes from 34.2% to 44.1%, a very significant increase. Inversely, the amount of bifurcations decreases from 54.7% to 47.6%. The rest of the change due to points and unknown minutiae is negligible in comparison. Concerning the classification of the fingermarks in terms of exploitable and identifiable in can be observed that the mean number of exploitable nearly doubles form novice to experienced, while it does more than double for the identifiable condition. In summary, results illustrate that participants do see more minutiae for the different fingermarks while showing a greater consensus in their observations.

3.2. Test II: observational biases

The aim was to observe possible differences in the annotation of minutiae between the two groups due to the manipulation of the factors: (1) availability of a known

<table>
<thead>
<tr>
<th>Group</th>
<th>Reference</th>
<th>Availability of known print</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Same finger mark</td>
<td>None</td>
<td>Matching print</td>
</tr>
<tr>
<td>B</td>
<td>Matching print</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Finger mark 1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 1

Illustration of the design for the test II on stimuli for potential observational biases

Graph 1. Variation of the mean number of minutiae annotated by novice and more experienced participants on 12 fingermarks.
effects of stimuli inducing observational biases, no effect of
persists.

comparison print and (2) differing background case informa-
tion. For the mean number of minutiae over all fingerprints
(group A: 138, 01; group B: 137, 28) no difference between the
groups A and B was noted. The same observation applies for all
other parameters studied (see Graph 2 for an illustration of the
difference in mean number of minutiae found by the two
groups).

More specifically, the first fingerprint which was used as
reference does not show any significant difference between the
groups, be it the amount of minutiae found, the mean or any
other factor compared. The same results apply to all other 10
fingerprints. Considering the presence/absence of a known
fingerprint, be it the matching or the non-matching condition,
no significant difference can be observed neither. The same
result does apply to the low/high-profile condition for all data
observed. The type of minutiae observed does not vary between
the two groups, nor does the qualification of the fingerprints as
exploitable or identifiable.

4. Discussion

4.1. Test I

In agreement with our expectations for test I a pronounced
effect (an increase of the total number of designated minutiae)
of training on the analysis stage of fingerprint individualisation
was found for test I. It might be inferred that participants have
gained knowledge on the observation of minutiae. A certain
move towards consensus of the way a fingerprint is perceived
can be observed as well. However, this effect is limited, as still
quite important variations do subsist between examiners. This
is in accordance with the observations of Evett and Williams
[11]. Furthermore, training does affect the type of minutiae
found, as bifurcations are less often attributed. Though these
variations are smaller for the more experienced participants in
comparison with their novice performance, the abovementioned
tests show that differences will still persist. Indeed, observation is linked to individuals, which might differ slightly
in what they observe. Therefore, a clear subjective element
persists.

4.2. Test II

Contrary to our initial expectations for test II on the potential
effects of stimuli inducing observational biases, no effect of
availability of known print nor context information has been
observed. This was true for all fingerprints used in the test.
These results do, to a certain degree, contradict previous
findings or hypotheses, for instance Risinger et al. [3] and their
overview of studies on the detrimental effects of expectation on
reasoning and perception. Also the study by Dror et al. [7],
showing with a within-subject design that context information
might influence the conclusion drawn from a fingerprint as to
its identification or not. However, these studies focus on the
outcome, the moment when observations have to be evaluated
in terms of evidential value. It is argued here, that not all stages
of the process of ACE-V are similarly vulnerable to
observational biases. Indeed, the less decisional tasks are
involved in a stage, the less the risk of “oversimplifying”
information in order to reach a decision might be [9]. Thus,
individuals might quite correctly observe minutiae if the task is
only to designate them, but they might not be able to undertake
this task correctly, if forced to compare with reference material
and draw a conclusion from that all encompassing process.

The study by Kerstholt et al. [8], although based on shoe
print examinations by experienced examiners and a slightly
different design, does come to a similar conclusion as far as the
effects of expectation and complexity go. They did not observe
any difference due to the manipulation of the background,
which should modify expectation. They explain this result
partly by the presence of a formal guideline employed by the
examiners, but also by the potential positive influence of
experience on vulnerability to these potential sources of biases.
As the experience of our participants is limited, this second
thesis does not explain our results. The first aspect, the presence
of a guideline, or a “structured” approach to fingerprint
examination might be a better way into explaining the results.
Indeed, as long as the guidelines are followed properly, there is
a small risk of drift happening.

Although further research is needed for a better understand-
ing of how the ACE-V protocol could be influenced by
potential observational biases, especially in the phases
following the analysis—the results presented here tend to
show the robustness of the analysis phase.

Acknowledgment

This research was financed by the Swiss National Science
Foundation (SNSF) 100012-105817/1.

References

Proceedings of the 16th International Conference of the International
1–14.
Kumho implications of observer effects in forensic science: hidden
1–56.
of us: the effect of contextual top-down processing on matching finger-


The current status of forensic science laboratory accreditation in Europe

Ekrem Malkoc a,*, Wim Neuteboom b

a Gendarmerie Forensics Department, Jandarma Kriminal Daire Baskanligi (JKDB), 06835 Beytepe-Ankara, Turkey
b The Netherlands Forensic Institute (NFI), Laan Van Ypenburg 6, NL-2497 GB The Hague, The Netherlands

Received 9 June 2006; accepted 14 June 2006
Available online 28 July 2006

Abstract

Forensic science is gaining some solid ground in the area of effective crime prevention, especially in the areas where more sophisticated use of available technology is prevalent. All it takes is high-level cooperation among nations that can help them deal with criminality that adopts a cross-border nature more and more. It is apparent that cooperation will not be enough on its own and this development will require a network of qualified forensic laboratories spread over Europe.

It is argued in this paper that forensic science laboratories play an important role in the fight against crime. Another, complimentary argument is that forensic science laboratories need to be better involved in the fight against crime. For this to be achieved, a good level of cooperation should be established and maintained. It is also noted that harmonization is required for such cooperation and seeking accreditation according to an internationally acceptable standard, such as ISO/IEC 17025, will eventually bring harmonization as an end result.

Because, ISO/IEC 17025 as an international standard, has been a tool that helps forensic science laboratories in the current trend towards accreditation that can be observed not only in Europe, but also in the rest of the world of forensic science. In the introduction part, ISO/IEC 17025 states that “the acceptance of testing and calibration results between countries should be facilitated if laboratories comply with this international standard and if they obtain accreditation from bodies which have entered into mutual recognition agreements with equivalent bodies in other countries using this international standard.” Furthermore, it is emphasized that the use of this international standard will assist in the harmonization of standards and procedures.

The background of forensic science cooperation in Europe will be explained by using an existing European forensic science network, i.e. ENFSI, in order to understand the current status of forensic science in Europe better. The Council of Europe and the European Union approaches to forensic science will also be discussed by looking at the legal instruments and documents published by these two European organizations. Data collected from 52 European forensic science laboratories will be examined and findings will be evaluated from a quality assurance and accreditation point of view. The need for harmonization and accreditation in forensic science will be emphasized. The steps that should be taken at the European level for increasing and strengthening the role of European forensic science laboratories in the fight against crime will be given as recommendations in the conclusion.

© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Forensic science; Accreditation; ISO/IEC 17025

1. Introduction

In 2004–2005, in the United Kingdom, 571 rape cases, 165 murder/manslaughter cases, 4465 theft-from-a-vehicle cases, and 16,059 burglary cases were linked to one or more suspects, thanks to forensic DNA technology.1 It is quite clear that approximately 48 rape, 14 murder, 372 theft-from-a-vehicle, and 1338 burglary cases were solved each month. But, are the measurements and calibrations used in such technology scientifically valid? Are laboratories using this technology competent and able to generate technically valid results? What criteria should be applied for European forensic science laboratories so as to ascertain their competence in an area of freedom, security, and justice?

The above figures taken from the United Kingdom’s National DNA database annual report clearly show that forensic DNA technology plays an active role in the fight
and their respective countries are not given. The European Network of Forensic Science Institutes (ENFSI) has amended its framework for membership in 2004 and stated that, in order to be an eligible applicant, “the forensic institute shall have achieved an accreditation or documented progress in quality assurance with a clear plan to obtain accreditation in the near future.”

Then, in 2005, ENFSI made the existing requirements in the “policy document on the standards of accreditation” more severe, by stating that “all member laboratories should have achieved or should be taking steps towards ISO/IEC 17025 compliant accreditation for their laboratory testing activities.”

According to a recent survey by the ENFSI standing committee on Quality and Competence (QCC), there are 13 accredited ENFSI laboratories. The most common standard is the ISO/IEC 17025 issued by International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC). The fact that there are still 40 non-accredited ENFSI laboratories shows that some sort of a harmonized path towards accreditation remains to be established.

It is important to emphasize here that this paper deals with the accreditation of forensic science laboratories only, not of forensic scientists/experts. Although being a frequently debated and interesting topic among forensic science communities across Europe, it might be another field of research unto itself. Rather, this study will try and look into how far are the European forensic science laboratories into the issue of accreditation. At the end, some recommendations, at European level, will be formulated to help increase the number of accredited European laboratories and, therefore, strengthen the role of forensic science laboratories in the fight against crime.

2. Materials and methods

In order to better understand the current situation of forensic science laboratories in Europe, answers given to a questionnaire that was sent out to members of ENFSI in 2004 were used. It is important to note here that the questionnaire was originally designed to find out about the current status of ENFSI member laboratories by the ENFSI Secretariat. Therefore, we did not have any control on the design of this questionnaire. Rather, the data compiled from the replies to the questionnaire were used for analysis.

Out of 53 ENFSI member forensic science laboratories, replies from 52 were used in this study since one answer had not been received until data analysis began. For the issues of privacy and secrecy, names of the laboratories and their respective countries are not given.

The questionnaire asked about several issues including: status in criminal justice system, availability of a quality assurance system, status/expected date of accreditation, name of the accreditation standard, number of staff and number of cases examined in 2004. The data based on the issues listed here became the variables that were chosen for this study.

Statistical package for the Social Sciences (SPSS) software was used to manage and summarize these data. All the data from the questionnaires were compiled in Microsoft Excel data sheets first. Then, the variables we chose to examine in depth and compare were transferred to SPSS for further analysis based on a coding scheme.

3. Results and discussion

After having formulated – as the formal ENFSI policy – the importance of being accredited, it might be insightful to verify what position the ENFSI member laboratories take. In Table 1, we see that more than 94% of the ENFSI member laboratories already have a system of quality assurance or are trying to develop one. This is a good signal especially if we consider that, following its constitution and the policy document, ENFSI is coaxing its members into achieving accreditation. The trend here looks promising and shows how willing and motivated the laboratories happen to be towards using a quality assurance system. Assuming that these laboratories will go one step ahead and seek accreditation, it should also be taken into consideration that one of the first necessities when one laboratory decides to take the path towards accreditation is not a big budget or lots of personnel, but motivation and determination. Nevertheless, before taking the motivation level we derived from Table 2 at face value, we need to take a look at the intended and estimated date of achieving accreditation. According to Table 2, it seems that 34.6% of the 52 laboratories intend to achieve accreditation by 2009. 32.7% of the laboratories are already accredited; another 32.7% did not provide any schedule for their accreditation plans. This fact can be regarded as some form of hesitation, which is inconsistent with our finding of motivation based on Table 1. Since the questionnaire did not ask why the laboratories do not have a plan for accreditation, but develop quality assurance system all the same, we cannot tell for sure if this fact is attributed to just a simple hesitation or some other factor like budget constraints or staff/management resistance. A more detailed questionnaire probing such concerns might bring some more insight on this issue.

Table 1

<table>
<thead>
<tr>
<th>Quality assurance system availability</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>26</td>
<td>50.0</td>
</tr>
<tr>
<td>No</td>
<td>3</td>
<td>5.8</td>
</tr>
<tr>
<td>Being developed</td>
<td>23</td>
<td>44.2</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>100.0</td>
</tr>
</tbody>
</table>

2 Ref. [9].
3 Ref. [10].
4 Ref. [19].
5 Ref. [17].
Before coming to an ambiguity in our findings in Table 3, it might be better to address some of the differences between certifications of quality assurance and accreditation. Actually, these certifications are two different, but, to some extent related concepts. For example, in the introduction part, ISO/IEC 17025 states that “testing and calibration laboratories that comply with this international standard will therefore also operate in accordance with ISO 9001 or ISO 9002”. It also states that “certification against ISO 9001 and ISO 9002 does not, of itself, demonstrate the competence of the laboratory to produce technically valid data and results”. The purpose of ISO 9000 series, on the other hand, is to describe fundamentals of quality management for organizations seeking advantage through the implementation of a quality management system. Therefore, it will not be wrong to deduce that quality assurance and accreditation are two phenomena in which having a quality management system becomes possible without accreditation whereas accreditation becomes impossible without having a quality management system. This, also, could easily be attributed to “lack of awareness syndrome”.

Another striking fact in this table is that five laboratories are referring to national regulation and it is not clear if these regulations have anything to do with either ISO/IEC 17025 or ISO 9001/9002. This table shows that the way accreditation is perceived by ENFSI member laboratories also shows some diversity. Therefore, the terms and meanings attributed to quality assurance and accreditation should be more clearly defined at the member laboratories level. Unless there becomes a common understanding of accreditation, its purpose, and how it should be pursued, all efforts would obviously be in vain. The misperception observed in this table might also be considered as a sign of the level of importance that has been attached to the idea of accreditation. It is clear that not all member laboratories of ENFSI are taking this issue equally seriously.

Since we are interested in the accreditation status of not only all European forensic laboratories, but particularly in those from the EU member states as well, we also needed to check the status of accreditation of the forensic science laboratories in the EU-countries. It is quite clear from Fig. 1 that more than 60% of the forensic science laboratories in the EU-countries are not accredited. Additionally, none of the ENFSI laboratories from EU candidate countries is accredited. If we take a look at the

Table 2
Accreditation time frame

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Cumulative percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of accredited labs until 2004</td>
<td>17</td>
<td>32.7</td>
<td>32.7</td>
</tr>
<tr>
<td>Expected increase in 2005</td>
<td>2</td>
<td>3.8</td>
<td>36.5</td>
</tr>
<tr>
<td>Expected increase in 2006</td>
<td>6</td>
<td>11.5</td>
<td>48.1</td>
</tr>
<tr>
<td>Expected increase in 2007</td>
<td>6</td>
<td>11.5</td>
<td>59.6</td>
</tr>
<tr>
<td>Expected increase in 2008</td>
<td>3</td>
<td>5.8</td>
<td>65.4</td>
</tr>
<tr>
<td>Expected increase in 2009</td>
<td>1</td>
<td>1.9</td>
<td>67.3</td>
</tr>
<tr>
<td>No response</td>
<td>17</td>
<td>32.7</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

Table 3
Accreditation standards implemented by the laboratories

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISO/IEC 17025</td>
<td>9</td>
<td>17.3</td>
</tr>
<tr>
<td>ISO 9001:2000</td>
<td>1</td>
<td>1.9</td>
</tr>
<tr>
<td>ISO 9002</td>
<td>1</td>
<td>1.9</td>
</tr>
<tr>
<td>National regulation</td>
<td>5</td>
<td>9.6</td>
</tr>
<tr>
<td>Standard not specified</td>
<td>1</td>
<td>1.9</td>
</tr>
<tr>
<td>Not accredited</td>
<td>35</td>
<td>67.3</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Fig. 1. Accreditation status according to EU membership.

10 Ref. [16].
12 For the purpose of this study, the EU candidate countries at the time of our data analysis were taken into consideration. These are Bulgaria, Croatia, Romania, and Turkey, all of which have at least one ENFSI member forensic laboratory.
situation of quality system availability in the EU and candidate countries in Fig. 2 we see that a system of quality assurance is at the stage of development in approximately 40% and does not exist in approximately 5% of the EU forensic laboratories. For candidate countries’ forensic science laboratories, approximately 12% do not have a quality assurance system, whereas more than 80% are at the stage of development. Findings in these two figures might be regarded as a sign showing that the issue of forensic science laboratory accreditation in Europe needs a concerted European level approach.

According to the results of data analysis, in Table 3 we can see that approximately 67% of European forensic science laboratories examined in this study are not accredited. This number is approximately 65% in the case of EU forensic science laboratories in Fig. 1. It is clear from this picture that accreditation is a big issue that should be tackled to facilitate cooperation in police and criminal matters and help involve forensic science better in the fight against crime. It is not only the high percentage of non-accredited laboratories that shows the need for a well-coordinated action at European level for the accreditation of forensic science laboratories. Additionally, we see in Fig. 2 that almost 10% of candidate countries’ forensic science laboratories do not have a quality assurance system whereas more than 80% are at the stage of development. This is another sign that points to the need of a concerted and supportive action at the EU level.

If we look at the situation of already accredited laboratories in Table 3, we see that the standard that some of them base their accreditation is not actually an accreditation standard. This shows that the idea of quality assurance and accreditation brings about confusion among some forensic science laborato-

tories in Europe. That is why, it is foreseen at the end of this study that some form of an initiative at the European level might be a good start in the field of forensic science due to its gradually rising importance in cooperation in criminal and police matters.

4. Conclusion and recommendations

In an era where rapid response is a very effective proactive weapon against all forms of crime, the role of forensic science cannot be denied. The increase in the acts of terrorism require an all-fronts fight in which every possible way of enforcing the rule of law and protecting security interests of all citizens should be re-evaluated. Individualization of traces of evidence left at the scenes of crime, especially those based on fingerprints and DNA, can be made very quickly and transferred to any nation on earth in approximately the same amount of time, thanks to the state of the art technology science provides. All it takes is high-level cooperation among nations, especially in the areas of police and criminal matters.

Harmonization is needed at many different levels where cooperation is meant to be achieved. For the scope of this paper, we need to consider harmonization in forensic science; which might produce results from different forensic science laboratories in Europe to be valid for use in other European countries. For example, all countries must use the same DNA-markers. If not, it will be impossible to exchange DNA-profiles between countries, no matter how fast or reliable the technology gets.

It is obvious that cooperation in criminal matters should be extended down to the level of forensic science institutes, which are an integral part of the specialized law enforcement services. Title VI, Article 30 of EU Treaty of Amsterdam gives these services a special importance in operational cooperation between the competent authorities in relation to the prevention, detection and investigation of criminal offences. A well-grounded co-operation does indeed require a good harmonization process along with accreditation, without which the outcome of operational cooperation between the competent authorities, including the police, customs and other specialized law enforcement services, becomes nearly useless. This, naturally, places a heavy burden on the budgets of the forensic science institutes. Manpower will be needed in order to perform all this work in addition to the normal casework, quality assurance, etc. Investment in new or improved equipment might often be necessary [1].

At the same time, we must keep in mind that more objective and harmonized methods can reduce the operational costs of a forensic institute. The benefits can even be greater if we take into account the fact that automated and computerized investigation techniques can accelerate investigations and help guide the police in the direction of the most promising aspects of investigation.

However, when we look at the activity regarding the role of forensic science at the level of policy making at European level,
what we see is nothing more than council decisions\textsuperscript{13}/ resolutions\textsuperscript{14,15} recommendations\textsuperscript{16,17,18} or common action plans\textsuperscript{19,20} which do not have any legal binding effect on EU member states. In order for European forensic science laboratories to assume their role as an effective, preventive and proactive tool against transborder crime, more concrete measures are needed. One such proposed measure that did not seem to find its target is the establishment of a forensic science center under the auspices of Europol. According to Kube and Dahlenburg:

“...Europol should assume an important initiating, supportive and coordinating function in the field of forensic science as well. It is also essential to ensure that effective existing international networks, such as ENFSI in particular, are taken into account in the strategic planning of this Europol function...”\textsuperscript{21}

The justification of establishing a forensic science center inside Europol, an already existing EU Justice and Home Affairs agency, was clearly defined and structured in this article that was published in 2000. Since no such center for forensic science has been set up inside Europol or any other EU agency so far, we are free to speculate that forensic science does not get the attention it deserves in the eyes of EU policy makers.

As also pointed out in the article mentioned above, with its current status and possibilities, ENFSI is in no place to enforce accreditation through the ISO/IEC 17025 standard document. In order for ENFSI “to ensure that the quality of development and delivery of forensic science throughout Europe is at the forefront of the world” as described in its Constitution, it is inevitable to help its member institutes get accredited in the quickest and healthiest means possible. The first priority could be given to the member institutes from EU member and candidate states. For this feat to be realized, it is clear that ENFSI needs some high-level assistance. Henceforth, it is considered at the end of this research study that the following initiatives may prove useful for providing better involvement for European forensic science laboratories in the fight against crime:

i. In the short term, the commission shall prepare a “proposal for a Council Decision on financing and supporting the EU forensic science laboratories” accreditation activities in connection with the operational cooperation between specialized law enforcement services”\textsuperscript{22}

ii. In the middle term, a feasibility study shall be initiated by the EU Commission in collaboration with ENFSI, in order to see if a “European Forensic Science Agency” needs to be set up for the purpose of producing a higher European standard for forensic science and its effective uses in crime prevention. If it is concluded that there is actually a need for such an agency, the commission shall initiate the secondary legislation procedure that will result in a “European Forensic Science Agency” being established.

Probably not as an alternative to the actions above, but as an additional measure that can help speed up quality assurance and accreditation efforts of forensic science laboratories, compliance specifically with ISO/IEC 17025 could be made compulsory, for example in the Council Decision mentioned in “i” above. This would not be the sole example of such compliance since the Commission Recommendation of 1 March 2005 concerning a coordinated programme for the official control of foodstuffs for 2005 does the same on a different topic. According to this recommendation, it is provided that the official laboratories in member states, as referred to in Article 7 of Directive 89/397/EEC, are to comply with the criteria set out in European Standard EN 45000 series, which is, now, replaced by ISO/IEC 17025:2005.\textsuperscript{23} This example can also be applied to forensic science laboratories within the EU.

The diverse structure among European forensic science laboratories, as illustrated in our findings based on empirical data gathered from 52 different European forensic science laboratories, might pose as an obstacle before establishing a harmonized action plan for accreditation and quality assurance. Besides, the data analyzed in this study cover 52 European forensic laboratories, which are ENFSI members. These laboratories are among the larger and better organized of all European forensic laboratories, the correct number of which is not known. Therefore, we can deduce that the issue of quality assurance and accreditation will be a bigger problem for these other European forensic science laboratories.

This diversity mentioned in the previous paragraph also comes forth as another reason why accreditation of forensic science laboratories should not only be supported, but also coordinated in a rightful manner. This obstacle can also be used as an opportunity like in the example of the European Union with its well known motto: “unity in diversity” in which unity is promoted while diversity has been preserved. The same philosophy can be applied in the case of forensic science laboratories, which operate on many different traditions, just like Europe itself. However, it is apparent that low-level, non-binding legislation that can be observed in CoE and EU approach to the role of forensic science in the fight against crime so far has not been fruitful. For a better involvement of forensic science in the fight against crime, the quality of European forensic science should achieve higher standards and, for this to happen more effectively, binding legislation at the European level is definitely and urgently needed.

\textsuperscript{13} Ref. [15].
\textsuperscript{14} Ref. [12].
\textsuperscript{15} Ref. [13].
\textsuperscript{16} Ref. [11].
\textsuperscript{17} Ref. [6].
\textsuperscript{18} Ref. [7].
\textsuperscript{19} Ref. [4].
\textsuperscript{20} Ref. [5].
\textsuperscript{21} Ref. [18].
\textsuperscript{22} Ref. [3].
\textsuperscript{23} Ref. [2].
5. For further information

Please feel free to contact e.malkoc@superonline.com or Wim.Neuteboom@nfi.minjus.nl for any further requests of information. You can access the pdf version of a poster based on this study, presented during the 4th European Academy of Forensic Science Congress, Helsinki, Finland, on 13–16 June 2006, as well as the SPSS coding scheme online at http://ekremalkoc.tripod.com.

Acknowledgements

This study could not have been done without the help and support of ENFSI, especially at the laborious stage of data collection and analysis. Furthermore, NFI and JKDB, behind the scene institutes that made this study possible, deserve our gratitude as well.

References


[3] Common action in the field of police cooperation, Title VI, Article 30 of the Treaty of Amsterdam.


Application of Computational Fluid Dynamics modelling in the process of forensic fire investigation: Problems and solutions

O. Delémont *, J.-C. Martin

Institut de Police Scientifique, Ecole des Sciences Criminelles, Bâtiment Batochine,
University of Lausanne, CH-1015 Lausanne, Switzerland

Received 8 June 2006; accepted 14 June 2006
Available online 28 July 2006

Abstract

Fire modelling has been gaining more and more interest into the community of forensic fire investigation. Despite an attractiveness that is partially justified, the application of fire models in that field of investigation rises some difficulties. Therefore, the understanding of the basic principles of the two main categories of fire models, the knowledge of their effective potential and their limitations are crucial for a valid and reliable application in forensic science.

The present article gives an overview of the principle and basics that characterise the two kinds of fire models: zone models and field models. Whereas the first ones are developed on the basis of mathematical relation from empirical observations, such as stratification of fluid zones, and give a relatively broad view of mass and energy exchanges in an enclosure, the latter are based on fundamentals of fluid mechanics and represent the application of Computational Fluid Dynamics (CFD) to fire scenarii. Consequently, the data that are obtained from these two categories of fire models differ in nature, quality and quantity.

First used in a fire safety perspective, fire models are not easily applied to assess parts of forensic fire investigation. A suggestion is proposed for the role of fire modelling in this domain of competence: a new tool for the evaluation of alternative hypotheses of origin and cause by considering the dynamic development of the fire. An example of a real case where such an approach was followed is explained and the evaluation of the obtained results comparing to traces revealed during the on-site investigation is enlightened.

© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Forensic science; Fire modelling; Field models; Hypothesis evaluation; Results interpretation

1. Introduction

Fire investigation, as it is traditionally undertaken, consists in the search for traces and their scientific evaluation in order to determine two “static” parameters that explain the start of the fire. First, the on-site investigation has to delimitate the origin of the fire, representing the location of the initial seat of the fire. Then a systematic and scientific methodology must be applied in order to establish the cause of the fire. The latter assumes an evaluation of the combustible that was first involved and, most of all, a determination of the nature of the heat source that allowed the start of the fire.

In some cases, the investigation needs to study the dynamic aspect of a fire: its development, its propagation from the initial flaming to its final extent, considering both the spatial and time evolutions of the phenomenon. These aspects are generally apprehended by collecting and analyzing testimonies and by reconstructing a chronology of the event. Despite providing a precious help in the evaluation of the possible scenarii of cause of the fire, such an approach is not sufficient to assess a conjecture of fire propagation arising from the fire investigation or to discriminate between two or more alternative hypotheses of cause of the fire.

The evolution of computer science in the past decades has brought a decisive improvement in calculation capacities allowing the set up of systems able to tackle very complex integrations of mathematical relations [1]. Among them, one specific development has had a great impact on the study of the dynamic evolution of fires: fire modelling. First developed for fire safety studies, so-called zone models and field models have been sporadically used in a more “forensic” way since the beginning of the 90s, with their first major applications being
most probably for the explanation of the 62 Watts Street Fire in New York (March 28, 1994) [2] and the King’s Cross Station Fire in London (November 18, 1987) [3–8].

In the first part of this paper, a brief description of the two different kinds of fire models will be presented. This will be followed by a proposition of use of fire modelling in a forensic fire investigation process, illustrated by an example of application. Finally, a reflection of the role and the future of fire modelling in forensic sciences will be proposed as a concluding part.

2. Computer fire models

Two large categories of fire modelling systems based on computer calculations are to be distinguished: the zone models and the field models. They differ by the general approach on which they are based.

2.1. Zone models

As clearly mentioned by Walton, ‘‘Zone fire models are computer programs designed to predict the conditions resulting from a fire in an enclosure’’ [9]. The zone models are built on the assumption that corresponds to an empirical recurrent observation of fire happening in enclosures. As the fire grows, the hot products of combustion are entrained in the plume of the fire in the upper part of the enclosure under the influence of convection flows. In a macroscopic point of view, a zone containing the hot gases and chemical species produced during the combustion invades the upper part of the environment as a cooler zone, free of smoke and soot, remains in the lower part; a stratification occurs in the enclosure where these two regions can be clearly distinguished. The zone models consist of mathematical relations applied to this empirical statement. The volume of the enclosure on which calculations are conducted, called computational environment, is divided in large homogenous regions. Four different zones are generally distinguished although more can be specified: the fire, the plume, the hot upper zone and the cold lower zone. The fire is considered as a volumetric source of heat and species; it is therefore described by heat and mass release rates. The heat and species produced are entrained by the plume, which acts like a pump, into the hot upper zone, whereas the cold lower zone remains close in property and composition to fresh air. Into each zone, the temperature and the concentration of the different chemical species are considered spatially uniform but vary with time.

After discretisation, the system of equations that constitutes the core of a zone model is solved by computing resources; the time evolution of the different variables in a given zone and the flows of energy and mass between two adjacent zones are then calculated [10–12].

Due to their relative simplicity and the empirical assumption of stratification on which they rely, zone models suffer from consequent limitations. Many of them were discussed in the past [13,14], enlightening the risk of obtaining non-valid or even non-realistic results for certain geometrical configurations. But among all limitations, the major one consists most probably in the requirement for an a priori knowledge of the structure of the flows that will take place in a given environment [15]. Since zone models rely on very constraining empirical assumptions, it must be assessed that the geometry and the configuration submitted to the model fulfil these assumptions. In a very pragmatic way, this limitation involves that a scientifically valid application of zone models cannot be carried out independently of scale or real size experiments.

2.2. Field models

The only characteristic that field models share in common with zone models relies in the use of computer systems for the solving of a set of equations that describe time and space evolution of scalars and flows in a given volume. For all other aspects, field models differ significantly. Instead of being separated in large homogenous zones, the environment is divided in a large amount of small cells which size must be related to the combination of geometric complexity, time discretisation and the magnitude of flows passing through opponent faces of the cell. The complete arrangement of cells composes the computational environment and is usually described as mesh or grid. On the basis of this mesh, the fluid flows induced by a fire are calculated by applying the basic equations of fluid dynamics, generally known as Navier-Stokes equations. These equations establish the balance of flows for each cell of the environment: the flows entering and leaving the cell, the residence time of flows in the cell, the rate of creation and destruction of flows in the cell are considered. The application of such equations describing fluid dynamics into models that solved them by means of computer is generally known as Computational Fluid Dynamics (CFD) [16–18]; field models correspond to the application of CFD codes to the simulation of fire scenarii. The principle of ascertaining a balance by means of computational calculations seems quite similar to the approach of zone models. But since the space discretisation is much finer, the number of distinguished volumes composing the complete environment is far greater and thus the data for heat and mass transfers are much more informative. With such an approach, the transfer of mass and energy that takes place by convection can be quite precisely calculated. The integration of other heat transfer mechanisms involved in a fire, conduction and most of all radiation, requires the set up of a parallel application of sub-models that interact with the core model of fluid dynamics [19,20]. Moreover, other sub-models must be integrated to take account of other phenomena that cannot be correctly represented with the core model, such as turbulence, which is often simulated with the k–ε model [19,21,22].

The application of CFD requires the delimitation of the environment on which calculations are performed. The computational environment is thus not a infinite medium and therefore boundary conditions must be applied to its limits as well as to each location where a geometrical discontinuity appears, such as walls or apertures. These boundary conditions refer to routines that guide the model in the calculations process at specific locations of the computational environment.
Moreover, the user has to set initial conditions to the CFD model in terms of geometrical measurements, physical and chemical properties of materials and fluids and computational parameters. Although frequently considered only as a preliminary work, the set up of boundary and initial conditions as well as the careful construction of a mesh take part of the process of modelling and often constitute the issues that have the most crucial influence on the results.

Despite their ability to give more useful and reliable data about fire development than zone models, especially in complex environment, field models, and more generally CFD models, suffer from some limitations that unavoidably introduce some errors and uncertainties in their results [15]. These limitations are multiple. They arise for example from the treatment of turbulence and radiation, from simplifications in the representation of the real environment, from the treatment of combustion and the propagation of flame and, most of all, from the choice of boundary and initial conditions that do not perfectly reflect reality. The different kinds of limitations must be clearly identified and understood in order to provide an accurate and scientific exploitation of the results arising from field models.

3. Suggested use of fire computer models in forensic sciences

Fire computer models, both zone models and field models, have been used first for fire safety purposes. Risk of flame and smoke propagation have been studied with such tools by performing simulation on the base of “worst-case scenarios”: conditions corresponding to a powerful and fully developed fire are applied as initial conditions and time evolution of the products of combustion is assessed. On the basis of the results obtained for such an application, improvement in architecture of enclosure or in selection of materials can be achieved and egress strategies can be evaluated. The level of safety towards fire has without any doubt improved with the use of fire models. Such tools are in fact really accurate for an application in the field of fire safety. Simulations undertaken do not intend to reproduce the reality of a specific event but to evaluate the response of a given environment to several scenarios that are plausible or possible. The limitations inherent to fire computer models have a relatively low influence since the results are not considered for their accuracy or correlation to reality but for their ability to reproduce a general tendency. The level of certainty and accuracy expected for fire safety application is not very high and therefore fire models represent a very useful tool, allowing to save time and money.

The application of fire computer models in forensic sciences should logically follow a comparable philosophy. Simulations are able to study a general tendency of development of fire in a specific event but they cannot reproduce the reality of a specific fire that happened. It is therefore erroneous to expect from the application of fire computer models to explain the dynamic evolution of a fire. It is commonly believed that since these models are based on fundamental equations of physics and that calculations are made by computers, the results they provide must correspond to reality. In fact, as previously emphasized, the data given by zone and field models rely strongly on the initial and boundary conditions chosen by the user. It is unquestionable that the calculations performed by these models are correct – this has been widely proved – but the way these calculations are set up, which relies on the user, can lead to tremendous errors.

Being aware of the inability of fire computer models to give results that correspond to reality independently from the data provided by the user, forensic fire investigators must develop strategies of application of fire computer modelling.

After studying different alternatives, one of the role for fire modelling which seems to be the most appropriate is to act as an experiment for the testing of hypotheses. The scheme of reasoning that is generally adopted for the forensic investigation of fires allows to raise a certain amount of hypotheses concerning the cause of the fire. The collection and exploitation of traces on the scene and in the laboratory, the confrontation with testimonies, the chronology of events as well as the consideration of scientific reasoning relying on chemistry and thermodynamics tend first to exclude some of these hypotheses. The remaining ones are then evaluated and balanced on the basis of the evidence collected. At this stage, different approaches can be considered, from previous case based inference to the use of Bayesian networks [23, 24]. At the end, if possible, an hypothesis is enlightened as the most plausible one, generally without formally excluding the other alternatives. In some situations, two or more alternative hypotheses remain comparatively probable or one hypothesis is dominant but the doubt associated to another one is too large for the justice to take a decision. In such cases, the application of fire modelling can be profitable by setting up simulations with initial conditions corresponding to the different hypotheses and confronting their results with the evidence revealed on the fire scene. In other situations, where a weird development of the fire occurred, fire modelling can help in explaining the factors that contributed to the specific propagation of the fire. Such an application of fire modelling was undertaken during the investigations following the King’s Cross Station Fire with the CFD code FLOW3D running on a CRAY-2 supercomputer [4].

The following example emphasizes the suggested role of fire computer models as an evaluation tool for alternative hypotheses.

4. Example of application: fire in a historical edifice

A vast fire damaged an historical edifice composed of several building fitted in a very complex architecture. The fire started in a religious structure and then propagated in different ways to the adjacent constructions. At the time of the fire, nobody was present in the edifice where the fire started but renovation work was in progress during the past few hours. Due to this work, the interior of the religious edifice was separated in four large volumes by horizontal wood panels that were erected at three different heights and that extended on the whole surface of the structure. Passage from one volume to another was possible through centred openings cut into the panels.
After a first round of judicial investigations, led by four different experts, two alternative hypotheses of cause of the fire remained. The first one considered that the seat of the fire was above the first horizontal wood panel and its cause was associated to a malfunction of electrical origin. The alternative hypothesis was arson, perpetrated on a gallery, approximately at mid-height under the first wood panel. The judicial authorities designated a second pool of experts that agreed on the first hypothesis, without being able to exclude formally the latter one.

An extensive study was then undertaken at the Institut de Police Scientifique of the University of Lausanne in order to assess the possibility to evaluate in a valid way the likelihood of the two alternative hypotheses by (1) setting up fire modelling simulations with a control of validity and significance of the results and (2) confronting the data obtained with traces revealed during the on-site investigations. The purpose of the research was limited to the first stage of the fire in the interior of the religious edifice before its extension to the others damaged buildings. Modelling were performed with the general CFD package CFX-4, version 4.3. This corresponds to the most advanced version of the CFX structured mesh solver that was preferred to unstructured solver (such as CFX-5 for instance). The mesh generator was the embedded software CFX-BUILD and calculations were performed on a DEC calculation server running CFX-4 solver. Post-processing was carried out with the software FieldView (Intelligent Light) and data were treated with Microsoft Excel, Kaleidagraph and the statistical software S-PLUS.

4.1. Set up and validation of the geometrical and computational configurations

The first step of the research was dedicated to a series of small studies on the mesh generation and the adjustment of the different sub-models and their parameters. The aim of these studies was to set up a geometrical and a computational configuration that were accurate for the subsequent full modelling of the two alternative hypotheses. This series of studies was very time-consuming and the effort required in terms of amount of calculations was huge but such an approach was absolutely necessary in order to fully understand the influence of the different parameters and thus be able to appreciate the real significance of the following results. Among others, the principal objects of the studies were the simplification of the complex real geometry, the definition of the fire as a heat or a heat and mass source in the model, the treatment of radiation and the determination of time steps for calculations in accordance to the geometry and the convective flows in the different parts of the environment. A complete description of all the studies performed is exposed in Ref. [25]. The most effort-consuming task was the generation of an accurate mesh. The creation of a mesh, in itself, is not very difficult but it must be verified that the resulting grid does not affect the modelling. As the mesh fixes the spatial discretisation, it has a strong influence on the calculation process and can thus influence the results.
obtained. The Fig. 1 shows time evolution of hot air (without heat source) entering the religious edifice through six peripheral apertures included in the middle of its height. Three different mesh structures are presented; calculations were achieved with the same parameters. The results between the three configurations indicate that a refinement of mesh changes the general evolution of flows. It is therefore necessary to set up an entire mesh that is sufficiently fine not to affect the results of calculations. But in the same time, a finer grid involves a greater amount of calculations and a reduction of the time increment between two consecutive time solving steps; as a consequence, the increase of calculations can rapidly become too large to be completed, even by powerful computer resources. The major difficulty of mesh generation is thus to find an accurate balance between sufficient mesh refinement to lead to trustable results (i.e., not affected by the grid refinement) and limitation of the associated increase in time of calculation to allow a reasonable application of the modelling.

4.2. Fire modelling of the two hypotheses of fire origin and cause

After having assessed several different types of geometry structures and associated meshes, two different grids were generated, one for each fire seat corresponding to the two hypotheses of fire origin: the geometry with the fire situated above the first wood panel consisted of a multi-blocks structured mesh of 1’002’080 cells and the geometry with the fire localised on a gallery under the first wood panel was a multi-blocks structure mesh of 872’800 cells. These two geometries are illustrated in Fig. 2.

The different sub-models applied either as part of the solver (buoyancy for instance) or as external modules run in conjunction with it (radiation for example), the boundary and initial conditions, as well as the computational parameters were adjusted in reference to the preliminaries studies and compiled in a command file submitted with the geometry file as inputs to the solver. The main information concerning the

<table>
<thead>
<tr>
<th>Physical submodels</th>
<th>Boundary conditions</th>
<th>Initial conditions and computational parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$–$\varepsilon$ model for turbulence, with the application of the constants proposed by Novozhilov [15] and modifications in the cross terms of diffusion</td>
<td>Wall treatment: adiabatic with zero normal velocity</td>
<td>Ambient temperature: 288.00 K</td>
</tr>
<tr>
<td>Modelisation of buoyancy</td>
<td>Atmospheric pressure conditions at the surface of the six peripheral apertures</td>
<td>Initial temperature: 288.00 K</td>
</tr>
<tr>
<td>Treatment of flows as weakly compressible</td>
<td>Fire: constant volumetric source of heat (1.79 MW, combustion of 100 g/s of wood [26]). Dimensions [m]: $1.5 \times 1.5 \times 1.5$</td>
<td>Zero fluid velocity (air), atmospheric pressure</td>
</tr>
<tr>
<td>Radiation treatment: Discrete Transfer Model</td>
<td>Time step: 0.5 s</td>
<td>Total simulation time: 30 min (3600 steps)</td>
</tr>
</tbody>
</table>
Fig. 3. Temperature evolution (coloured zones) in the plane x–y cutting the religious edifice in its centre. From left to right and top to bottom: temperature profile after 120, 360, 600, 840, 1080, 1320, 1560 and 1800 s for the fire seat above the first wood panel. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).
Fig. 4. Temperature evolution (coloured zones) in the plane $x$-$y$ cutting the religious edifice in its centre. From left to right and top to bottom: temperature profile after 120, 360, 600, 840, 1080, 1320, 1560 and 1800 s for the fire seat on a gallery in the volume under the first wood panel. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).
applied parameters are listed in Table 1; further details can be found in Ref. [25].

The time step selected was the same for both simulations: increment of time of 0.5 s was applied between two consecutive time sets of calculations. The total amount of development that was calculated was for 30 min; with time increment of 0.5 s, the total amount of time steps for each simulation was 3600. For the modelling of the first hypothesis (fire seated above the first wood panel) 1.224 millions of seconds (approximately 14.2 days) of processing time (CPU time) were necessary for the completion of the calculation. The second simulation (fire on a gallery under the first wood panel) was completed in 1.079 millions of seconds (approximately 12.5 days) of CPU time.

A three-dimensional analysis of the results for the simulation run on the basis of the hypothesis of a fire starting on the above surface of the first wood panel [hypothesis 1] was undertaken, considering flows distribution and scalar evolution, especially temperature. This analysis indicates that the majority of the heat released by the fire propagates into the volume delimited by the first and the second wood panel. A portion of this energy dissipates in the external surrounding through the peripheral apertures present in this volume. After approximately 20 min, a quasi steady-state is attained in the intermediate volume: the amounts of hot and cold air do not change significantly anymore. The transfer of energy continues into the upper volumes of the edifice and the temperature above the second and the third wood panel keeps on rising for the whole duration of the simulation. The time evolution of spatial distribution of temperature on a plane in the centre of the construction for this simulation is presented in Fig. 3.

A comparable analysis was carried out with the data obtained from the modelling of the alternative hypothesis specifying the location of the fire on a gallery in the lower volume of the religious edifice. For this configuration, the greater part of heat released remains in the lower section of the construction. The temperature in this volume increases rapidly, especially under the first wood panel where hot gases accumulates due to convection flows. Under the conditions applied to this simulation, the temperature of auto-ignition of wood (568 K for white pine [27]) is achieved at the lower surface of the panel in less than 15 min. A minor portion of the heat produced by the fire propagates into the volume delimited by the first and the second panels where another hot layer grows in its upper part. The temperature reached in this zone is nevertheless far under the values reached in the lower volume. During the second half of the simulation, the heat keeps on propagating and penetrates into the upper volumes of the edifice, above the second and the third panels. The time evolution of spatial distribution of temperature on a plane in the centre of the construction for this simulation is presented in Fig. 4.

Despite the fact that the simulations performed do not take into account the extension of the flames, the results of the modelling for the two hypotheses provided very useful information. Taking into consideration a fire location above the first panel [hypothesis 1], the damages in the edifice will essentially occur in the upper part of the construction, above the first wood panel. Of course once the fire is fully developed, this panel will burn down as well, allowing the extension of the fire in the lower volume. Considering the other hypothesis (fire on a gallery under the first panel), carbonisations will necessarily be intense in the lower volume. After burning the wood panel, the fire would propagate extensively to the upper sections of the edifice.

4.3. Comparison with traces revealed during the investigation

The on-site investigation of this event revealed that only superficial traces of calcinations were present on the combustible elements in the volume under the first wood panel, mainly on decorating wood objects. The information obtained from the interpretation of the data arising from the modelling was then confronted to these traces recorded during on-site investigations and the potential significance of this procedure was assessed by considering the limitations inherent to fire modelling. Despite uncertainty and imprecision that unavoidably characterise the outcomes of the simulation, it was possible to determine that the second hypothesis was definitely not consistent with the traces left by the fire on the scene: a fire starting below the first wood panel would manifestly involve intense carbonisations on the combustibles present in this volume but such damages were not observed. Thus, for the specific case and the specific problematic involved, fire modelling proved to be a very useful tool for the assessment of alternative hypotheses.

5. Discussion and conclusion

The previous example provides an illustration of the suggested role that computer modelling must play in a general forensic investigation of a fire. It constitutes a convenient instrument to distinguish between alternative hypotheses on the basis of fire development. The general process is the following: (1) taking one hypothesis as guide for the determination of boundary and initial conditions, a simulation is run. (2) Then the process is repeated with the next hypothesis. (3) The general tendency of fire development proposed by the different simulations can finally be compared to evidence left by the fire on the scene.

However, it is crucial to bear in mind that fire models do not provide a reconstruction of the reality of an event. They are simplified representation of reality that will always suffer from a certain lack of accuracy and precision. Under the condition that the user is fully aware of this status and has an extensive knowledge of the principles of the models, their functioning, their limitations and the significance attributed to their results, fire modelling becomes a very powerful tool. Another consequence of these restrictions is that results provided by fire models can be helpful to study a problematic only where the different hypotheses submitted involve dissimilarities in fire development, either in time or space. If the fire propagation schemes proposed by the two hypotheses are too similar, fire modelling will not be very useful due to the degree of uncertainty that characterise its results.
But the strength of the fire models can as well turn out to be their threat: fire models will provide data even when inaccurately used or poorly configured. One of the major difficulties is to discriminate afterwards between reliable information and invalid statement inferred from the results of fire modelling. In this perspective, the current tendency to enlarge the number of applications of fire modelling is not entirely positive. Of course, the fact that easier, cheaper and very attractive software of modelling becomes available even to non-specialists constitutes a progress, especially in the field of fire safety. But considering the important risk of misuse of fire models or of misinterpretation of their results, forensic scientists should be very careful before taking up this tool. The development of guidelines advising procedures for fire modelling application and the elaboration of strategies for the validation and the integration of fire modelling results in the process of forensic fire investigation are major tasks that must be considered in a close future. The implication of forensic scientists in fire modelling should grow in parallel with an intensification of communication and collaborations with engineers specialised in the field of Computational Fluid Dynamics.

Acknowledgements

The authors wish to acknowledge Dr. Jacques-André Hertig and Dr. Pascal Goulpié, Ecole Polytechnique Fédérale de Lausanne (EPFL), for their expertise and their support with this study.

References

Applications of ENF criterion in forensic audio, video, computer and telecommunication analysis

Catalin Grigoras*

Forensic Examiner, 18 Magheru Blvd., Ap 43, Bucharest S1, Romania

Received 6 June 2006; accepted 14 June 2006

Available online 1 August 2006

Abstract

This article reports on the electric network frequency criterion as a means of assessing the integrity of digital audio/video evidence and forensic IT and telecommunication analysis. A brief description is given to different ENF types and phenomena that determine ENF variations. In most situations, to reach a non-authenticity opinion, the visual inspection of spectrograms and comparison with an ENF database are enough. A more detailed investigation, in the time domain, requires short time windows measurements and analyses. The stability of the ENF over geographical distances has been established by comparison of synchronized recordings made at different locations on the same network. Real cases are presented, in which the ENF criterion was used to investigate audio and video files created with secret surveillance systems, a digitized audio/video recording and a TV broadcasted reportage. By applying the ENF Criterion in forensic audio/video analysis, one can determine whether and where a digital recording has been edited, establish whether it was made at the time claimed, and identify the time and date of the registering operation.

Keywords: Electrical network frequency criterion; Forensic audio; Forensic video; Audio authentication; Video authentication; Digital audio/video recordings

1. Introduction

Recent years have seen a significant increase in the number of attempts to use digital audio/video evidence in every sector of litigation and criminal justice. And, whilst there are well-established and effective procedures for evaluating the authenticity of analogue recordings (see [1–5] on the analysis of switching transients and magnetic patterns), it is well-known that digital recordings may be edited or falsified in ways that defy the attentions of the forensic examiner (see [6]). This article proposes a new method of analysing digital recordings. It is a method that will reveal recordings as falsifications or forgeries in cases where they might hitherto have ‘passed muster’ as authentic records of events (Fig. 1).

2. The electrical network frequency

When digital equipment (computer, mini-disc recorder, audio recorder, video camera, secret surveillance system, etc.) is used to record a conversation or event, it captures not only the intended speech but also the 50/60 Hz electrical network frequency (ENF).

This is so if the recording device is mains-powered and used in the absence of an ideal voltage regulator. However, the equipment may also capture the ENF if it is battery-powered and used in a built-up environment or in proximity to other mains-powered equipment or transmission cables.

In a real electric network the ENF is not fixed at precisely 50 Hz. Over time, frequency variations inevitably occur, principally because of differences between produced and consumed power. In Fig. 2 a real ENF analysis is presented. Although the waveform and FFT can be used to examine the signal’s periodicity in short time windows, more information on variation over time can be derived from the 3D-spectrogram (so-called ‘waterfall’) and 2D-spectrogram. Here one can observe a continuously updated history of peak values over time, i.e. a running representation of temporal variation in the ENF.

Having analysed ENF variation over a period of six months, it is clear that there is no predictable patterning; the fluctuation over time around 50 Hz is purely random. Exceptions from this can be considered some events during
the evening (about 11:00 p.m.), night (about 1:00 a.m. GMT) and morning (5:00, 6:00 and 7:00 a.m. GMT), generated by maintenance operations or network components switch on/off. Even so, their shapes or values cannot be predicted. In these situations only their presence can be predicted, during working days, as they are illustrated in Fig. 3.

The power grid can have an instantaneous frequency variation of up to \( \pm 0.6 \) Hz for up to a 10 s interval. When measured over an 8 h or even 24 h interval, the frequency tolerance is much tighter.

At any given moment, the mathematical form of the ENF can be represented as:

\[
f = [50 \pm \Delta f] \text{ Hz}
\]

where \( \Delta f \) represents the deviation between the instantaneous frequency and the set point frequency.

Owing to electromagnetic wave propagation, the entire electric network formed by the interconnected systems, including all sources (power stations) and loads (shops, offices, houses, factories, etc.), will have the same frequency, irrespective of point on the network from which it is measured. From a forensic viewpoint, it is the continuing variation in frequency over time together with its stability over distances that provide a powerful resource for carrying out authenticity examinations of digital recordings.

According to the Union for Power Production and Transport Coordination (see [7]) recommendations, under normal conditions the ENF is maintained within strict limits and the set point frequency is around 50 Hz. There are three types of operating conditions, based on \( \Delta f \).

If \( \Delta f \leq 50 \text{ mHz} \), the conditions are considered to be normal. If \( \Delta f \) is between 50 and 150 mHz operating conditions are deemed to be impaired but with no major risk. If \( \Delta f > 150 \text{ mHz} \), operating conditions are deemed to be severely impaired and there are significant risks of malfunction of the electric network.

From an engineering instrumentation point of view, all electric and electronic systems contain noise sources (e.g. white noise, pink noise, etc.). In order to analyse the ENF, only high

---

Fig. 1. The waveform (a), spectrum (b), waterfall (c) and spectrogram (d) of an ideal 50 Hz signal.

Fig. 2. The waveform (a), spectrum (b), waterfall (c) and spectrogram (d) of a real 50 Hz signal.
quality sampling devices with very low levels of noise and distortion are to be recommended. In practice, a sound card with a signal to noise ratio less than $-94$ dB and total harmonic distortion of less than 0.003% will allow the examiner to capture ENF without significant errors and distortions.

The signal may be captured via a band-pass filter with the cut-off frequencies below 49 Hz and above 51 Hz. This filter eliminates the other, non-relevant signals but leaves the ENF component unmodified and intact for database storing and analysis (see [8–10]).

3. Methods to extract and analyse ENF

Three methods to extract ENF are reported:

1. ‘time/frequency domain’ spectrograms method consists on computing spectrograms (see Figs. 2d, 3 and 4a) and visually compare questioned versus database ENF;
2. ‘frequency domain’ method means to compute FFT over short time windows, extract the maximum magnitude value around 50 Hz (see Fig. 4b) and compare questioned samples versus database ENF;
3. ‘time domain’ analysis consists on zero-crosses measurement (see Fig. 4c) and questioned versus database ENF comparison.

The spectrograms method is the fastest one, is easy to be implemented, reveals the ENF components number and is useful especially for questioned versus database ENF date and time verification. It is also suitable for longer than 10–15 min recordings. In most cases the presence of minimum two ENF (see Fig. 5) components indicates a non-authentic or non-duplicate evidence recording. This method must be used before the other two in order to find out the ENF components number.

Fig. 3. A 1 week analysis, from 0:00 to 8:00 a.m. GMT, vertical span 49.9–50.1 Hz.

Fig. 4. Three methods to extract ENF: spectrogram (a), frequency domain (b) and time domain zero-crosses (c).

Fig. 5. Non-authentic recording with two ENF components.

Fig. 6. Non-authentic digital recording with two ENF components ENF1 and ENF2 (a), and their separation by filtering: ENF 1 (b) and ENF 2 (c).
The frequency domain method can be applied to one ENF component recordings for questioned versus database ENF verification or for database searching in order to identify the questioned ENF recording date and time. This method can also be used on more than one ENF component recordings in order to separate the ENF components and detect their chronological order of entry.

The time domain method can be used on one ENF component recordings only. It is the deepest ENF extraction and analysis method. For bigger than 44 kHz sampling frequencies and by implementing zero-crosses interpolation can become sampling frequency independent. A particular attention must be paid to averaging operations over different length windows. For questioned versus database analysis based on this method, because of comparison samples synchronization delays, non-identical ENF shapes can be obtained. To minimize errors DC removal and 50 Hz band pass filter are requested (Fig. 6).

In certain circumstances, a time domain analysis can offer more useful detail than a frequency domain analysis. By applying a band-pass filter with 49–51 Hz cut frequencies, without down sampling the signal, one can separate the ENF waveform from the rest of the recording. Computing all the zero-crosses of the remaining ENF signal and measuring the time differences between consecutive zero values, the results when plotted can show micro-variations of the ENF. Figs. 7–10 represent detailed waveform analyses.

Also, by amplifying low-level signals containing ENF it is possible to analyse the original quantification levels of the signal. In some situations, if there is a link between two different audio segments coming from different digital recordings made on different equipment, it is possible to see the differences in quantification levels as a mark of manipulation of the original recording.

Although micro-variations are observable across Figs. 8–10, the greatest degree of correlation will be observed between Figs. 9 and 10.

Figs. 9 and 10 represent contemporaneous recordings made at different points on the same network; Fig. 8 represents a recording made at a different point in time.

4. ENF database

In May 2000 the author began to build an ENF database which involved continuous, long-term round-the-clock recording of ENF in Romania. The digital storage space necessary to save the ENF reference recordings as WAV PCM files at 120 Hz sampling rate 16 bit resolution is set out in Tables 1 and 2.

In order to save the reference material over one month less than 630 MB are needed. This can be saved on a single CD and for 1 year the storage therefore requires only 12 CDs. Another

Fig. 7. Three consecutive zero-crosses corresponding to two different semi-periods D6,7 and D7,8.

Fig. 8. ENF time domain analysis: waveform, histogram and ZCR variation.

Fig. 9. ENF time domain analysis: waveform, histogram and ZCR variation.

Fig. 10. ENF time domain analysis: waveform, histogram and ZCR variation.
solution would be to store the information on double-sided DVDs, in which case a single disc would be enough to store a year’s data.

The task of establishing the date and time when a digital audio recording was created potentially involves the examiner in many hours of searching. In view of this, the author would propose an automatic system, based on a signal correlation method, along the lines presented in Fig. 11. Properly configured, such a system would search and find the most closely correlated ENF reference signal for any evidential recording within a time period circumscribed by the examiner.

5. ENF sources

ENF shapes from three different sources are presented in Fig. 12: European electric network on 22 May 2006 between 20:00 and 20:30 GMT (Fig. 12a), an uninterruptible power supply (Fig. 12b) and 12 V dc to 220 V ac/50 Hz inverter (Fig. 12c).

The ENF component is mostly found on fixed, buildings secret surveillance audio/video recordings made with mains-powered equipments having no ideal voltage regulators.

The UPS component can be found on digital audio/video recordings coming from both fixed and mobile secret surveillance units.

The Inverter component is generally coming from mobile secret surveillance units installed on cars, vans, etc.

The spectrograms method application can reveal not only the ENF components number but their type too. It is possible to find out that an digital evidence has been previously created through an inverter powered device and than ENF powered equipment edited.
6. Experiments: intra-network ENF stability

First set of tests were undertaken between summer 1998 and autumn 2000 at different locations on the Romanian electric network (in the author’s laboratory, elsewhere in the same building, elsewhere in the same town and in different other cities). The results showed that for all consumers, at any moment in time, the ENF value is the same.

An electronic device based on a transformer with a maximum peak-to-peak voltage of 100 mV was connected to a soundcard on a personal computer. In order to record and analyse the audio signal the software package DCLive Forensics was used. The spatial relationships between the three cities sampled are represented in Fig. 13.

The distances between the cities are more than 300 km. Despite this, the results obtained showed a high degree of correlation. An example is provided in Fig. 14.

The overall shapes of the spectrograms (CJ, IS and B) are very similar. Even though a more detailed analysis of the ENF in the time and spectral domains did reveal small differences, these were considered to have been caused by noise extraneous to the ENF and by differences between the sampling boards used in the three sampling computers (see [11]).

Second set of tests were undertaken between summer 2000 and spring 2006 at different locations on the European electric network, including: Wien, Austria; Paris, France; Nice, France; Munchen, Germany; Saarbrucken, Germany; Wiesbaden, Germany; Athens, Greece; Budapest, Hungary; Milan, Italy; Rome, Italy; Hague, The Netherlands; Krakow, Poland; Madrid, Spain. The overall shapes of the spectrograms are also very similar and the spatial relationships between Bucharest and different cities are represented in Figs. 15–18. These graphically results are from inter-laboratories “ENF criterion” validation tests conducted by author.

The distances between the cities are more than 300 km. Despite this, the results obtained showed a high degree of correlation. An example is provided in Fig. 14.

The overall shapes of the spectrograms (CJ, IS and B) are very similar. Even though a more detailed analysis of the ENF in the time and spectral domains did reveal small differences, these were considered to have been caused by noise extraneous to the ENF and by differences between the sampling boards used in the three sampling computers (see [11]).

Second set of tests were undertaken between summer 2000 and spring 2006 at different locations on the European electric network, including: Wien, Austria; Paris, France; Nice, France; Munchen, Germany; Saarbrucken, Germany; Wiesbaden, Germany; Athens, Greece; Budapest, Hungary; Milan, Italy; Rome, Italy; Hague, The Netherlands; Krakow, Poland; Madrid, Spain. The overall shapes of the spectrograms are also very similar and the spatial relationships between Bucharest and different cities are represented in Figs. 15–18. These graphically results are from inter-laboratories “ENF criterion” validation tests conducted by author.

![Fig. 12. Three different ENF sources: electric network (a), uninterruptible power supply (b) and 12 V dc to 220 V ac/50 Hz inverter (c).](image)

![Fig. 13. Spatial relationships of the three cities sampled (B: Bucharest; CJ: Cluj-Napoca; IS: Iasi).](image)

![Fig. 14. Spectrographic comparison of temporally co-extensive samples from three Romanian cities: Cluj-Napoca (CJ), Iasi (IS) and Bucharest (B).](image)

![Fig. 15. Spectrographic comparison of temporally co-extensive samples from Bucharest and Madrid, on 23 January 2006, 12:29–13:35 GMT.](image)
7. Casework examples

7.1. Casework example 1

In a forensic case the author was asked to analyse a digital recording in order to establish its authenticity and originality. The recording consisted of an audio file on a CD in WAV PCM format sampled at 8 KHz with 16-bit depth. It comprised more than 3 h of conversation between two people (speakers A and B) and had been made using a recording system installed in a room in an office building. The system allowed the audio signal to be picked up via a small electret microphone. The signal was then amplified and stored on the hard disk of a PC notebook.

Speaker A, who had previously been recorded by speaker B with this same system, claimed that the recording was altered and also that the claimed date of recording was incorrect; he indicated another date. Both possible dates were given to the author to investigate.

Checks were made to establish that the mains electricity supply had been continuous over the period of time concerned, and the ENF signal present in the recording was compared with the ENF database to determine whether there was any match.
The audio file was analyzed using DCLive Forensics software. It was down sampled to 120 Hz and band-pass filtered (LF = 49 Hz, HF = 51 Hz, slope = 24 dB/octave, Butterworth). A spectrogram was computed (4096 points FFT) and inspected using a vertical zoom around 50 Hz. The resulting 2D-spectrograms are presented in Fig. 19.

The conclusion is that the evidence had not been recorded on the date claimed by speaker B and there were significant matches with the date and time claimed by speaker A. When the ENF continuity was checked against the reference material more than ten major discontinuities were revealed.

Speaker A was asked to listen to the recording and to indicate in an accompanying transcript where he considered there to be any irregularities. He indicated more than twenty deletions of words and expressions. The discontinuities in ENF that I had discovered corresponded to the points he indicated.

7.2. Casework example 2

In a recent case the author was asked to analyse a digital audio/video recording in order to establish its authenticity. The recording consisted of an MPG file on a CD in MPEG2 format, 720 × 576 resolution, 25 fps, audio 48 kHz, 16 bits, stereo. It comprised about 19 min audio and video recording made in a village called Ciurea – Iasi, located in the North-Eastern part of Romania.

It was known that the original recording has been made on a Video 8 Sony camera and then digitized into the questioned audio/video file. It was claimed that no deletions have been done before, during and after digitization, and the evidence represents a copy of an original recording. No analog video tape neither camera or other details have been available for a complex examination. Fig. 20 shows two frames of the questioned video recording.
The expert presented his point of view about forensic video analysis and the limits encountered in similar situations and the final question was established “to identify the date of digitization process”.

The expert asked for the date and time of the original recording and the digitization process. Both possible dates were given to the author to investigate.

The methodology to examine the evidence was the same as in previous case. After extracting the audio channels the analyses were carried out using DC LIVE/Forensics v6.14 running on a PC Windows XP platform. For the audio recording, phase analysis indicated the left and right channels are identical (Fig. 21) and the left channel was used for further examination.

The resulting 2D-spectrograms are presented in Fig. 14 and there were significant matches with the date and time claimed by the side who made the recording. When the questioned ENF continuity was checked against the reference material (from the author’s database) no discontinuities were revealed and the expert opinion was that the questioned recording had been digitized on the date claimed by side (Fig. 22).

7.3. Casework example 3

In this case the evidence consisted on a digital audio/video recording of a TV broadcasted reportage from 16 February 2003. The question was to establish if it is an authentic video tape broadcasting.

In this case zero-crosses analysis on time domain is not effective and the frequency domain spectrogram was used (see Fig. 23). Three ENF components can be observed: ENF1 and ENF2 very close to 50 Hz and ENF3 with about 1 Hz elongation around 50 Hz.

Knowing some details of the case, it is the expert opinion that:

- ENF1 was introduced by the digital video grabbing system during the acquisition and recording operation;
- ENF2 is probably coming from the broadcasting company;
- ENF3 is coming from the questioned tape containing the video transmitted signal;
- the original video tape has been recorded by using a 220 V ac/50 Hz generator, not a public electric network.
In all cases the forensic examiner respected G8 Proposed Principles for the Procedures Relating to Digital Evidence [12] and Video and Audio Systems Principles, Practices and Procedures [13].

8. Applicability of the ENF Criterion to analogue recordings

Currently, the ENF criterion cannot readily be applied to analogue tape recordings owing to the presence of both enduring long-term and fluctuating short-term (‘wow and flutter’) variations in tape transport speed. While the problem of relatively stable long-term variations might be overcome by an overall upshift or downshift in the frequency settings of the spectrograms for analysis, wow and flutter variations present a far greater problem and one which can result in severe and so-far irremediable distortions of the original ENF signal.

9. Conclusion

The electric network frequency criterion is a tool that can be used to analyse digital audio and video recordings, checking their integrity, verifying or determining the point in time when they were created and appreciate the area they are coming from (50 or 60 Hz). This is accomplished by using a reference frequency database recorded in a laboratory or obtained from the electrical network company. The ENFC can also reveal the mains-powered recording type: ENF, UPS, inverter, etc.

Particular attention should be given to the quality of the reference that is critical in allowing an accurate comparison, as well as to the entire methodology of analysis and evaluation of the results. It is recommended that the ENF criterion be used in conjunction with other methods, like classical forensic audio and video techniques, and/or impulse response of a room (see [9]) and IT investigations, that must be improved upon and incorporated into a standardized analysis protocol for forensic audio/video examiners.

The ENF Criterion represents a significant breakthrough in the quest for forensic techniques that will defeat the efforts of the audio/video forger operating in the digital domain.

Acknowledgements

The author would like to thank Jos Bouten, Netherlands Forensic Institute, Ministry of Justice, and Peter French, J. P. French Associates, UK for helpful suggestions and comments on this article. I recognize the assistance I have received from Professor Brandusa Pantelimon, Electrical Department, Politehnica University of Bucharest, Romania.

I would also like to thank Andrea Paoloni, Fondazione Ugo Bordoni, Rome, Italy; Joaquin Gonzalez Rodriguez and Francisco Javier Simon del Monte, Universidad Autonoma de Madrid, Spain; Jose Juan Molina, Guardia Civil, Madrid, Spain; Carlos Delgado, Comisaria General de Policía Cientifica, Madrid, Spain; John D. Makris, Hellenic Police, Forensic Sciences Division, Athens, Greece; Mateusz Kajstura, Institute of Forensic Research, Cracow, Poland, for entering the inter-laboratory validation tests of ENF Criterion.

The discussions with Alan Cooper, Metropolitan Police Service, Department of Information, London, UK; Tom Owen, The New York Institute of Forensic Audio, USA and Curtis Crowe, Tracer Technologies, Inc., USA are also appreciated. This manuscript could not have been completed without their help.

References

Superimposition and projective transformation of 3D object

Hana Eliášová, Pavel Krsek

Institute of Criminalistics Prague, P.O. Box 62/KUP, Strojnicka 27, 170 89 Praha 7, Czech Republic
Centre for Applied Cybernetics, CTU, Faculty of Electrical Engineering, Technicka 2, 166 27 Praha 6, Czech Republic

Received 9 June 2006; accepted 14 June 2006
Available online 1 August 2006

Abstract

Superimposition is an efficient method for evaluation of coincidence between a skull and a photo portrait. The principle of superimposition method lies in the projection of the skull into the face image. During the projection of an object with a perspective camera, the mapping of a three-dimensional object into a two-dimensional image takes place. The acquired images of the same object are more or less distorted due to various photographic conditions, due to extrinsic and intrinsic parameters of the camera. The distortions have important influence onto reliability of human identification by the superimposition method.

Mathematically we can describe most of the distortions. On the basis of the description the divergences could be simulated and in some cases eliminated by geometric transformation of the compared images. We are presenting a mathematical model of the standard projective camera and the mathematical description of distortions which are important for the superimposition process. The results show the distortions and the elimination of the distortions by means of the projection model.

Keywords: Superimposition; Projective transformation; Homography

1. Introduction

Identification of human skeletal remains pertains to priority expertises in the forensic practice. The identification process includes also the craniofacial analysis and the superimposition. The skull-photo superimposition is an efficient method for the final confirmation of the unknown skull identification because facial photographs are usually available.

During the superimposition process a three-dimensional object (skull) is projected via an optical system into a two-dimensional image (portrait). To compare a portrait photograph with an unknown skull we need to carry out the projection of the skull into the image in the same manner as the original projection of the portrait was. Only the image of the skull and the image of portrait, captured in the approximately same geometric conditions, can be compared. In practice, to ensure the same conditions is very difficult. The prerequisite of the relatively accurate superimposition are both anthropological and the photographic craniofacial analyses. The photograph of the missing person should be of a good quality, a high resolution, a sufficient contrast, suitable positions of head (en face, profile), minimal image distortions.

Morphological and metric divergences of images from the real object (the skull or the face) can occur during taking pictures due to various photographic conditions, due to extrinsic and intrinsic parameters of the camera (Fig. 1). These facts explain why the superimposition setting of the skull into the portrait sometimes shows deviations, even though other, stronger evidence may confirm the identity of the person.

The geometric variability of the different images of the same object can be described by mathematical model of the camera. The pinhole camera model is a simple model of the standard camera without optical distortions. In our case the optical distortions are negligible. Therefore, the pinhole camera model is sufficient for the description of the image composition.

The following objectives were determined:

• to evaluate the character and the degree of the distortion of skull images during photographing under various, but exactly defined conditions;
• to explain the reasons for image deformation changes.
2. Material and methods

The distortions of the skull images were studied. The skull was fixed on a craniophore, oriented in the en face position in a standard way according to Frankfurter horizontal; the height of its position was exactly determined. The camera Canon EOS D1 MARK II (17–400 mm/4.0 L lens, taken at 1/8 s / f 9.0, zoom setting range 17–400 mm; chromatic temperature TUNGSTEN) was placed on a tripod with a special position settable panoramic head. An angle of view (0 ± 40°) and the distance from the lens 280–4000 mm were determined. We capture two sets of images. The first set consists of images taken with the camera placed in the stable distance from the object. The images differ by the angle of view in range 0 ± 40° (Fig. 2a). We take images with different focal length of a lens. The second set represents the images which were captured by the moving camera with the range from 4000 to 280 mm (Fig. 2b). The different focal lengths of the lens were applied.

Morphometrical method was selected for the effective analysis of the relation between the skull images distortions and the conditions of photographing. The morphometry transforms qualitative information into quantitative information and thus converts subjective evaluation to more objective evaluation. Numeric expression of shape facilitates orientation in a numerous sets and enables to compare the shape changes in dependence on a series of factors.

Landmark methods (thin-plate spline and Procrustes superimposition) elucidated the deformation on the basis of defined homologous points configuration changes on the images of cranium. Using the tps Dig software [1].

![Fig. 1. Two images of the same skull (various photographic conditions).](image1)

![Fig. 2. Scheme of the image distorsion owing to various (a) angles of view and (b) lens distances.](image2)
points were digitalized in exactly defined (marked) places in the same order on individual skull images. The choice and the position of the homologous points are shown in Fig. 3. The ellipses represent an error of a manual localization of the points.

Procrustes superimposition method [2] was used to rotation, translation and scaling of the coordinates of homologous points. Procrustes superimposition (GLS algorithm) of the homologous points of the individual set images of the skull provided us a visual imagery of a variation range of the image deformation changes. The size and directions of the point shifts responsible for the particular distortions are usually visually represented using vectors starting from centroid and expressing for each point deviations in the position of landmarks of individual images from the reference configuration. The degree of shape changes between the chosen images and the reference image was quantified by computing Procrustes distances. The evaluation of shape changes in the form of deformation grids was done using the package of tps SW and PAST SW, v.1.09, [3].

3. Results

- With the first set, the morphometric method have confirmed the considerable distortion in the images farthest from the reference object. That means, the higher is the angle of the lens axis in relation to the line passing through the reference object centroid perpendicularly to the frontal plane, the higher is the image distortion. The most striking changes, which were expressed by vectors, are visible in the skull contour area, especially of the set photographed with $f$ 17 mm. The area of central splanchnocranium showed lesser changes in all cases, minimal changes were found in the area delimited by the lower part of apertura piriformis nasi and in the area surrounding infraorbital points. Differences between the reference images and the lateral images (especially in corner positions) were quantified by Procrustes distances. The values of Procrustes distances decrease with the increasing focal length. The images located accurately above and below the reference image show smaller changes with the increasing focal length.
- With changing the focal length, the size of the object is changed. The evident distortion of the images results partially from the depth dimension of the object, from the lower lens distances (the higher angle of view) and from the various defects of the camera optical system. In case of the images acquired from the same distance but under the different lens axis angle, these differences are caused by the fact that the camera is not rotating around the centre of projection but around the centre point given by mounting the camera.
- Changing the distance from the lens is documented by the second image set. Procrustes superimposition of the images photographed from the different distances represents the striking deformation changes on lateral neurocranium contours and in the frontal sector of os frontalis. The central part of the face, the lower part of apertura piriformis nasi and the area surrounding infraorbital points are affected by the distortions very little. The differences of the images related to the reference skull image photographed from the distance of 4000 mm, were quantified by Procrustes distances. An exponential dependence of Procrustes distances on the distance from the lens was proved. When the lens distance is bigger than 2500 mm, the change of relative size of the images is already smaller than 3% (Fig. 4).

3.1. Camera model

Most of the observed distortions of the skull images could be explained by a standard mathematical camera model. The pinhole camera is the simplest hardware realization of the camera which is defined by the relatively simple model. The mathematical model of the pinhole camera is based on a central projection [4]. The projection is called perspective projection. The perspective projection is a useful approximation for all cameras with standard.

Fig. 5 shows a basic set-up of the pinhole camera model. The camera consists of an image plane and a focal point which is above the image plane. The distance between the image plane and the focal point is called the focal length $f$. The point $X$ is projected through the focal point onto the image plane.

![Fig. 4. Exponential dependence of procrustes distances on the distance from the lens.](image-url)
World Euclidean coordinate system is a basic coordinate system. The camera coordinate system is placed in the focal point with an axis $Z_C$ in an optical axis. This coordinate system is the fundamental coordinate system for the mathematic description. The position of the camera in the world Euclidean coordinate system is defined by translation vector $t = [x_t, y_t, z_t]^T$ and rotation matrix $R$ ($3 \times 3$). The relationship between the coordinates of scene point $X$ in the world coordinate system $X_W = [x_W, y_W, z_W]$ and the coordinates of the same point in the camera coordinate system $X_C = [x_C, y_C, z_C]$ is done by the system of linear equations.

$$X_C = R(X_W - t)$$

$$\begin{bmatrix} x_C \\ y_C \\ z_C \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x_W \\ y_W \\ z_W \end{bmatrix} - \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix},$$

where $a, b, c$ define the size of sensor elements, parameter $b$ describes an angle between the axes of the sensor and $[u_0, v_0]$ are coordinates of the principle point.

The translation vector $t$ and the rotation matrix $R$ determine the position of the camera. The equation $\hat{u} = a\hat{u}$ is valid in the homogenous coordinate system. Therefore, the value $z_C$ could be eliminated and the equation of the projective transformation is rewritten into a compact form.

$$\hat{u} = [K]_c [KR][X_W - t]$$

A matrix $K$ is an internal matrix of the camera. The matrix keeps internal parameters of the camera. A parameter $f$ is the focal distance, parameters $a, c, d$ define the size of sensor elements, parameter $b$ describes an angle between the axes of the sensor and $[u_0, v_0]$ are coordinates of the principle point. The translation vector $t$ and the rotation matrix $R$ determine the position of the camera. The equation $\hat{u} = a\hat{u}$ is valid in the homogenous coordinate system. Therefore, the value $z_C$ could be eliminated and the equation of the projective transformation is rewritten into a compact form.

$$\hat{u} = [K]_c [KR][X_W - t]$$

### 3.2. Restricted perspective projection

Let us focus on the relation between two images of the same 3D object from the different view point of a camera set-up. The relation is defined by two different perspective projections of
the same 3D object.

$$\mathbf{u}_1 = K_1 \mathbf{R}_1 (X - \mathbf{t}) \quad \text{and} \quad \mathbf{u}_2 = K_2 \mathbf{R}_2 (X - \mathbf{t})$$

In general, the process of perspective projection could not be inverted. But in special cases of the camera movement, the relation between images is done by linear transformation in the homogenous coordinate system called homography.

$$\mathbf{u}_1 = \mathbf{H} \mathbf{u}_2$$

Homography is defined by a homography matrix $\mathbf{H}_{(3 \times 3)}$. The differences between the images are just the distortions of the images. One of the images can be artificially generated on the basis of the second image and of the known homography. Homography could be computed when more than two corresponding points on the images are known.

Homography represents the relation between the images obtained by the camera, which rotates around the focal point. The rotation of the camera on the tripod is a physical approximation of the rotation. The same case is the different position of the object on the image. An example of the homography projection is shown in Fig. 6. In other case of homography is the relation between the two images which differ by internal parameters of the camera. The focal length (zoom) is only one internal parameter that is changed in practical cases.

The image changes arising during the focal point movement in the world coordinate system can be described by full projective transformation of the 3D object. Therefore, the image changes depend on the shape of the object. There is no direct transformation between the images without known object shape.

**Fig. 6.** Images of a skull distorted by homography. The distortion was realized by different positions of the same object on the image plane (camera rotation around focal point). Grids demonstrate the deformation of planar object. Images were captured by camera with lens $f = 17$ mm, object distance 1050 mm. Camera angle difference is around 40°.
3.3. Experiments

We described the pinhole camera model and showed that images captured by the camera which rotates around the focal point are distorted by the homography projection. For the experiment we used sets of images captured by the camera with the focal length 17 mm. Forty-six homologous points (Fig. 3) were used for the estimation of homography as the relation between the reference image and the border images. The corresponding points define an overdetermined system of the linear equations. The solution of the system is the desired homography.

The set consists of nine images—one central image and eight border images where the view angle is maximal. The distance between camera and the object was 1050 mm. The captured images differ by the rotation of the camera. The rotation angle is around ±40°. For each border image we computed homography, that projects central image onto border images. Then we applied homography on the central image. We visualized an error of the projection as the difference between the border image and the transformed central image. Differential images which were obtained as the difference between the border image and the central image transformed by homography are illustrated in Fig. 7. Table 1 presents the differences between the border images and the central image transformed by the homography projections into the coordinate system of the border images.

Table 1 shows errors of Procrustes superimposition. The method of Procrustes superimposition minimizes the distance between a pair of intrinsic points by translation, rotation and scaling. The homography is more complex.

Table 1 documents that homography is suitable transformation which correctly describes the rotation of the camera around

Fig. 7. Differential images which were obtained as difference between border image and central image transformed by estimated homography. Images were captured by camera with lens \(f = 17\) mm, object distance 1050 mm. Camera angle difference is around 40°.
the focal point. Greater errors appear with lenses of a small focal length and in short object distance from the lens. Here, the whole field of view is utilized, and thus the observed error can be explained by the optical system error that does not fully correspond to the perspective camera model. However, the errors could be partially explained by the rotation of the camera around a tripod joint which does not exactly correspond to the focal point.

4. Discussion

Within the forensic portrait identification, namely superimposition, we encounter cases where a false negative output of the expertise is caused by the unsuitable photograph both of the missing person and of the skull. This can be caused by the distortion arising during the transformation of 3D object to 2D image.

The morphometric methods can be selected to analyse the distortion of the images. Whereas the classical morphologic evaluation of the shape variability using a verbal description or a categorization or a mere metric evaluation are often subjective, the morphometric image analysis represents a tool having no limitation regarding a distinguishing level and a subject of investigation. It provides more information about the image based on the changes of the configuration of the homologous points. Numerous studies proved the validity of that geometric morphometry during the last 20 years [6–8].

We used the morphometric, landmark methods which were selected so that they could effectively capture the monitored complex as well as the partial image changes. The digitalization of landmarks and semilandmarks was performed by tps Dig SW [1]. With the help of the configuration of landmarks, Bookstein [9] presented the variability of head deformities related to Apert syndrome. According to authors [10,2], it is suitable to implement Procrustes analysis, GPA algorithm, to evaluate the landmark configuration changes in compared objects.

Lele and Richtsmeier [11] used a morphometric analysis to evaluate a craniofacial dysmorphology (craniosynostosis—Crouzon sy and Apert sy). The analysis of the skull shape deformation in moles (Talpa) was performed on the basis of landmark configuration [12]. Consensual configurations were compared using the relative deformations method (TPSRW). A visualization was performed in the form of the position change vectors of the homologous points and the deformation grids. The deformations connected with the rotation of vomer, the length of the preorbital region, and the changing height of linea supraoccipitales were evaluated using the deformation grids including eight crucial landmarks on neurocrania [13]. Singh et al. [14] presented results of a study in which, based on a combination of the morphometric and the cephalometric analysis, they evaluated the morphology of the middle face in a set of individuals with a diagnosis of retrognathia. Procrustes analysis of semi-landmarks was successfully applied in case of the comparison of the frontal bone profiles of Homo sapiens, Homo heidelbergensis, and Homo neanderthalensis [15]. Penin et al. [16] determined ontogenetic, shape-related changes on human and chimpanzee skulls using Procrustes superimposition utilizing statistical methods (PCA, multivariate regression and discriminant function).

The distance of the photographed skull in case of superimposition should not be lower than 2000 mm, as stated by Miyasaka et al. [17]. Šimková and Šimek [18] emphasize in their work that the lens distance lower than 2000 mm causes the object image deforming changes. They recommend the minimum lens distance to be approximately 15 times the focal length. Results of our study confirmed these conclusions. On the contrary Lan and Cai [19] state that the optimal lens distance from the skull for superimposition is 1000 mm.

Using photographic lens with the shorter focal length, the significant distortions of the images of the face or of the skull arise [18]. The dependence of the focal length on the image deformation is presented by Fetter et al. [20]. The outputs of our research confirmed the mentioned conclusions.

Our study is differs from the studies of cited authors by mathematical description of the substance of the image distortions. This mathematical description is our original contribution to the methods used in the superimposition process.

5. Conclusion

The application of the morphometric analysis to the skull images as well as the theoretical and experimental analyses of mathematic rules of the standard camera model have resulted into the following conclusions:

- The morphometric analysis proved the distortion variability of the photographic skull images depending on various photographic conditions (lens distance, focal length, the distance from the medial plane, etc.). We have determined the character and the trend of the distortions of the photographic images.
- Changes in the skull images which can be simulated using homography (with a computer calculation) were documented.
- The image distortions were explained by means of the theoretical rules of the mathematical model of the classical photographic camera.

Table 1

<table>
<thead>
<tr>
<th>Image</th>
<th>Procrustes superimposition</th>
<th>Homography transformation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6.25</td>
<td>3.76</td>
</tr>
<tr>
<td></td>
<td>3.14</td>
<td>2.36</td>
</tr>
</tbody>
</table>
Acknowledgement

The work was supported by The Czech Ministry of Education under project 1M0567.

References

Multimodal biometrics for identity documents (MBioID)

Damien Dessimoz a,*, Jonas Richiardi b, Christophe Champod a, Andrzej Drygajlo b

a Institut de Police Scientifique, École des Sciences Criminelles, Université de Lausanne, Switzerland
b Speech Processing and Biometrics Group, Signal Processing Institute, École Polytechnique Fédérale de Lausanne, Switzerland

Received 9 June 2006; accepted 14 June 2006
Available online 4 August 2006

Abstract

The MBioID initiative has been set up to address the following germane question: What and how biometric technologies could be deployed in identity documents in the foreseeable future? This research effort proposes to look at current and future practices and systems of establishing and using biometric identity documents (IDs) and evaluate their effectiveness in large-scale developments.

The first objective of the MBioID project is to present a review document establishing the current state-of-the-art related to the use of multimodal biometrics in an IDs application. This research report gives the main definitions, properties and the framework of use related to biometrics, an overview of the main standards developed in the biometric industry and standardisation organisations to ensure interoperability, as well as some of the legal framework and the issues associated to biometrics such as privacy and personal data protection. The state-of-the-art in terms of technological development is also summarised for a range of single biometric modalities (2D and 3D face, fingerprint, iris, on-line signature and speech), chosen according to ICAO recommendations and availabilities, and for various multimodal approaches. This paper gives a summary of the main elements of that report.

The second objective of the MBioID project is to propose relevant acquisition and evaluation protocols for a large-scale deployment of biometric IDs. Combined with the protocols, a multimodal database will be acquired in a realistic way, in order to be as close as possible to a real biometric IDs deployment. In this paper, the issues and solutions related to the acquisition setup are briefly presented.

# 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Biometrics; Multimodality; Identity documents; Electronic passport; Acquisition protocol; Evaluation protocol

1. Introduction

The issues associated with identity usurpation are currently at the heart of numerous concerns in our modern society. Establishing the identity of individual person is recognized as fundamental to the numerous administrative operations. Identity documents (IDs) are tools that permit the bearers to prove or confirm their identity with a high degree of certainty. In response to the challenges posed by theft or fraudulent use of IDs and security threats, a wide range of biometric technologies is emerging, covering, e.g. face, fingerprint and iris recognition. They are also proposed to enforce border control and check-in procedures. These are positive developments and they offer specific solutions to enhance document integrity and ensure that the bearer designated on the document is truly the person holding it. Biometric identifiers – conceptually unique attributes – are often portrayed as the panacea for identity verification.

In many countries, IDs is increasingly associated with biometry. Most modern identity cards are proposed associated with chips embedding biometric identifier. Under the impetus of the United States of America, a large number of countries (all EU countries) are developing and piloting if not delivering biometric passports. The International Civil Aviation Organization (ICAO, a United Nations specialised agency) issued specific recommendations for travel documents inviting members to use facial images and optionally fingerprint or iris as biometric modalites. The Swiss government is currently conducting a pilot study testing and evaluating the next generation of passport developed according to the ICAO recommendations.

2. Purpose of the initiative

This project has been triggered by the frenetic technological promises and claim of simplicity of biometric technology
applied to IDs. We believe that the deployment of such technology is a complex task that should receive proper research attention from various perspectives (technological, economical and legal). The MBioID research initiative aims at addressing the following germane question: What and how biometric technologies could be deployed in identity documents in the foreseeable future? This research proposed to look at current and future practices and systems of establishing and using IDs and evaluates their effectiveness in large-scale deployments it takes advantage of an acquired multimodal database specifically designed for exploring IDs biometric recognition efficiency. Indeed, most research today has been focused on studying single modalities independently, making difficult comparisons between various biometric solutions; a multi-modal approach is favoured in this initiative.

3. MBioID research report [1]

At the outset of the initiative, it was felt that all relevant information should be gathered in a review document, in order to establish the current state-of-the-art related to the use of multimodal biometrics in an IDs application. In such a rapidly evolving field, it is of paramount importance to conduct a state-of-the-art review to guide our next steps into the elaboration of acquisition and evaluation protocols and the establishment of a multimodal biometric research database. The MBioID research report [1] gives the main definitions, properties and the acquisition and evaluation protocols and the establishment of a multimodal biometric research database. The MBioID research report [1] gives the main definitions, properties and the framework of use related to biometrics, an overview of the main standards developed in the biometric industry and standardisation organisations to ensure interoperability, as well as some of the legal framework and the issues associated to biometrics, such as privacy and personal data protection. The state-of-the-art in terms of technological development is also summarised for a range of single biometric modalities (face, fingerprint, iris, on-line signature and speech), chosen according to ICAO standards, such as those proposed by the International Civil Aviation Organization [4], are essential so that interoperability and data exchange between applications and systems, thus avoiding problems and costs stemming from proprietary systems. For IDs such as passports, international standards, such as those proposed by the International Organization for Standardization (ISO) [2,3] and the International Civil Aviation Organization [4], are essential so that multimodal fusion with partial templates, at the score and decision levels, to provide better privacy protection to the enrolled users, as partial templates by themselves (i.e. not in combination) would yield very low identification power.

3.2. Multimodality

For IDs application, multimodality may be an effective tool to reduce the Failure to Enroll (FTE) rate. The sequential use of multiple modalities guarantees that the non-enrollable population is reduced drastically. Furthermore, sequential use of modalities permits fair treatment of persons that do not possess a certain biometric trait. We also aim at investigating multimodal fusion with partial templates, at the score and decision levels, to provide better privacy protection to the enrolled users, as partial templates by themselves (i.e. not in combination) would yield very low identification power.

3.3. International standards

International standards relating to biometrics are maturing quickly and many are already available. They support interoperability and data exchange between applications and systems, thus avoiding problems and costs stemming from proprietary systems. For IDs such as passports, international standards, such as those proposed by the International Civil Aviation Organization [4], are essential so that biometric verification can be performed, independently to the location of the transaction.

3.4. Privacy, legal and societal aspects

As part of the feasibility study and before deployment, a full privacy impact assessment should be carried out, under the leadership of the states’ data protection commissioners.

The collection and the process of biometric data should be only conducted in accordance with the requirements of data protection basic principles (e.g. lawfulness, good faith, purpose-link, data security, proportionality and rights of persons concerned). An important issue to address is whether or not biometrics in IDs should serve for verification or their

<table>
<thead>
<tr>
<th>Criterion</th>
<th>2D face</th>
<th>FP</th>
<th>Iris</th>
<th>Signature</th>
<th>Speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error rates</td>
<td>M-H</td>
<td>L</td>
<td>Very L</td>
<td>L</td>
<td>M-H</td>
</tr>
<tr>
<td>Inter-session variability</td>
<td>M</td>
<td>L</td>
<td>Very L</td>
<td>M</td>
<td>M-H</td>
</tr>
<tr>
<td>Universality</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>M-H</td>
<td>M-H</td>
</tr>
<tr>
<td>Risk of failure to enroll</td>
<td>L</td>
<td>M-L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Noise sensitivity</td>
<td>H</td>
<td>M-L</td>
<td>M-L</td>
<td>L</td>
<td>H</td>
</tr>
<tr>
<td>Time to enroll</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Distributed templates</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>Sensor cost</td>
<td>M-L</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>L</td>
</tr>
<tr>
<td>Choice of vendors</td>
<td>H</td>
<td>H</td>
<td>Very L</td>
<td>M</td>
<td>H</td>
</tr>
<tr>
<td>Match-on-card implementation</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Covert acquisition</td>
<td>H</td>
<td>M-H</td>
<td>L</td>
<td>M-L</td>
<td>H</td>
</tr>
</tbody>
</table>
3.5. Integration to identity documents

Technical requirements for the integration of biometrics to IDs were proposed by the International Civil Aviation Organization [5–9] and the National Institute of Standards and Technology (NIST) [10,11]. Furthermore, the European Council has adopted since 2003 six regulations and proposals related to the introduction of biometrics in IDs [12–17]. All European countries have to follow these requirements for their biometric IDs projects.

4. MBioID database

The second objective of the MBioID project is to propose relevant acquisition and evaluation protocols for a large-scale deployment of biometric IDs. The MBioID acquisition protocol has been adapted to a multimodal database size where the number of transactions and enrollment acquisitions has been chosen in order to estimate reliably the error rates of each modality to a certain level of confidence. Furthermore, the acquisition procedure, which depends on the acquisition environment, is also conducted in a realistic way, in order to be as close as possible to a real biometric IDs deployment. Finally, when consistent with our usage scenarios, the acquisition devices and the acquisition procedure of the MBioID database were chosen in order to be as interoperable as possible with publicly available databases. Indeed, many biometric databases exist, with virtually every research laboratory in the field defining and recording their own dataset. Due to limited resources, this approach has led to the development of a vast number of small databases, which taken on their own, can often not be used to predict performance on large populations due to the large confidence intervals induced by small sample sizes. The definition of this database follows the principle established by the MyIDea database [18] that interoperability with publicly available databases is a good way to increase the number of subjects with very limited investment. Therefore, where possible, the MBioID database has been made compatible with leading multimodal and monomodal databases (such as CIxTER, MyIDea, XM2VTS, CASIA, BIOMET and MCYT) so that these can be extended with our data and vice-versa.

The political choice of using biometric information in travel documents of a specific nation is echoed on the whole population of this nation. In order to evaluate this large-scale introduction, a pilot project, on a smaller scale, has to be conducted [19]. Two databases are needed for such an evaluation: a background and a search database, for simulating the system database and the applicants, respectively. The background database has to contain at least several hundred thousand subjects. Both data sets have to be statistically representative of the relevant population (in this case the whole population of the nation) and be large enough in order to infer validly the results of the test evaluation to the potential population of the future application. In order to satisfy this requirement, either a random sampling of subjects from the relevant population should be performed or a sampling using quota (i.e. sex, age, social and professional category, demography, etc.) which represents the characteristics of the population. For the MBioID database, all these restrictive recommendations cannot be met; most of our contributors are staff of student volunteers. However, we have chosen a priori the number of subjects and the number of acquisition per subject we need for evaluating specific error rates, in a specific confidence level, following the methodology proposed by Schuckers [20].

Schuckers proposes a sample size calculation in function of the error rate to estimate, the maximum error rate allowed, the number of acquisitions per subject, the confidence level and a correlation parameter. This correlation parameter is estimated for face and fingerprint from the data in Ref. [21]. Those for 3D face, iris, signature and speech are guessed from our experience. The program (and the documentation related to this tool) is publicly available.²

The acquisition procedure is conducted in two different sessions, separated in time, where both enrollment and transaction data will be acquired. According to the methodology described above, the level of performance that we want to achieve for each modality and the fact that about 120 subjects will be able to be acquired, Tables 2 and 3 presents for each modality the content (enrollment and transaction data) of the MBioID database, as well as the expected and maximum expected errors for a specific confidence level in transaction conditions.

4.1. Environmental conditions

The enrollment procedure, as it is the case with the Swiss biometric passport, will be set in enrollment centres, in order to obtain biometric data of best quality. Indeed, each biometric data can be acquired in standardised conditions, the same standardised conditions in each enrollment centre. Furthermore, “active” operator will probably be present during the acquisitions, in order to avoid any variability introduced by

---

1. Indeed, “background databases smaller than a few hundred thousand people are not suitable for reliable speed/throughput extrapolation” [19].

each subject. Here are some examples of standardised conditions: position, distance to sensor, environmental noise, acquisition noise, etc.

The transaction procedure will be set at border controls, such as airports, ports and embassies. The environmental conditions of these places will be less controllable than the enrollment centres. “Passive” operators will also be present during the acquisition procedure, but will probably not have a sufficient skill level for controlling this stage. Furthermore, some environmental conditions cannot be standardised at border control, such as position, distance from sensor, environmental noise, acquisition noise, etc.

In order to acquire the data of the MBioID database in a realistic way (mimicking the operational scenario described above), the environmental conditions will be recreated to some extent. The enrollment and the transaction acquisitions will be performed in a same room, in which these environments will be recreated. For the enrollment acquisitions, an “active” operator will give all the necessary instructions during the acquisition procedure to the subjects. Each acquisition will be standardised in a particular way, identically for all subjects during the enrollment procedure: seated, without background noise, standardised illumination, without glasses, etc. For the transaction acquisitions, the “passive” operator will not give any instruction during the acquisition procedure, but all devices will have a display board with all necessary written instructions. Each of these latter acquisitions will be less controlled, but conducted identically for all subjects during the transaction procedure: standing, with jacket, window curtain open, with glasses, windows and door opened, etc.

4.2. Acquisition devices

In a large-scale biometric deployment, such as IDs, each acquisition device should follow the standards proposed by international organisations, in order that all data acquired be interoperable with the acquisition devices selected by other countries. The acquisition devices used for the MBioID project, and thus the acquired data, meet the standards and requirements of international organisations (e.g. ISO and ICAO). Table 4 presents the acquisition devices used for the MBioID database.

### 4.3. Acquisition protocol

Biometric data are personal data and have thus to be treated in an appropriate way. A “personal data protection document” is signed by each subject in order to give them the guarantee that their data will be anonymised, that these data will only be transmitted to other institutions for research purposes only if the concerned country benefits from a law on data protection at least equivalent to that existing in Switzerland and that they have a challenge right if they want to be removed from the database. The MBioID acquisition protocol for each modality and each session is as follows.

#### 4.3.1. 2D face

- **Enrollment session**: Five frontal face shots will be taken using a high-quality Fuji Finepix S2 pro camera. Three flashes will be used to remove shadows and ensure even lighting. The background, distance, facial expression, head pose and illumination will be strictly controlled. The shots will comply with ISO photograph regulations.

- **Transaction session**: Five frontal face shots will be taken using the same camera. The background will not be controlled, the distance to the camera will be roughly indicated using a mark on the ground, no pose requirements will be in place and the illumination will be provided using the standard energy-saving light bulbs in use in the room.

#### 4.3.2. 3D face

- **Enrollment session**: Every subject will be enrolled five times with the A4Vision technology, in the best possible conditions (e.g. sitting position, without glasses, hat, etc.). The raw data corresponding to these enrollments will be automatically stored.

- **Transaction session**: Five raw data recordings (called “attempts”) per subject will be acquired additionally in a less controlled way. These attempts will be used during the evaluation process as transaction data. If the subject is wearing glasses, five additional recordings, with glasses, will be acquired during the session.
4.3.3. Fingerprint
- **Enrollment session**: For both indexes and thumbs of each subject, five samples will be acquired in controlled conditions (e.g. sitting position, indication for the finger position on the acquisition device, support cleaning after each subject, etc.) with the Acco 1394 device of Smith & Heinmann Biometrics.
- **Transaction session**: Five additional samples of these fingers will be acquired in a less controlled way (e.g. standing position, no oral indication for the finger position, support not cleaned after each subject, etc.).

4.3.4. Iris
- **Enrollment session**: Five samples per eye will be acquired in controlled conditions (e.g. sitting position, controlled illumination conditions, etc.) for every subject with the BM-ET 300 camera of Panasonic.
- **Transaction session**: Five other samples per eye, in less controlled conditions (standing position, without any oral indication from the operator for the positioning, etc.), will be acquired. If the subject is wearing glasses, five additional transaction data, with glasses, will be acquired during each session.

4.3.5. On-line signature
- **Enrollment session**: Ten signatures will be required with a Wacom Intuos 2 A4 tablet with two inking pen. The angle of the writing tablet, as well as the seating height, can be freely adjusted by each subject to “what feels comfortable for writing”. A sheet of paper is placed on the tablet to allow visual feedback and to give a usual feel to the friction between pen-tip and paper. Additionally, expert forgers will be trained to forge enrollment signatures using purpose-built software from the University of Fribourg [22].
- **Transaction session**: Ten signatures will be required. The angle of the writing tablet, as well as the height of the tablet, is fixed. The subject has to stand up to sign, and is told to keep his or her jacket on.

4.3.6. Speech
- **Enrollment session**: A group of 10 phonetically balanced sentences in French, followed by 2 PINs in French and 2 PINs in English will be recorded by an audio-technica AT3031 microphone. The user is seated, and the directional microphone is positioned laterally with respect to the mouth to prevent excessive pressure gradients on plosives. The distance to the head is kept constant during recording. The acquisition room is padded with absorbing foam to decrease its reverberation time.
- **Transaction session**: The same speech material is recorded, but the door (opening into an office corridor) and the window (opening on a busy road) are opened, and the distance to the microphone is not controlled.

4.4. Acquisition room

The MBioID acquisition room is illustrated below. Fig. 1 presents the relative positioning of the acquisition devices in the room and between them.

5. Database evaluation

The comparison process will be made offline, after the acquisition procedure. Commercial recognition systems and publicly available algorithms will be used for the evaluation of the MBioID database. The performance metrics which will be used are those proposed by Mansfield and Wayman [23].
6. Conclusion

The MBioID initiative proposes to look at current future practices and systems of establishing and using IDs and evaluates their effectiveness in large-scale deployments. At the outset of the initiative, it was felt the need for gathering, all relevant information for establishing the current state-of-the-art related to the use of multimodal biometrics in an IDs application. This research report [1] is publicly available on the European Biometrics Portal (http://www.europeanbiometrics.info). The proposal of relevant acquisition and evaluation protocols for large-scale deployment of biometric IDs is the second milestone of the initiative. Indeed, there is a need of realistic acquired data for the scientific community in order to impact the political choice about the modality(ies) choice, the performance requirements and the system architecture of a biometric IDs deployment. For this purpose, the MBioID project has started an acquisition procedure for a multimodal biometric database, conduced in a realistic way. These data will be available for the scientific community at the end of the acquisition step.

Acknowledgements

We take this opportunity to thank the University of Lausanne, the EPFL and the Foundation BCV, for the financial support to this research initiative. We also wish to thank Eric Dürst and Eric Sapin, for the setup of the acquisition room.

References


Recent documents dating: An approach using radiocarbon techniques

D. Zavattaro\textsuperscript{a,}\textasteriskcentered, G. Quarta\textsuperscript{b}, M. D’Elia\textsuperscript{b}, L. Calcagnile\textsuperscript{b}

\textsuperscript{a} Raggruppamento Carabinieri Investigazioni Scientifiche, Roma, Italy
\textsuperscript{b} CEDAD, Department of Engineering of Innovation, University of Lecce, Lecce, Italy

Received 9 June 2006; accepted 14 June 2006
Available online 26 July 2006

Abstract
The possibility to develop an absolute technique, independent from the paper conservation conditions, to date recent paper documents (i.e. less than 50 years old) for forensics purposes is discussed. We suggest the possibility to use the curve representing the strong increase in the atmospheric radiocarbon concentration induced in the last 50 years by nuclear weapons tests as reference to date paper documents, with a resolution down to a few months. The results obtained in the analysis of two known age documents are presented together with a first order mathematical model developed in order to take into account the contributions of the different tree rings employed in the paper production.

#2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Forensic; Questioned documents; Radiocarbon; Paper dating

1. Introduction

The development of a reliable technique for the absolute dating of paper documents is still now a relevant, unsolved, question for forensic purposes. Standard methods, actually, are too sensitive with respect to the document storing conditions. Humidity, light exposure and heat strongly affect all the chemical/physical variables usually employed in measurements, both in dating inks and paper [1].

Does paper possess a feature, age-dependent, not affected by conservation conditions? If so, this feature would be related with the production moment and would be invariant versus all other agents involved during the paper life.

Following this aim we investigated the possibility to date paper documents by using the Accelerator Mass Spectrometry (AMS) radiocarbon dating method [2].

In fact, the radiocarbon dating technique is usually limited in the analysis of samples older than 200 years and with uncertainties in the range between 20 and 40 years [3], resulting in a limited importance of the method for forensics purposes. However, the well documented, strong increase (bomb peak) in the radiocarbon concentration induced by the nuclear detonation test carried out in atmosphere after the second world war, has given the possibility to develop a reliable method to date samples younger than 50 years and with uncertainties often better than 1 year. The potentialities of this “bomb peak dating” method has been shown recently in several forensics applications such as the dating of illicit drugs, food stuff, wines and human rests [4]. Fig. 1 shows the temporal variation of the atmospheric $^{14}\text{C}$ concentration as measured in atmospheric CO$_2$ and tree rings, expressed through the $\Delta^{14}\text{C}$ term defined as the relative difference with the undisturbed 1950 value [5,6]. In the Northern hemisphere the radiocarbon concentration started to increase in 1955, reached its maximum level in 1963–1964 and started to decrease in 1962–1963 as the result of the Nuclear Test Ban Treaty due to exchanges with the biosphere and the hydrosphere.

Nevertheless the application of the method has been generally limited, so far, to “short living” samples (i.e. with an annual life cycle) where the measured radiocarbon concentration reflects the atmospheric value in a single year. In this case, in fact, the measured radiocarbon concentration can be easily converted in a calendar year using the “bomb peak” curve as reference. This is not, however, the case of the dating of paper documents when the measured radiocarbon content is an average value resulting from the contribution of different tree rings, each one having a radiocarbon concentration characteristic of the year of growth. This makes the application of the bomb peak dating method to paper documents not a straightforward procedure and requires the
development of a mathematical model to correct the measured radiocarbon concentration in the paper in calendar years (see Fig. 1). 

2. Methods

In our experiment 2 known-age documents have been investigated. They are two similar postal bulletins (LTL841A and LTL840A), issued in the same Italian town, respectively, in 1992 and 1993.

Since the determination of the carbon isotopic ratios by mean of the Accelerator Mass Spectrometry (AMS) technique requires between 20 and 50 mg of paper material only few mm² of paper are needed for the analyses resulting in a reduced destructiveness of the method. In this work two kinds of sample were extracted from each document, one ‘paper only’ and one ‘paper and toner’ in order to estimate whether the presence of the ink could influence the radiocarbon determinations or not. The samples were treated by alternate acid/alkali/acid attacks in order to remove contaminations (i.e. inks and additives), they were combusted at 900 °C to carbon dioxide in sealed quartz tubes together with copper oxide and silver wool and then reduced at 600 °C to graphite by using H₂ as reducing agent and iron powder as catalyst [7]. The radiocarbon concentration in the obtained graphite was then measured by the Accelerator Mass Spectrometry technique at Centro di Datazione e Diagnostica (CEDAD), University of Lecce, Italy using standard materials supplied by International Atomic Energy Agency (IAEA) as reference. The accuracy of the measurements was monitored by preparing and measuring, together with the samples, IAEA C3 cellulose standards with a nominal, certified radiocarbon concentration of 129.41 pMC.

3. Results

The results of the AMS measurements are summarized in Table 1 for both the $^{14}$C/$^{12}$C ratio and the $\delta^{13}$C term.

The comparison of the radiocarbon concentration measured for the IAEAC3 cellulose with the nominal, consensus value shows a relative difference of 0.4%, suggesting an high degree of accuracy of the overall process. No statistically significant differences were measured between the ‘paper only’ and ‘paper and toner’ samples indicating an effective removal of ink contamination by the followed sample processing procedures.

The calibration of the raw data was obtained by using the Northern Hemisphere Bomb $^{14}$C dataset and the software Calibomb [8] (Fig. 2) and resulted in an “apparent” date of production for the two documents of 1981 and 1982, for the two samples LTL841A and LTL840A, respectively, with a one
standard deviation uncertainty of 6–7 months. Although these “apparent” ages are different, as was expected, from the real ones the method appears, without applying any correction, to give consistent results at least with respect to the relative age of the samples.

4. Discussion

To obtain a real absolute age for the samples a deeper analysis and interpretation of these raw data is needed by taking into account the number of rings and so the age of the tree employed by the producers and the radiocarbon concentration in each ring. We propose a mathematical model expressing the $^{14}$C concentration ($C$) in the samples given the number $N$ of rings forming the tree and the shape $f_{^{14}C}(t)$ of the “bomb peak” curve.

Starting from the first-order hypothesis that the distance $d$ between each ring was a constant inside the same tree, we can calculate the amount of superficial $^{14}$C content into the $n$th-ring of the tree:

$$Q_n = \left(\frac{2}{C_0}\right) \pi d^2 \int_n^{n+1} f_{^{14}C}(t_n) \, dt$$

(1)

to obtain concentration we have to sum over the $N$ rings forming the tree and to divide by the total surface $\pi (Nd)^2$, so the general formula becomes independent of the parameter ‘$d$’:

$$C = \frac{1}{N^2} \sum_{n=1}^{N} (2n-1) \int_n^{n+1} f_{^{14}C}(t_n) \, dt$$

(2)

By using this equation, known (by measurements) the concentration ‘$C$’ of the two documents as given in Table 1, the number $N$ can be calculated. For both the samples we obtained $N = 25$, which would be the average age of the trees used to produce the paper pulp. The agreement of this value with the information about the ages of the tree employed in the paper production as supplied by the manufactures, points towards a consistency of the method.

5. Conclusions

We suggest a technique, based on the “$^{14}$C bomb peak” dating method to absolutely date documents younger than 50 years for forensics applications with a resolution down to a few months.

However, preliminary analyses carried out on documents of known age suggest that a careful interpretation of the radiocarbon data is needed. In particular we developed a first order correction model to express the overall radiocarbon concentration measured in the paper sample as a function of the concentration in each tree ring. By using this model we found that, if the age of the trees used to produce the paper is known, a correct value of the year of production can be estimated.

Further studies are in progress in order to confirm the obtained results also analyzing samples of different origin and eventually developing suitable databases. Nevertheless the application of the method to absolutely date cotton-pulp paper (i.e. currencies) is straightforward and no complex corrections model are needed thank to the short life of the plants used in the production.

References

Max Frei theory revisitation: Does really strokes depth change along time?

Giuseppe Schirripa Spagnolo\textsuperscript{a,}\textsuperscript{*}, Davide Zavattaro\textsuperscript{b}, Michele De Santis\textsuperscript{a}, Luca Gennaro Ferrillo\textsuperscript{a}

\textsuperscript{a}Dipartimento di Ingegneria Elettronica, Università degli Studi “Roma Tre”, Via Della Vasca Navale 84, I-00146 Roma, Italy
\textsuperscript{b}Raggruppamento Carabinieri Investigazioni Scientifiche, Viale Tor di Quinto 119, Roma, Italy

Received 8 June 2006; accepted 14 June 2006
Available online 2 August 2006

Abstract

A serious problem in questioned document examination is to establish the age of written lines. With respect to paper dating, in the past, Max Frei theory (based only on microscope analysis) claimed that strokes are time-dependent. Therefore, according to this theory, it has been asserted that from the analysis of the strokes depth changes it is possible to try to date the handwritten document (to find out that the document is older than \ldots).

In the present work, we investigate the strokes depth change by a laser profilometer considering not only the time but also microclimatic variations. First, we analyze the stability of stroke characteristics along the time. In particular, we demonstrate that if the document is preserved without change of temperature and humidity, the depth of the strokes has not appreciable changes. In this way, we have the purpose to verify the real possibility of documents dating by means of Max Frei theory. Subsequently, we test how the 3D profile of strokes changes in connection with the microclimatic variations. In particular, we test humidity variations. With this experiment, we show that humidity variations reduce the strokes depth. Moreover, this effect shows a non-linear trend, leaving a hysteresis on the depth.

Finally, we show that the analysis of 3D stroke profile is unable to determine the age of documents.

\textsuperscript{*}Corresponding author. Tel.: +39 06 55177046.
E-mail addresses: schirrip@uniroma3.it (G. Schirripa Spagnolo), dzavattaro@carabinieri.it (D. Zavattaro).

\section{1. Introduction}

A difficult forensic science problem, in questioned document examination, is to establish the age of written lines. In fact, at the moment, there are no accepted techniques available to allow absolute ink’s dating [1], specially because all of them are too sensitive with respect to document storing conditions. Therefore, it is often necessary to utilize other features pertaining to the creation and history of the documents. Because the signature, produced with a pen, is often the only handwritten trace in a questioned document, it represents the most important element to date it.

Handwriting on a common paper sheet it is possible to observe how the pen-tip, besides releasing the ink, deforms the paper. This is because, the writing pressure leaves more or less deep impressions, according to:

\begin{itemize}
  \item writing pressure (amount of pressure exerted over the paper during the act of writing);
  \item underlying material (sheet of paper lying on a metal surface or on a paper block);
  \item writing material (fountain, pencil, ballpoint pen interact in different ways with the paper);
  \item type of paper used (the production process determines the size and the morphology of the fiber’s layers in the paper).
\end{itemize}

Therefore, three-dimensional analysis of handwriting can be used to have information on stroke impression and on the pressure applied to the paper during writing.

Recently, 3D laser profilometry has been introduced to transform seemingly flat pen strokes into landscapes of hills and valleys [2–4]. The resulting 3D profile shows the pen-tip strokes as an impression in the paper. This method allows measurements in z-axis (uncertainty less than 1 \textmu m) in every point of the map.

With respect to paper dating, in the past, Frei–Sulzer theory (based only on microscope analysis) [5] claimed that strokes are time-dependent, due to the (supposed by his observations) 3-year paper elastic release time. According to this theory, it is
possible to try to date the handwritten document studying depth change over the time.

Today, by means of 3D reconstruction of stroke profile, it is possible to verify if the Frei–Sulzer theory is really applicable for documents dating.

In this paper, we investigate the strokes by a laser profilometer considering not only the time but also microclimatic variations.

The paper is organized as follows.

First we analyze the stability of stroke characteristics along the time. In particular we demonstrate that if the document is preserved without change of temperature and humidity, the depth of the strokes has not appreciable changes, within the sensitivity of the used profilometer. In this work, the stability of three-dimensional features is tested along 3 years. It is valid the assumption that the document is preserved with stable microclimatic conditions.

Subsequently, we test how the 3D profile of strokes changes in connection with the microclimatic variations. In particular, we have studied humidity variations. With this experiment, we show that an increase of humidity reduces the strokes depth.

2. Material and methods

To test the stability of stroke characteristics along the time we have performed some different experiments. Firstly, we have made some specimens with six different strokes. Three strokes, approximately rectilinear, have been carried out with different pressure using a ballpoint pen and soft underlying material. The other three strokes have been realized with the same precedent conditions but using hard underlying material. Subsequently, one specimen has been sealed up inside a box (see Fig. 1). In this way, the specimen is preserved without moisture change. Besides, if it is stored with stable temperature, there are no variations of relative humidity. To have the possibility to use a laser, to determine the depth of the stroke, an optical window is present in the box.

In order to vary the writing pressure in an objective and quantified way, an experimental setup as shown in Fig. 2 was constructed. The ballpoint pen is kept in a pen support that is translated, in X–Y coordinates, by means of an electromechanical device. Two kinds of underlying material were considered. On the one hand, the strokes were made on a “soft” underground consisting of a pile of 0.1 cm of 80 g/m² paper; on the other hand, strokes were made on the “hard” underground of a metal plate.

In order to obtain controlled writing pressure, weights of lead were loaded on the pen support (weights of 250, 500 and 750 g were used). Strokes written with a weight of 500 g were found to have a similar impression as normal handwriting.

The experimental setup and the procedure used to perform the writing are very similar to that proposed in Ref. [3].

To determine the stroke depth, in this work, we propose the use of the 3D laser profilometry, realized by means of the conoscopic holography. It is suitable to obtain 3D micro-topography with high resolution, better than 1 μm in height (z-coordinate), also on surfaces with uneven reflectivity (this situation is usual on the surface of the handwritten document). The technique is able to obtain 3D profiles in non-invading way. Therefore, the system leaves the investigated surface unaltered so that the questioned document can be studied by means of other destructive or non-destructive techniques in different moments, also in the case of forensic analysis with the necessity to preserve the original sample.

A second specimen, made in contemporary to the precedent, is positioned in the laboratory without protection. In this way, it has suffered all microclimatic variations present in the laboratory. It is important to note that the temperature of the laboratory was nearly constant during all experiments (about 22 °C). On the contrary, the relative humidity was variable.

The second experiment, to verify the time stability of strokes, consists to compare the depth of the same strokes measured with a distance, in time, of 32 months.

Finally, to check the humidity effect, other samples (carried out in similar way to the precedents) have been exposed to cyclical variations of humidity (during all the experiments we use constant temperature).

3. Results and discussion

To control the strokes depth stability, along the time, the sealed box was monitored, with 3D laser profilometry, for about 8 months. To avoid the uncertainties of box repositioning, under the laser profilometry, strokes profiles are averaged along the stroke direction (see Fig. 3). Fig. 4 shows the percentage of depth variation versus time. From this graphic, it is possible to affirm that without microclimatic variations the depth of strokes is stable along the time. In this case, the percentage variation of 1.5% corresponds to less of a micron of variation of depth (this variation is less than the sensibility of the used profilometer considering not only the time but also microclimatic variations.

Fig. 2. Experimental setup by means X–Y electromechanical device with the ballpoint pen kept in the pen support.

profilometer). The profile of reference, in comparison with which the variations have been determined, has been performed 3 days after the writing. The box has been sealed immediately after tracing the strokes. Three days have been necessary to stabilize the microclimate inside the box (evaporation of the solvents used for the sealing).

Fig. 4 is relative to the stroke writing using a weight of 500 g on the pen support and using a soft underlying support. However, all the other strokes have analogous course. An other experiment, to verify the time stability of strokes, was performed on a linear stroke. In this experiment we have determined, the first time, the average depth of a stroke line made in September 2003. Subsequently, this specimen has been preserved inside a plastic envelope. In May 2006 we have repeated the test.

The comparison between the two tests is shown in Fig. 5. It is possible to note that in 32 months no appreciable variation in the stroke depth occurs.

In contemporary to the effected tests on the stroke contained in the sealed box, we have monitored the change in the stroke depth of the specimen positioned in the laboratory without protection. Fig. 6 shows the percentage of stroke depth variation versus time. In this sample, on the contrary of that contained in the box, we observe some significative variation in strokes depth. In particular, it is possible to note a reduction during the first 20 days. This is related to the average value of relative humidity increase (from 45 to 65% RH) along the period from the 15 October to the 15 November.

Subsequently, with the activation of the winter heating, inside the laboratory, we had ~35% RH. The reduction of relative humidity causes a “regeneration” of the original stroke depth. Finally, with the turning off of the winter heating, and the increase of the relative humidity (inside the laboratory ~55%), the stroke depth reduced again.

Analyzing Fig. 6 it is possible to note that, on the contrary of time, humidity has an important role on the stroke depth. To check the humidity effect, we have induced, by means of a climatic chamber, cyclical variations of humidity (with constant temperature).

Initially, using a sample realized with the same procedure previously exposed, we have effected the following variation of relative humidity. Starting from ~40% RH, we have increased it to the value of 90% RH for 24 h. Later on, we have brought back the relative humidity (to ~40% RH). With this variation of relative humidity, the strokes depth has

![Fig. 3. Typical strokes used in our experiments. The strokes are acquired with laser profilometer, subsequently they are averaged along line direction.](image1)

![Fig. 4. Stroke depth variation, in percentage, vs. time. The figure shows that, in 7 months, there are not appreciable variations in stroke depth.](image2)

![Fig. 5. Comparison of stroke depth after long-time. It is possible to note that in 32 months no appreciable variation in the stroke depth occurs.](image3)

![Fig. 6. Percentage of stroke depth variation vs. time for a sample positioned in the laboratory without protection.](image4)
reduced to the 80% of its initial size. Subsequently, we have kept constant the relative humidity for 16 days (as in the first 400 h in Fig. 7). In this period, we have not noticed variations in strokes depth. In particular, the initial variation does not change. Subsequently, we have effected another humidity variation (40% RH → 90% RH → 40% RH). In this test, it is possible to notice that the depth of the strokes reduces. Besides, if we bring the document back to the initial conditions of humidity, we do not have a complete restoration of the initial depth.

Finally, we have effected another humidity variation (40% RH → 85% RH → 40% RH). Even in this test, the depth of the strokes reduces.

Fig. 7 shows the percentage of depth variation versus time in the test with cyclical variations of humidity. From Fig. 7 it is possible to note that variations of relative humidity (if implicated high values of relative humidity) produce depth variation with hysteresis. In other words, the cyclical variations of humidity reduce progressively the strokes depth.

In all experiments, we have strokes depth between 10 and 50 μm. Besides, all the percentage variations, in strokes depth, are similar for every type of stroke used in our experiments.

4. Conclusion

In this work we have introduce a certain number of experiment that shows that the long-time elastic paper release was not observed, while a relevant variation in the strokes depth was seen in connection with humidity changes.

In particular, we have proved that if a document is stored in stable microclime (constant relative humidity) the strokes depth is stable.

Instead, cyclical variations of relative humidity can cause a progressive reduction of the strokes depth. The most important aspect is that such process introduces hysteresis. Bringing a document back to the initial conditions of humidity, does not guarantee that the strokes depth returns to its initial size. Probably, this is what Frei–Sultzer has noticed.

Finally, we have to conclude that a strokes-depth study is not a suitable method to date documents.

References

Non-destructive testing techniques for the forensic engineering investigation of reinforced concrete buildings

Brian Hobbs a, Mohamed Tchoketch Kebir b,*

a Centre for Forensic Investigations, School of Science & Technology, University of Teesside, Middlesbrough, UK
b National Institute of Criminalistics and Criminology, DP/CGN, BP 53 Alger gare, 16000 Algiers, Algeria

Received 10 June 2006; accepted 14 June 2006
Available online 14 August 2006

Abstract

This study describes in detail the results of a laboratory investigation where the compressive strength of 150 mm side-length cubes was evaluated. Non-destructive testing (NDT) was carried out using ultrasonic pulse velocity (UPV) and impact rebound hammer (IRH) techniques to establish a correlation with the compressive strengths of compression tests. To adapt the Schmidt hammer apparatus and the ultrasonic pulse velocity tester to the type of concrete used in Algeria, concrete mix proportions that are recommended by the Algerian code were chosen. The resulting correlation curve for each test is obtained by changing the level of compaction, water/cement ratio and concrete age of specimens. Unlike other works, the research highlights the significant effect of formwork material on surface hardness of concrete where two different mould materials for specimens were used (plastic and wood). A combined method for the above two tests, reveals an improvement in the strength estimation of concrete. The latter shows more improvement by including the concrete density. The resulting calibration curves for strength estimation were compared with others from previous published literature.

© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Concrete; Compressive strength; Non-destructive testing (NDT); Impact rebound hammer (IRH); Ultrasonic pulse velocity (UPV); Formwork; Density; Combined method

1. Introduction

Failure of reinforced concrete buildings due to earthquake activity is a major problem in Algeria. It is believed that many such failures result from poor construction or inadequate materials. This paper presents an experimental investigation of the case of non-destructive testing methods for use as a part of the forensic investigation of such failures.

2. Test programme

2.1. Test programme design

2.1.1. Aim

The main advantage of non-destructive test methods is that they do not damage or affect the structural performance of building components. In addition, they offer simplicity and rapidity in use: test results are readily available on site, the possibility of testing concrete strength in structures where cores cannot be drilled and the application of less expensive equipment. However, too much reliance should not be placed on the calibration curve supplied with the equipment, since the manufacturer develops this curve using standard cube specimens and the mix used could be very different from the one being tested. Both, the impact rebound hammer (IRH) and the ultrasonic pulse velocity (UPV) tests, are only useful if a correlation can be developed between the rebound number/ultrasonic pulse velocity readings and the strength of the same concrete, or at least a concrete specimen with the same composition as the one in question. For that reason, an experiment was planned to achieve the following goals:

• To adapt the Schmidt hammer apparatus and the ultrasonic pulse velocity tester to the type of concrete used in Algeria, so concrete mix proportions that are recommended by the Algerian regulations is to be adopted (Table 1).
The resulting correlation curves for each testing technique is to be established by changing the level of compaction, water/cement ratio, mould material and concrete age of specimens. If the effect of one of these factors is proved to be significant on correlation, then the general correlation may be simplified to particular ones for different levels of this factor. The use of a particular correlation may be more reliable but requires more information about this factor.

A combined method for the above two tests should be established, as an improvement in the strength estimation of concrete would be expected, given the results of previous similar researches. More improvement in strength estimation should be expected by including the concrete density, given its proportional relation to the velocity of ultrasonic pulse through concrete on one hand and to the compressive strength on the other.

Unlike other researches, this paper would introduce a new factor which is formwork material. Its effect on concrete properties (surface hardness or ultrasonic pulse transmission) is to be determined. If it is found to be significant, each of the two adopted mould materials (metal and wood) should have a specific correlation curve in order to improve the estimation of compressive strength of concrete.

2.2. Variables investigated

The following factors were chosen—mould material, level of compaction and concrete original W/C ratio:

- To determine the separate effect of each factor on the resulting correlation for the estimation of compressive strength of concrete.

2.2.1. Mix proportions

A typical concrete mix proportions, which is specified in Algerian regulations was adopted as a reference in this experimental work (mix A). Its constituents’ proportions are summarised in Table 1. Two other mixes B and C were chosen where their W/C ratios were 0.6 and 0.7, respectively, and the same proportions for the other constituents as mix A.

2.2.2. Level of compaction

The recommended amount of tamping for fresh concrete is 10 times/10,000 mm² of concrete surface using a standard rod (NA 2610), i.e. 23 times of tamping are sufficient for a 150 mm side-length cubes. Three different levels of compaction were used (Table 2).

2.2.3. Type of mould material

The most common types of formworks used in Algeria are metallic and wooden formworks. Therefore, hard plastic and wooden moulds were used in the laboratory in order to simulate the respective effect of metallic and wooden formwork on concrete material.

2.3. Test specimens

2.3.1. Concrete mixes constituents

All specimens were cubes of 150 mm side-length. They were made from the following materials: Portland cement (42.5N), wet semi-crushed gravel (20 mm) and wet sharp sand.

2.3.2. Specimens

The above three concrete mixes A, B and C were prepared. Three additional specimens from each mix D and E were prepared later in order to check the consistency of the previous results. Specimens made from each mix are summarised in the following table:

### Table 1

Typical concrete mix proportions required by Algerian regulations [2]

<table>
<thead>
<tr>
<th>Constituents</th>
<th>Quantity (kg)</th>
<th>Equivalent weight (kg)</th>
<th>Proportion by weight</th>
<th>W/C ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cement (42.5N)</td>
<td>350</td>
<td>350</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>Gravel (20 mm) (wet)</td>
<td>800</td>
<td>1129</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>Sharp sand (wet)</td>
<td>400</td>
<td>499</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>175</td>
<td>175</td>
<td>0.5</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2

Different specimens used in the experiment

<table>
<thead>
<tr>
<th>Set</th>
<th>No. of specimens</th>
<th>Mix</th>
<th>W/C ratio</th>
<th>Compaction (tamping times)</th>
<th>Mould</th>
<th>Purpose of experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAC</td>
<td>12</td>
<td>A</td>
<td>0.5</td>
<td>Full (23)</td>
<td>Plastic</td>
<td>×</td>
</tr>
<tr>
<td>PBC</td>
<td>12</td>
<td>B</td>
<td>0.6</td>
<td>Full (23)</td>
<td>Plastic</td>
<td>×</td>
</tr>
<tr>
<td>PCC</td>
<td>12</td>
<td>C</td>
<td>0.7</td>
<td>Full (23)</td>
<td>Plastic</td>
<td>×</td>
</tr>
<tr>
<td>WCC</td>
<td>12</td>
<td>C</td>
<td>0.7</td>
<td>Full (23)</td>
<td>Wood</td>
<td>×</td>
</tr>
<tr>
<td>PCS</td>
<td>12</td>
<td>C</td>
<td>0.7</td>
<td>Semi- (11)</td>
<td>Plastic</td>
<td>×</td>
</tr>
<tr>
<td>PCN</td>
<td>12</td>
<td>C</td>
<td>0.7</td>
<td>No (0)</td>
<td>Plastic</td>
<td>×</td>
</tr>
<tr>
<td>PDC</td>
<td>3</td>
<td>D</td>
<td>0.55</td>
<td>Full (23)</td>
<td>Plastic</td>
<td>×</td>
</tr>
<tr>
<td>PEC</td>
<td>3</td>
<td>E</td>
<td>0.65</td>
<td>Full (23)</td>
<td>Plastic</td>
<td>×</td>
</tr>
</tbody>
</table>

Essentially, it is necessary to vary the strength of specimens in order to obtain a correlation between the test method results and the current concrete strength [1,3].
2.4. Test procedure

2.4.1. Correction of water proportion

All W/C ratios adopted in Table 2 were corrected taking into account the amount of water that already exists in aggregates as both gravel and sand used were wet. Aggregates were dried in an oven for 2 h at 60°C. The volumic weights of wet/dry aggregates and the amount of extra water to be corrected are as follows (Table 3):

<table>
<thead>
<tr>
<th>Aggregate</th>
<th>Volumic weight (kg/m³)</th>
<th>Extra-water in 1 m³ of aggregate (kg)</th>
<th>Extra-water in 1 m³ of concrete (kg)</th>
<th>(Extra W/C) ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wet</td>
<td>Dry</td>
<td>Extra-water in 1 m³ of concrete (kg)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gravel</td>
<td>1411.4</td>
<td>1369.1</td>
<td>42.3</td>
<td>66.3</td>
</tr>
<tr>
<td>Sand</td>
<td>1248.1</td>
<td>1166.9</td>
<td>81.2</td>
<td></td>
</tr>
</tbody>
</table>

3. Data presentation and analysis

Calibration curves for each test method IRH and UPV are drawn using regression analysis. The effect of each of the following factors: W/C ratio, level of compaction and the mould material on the correlation was represented by plotting the averages of rebound number/ultrasonic pulse velocity against the compressive strength of each of the three identical cubes.

3.1. Impact rebound hammer results

3.1.1. Graphical presentation

The best-fit line, which represents the relationship between the rebound number and the compressive strength of concrete, is a straight line which has the following equation:

\[ f_c(R) = 2.1459R - 27.22 \]  

where \( R \) is the rebound number. The number of data used in the correlation is \( n = 23 \). The \( r^2 \) value was found to be 0.9238, which indicates a significant correlation. The 95% prediction interval is quite narrow (\( f_c \pm 4.78 \text{ MPa} \)) where all the data values are within this interval. The standard error was found to be S.E. = 2.3895.

3.1.2. Analysis

3.1.2.1. Water/cement ratio effect. The results show that for any rebound number, there is only one unique compressive strength value. This conclusion is very important as there is no need to know the W/C ratio of concrete to predict its strength using the IRH test since this factor was considered in the establishment of the correlation. The range [0.5, 0.6] of W/C ratio is more sensitive than the range [0.6, 0.7] in terms of concrete compressive strength and the rebound number readings.

3.1.2.2. Level of compaction effect. The results show that for most rebound numbers there is an interval of uncertainty of 15 MPa of compressive strength. It is more useful to simplify the correlation to particular correlations for each level of compaction in order to reduce the error in strength prediction. However, the determination of the level of compaction for such building component is impracticable, without further research. Hence, there is no other way except plotting the whole data in the same correlation curve and taking this imprecision in strength estimation into consideration.

3.1.2.3. Mould material effect. The results show that the reduction in density after mould stripping for mix C, which was cast in plastic mould is about 1.21% (about 94 g of water loss/cube). This percentage is smaller than the one recorded for mix C which was cast wooden mould (2.40%), it is about 192 g/cube of water loss. This difference of water loss is evidenced by...
the increase in weight of the wooden moulds by an average of 96 g for each one. The results show two quite parallel lines where the surface hardness of cubes that were cast in wooden moulds is higher than the one of cubes that were cast in plastic moulds (by about 2 units in rebound number) for the same compressive strength of concrete. This difference is quite constant and seems to be significant. Hence, the data should not be plotted in the same correlation curve unless the formwork used during construction in such structure is unknown. Otherwise, an error of ±3 MPa is implicated in the evaluation of concrete strength.

3.1.2.4. The final IRH calibration curve. After including the two additional data values (two small circles in Fig. 1), the correlation was improved. This is evidenced by the increase in the coefficient of correlation value from $r = 0.9238–0.9252$ and the decrease in the standard error from S.E. = 0.3895–2.3334. The final calibration curve (Fig. 1) for the IRH test has the following equation:

$$f_c(R) = 2.168R - 27.747$$

where $n = 25$, $r^2 = 0.9252$ and S.E. = 2.3334.

$$f_c(R)_{\text{predicted}} = f_c_{\text{observed}} \pm 4.7 \text{MPa}.$$  

3.2. Ultrasonic pulse velocity results

3.2.1. Graphical presentation

The best-fit curve that represents the relationship has the following equation:

$$f_c(V) = 11.228V^2 - 39.075V + 1.4658$$

where $V$ is the ultrasonic pulse velocity. The number of data used in the correlation $n = 23$. The $r^2$ value was found to be 0.9034, which indicates a significant correlation. The 95% prediction interval is quite wider than the previous one ($f_c \pm 5.4 \text{MPa}$) where all the data values are within this interval. The standard error was found to be S.E. = 2.7579.

3.2.2. Analysis

3.2.2.1. Water/cement ratio effect. The results show that for any ultrasonic pulse velocity, there is only one unique compressive strength value. This conclusion is very important as there is no need to know the W/C ratio of concrete to predict its strength using the UPV test. Hence, the data should be plotted in the same correlation curve. The range [0.5, 0.6] of W/C ratio is more sensitive than the range [0.6, 0.7] in terms of concrete compressive strength and the ultrasonic pulse velocity measurements.

3.2.2.2. Level of compaction effect. The results show that for most ultrasonic pulse velocity measurements there is an interval of uncertainty for compressive strength ($f_c \pm 13.5 \text{MPa}$). Since the determination of the level of compaction for such building component is quite impossible, without further research. Hence, there is no other way except plotting the whole data in the same correlation curve and taking this imprecision into consideration.

3.2.2.3. Mould material effect. The results show that for any ultrasonic pulse velocity there is only one unique compressive strength value. This conclusion is very important as there is no need to know the formwork material used to predict its strength using the UPV test.

3.2.2.4. The final UPV calibration curve. After including the two additional data values (two small circles in Fig. 2), the correlation was improved. This is evidenced by the decrease in the standard error from S.E. = 2.7579–2.7164 even though the coefficient of correlation value is quite the same (0.9034 ≈ 0.9031). The final calibration curve for the UPV test (Fig. 2) has the following equation:

$$f_c(V) = 12.289V^2 - 49.024V + 24.271$$  

$$r^2 = 0.9031 \quad n = 25.$$
where \( n = 25 \), \( r^2 = 0.9031 \) and S.E. = 2.7164.

\[
f_c(V)_{\text{predicted}} = f_c \text{observed} \pm 5.4 \text{MPa}.
\] (6)

### 3.3. Combined analysis

#### 3.3.1. Analysis

A relationship between the concrete strength and both the rebound number and the pulse velocity together is found to be very useful in term of improving the reliability of the results.

#### 3.3.2. Final regression models

The final regression models with and without density after including the additional data \( (n = 25) \) are as follows:

(a) The final regression model for \( f_c = f(R, V) \): Different regression models were used [4] where the best one which fit the data is the following:

\[
f_c(R, V) = a + bV^2 + cV + eR.
\] (7)

where \( a = -173.033, b = -4.069, c = 57.693, e = 1.307, \) \( n = 25, r^2 = 0.9490 \) and S.E. = 1.849.

\[
f_c(R, V)_{\text{predicted}} = f_c \text{observed} \pm 3.7 \text{MPa}.
\] (8)

(b) The final regression model for \( f_c = f(R, V, D) \): Different regression models were used [4] where the best one which fit the data is the following:

\[
f_c(R, V, D) = (a + bD + cR + eV)D^f
\] (9)

where \( n = 25, a = -7949.461, b = 1.790, c = 75.303, e = 895.381, f = 0.542, r^2 = 0.9641 \) and S.E. = 1.549.

\[
f_c(R, V, D)_{\text{predicted}} = f_c \text{observed} \pm 3.1 \text{MPa}.
\] (10)

### 3.4. Discussion

It is obvious that the IRH best-fit line showed better correlation than the UPV one. The range of the 95% prediction interval is narrower than the one shown by the UPV best-fit line. The variations were relatively large and these were attributed to the fact that the tested cubes were all cured in dry conditions where the moisture content was not constant during the experiment.

The obtained regression model using the combined method is more accurate and gives closer results to the experimental ones than the results obtained from each method separately. The resulting regression models for strength evaluation \( f_c = f(R, V) \) which contains both variables: rebound number and pulse velocity yielded values for concrete strength within \((-3.7 \text{MPa})\) from the real strength of specimen. However, by including the specimens’ densities in the regression model \( f_c = f(R, V, D) \), more improvement in the strength evaluation of concrete was noticed; the estimation values are within \(\pm 3.1 \text{MPa}\) from the real value of concrete strength.

#### 4. Comparison with other published works

### 4.1. IRH regression model

Fig. 3 shows that the concrete strength estimated by the resulting calibration curve is higher than the ones obtained using the curves proposed by other researchers. It is the closest one to the calibration curve proposed by the manufacturer. However, this does not mean that the others are less confident but each one is intended to give more reliable results than others if the questioned concrete has the same properties as the one used to produce this curve.

### 4.2. UPV regression model

Fig. 4 shows that the estimated concrete strength by the resulting calibration curve is higher than the others. It is the furthest one from the calibration curve proposed by the manufacturer. However, this does not mean that it is less confident than others but the differences between them are due to the influence of different factors such as the type of aggregate used and/or proportions, original water/cement ratio, curing...
conditions, level of compaction and moisture content where each curve is used to predict the strength for a specific concrete mix.

4.3. Combined method regression model

Various regression models proposed by different investigators and others proposed by the author were tried for the data obtained from the current experimental work. The results show that the regression model defined by Eq. (7) (proposed by the author [4]) is the best representation for the data got from the experiment.

5. Conclusion

The development of new calibration curves to adapt both the IRH and the UPV testing techniques for typical concrete mixes required by the Algerian regulation revealed the following points:

- There is no need to know the water/cement ratio of concrete since for a specific concrete, each water/cement value yields only one value of strength which is associated with only one value of rebound number/ultrasonic pulse velocity.
- The effect of level of compaction on the correlations was proved to be significant but there is no other way of improving the accuracy of correlation without further research, since this factor is unquantifiable.
- The effect of mould material on surface hardness of concrete was proved to be significant. Hence, each of the two formwork types (metal and wood) should have a specific correlation curve rather than only one curve for both of them. This may improve strength estimation of concrete given that formwork material is often recognised.
- The use of ultrasonic pulse velocity method alone is not appropriate to estimate and predict the strength of concrete. High variations are obtained, making concrete strength evaluation to a certain extent not easy and the forensic engineering decision would be quite difficult. A reliable estimate of in situ strength can only be obtained if the correlation between cube crushing strength and pulse velocity is known for the particular concrete mix used in the condition in which it exists in the structure.
- When compared to the ultrasonic pulse velocity, the rebound number method seems to be more efficient in predicting the strength of concrete under certain conditions, and even that, the use of this method alone would not give an accurate estimation of the strength of concrete. The use of the impact rebound hammer for strength estimation of in situ concrete must never be attempted unless a specific calibration chart is available, and even then, the use of this method alone is not recommended.
- The use of combined methods produces results that are more reliable and closer to the true values when compared with the use of the above methods separately. In addition, an acceptable level of accuracy was achieved for strength estimation of concrete. Hence, the resulting regression model for strength evaluation could be used safely for concrete strength estimation for the forensic engineering investigation in Algeria.
- More improvement was obtained by including specimens’ densities in the correlation. The latter is very useful since the density of concrete can be determined easily.

References

Influence of rubber compound and tread pattern of retreaded tyres on vehicle active safety

Jakub Zebala,*, Piotr Ciepka, Adam Reza, Robert Janczur

* Institute of Forensic Research, Department of Road Accidents Analyses, Westerplatte St. 9, 31-033 Cracow, Poland
b Cracow University of Technology, Jana Pawla II St. 37, 31-864 Cracow, Poland

Received 5 June 2006; accepted 14 June 2006
Available online 4 August 2006

Abstract

The aim of the article is to present the results of road tests of original and retreaded tyres, analysed in the aspect of vehicle active safety. The tests covered emergency braking, steady-state circular tests and severe lane-change manoeuvre. The tests were performed in summer and winter conditions. Original Michelin tyres and Michelin tyres retreaded with summer and winter rubber compounds were used. The results of the tests proved that vehicle active safety is affected by retreaded tyres. The differences between braking deceleration of a car with original tyres and braking deceleration of a car with retreaded tyres confirmed the necessity of performing braking tests on the place of accident with the set of tyres with which the vehicle participating in the accident was equipped.

Keywords: Retreaded tyres; Steady-state circular test; Severe lane-change manoeuvre; Braking deceleration; Road holding ability

1. Introduction

One of the forms of utilisation of worn car tyres is to retread them. In this process the tread pattern of the used tyre is reconstructed by placing a layer of a rubber compound. The recently improved quality of tyre retreading together with lower prices of retreaded tyres as compared with brand new ones have made such tyres used fairly frequently. Vehicles with retreaded tyres often take part in road collisions and accidents. In reconstructing an event information on the contribution of the properties of retreaded tyres in traffic safety is very important. What is of particular interest is the knowledge of achievable braking intensity in given road conditions as well as the effect of tyres on road holding ability and vehicle manoeuvrability.

There is no up-to-date data available in literature that determine the influence of the properties of retreaded tyres on vehicles active safety. That is why research was undertaken at the Institute of Forensic Research in Cracow the aim of which was to compare car motion parameters obtained when retreaded tyres were used with corresponding parameters obtained for original tyres. The other aim was to determine the effect of the tread rubber compound composition and type of tread pattern of retreaded tyres on car motion parameters.

The tests performed covered emergency braking, steady-state circular test and severe lane-change manoeuvre. The manoeuvre of emergency braking was selected as the basic manoeuvre of defence, most frequently used by drivers to avoid accident. Steady-state circular test is one of the basic quasi-static tests in vehicle road holding ability and manoeuvrability tests. Severe lane-change manoeuvre is a dynamic test useful in assessing vehicle road holding ability and manoeuvrability tests. Severe lane-change manoeuvre is a dynamic test useful in assessing vehicle road holding ability and manoeuvrability, similar to defensive manoeuvre of by-passing an obstacle, often occurring in road traffic. Severe lane-change manoeuvre can be described as a type of vehicle anticipatory steering which helps to establish the response of the driver–vehicle system in traffic emergency [3,5].

2. Subject of research

The tests were done on a Renault Megane Classic car equipped with ABS. The tests were carried out for new and retreaded tyres. The new tyres were manufactured by Michelin, with XH1 tread (Photo 1).

The retreaded tyres used in the tests were hot retreaded. Since the tyres selected for retreading were worn Michelin tyres, the carcass structure of all the tested tyres was the same.

* Corresponding author. Tel.: +48 12 422 87 55; fax: +48 12 422 38 50. E-mail address: jzebala@ies.krakow.pl (J. Zebala).

0379-0738/$ – see front matter © 2006 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.forsciint.2006.06.051
The retreaded tyres used in tests differed in rubber compound. Besides three summer compounds, marked with symbols A–C, tyres with two different winter compounds, marked with D and E, were tested. Four types of tread were selected for tests: two summer treads—energy and sport, and two winter treads—MK770 and MK790. The energy tread was modelled on the Michelin energy XH1 tyre tread, while the sport tread corresponded with the Michelin pilot sport tyre tread. Two winter treads MK770 and MK790 were modelled on Continental 770 and Continental 790 (Photo 1).

Three sets of tyres of different rubber compounds (A–C) and the same tread pattern (energy), and two sets of tyres of identical rubber compound (C) and different tread patterns (energy, sport) were prepared (Table 1). Separate tests were done for winter tyres (D and E).

### 3. Instrumentation

The following instrumentation was used in tests (Table 2):

- CORREVIT S-CE head for non-contact measurement of longitudinal $v_L$ and lateral $v_Q$ components of the vector of car selected point velocity;
- piezoelectric vibratory gyroscope MURATA ENV 05A for measurement of yaw velocity $\psi$;
- converter of steering wheel angle for measurement of steering wheel angle $\delta_H$;
- decelerometer XL-Meter for measurement of braking deceleration.

#### 4. Emergency braking tests

4.1. Research methodology

During braking tests the braking system was applied violently. The employment of the same driver and operation of BAS resulted in good repeatability of time of braking deceleration increase in all the tests. The braking tests were run on a flat and even asphalt pavement in summer and winter conditions. In summer two series of braking tests were run on wet and dry pavement. In each series the car velocity at the beginning of braking was ca. 60 and 90 km/h. The temperature of air at the time was ca. $-5$ °C. The tests were done with the ABS both turned on and off.

For measurements of deceleration XL-Meter Pro Beta manufactured by Inventure Automotive Electronics Inc. was used [7]. On the basis of the recorded changes of vehicle braking deceleration as a function of time mean fully developed deceleration MFDD (mean fully developed deceleration) ($m/s^2$), defined in ECE regulations no. 143 and ECE 71/320, was calculated by dependence:

$$MFDD = \frac{\frac{v_B^2 - v_A^2}{25.92(S_B - S_A)}}$$

where: $v_B = 0.8v_0$, $v_A = 0.1v_0$, $v_0$ is the initial velocity and $S_B$ and $S_A$ are the distances covered during braking from velocity $v_B$ to velocity $v_A$. 

---

**Photo 1. Tread patterns of tested tyres.**

**Table 1** Specification of tyres according to tread pattern and type of rubber compound

<table>
<thead>
<tr>
<th>Tread pattern</th>
<th>Rubber compound</th>
<th>Size of tyre</th>
<th>Marking</th>
</tr>
</thead>
<tbody>
<tr>
<td>XH1</td>
<td>Michelin</td>
<td>185/60 R15</td>
<td>XH1-Michelin</td>
</tr>
<tr>
<td>Energy</td>
<td>A</td>
<td>195/65 R15</td>
<td>Energy-A</td>
</tr>
<tr>
<td>Energy</td>
<td>B</td>
<td>195/65 R15</td>
<td>Energy-B</td>
</tr>
<tr>
<td>Energy</td>
<td>C</td>
<td>195/65 R15</td>
<td>Energy-C</td>
</tr>
<tr>
<td>Sport</td>
<td>C</td>
<td>195/50 R15</td>
<td>Sport-C</td>
</tr>
<tr>
<td>MK790</td>
<td>D</td>
<td>195/50 R15</td>
<td>MK790-D</td>
</tr>
<tr>
<td>MK770</td>
<td>E</td>
<td>195/65 R15</td>
<td>MK770-E</td>
</tr>
</tbody>
</table>
The calculated values of MFDD were used in the analysis of braking intensity during emergency braking.

4.2. Results of braking tests

Car braking tests in the same conditions were repeated at least three times. Following each test series the mean value and standard deviation of braking deceleration MFDD were calculated. In total over 300 braking tests were performed. The results of tests have been shown in diagrams (Figs. 1–3).

4.3. Conclusions from braking tests

In all the braking tests in summer conditions with ABS turned on the braking deceleration values for retreaded tyres were lower than those for brand new tyres. The difference between the braking deceleration for brand new tyres and tyres retreaded with summer rubber compound did not exceed 1.5 m/s², while for tyres retreaded with winter compound the difference was up to 1.8 m/s².

In the braking tests in winter conditions the difference between the braking deceleration values for brand new tyres and tyres retreaded with summer compound did not exceed 0.4 m/s², while for tyres retreaded with winter compound the difference reached up to 1.3 m/s².

With ABS turned off on wet pavement in each braking test the deceleration values were lower for retreaded tyres than for brand new tyres. The difference in deceleration did not exceed 1.1 m/s².

With ABS turned off on dry pavement in summer and winter conditions no significant differences were observed between the deceleration for tyres retreaded with summer compound and brand new tyres. With ABS turned off on dry pavement in summer conditions in each braking tests the deceleration values were higher for tyres retreaded with winter compound than for brand new tyres. The difference reached up to 0.7 m/s². In winter conditions, in the majority of tests, the deceleration values were higher for tyres retreaded with winter compound than for brand new tyres. The difference reached up to 0.4 m/s².

During emergency braking in summer and winter conditions on tyres retreaded with the same compound but with different tread the deceleration values were higher for tyres with energy tread. The difference in summer did not exceed 0.7 m/s² for wet pavement and 0.4 m/s² for dry pavement. In winter conditions, however, the difference in deceleration did not exceed 0.5 m/s².

Table 2
Specification of instrumentation used in tests

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Quantity</th>
<th>Marking</th>
<th>Measurement range</th>
<th>Measurement accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correvit Corrsys® S-CE head</td>
<td>Longitudinal velocity</td>
<td>v_L</td>
<td>0–350 km/h</td>
<td>0–97 m/s</td>
</tr>
<tr>
<td></td>
<td>Lateral velocity</td>
<td>v_Q</td>
<td>0–225 km/h</td>
<td>0–62 m/s</td>
</tr>
<tr>
<td>Gyroscope Murata Gyrostar ENV-05A</td>
<td>Yaw velocity</td>
<td>ψ</td>
<td>±90°/s</td>
<td>±0.1°/s</td>
</tr>
<tr>
<td>Converter potentiometric MU 161</td>
<td>Steering-wheel angle</td>
<td>δ_H</td>
<td>±300°</td>
<td>1°</td>
</tr>
<tr>
<td>XL Meter pro Beta Inventure Automotive Electronics Inc.</td>
<td>Deceleration</td>
<td>MFDD</td>
<td>±12.7 m/s²</td>
<td>0.01 m/s²</td>
</tr>
</tbody>
</table>

Fig. 1. Mean values and standard deviation of braking deceleration MFDD reached during emergency braking at velocities of 60 km/h and with ABS turned on and off in summer on wet asphalt pavement.
The maximum difference in mean values of deceleration reached during emergency braking on energy tread tyres retreaded with different compounds was in summer 1.0 m/s² for wet pavement and 0.4 m/s² for dry pavement, in winter it was 0.5 m/s².

5. Steady-state circular tests

5.1. Methodology of steady-state circular tests

Due to the available test track, the steady-state circular test, with the constant radius of the circle \((R = \text{constant})\) was selected \([2,4]\). A circle of radius of 20 m was marked on the roadway. During the test the drive drove the car along the circle, making corrections of the ride trajectory so that any deviation off the marked circle did not exceed 0.3 m each side. The travelling speed at a given gear was changed from the lowest possible to such at which the driver was still able to drive the car along the assigned trajectory. The car travelling speed was changed at a constant rate at the acceleration not larger than 0.5 km/h/s.

The constant quantity that was determined prior to the tests was the position of centre of gravity relative of the vehicle axis. The quantities measured during the experiments, used in calculations and the quantities calculated on the basis of the recorded histories have been shown in Fig. 4.

The steady-state circular tests were performed in summer conditions. The car was loaded with the instrumentation, the driver and the instrumentation operator. The mass distribution during tests did not differ from that during regular operation.

5.2. Test results

The recorded results were processed using the software of the Institute of Automobiles and IC Engines, Cracow University of Technology together with programmes working
Fig. 4. Location of instrumentation in the tested car and scheme of quantities measured in tests, used in calculations and calculated.

Fig. 5. Dependencies of steering-wheel angle on lateral acceleration (left) and rear axle slip angle on lateral acceleration (right); XH1-Michelin tyres.

Fig. 6. Dependencies of steering-wheel angle on lateral acceleration (left) and rear axle slip angle on lateral acceleration (right); energy-A tyres.
in Matlab environment [6]. The histories of yaw velocity $\dot{\psi}$, longitudinal velocity $v_L$ and lateral velocity $v_Q$ were filtered with Butterworth low-pass filter. The recorded data were used in calculations of (Fig. 4):

- cog longitudinal velocity: $v_{Sx} = v_L + y_L \dot{\psi}$;
- cog lateral velocity: $v_{Sy} = v_Q + x_Q \dot{\psi}$;
- lateral acceleration: $a_c = v_{Sx} \dot{\psi}$;
- rear axle slip angle: $\alpha_2 = \arctan\left(\frac{-v_Q + x_Q \dot{\psi}}{v_{Sx}}\right)$.

Examples of dependencies of steering-wheel angle and rear axle slip angle as a function of lateral acceleration for brand new tyres and energy-A retreaded tyres have been shown in figures below (Figs. 5 and 6).

On the basis of motion parameters of the tested car the steer coefficient $K$ ($s^2/m^2$) was calculated:

$$K = \frac{A_{\delta_H}}{i_s l}$$

where $A_{\delta_H}$ is the steering-wheel angle gradient (derivative of steering-wheel angle $\delta_H$ relative of lateral acceleration $a_c$, defined from the diagram made on the basis of road experiments), $i_s$ the steering system transmission ratio of the tested car and $l$ the axle base of the tested car.

From a series of several steady-state circular tests the mean value of steer coefficient $K$ was determined (Table 3).

The car steerability characteristics was determined from dependence [3]:

$$i_s l \left(\frac{\dot{\psi}}{\delta_H}\right) = \frac{v_{Sx}}{1 + Kv_{Sx}^2}$$

For the tested car boundary acceleration of linearization range 4 m/s² was adopted [3,8].

On the basis of measurements of car motion parameters during steady-state circular tests repeated many times steerability characteristics at lateral acceleration of 4 m/s² was determined for each set of tyres (Fig. 7).

The experiments with all the tested tyres proved the car to have favourable steerability characteristics—moderate understeer. All the characteristics in Fig. 7 are within the characteristics band for a safe car RSV (research safety vehicle), restricted with the black lines: fine and broken.

5.3. Conclusions from steady-state circular tests

In case of XH1-Michelin tyres, treated as model ones, the dependence of steering-wheel angle and rear axle slip on lateral acceleration was almost linear. The lateral wheels slip occurred at lateral acceleration exceeding 6.5 m/s².

Nearly all the retreaded tyres (except energy-B) have lower cornering stiffness. During experiments the highest values of rear axle slip angle, even up to 5.2° (for MK790-D tyres) were reached at lateral acceleration of ca. 6 m/s², i.e. higher by ca. 1.1°–1.7° in comparison with XH1-Michelin tyres.

In case of energy-A, energy-B, energy-C, sport-C and MK790-D tyres the characteristics of steering-wheel angle change was observed to be strongly non-linear. This non-linearity points to poorer manoeuvrability in the range of lateral acceleration higher than 4 m/s².

Table 3
Specification of mean values of steer coefficients for particular sets of tested tyres

<table>
<thead>
<tr>
<th>Tyres</th>
<th>$K$ ($s^2/m^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>XH1-Michelin</td>
<td>0.0022</td>
</tr>
<tr>
<td>Energy-A</td>
<td>0.0014</td>
</tr>
<tr>
<td>Energy-B</td>
<td>0.0021</td>
</tr>
<tr>
<td>Energy-C</td>
<td>0.0020</td>
</tr>
<tr>
<td>Sport-C</td>
<td>0.0025</td>
</tr>
<tr>
<td>MK790-D</td>
<td>0.0028</td>
</tr>
<tr>
<td>MK770-E</td>
<td>0.0027</td>
</tr>
</tbody>
</table>

![Fig. 7. Characteristics of tested car steerability for different tyres (for $a_c = 4 m/s^2$) where $S = i_s l ( \dot{\psi}/\delta_H )$ and $v = v_{Sx}$.](image-url)
As far as car steerability is concerned, energy-A tyres change car properties the strongest and due to a significant rear axle slip angle they reduce understeerability as compared with a car with model XH1-Michelin tyres. At lateral acceleration higher than 4 m/s², the effect of retreaded tyres properties on car steerability can be even more pronounced. It should be noticed, however, that an ordinary driver rarely decides to negotiate a turn at lateral acceleration higher than 4 m/s², but such values are usually reached at undertaking defensive manoeuvres.

6. Severe lane-change manoeuvre tests

6.1. Research methodology

The severe lane-change manoeuvre is a dynamic procedure of fast transition from the initial straight trajectory to the parallel laterally shifted one, which is next followed by the return to the trajectory identical with the initial one. The test track on which the severe lane-change manoeuvre tests were performed was determined according to ISO/TR 3888-1975 standard [1]. The car speed during tests was constant, about 80 km/h. The experiments were run under load identical with that during the steady-state circular tests. During the tests the concrete pavement of the test section was dry. The car motion parameters recorded during tests were a basis for determination of time histories of steering-wheel angle and yaw velocity.

6.2. Presentation and discussion of results

To assess car behaviour during tests with different tyres two factors were adopted:

- gain: as maximum yaw velocity to maximum steering-wheel angle ratio in first lane-change manoeuvre;
- car response time: as time difference between reaching maximum yaw velocity and maximum steering-wheel angle reached.

The notation of respective quantities has been shown in Fig. 8. The values of gain and car response time are determined by the following dependencies:

\[
W = \frac{X_{\dot{\phi}, \text{max}}}{X_{\delta, \text{max}}}, \quad t_{rs} = T_{\psi, \text{max}} - T_{\dot{\delta}_{1, \text{max}}}
\]

where: \(X_{\delta, \text{max}}\) is the maximum steering-wheel angle to the left in the first lane-change manoeuvre (turn to the left), \(X_{\dot{\phi}, \text{max}}\) the maximum value of yaw velocity in the first lane-change

Fig. 8. Quantities necessary to determine car gain and response time.

Fig. 9. Time histories of steering-wheel angle and yaw velocity for XH1-Michelin tyres.

Fig. 10. Time histories of steering-wheel angle and yaw velocity for energy-A tyres.
manoeuvre (turn to the left), $T_{\dot{\psi},\max}$ the time necessary to reach maximum value of steering-wheel angle in the first lane-change manoeuvre (turn to the left) and $T_{\dot{\psi},\max}$ the time necessary to reach maximum yaw velocity in the first lane-change manoeuvre (turn to the left).

Examples of steering-wheel angle and yaw velocity curves as time function have been shown in Figs. 9 and 10. Fig. 11 shows the values of gain and response time for all the sets of tyres.

### 6.3. Conclusions from severe lane-change manoeuvre tests

The results of tests prove that the differences in the manner of driving a car with alternative tested tyres are relatively insignificant, particularly in reference to “gain” defined as maximum yaw velocity to steering-wheel angle ratio in the first stage of severe lane-change manoeuvre. The maximum relative differences of gain do not exceed 12% of the value obtained for model XH1-Michelin tyres. The highest value of gain, indicating lateral elasticity of tyres, was observed in energy-A tyre. Similar values were obtained for energy-A, energy-C, sport-C and MK770-E tyres. The lowest values of gain, on the other hand, were noticed for MK790-D tyres, but the difference in reference to model tyres (XH1-Michelin) did not exceed 5%.

The time of car response to steering-wheel angle can be considered a certain indicator for manoeuvrability assessment. The faster the response, i.e. the shorter the response time, the better manoeuvrability of the car (it follows the driver’s manoeuvres). During the experiments the vehicle response time for six sets of tyres (including the model XH1-Michelin tyre) was approximately 0.15 s–0.17 s and only in the case of energy-A tyre the response time exceeded 0.20 s and was longer by ca. 31% compared to the model tyres. Taking into account the value of gain it can be stated that energy-A tyres had the best lateral elasticity. With these tyres the driver found it most difficult to keep the straight trajectory after completion of the severe lane-change manoeuvre.

### 7. Final conclusions

The tests performed on retreaded tyres showed that different rubber compounds and tread patterns resulted in significantly different braking deceleration values.

The curved trajectory tests revealed that, compared with brand new tyres, the majority of retreaded tyres showed a lower cornering stiffness and strongly non-linear steering-wheel angle change. It can be stated, then, that the effect of retreaded tyres on vehicles active safety is noticeable.

The test results, and the differences in braking intensity in particular, have confirmed the necessity of performing braking tests on the place of accident with the same type of tyres with which the car involved in the accident was equipped.

### Acknowledgment

The tests were performed within the research project No 4T12C04826 financed by the Committee for Scientific Research in the years 2004–2005.

### References

Uncertainty of calculation results in vehicle collision analysis

Wojciech Wach*, Jan Unarski
Institute of Forensic Research, Department of Road Accident Analysis, ul. Westerplatte 9, 31-033 Krakow, Poland
Received 30 May 2006; accepted 14 June 2006
Available online 1 August 2006

Abstract
In the analysis of road accidents two types of calculation result uncertainty can be distinguished: modelling uncertainty and uncertainty in calculation results [R.M. Brach, M. Brach, Vehicle Accident Analysis & Reconstruction Methods, SAE International Publisher, Warrendale, 2005]. The problem becomes very important first of all when minor modifications of input parameters or application of different models of the phenomenon lead to a fundamentally different answer to the question posed by the court.

The aim of the paper was to prove the necessity of including the problem of uncertainty in calculations related to vehicle collision mechanics and to justify the application of different error analysis methods recommendable in vehicle collision reconstruction.

The data file from crash test No. 7 [H. Burg, M. Lindenmann, Unfallversuche, Verlag Information Ambs, Kippenheim, 1982] was used, the selection restricted to the range typical of average police records of collision place.

Collision speeds were calculated using two methods: reconstruction and simulation.

The analysis of uncertainty was carried out. Maximum and mean square uncertainty were calculated by means of total differential of relevant forms. Since the reconstruction resulted in very broad error intervals of uniform distribution, additional calculations were performed by the Monte Carlo method using algorithm described in [W. Wach, J. Unarski, Determination of vehicle velocities and collision location by means of Monte Carlo simulation method, Special Publication Accident Reconstruction SP-1999, SAE Paper No. 2006-01-0907, 2006].

© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Uncertainty; Collision; Reconstruction; Simulation; Monte Carlo

1. Introduction
The total uncertainty of calculation results of road accidents analysis is biased, in general terms, with uncertainty in measurements, calculation and modelling. Although in the literature on the subject various synonyms of the idea of “uncertainty” can be found such as “error”, “fluctuation”, “disturbance”, “variation”, etc., the experience of court proceedings prompts one to avoid using the word “error” since its popular, suggestive meaning may lead to unnecessary misunderstandings.

Uncertainty in measurements: The authors of ref. [4] point that the majority of typical measurements taken by the police officers on the place of collision corresponds to normal distribution. They illustrate this by standard deviations of measurements with a roller wheel and total station as functions of distance. The problem is that the expert does not know at which point of the bell-curve the actual single measurement was taken. It is assumed then that uncertainty in measurement is connected with the accuracy of the instrument and equals, e.g., half of the smallest scale interval even if the policeman’s measurement was wrong by 2 m (of which, obviously the expert is unaware).

Uncertainty in calculation refers to the scatter of the values of the parameters that were not measured directly but taken from tables, e.g., driver’s response time, lag time, pedestrian’s velocity, and friction coefficient. The values can frequently be adopted as normal distribution [4].

Modelling uncertainty occurs when two or more mathematical models give different results of the same physical problem [1]. Modelling uncertainty is separate and distinct from the uncertainty in the statistical nature of input variables for a given model.

In the article attention was focused on the problem of the effect of input data uncertainty on calculation results uncertainty, presenting modelling uncertainty very briefly in 5.3 Monte Carlo simulation using a vehicle dynamic model.
The simplest way of determining calculation results uncertainty is to apply the method of limits in which double calculations are performed—first time for the minimal values of each input parameter, and the second time for the maximal values. An equivalent method is the application of total differential, with each number within the given interval of equal probability in both cases.

In case of functions containing quotient and subtractable terms this simple evaluation may not be possible, and it is necessary to use more sophisticated methods, e.g., Monte Carlo simulation.

2. Sensitivity to uncertainty

A change in the value of any of the measured quantities will have some effect on the calculated value. This difference is the “sensitivity” of the equation to the uncertainty.

As an example, a single-parameter function is given

\[ y = f(x) = \frac{1}{x^2} \]  

which is plotted in Fig. 1. As can be easily seen, for the same uncertainty \( \Delta x = \text{const} \) the calculation result range \( \Delta f(x) \), corresponding to it, changes dramatically, depending on \( x \), which means that function (1) is “sensitive” to \( x \). Sensitivity (the sensitivity function, to be precise) is then an absolute value of the derivative of \( f(x) \) and for (1) is expressed by formula

\[ w(x) = \left| \frac{d f(x)}{d x} \right| = \left| - \frac{2}{x^3} \right| \]  

which is plotted with dashed line in Fig. 1.

In a general case of a function of \( n \) independent arguments

\[ y = f(x_1, \ldots, x_n) \]

sensitivity can be presented as a sensitivity vector of function \( y \) to particular parameters \( x_i, i = 1, \ldots, n \) (these can be coefficient of friction, braking distance, pedestrian’s velocity, etc.)

\[ w = \left[ \frac{\partial y}{\partial x_1} \right., \ldots, \left. \frac{\partial y}{\partial x_n} \right]^T = [w_1, \ldots, w_n]^T. \]  

(4)

Since each element of vector \( w \) is expressed in a different unit, to evaluate the contribution of particular variables in the total sensitivity it is necessary to present it as a vector of relative sensitivity

\[ w_{rel} = \left[ \frac{\partial y}{\partial x_1} \frac{x_1}{y}, \ldots, \frac{\partial y}{\partial x_n} \frac{x_n}{y} \right]^T \]  

(5)

and next normalize the vector \( w_{rel} \)

\[ w_{rel,norm} = \frac{1}{\|w_{rel}\|} w_{rel}, \]  

(6)

with e.g., the Euclidean norm

\[ \|w_{rel}\| = \sqrt{\sum_{i=1}^{n} \left( \frac{\partial y}{\partial x_i} \frac{x_i}{y} \right)^2}. \]  

(7)

Having adopted actual data for calculations

\[ x_0 = [x_{10}, \ldots, x_{n0}]^T \]

the value of function \( y \) can be calculated

\[ y_0 = f(x_{10}, \ldots, x_{n0}), \]  

(9)

vector of sensitivity coefficients \( w_0 = w|_{x=x_0} \) (4), vector of coefficients of relative sensitivity \( w_{rel0} = w_{rel}|_{x=x_0} \) (5) and normalized vector of relative sensitivity coefficients \( w_{rel,norm0} = w_{rel,norm}|_{x=x_0} \) (6) at point \( x_0 \).

3. Maximum uncertainty and mean square uncertainty

Maximum uncertainty of calculation result can be expressed as scalar product

\[ \Delta y_{\text{max}} = w_0^T \Delta x_0 = \sum_{i=1}^{n} (w_{i0} \cdot \Delta x_{i0}), \]  

(10)

where \( w_0 \) is the vector of sensitivity coefficients \( w_0 = w|_{x=x_0} \) (4) and \( \Delta x_0 \) is the vector of uncertainties of particular variables around nominal values

\[ \Delta x_0 = [\Delta x_{10}, \ldots, \Delta x_{n0}]^T. \]  

(11)

Maximum relative uncertainty

\[ \Delta y_{\text{max,rel}} = \frac{\Delta y_{\text{max}}}{y_0}. \]  

(12)

In some cases maximum uncertainty can be so high that it is profitable to restrict its range by means of mean square uncertainty

\[ \Delta y_{\text{sq}} = \|w_{10} \cdot \Delta x_{10}, \ldots, w_{n0} \cdot \Delta x_{n0}\|^T \|

= \sum_{i=1}^{n} \sqrt{(w_{i0} \cdot \Delta x_{i0})^2}. \]  

(13)
The calculations result certainly is comprised within the range $y_0 \pm \Delta y_{\text{max}}$ and with “great” probability within the range $y_0 \pm \Delta y_{\text{sqr}}$. This method is frequently used in engineering practice.

A comprehensive description of the analytical and statistical methods is beyond the scope of this paper, but can be found in most mathematics and statistics books, e.g. [5].

4. Monte Carlo simulation

Many researchers pointed out the great potential in using probabilistic methods, in particular Monte Carlo simulation, in the analysis of uncertainties of calculation results [1,6–11]. The Monte Carlo method typically is applied to an existing, deterministic mathematical model of a physical system—Monte Carlo method randomly and systematically executes the model and determines the results of the simulation for a large number of randomly selected input parameters. It is assumed that these parameters are statistically uncorrelated and the statistical distribution of the input parameters must be known or assumed.

The calculations are performed at least several hundred thousands times, due to which the calculation result is given as a probability distribution, usually in the shape similar to normal distribution.

5. Example and discussion

Accident reconstruction calculations were performed the aim of which was to determine the collision velocities of cars. The data file from crash test No. 7 [2] were used, the selection restricted to the range typical of average police records of collision place (Fig. 2).

The example is to illustrate calculations results uncertainty caused by the input data uncertainty employing different methods of analysis.

5.1. Data

The input data have been given in Table 1. The table does not include the uncertainty of data which were not covered in calculation models, and the effect of which on the result is negligibly small.

5.2. Calculations with classical models

5.2.1. Separation speed

Vehicles separation speeds were determined using McHenry–Marquard’s theoretical–empirical dependencies, derived from energy–work equivalence principle [12,13]. Tire frictional impulses are ignored during the impact phase. At some time instant during the impact, the contact point on both vehicles reaches a common velocity. The wheels in the post-crash movement remain in contact with the ground. The vehicles spin out to rest with constant rolling resistances, no active steering, and over a single friction surface.

Vehicles angular velocities at separation

$$\omega_j^{(\text{sep})} = \text{sgn} \Delta \psi_j$$

$$\cdot \sqrt{\left(I_j/(m_j \cdot L_j) \cdot |\Delta \psi_j| \cdot (1 - f_j) + (s_j/1.7)\right)}$$

$$j = 1, 2,$$

where $j$ is the number of vehicle, $\Delta \psi_j$ the total angle of rotation in post-impact motion (rad), $\mu_j$ the friction coefficient, $g$ the
gravity coefficient, 9.81 m/s², \( I_j \) the moment of inertia (kg m²), \( m_j \) the real mass (m), \( L_j \) the wheel base (m), \( f_j \) the coefficient of post-impact motion drag (from 0 for all wheels freely rotating to 1 for all wheels locked) and \( s_j \) is the distance between CGs of vehicles in impact and rest positions (m).

Speeds of vehicles’ CGs at separation

\[
v_j^{\text{sep}} = 1.7 \cdot \left[ \mu_j \cdot g \cdot |\Delta \varphi_j| - \frac{I_j \cdot |\omega_j^{\text{sep}}| \cdot (1 - f_j)}{m_j \cdot L_j} \right],
\]

\( j = 1, 2. \)

Substituting the nominal values from Table 1 the nominal values of separation speeds are

\[
\begin{align*}
\omega_1^{\text{sep}} &= 1.0 \text{ rad/s}, \\
\omega_2^{\text{sep}} &= -0.2 \text{ rad/s}, \\
v_1^{\text{sep}} &= 8.2 \text{ m/s} = 29.5 \text{ km/h}, \\
v_2^{\text{sep}} &= 12.9 \text{ m/s} = 46.4 \text{ km/h}.
\end{align*}
\]

5.2.2. Pre-impact speed

To avoid too much complication, impact modelling has been limited in the article only to the principle of momentum conservation. However, an analogous way can be applied to perform the calculations with the principle of momentum-impulse, taking into account the restitution coefficient and sliding energy loss over the intervehicular surface during contact duration.

Angular velocities at impact are 0.

Pre-impact speed of vehicle 1

\[
v_1^{\text{imp}} = \frac{m_1 \cdot v_1^{\text{sep}} \cdot \sin(\alpha_1^{\text{sep}}) - \alpha_2^{\text{imp}} - \alpha_1^{\text{imp}})}{m_1 \cdot \sin(\alpha_1^{\text{imp}}) - \alpha_2^{\text{imp}}},
\]

where \( \alpha_j^{\text{imp}}, j = 1, 2 \) is the direction of vehicle \( j \) CG separation velocity vector and \( \alpha_j^{\text{sep}} \) is the direction of vehicle 1 CG velocity vector at impact.

After calculations the nominal values of impact speeds are

\[
\begin{align*}
v_1^{\text{imp}} &= 24.1 \text{ m/s} = 86.8 \text{ km/h}, \\
v_2^{\text{imp}} &= 0.3 \text{ m/s} = 1.2 \text{ km/h}.
\end{align*}
\]

For comparison, the real values measured in crash test were 87.7 km/h and 0 km/h, respectively (cf. Table 1).

5.2.3. Uncertainty

Incorporating (14) and (15) in formulae (16) and (17), impact speeds are functions of the following independent variables:

\[
\begin{align*}
v_1^{\text{imp}} &= v_1^{\text{imp}}(\mu_1, I_1, m_1, f_1, s_1, \mu_2, I_2, m_2, f_2, s_2, \alpha_1^{\text{sep}}, \\
& \hspace{1em} \alpha_2^{\text{imp}}, \alpha_1^{\text{imp}}, \alpha_2^{\text{imp}}), \\
v_2^{\text{imp}} &= v_2^{\text{imp}}(\mu_1, I_1, m_1, f_1, s_1, \mu_2, I_2, m_2, f_2, s_2, \alpha_1^{\text{sep}}, \\
& \hspace{1em} \alpha_2^{\text{imp}}, \alpha_2^{\text{imp}}).
\end{align*}
\]

As can be easily noticed, with such complex formulae performing double calculations, first for the minimum, second for the maximum values of the input parameters from Table 1 will not produce information on the maximum uncertainty of the result. Therefore, an error analysis is necessary, employing the sensitivity coefficients.

Following differentiation of (18) after formula (4) (with \( x = x_0 \), vectors of sensitivity coefficients: \( w_{10} \) for \( v_1^{\text{imp}} \) and

| Table 1
<table>
<thead>
<tr>
<th>Input data (test No. 7 in [2])</th>
<th>1. Ford Taunus 1300</th>
<th>2. Opel Commodore GS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length, ( d_j ) (m)</td>
<td>4.28</td>
<td>4.54</td>
</tr>
<tr>
<td>Width, ( B_j ) (m)</td>
<td>1.705</td>
<td>1.765</td>
</tr>
<tr>
<td>Wheel base, ( L_j ) (m)</td>
<td>2.58</td>
<td>2.672</td>
</tr>
<tr>
<td>Track front, ( l_j ) (m)</td>
<td>1.422</td>
<td>1.424</td>
</tr>
<tr>
<td>Track rear, ( l_j ) (m)</td>
<td>1.422</td>
<td>1.420</td>
</tr>
<tr>
<td>Distance from front axle to CG, ( c_j ) (m)</td>
<td>1.14 ± 0.10</td>
<td>1.24 ± 0.10</td>
</tr>
<tr>
<td>Friction coefficient, ( \mu_j )</td>
<td>0.7 ± 0.1</td>
<td>0.7 ± 0.1</td>
</tr>
<tr>
<td>Post-impact motion drag coefficient, ( f_j )</td>
<td>0.7 ± 0.1</td>
<td>0.6 ± 0.2</td>
</tr>
<tr>
<td>Mass, ( m_j ) (m)</td>
<td>990 ± 30</td>
<td>1220 ± 30</td>
</tr>
<tr>
<td>Moment of inertia, ( I_j ) (kg m²)</td>
<td>1387 ± 200</td>
<td>1877 ± 200</td>
</tr>
<tr>
<td>Total rotation angle in post-impact motion, ( \Delta \varphi_j ) (°)</td>
<td>42°</td>
<td>−13°</td>
</tr>
<tr>
<td>Distance between CG in initial and final position, ( s_j ) (m)</td>
<td>6.0 ± 0.2</td>
<td>14.3 ± 0.2</td>
</tr>
<tr>
<td>End speed, ( v_j^{\text{end}} ) (km/h)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Direction of vector of vehicle 1 impact velocity, ( \alpha_1^{\text{imp}} ) (in ( {1} )), °</td>
<td>0° ± 2.0°</td>
<td>0° ± 2.0°</td>
</tr>
<tr>
<td>Direction of separation velocity vector, ( \alpha_j^{\text{sep}} ) (in ( {1} )), °</td>
<td>1.68° ± 3.0°</td>
<td>0.63° ± 3.0°</td>
</tr>
<tr>
<td>Pre-impact speed ( v_j^{\text{imp}} ) measured in test (km/h) (for comparison)</td>
<td>87.7</td>
<td>0</td>
</tr>
</tbody>
</table>
For vehicle 1, for example 1 (see Fig. 3)

\[ g(v_{1\text{imp}}) = \begin{cases} 
0.046 \text{ h/km}, & v_{1\text{imp}} \in (76.0, 97.6) \text{ km/h} \\
0, & v_{1\text{imp}} \notin (76.0, 97.6) \text{ km/h}.
\end{cases} \quad (20) \]

Dimensionless, normalized coefficients of relative sensitivity \( w_{1\text{rel.norm0}}^{(\text{imp})} \) and \( w_{2\text{rel.norm0}}^{(\text{imp})} \) make it possible to evaluate the effect of particular variables on the results of calculations of pre-impact speeds, which was impossible with the corresponding vectors \( w_{10}^{(\text{imp})} \) and \( w_{20}^{(\text{imp})} \). For instance, the effect of distance of CG \( s_{1j} \) on speed \( v_{j\text{imp}}^{(\text{imp})} \) is 0.165/0.006 = 28 times stronger than that of moment of inertia \( I_{1j} \), hence this parameter should be measured with great care. Table 2 presents calculations results.

For comparison, the real pre-impact speeds of vehicles 1 and 2 were 87.7 km/h and 0 km/h, respectively.

The probability density function of impact speed within the ranges defined by deviation from the expected value \( v_{j\text{imp}}^{(\text{imp})} = v_{j\text{max}}^{(\text{imp})} \), \( j = 1, 2 \) is uniform and described by dependence

\[ g(v_{j\text{imp}}^{(\text{imp})}) = \begin{cases} 
\frac{1}{2\Delta v_{j\text{max}}^{(\text{imp})}}, & v_{j\text{imp}}^{(\text{imp})} \in (v_{j\text{min}}^{(\text{imp})} - \Delta v_{j\text{max}}^{(\text{imp})}, v_{j\text{min}}^{(\text{imp})} + \Delta v_{j\text{max}}^{(\text{imp})}) \\
0, & v_{j\text{imp}}^{(\text{imp})} \notin (v_{j\text{min}}^{(\text{imp})} - \Delta v_{j\text{max}}^{(\text{imp})}, v_{j\text{min}}^{(\text{imp})} + \Delta v_{j\text{max}}^{(\text{imp})})
\end{cases} \quad (19) \]

The impact speed of vehicle 1 certainly is comprised within the limits \( v_{1\text{imp}}^{(\text{imp})} = v_{1\text{max}}^{(\text{imp})} \), i.e., 76.0–97.6 km/h, while the likelihood for it to be comprised in the range, e.g., from 80 km/h to 85 km/h is 5 × 0.046 = 0.23 (the area below the corresponding part of graph \( g(v_{1\text{imp}}^{(\text{imp})}) \)).

Analogous dependencies are valid for mean square uncertainty.

### 5.2.4. Monte Carlo Simulation

Each speed within the ranges defined by the maximum uncertainty \( v_{j\text{imp}}^{(\text{imp})} = \Delta v_{j\text{max}}^{(\text{imp})} \), \( j = 1, 2 \) or mean square uncertainty \( v_{j\text{imp}}^{(\text{imp})} = \Delta v_{j\text{max}}^{(\text{imp})} \), \( j = 1, 2 \) is equally probable. This is why Monte Carlo simulation was performed, employing the algorithm described in [3]. Since the same deterministic, mathematical
model of impact was used (18), the expected values of speed \( v_{\text{imp}}(j) \), \( j = 1, 2 \) are obviously identical as before. However, probability density functions in the form of bell-curves were obtained of shape close to normal distribution (Figs. 4 and 5). These distributions narrow down the area of uncertainty of speed calculations results \( v_{\text{imp}}(j) \), \( j = 1, 2 \), which enables their statistical interpretation (e.g. determination of probability with which the speed was within each range the court might ask about).

Presentation of results in one-dimensional charts can be used improperly (e.g. to justify one’s reasoning in court) by treating arbitrarily chosen values selectively, in a non-correlated way. This is why it is proper to present 2D charts in which speeds \( v_{\text{imp}}(1) \) and \( v_{\text{imp}}(2) \) are interdependent (Figs. 6 and 7).

The probability for impact speed of vehicle 1 to be comprised in the limits from \( a \) to \( b \), and from \( c \) to \( d \) for vehicle 2 is

\[
P_{v_1,v_2} = \int_a^b \int_c^d f\left(v_{\text{imp}}^{(1)},v_{\text{imp}}^{(2)}\right) \, dv_{\text{imp}}^{(2)} \, dv_{\text{imp}}^{(1)}. \tag{21}
\]

The results were presented also as 2D distribution function (Fig. 7), which facilitates interpretation—for instance, from this chart it follows that the probability of vehicles 1 and 2 impact speeds of less than 88 km/h and 3 km/h, respectively, is 0.7, i.e. 70%.

### 5.3. Monte Carlo simulation using a vehicle dynamic model

Calculations have been performed in PC-Crash program using a dynamic model of a vehicle of six degrees of freedom [14]. The same data have been used as before with a set of additional necessary parameters.

---

**Table 2**

<table>
<thead>
<tr>
<th>No. of vehicle ( j )</th>
<th>Nominal value ( v_{\text{imp}}^{(j)} )</th>
<th>Maximum uncertainty after (10) ( \Delta v_{\text{imp}}^{(j)_{\max}} )</th>
<th>Maximum relative uncertainty after (12) ( \Delta v_{\text{imp}}^{(j)_{\max,\text{rel}}} )</th>
<th>Mean square uncertainty after (13) ( \Delta v_{\text{imp}}^{(j)_{\text{sqr}}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24.1 m/s = 86.8 km/h</td>
<td>3.0 m/s = 10.8 km/h</td>
<td>0.12</td>
<td>1.4 m/s = 5.0 km/h</td>
</tr>
<tr>
<td>2</td>
<td>0.3 m/s = 1.2 km/h</td>
<td>1.1 m/s = 4.0 km/h</td>
<td>3.66</td>
<td>0.7 m/s = 2.7 km/h</td>
</tr>
</tbody>
</table>

---

---

---
The idea of Monte Carlo simulation in this case was to perform iterative calculations which included pre-collision velocities prediction, calculation of separation speed and simulation of vehicle motion. The adopted criterion of calculation correctness was the difference between the real and simulated traces and end positions of vehicles. The method described in [15] was applied. The result was the quality function versus pre-crash speeds (Fig. 8) (interpretation, see Table 3).

The vehicle dynamic model gives an incomparably more comprehensive view on the effect of various parameters on impact speeds than a simple equation of momentum taking into account Marquard–McHenry quasi-empirical dependencies.

The fact that the expected value of vehicle 1 speed obtained in classical calculations (86.8 km/h) is closer to the real value (87.7 km/h) than that obtained using PC-Crash (84.0 km/h) should be treated as accidental.

The problem of modelling uncertainty has been mentioned here only briefly, its analysis exceeding the scope of the present article.

6. Conclusion

(1) The simplest way to determine calculations result uncertainty is to apply the method of limits. In case of functions containing quotient and subtracting elements it is necessary to apply methods that use sensitivity coefficients. The probability distribution of the results in both cases is uniform.

(2) It is advantageous to apply the Monte Carlo method, which gives the results as the distribution of probability density in the shape similar to normal distribution.

(3) The results of calculations of impact speeds of both vehicles are correlated and that is why they should be presented as 2D probability density function or 2D distribution function.

(4) Calculations are only part, not the crucial one at that, of road accident analysis. Their importance is significant only if
they can help to identify the reference points necessary to evaluate human behaviour. The problem of uncertainty in calculations, measurements and modelling is undoubtedly very important, however, what experts differ about most is not controversial calculations, but the evaluation of drivers’ behaviour.

Acknowledgement

The study was financially supported by the Institute of Forensic Research in Kraków with in project No II/W.

References

Evidence evaluation in fingerprint comparison and automated fingerprint identification systems—Modelling within finger variability

Nicole M. Egli *, Christophe Champod, Pierre Margot

Ecole des Sciences Criminelles, Institut de Police Scientifique, Batochime, Quartier Sorge, Université de Lausanne, CH-1015 Lausanne-Dorigny, Switzerland

Received 9 June 2006; accepted 14 June 2006
Available online 17 August 2006

Abstract

Recent challenges and errors in fingerprint identification have highlighted the need for assessing the information content of a papillary pattern in a systematic way. In particular, estimation of the statistical uncertainty associated with this type of evidence is more and more called upon.

The approach used in the present study is based on the assessment of likelihood ratios (LRs). This evaluative tool weighs the likelihood of evidence given two mutually exclusive hypotheses. The computation of likelihood ratios on a database of marks of known sources (matching the unknown and non-matching the unknown mark) allows an estimation of the evidential contribution of fingerprint evidence.

LRs are computed taking advantage of the scores obtained from an automated fingerprint identification system and hence are based exclusively on level II features (minutiae). The AFIS system attributes a score to any comparison (fingerprint to fingerprint, mark to mark and mark to fingerprint), used here as a proximity measure between the respective arrangements of minutiae. The numerator of the LR addresses the within finger variability and is obtained by comparing the same configurations of minutiae coming from the same source. Only comparisons where the same minutiae are visible both on the mark and on the print are therefore taken into account. The denominator of the LR is obtained by cross-comparison with a database of prints originating from non-matching sources. The estimation of the numerator of the LR is much more complex in terms of specific data requirements than the estimation of the denominator of the LR (that requires only a large database of prints from a non-associated population). Hence this paper addresses specific issues associated with the numerator or within finger variability.

This study aims at answering the following questions: (1) how a database for modelling within finger variability should be acquired; (2) whether or not the visualisation technique or the choice of different minutiae arrangements may influence that modelling and (3) what is the magnitude of LRs that can be expected from such a model. Results show that within finger variability is affected by the visualisation technique used on the mark, the number of minutiae and the minutiae configuration. They also show that the rates of misleading evidence in the likelihood ratios obtained for one of the configurations examined are low.

# 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Fingerprint evaluation; Likelihood ratio; Within-variability

1. Introduction

In law enforcement, the fingerprint used to be the most sought after evidence from crime scene investigation, for several reasons. Firstly, fingerprints are extremely variable between individuals, so that even the partial and smudged impressions lifted from crime scenes can be used to identify the donor of these marks. Secondly, chemical and physical detection techniques applied on relevant exhibits have allowed through their selectivity and sensitivity to increase the number and quality of marks detected and useful for comparison purposes. Finally, classification of reference collections of prints was quite straightforward, and allowed the identification of known suspects previously fingerprinted and from the 1970s algorithmic research has brought to the field efficient search algorithms that have been implemented in most nations, in the form of automated fingerprint (or palmprint) identification systems (AFIS). These systems allow the successful search of the finger or palm area having left a partial mark among millions of ten-print cards. AFIS produce a ranked list of ‘best matches’ candidates, based on proximity measures called scores. Their computation is generally based on the positions of minutiae as well as their directions.
These scores will be used in this study. Our system outputs higher scores indicating a better ‘match’ (or a smaller algorithmic distance) between two compared configurations of minutiae. When AFIS systems are used operationally, the computed score is only used as a generator of a promising list of say 15 potential donors. The operator then carries out a usual fingerprint comparison from this suggested list as if these potential donors had been suggested through conventional police inquiry.

Recent cases of misidentification, combined with a more strident legislative scrutiny on identification evidence fields (mainly as a consequence of the Daubert admissibility criteria), have led commentators to call for a more articulated underpinning, including from a statistical perspective, of the identification process [1–3].

In this paper, we propose to move towards a probabilistic approach [4] to fingerprint evidence by deriving benefit from the score information from an AFIS system.

Several statistical studies have already been carried out, and are commented in Ref. [5]. The first step for modelling probabilities associated with level II features is to measure distances (in some sense) between minutiae configurations. For that purpose, AFIS scores will be used a proxy for the measure of similarity between two configurations. Secondly, in any model for fingerprint feature evaluation, the fact that no two appositions of a given papillary surface are exactly alike needs to be considered very closely. It is only the consideration of this within mark variability combined with the probability of a random association with a unrelated source that allows to assess the weight of evidence associated with a set of matching level II features. This is fully achieved when adopting a likelihood ratio approach such as proposed in recent model development [6].

The evaluative tool investigated here is applicable in a very precise context: a mark has been compared to a fingerprint, and no discordance, which should lead to exclusions, has been noted. In this context, the minutiae, which have been found in concordance between the mark and the print, are considered as known, and it is the evaluation of this concordance that can be carried out by the present approach. The number of minutiae and their topological configuration is therefore known.

The likelihood ratio (LR) can be expressed by the following ratio of probabilities:

\[ LR = \frac{P(E|H, I)}{P(E|\overline{H}, I)} \]

where \( E \) is the evidence or findings, observed between a mark and a given print; \( H \) the proposition that the same finger is at the origin of the print as well as the mark and; \( \overline{H} \) is the alternative, that the mark originates from another unknown finger than the finger at the source of the print. \( I \) is any relevant background information which may impact on the probability of the evidence, such as for example information regarding the gender or ethnicity of the donor of the mark.

In the approach presented here, an AFIS is used in order to obtain the ‘evidence’. The scores employed by the system are the proximity measure between two minutiae configurations.

The likelihood ratio therefore becomes:

\[ LR = \frac{p(s|H, I)}{p(s|\overline{H}, I)} \]

where \( s \) is now the score obtained for the comparison between the mark and the considered print. \( H, \overline{H} \) and \( I \) have the same definition as before. Note here that we are assessing the LR using a ratio of densities for the score value obtained. The estimation of these two density functions is at the heart of this approach and this paper will deal firstly with the numerator density and secondly with the denominator, starting with a configuration of 6 minutiae.

This study aims at answering the following questions: (1) how a database for modelling within finger variability should be acquired; (2) whether or not the visualisation technique or the choice of different minutiae arrangements may influence that modelling and (3) what is the magnitude of LRs that can be expected from such a model.

2. Material and methods

To model within finger variability and address question (1) and (2), three sources of images have been considered:

– 704 livescan images of a right thumb have been acquired using a Single Finger Scanner (livescan) ACCO 1394 from Smiths Heimann Biometrics. These images have been acquired to cover extreme distortion allowed by the flexibility of the skin.
– 88 pseudo-marks have been acquired of the same right thumb, where 22 were visualised using each of four frequently used detection methods: DFO (1,8-diaza-9-fluorenone), ninhydrin, cyanoacrylate fuming and powder.
– 15 ten-print cards of the same donor have been used.

Fig. 1. Illustration of the 6 peripheral minutiae.
On all these images, 6 minutiae have been selected initially positioned in the periphery of the fingerprint pattern (Fig. 1).

In a second step, the number of selected minutiae has been increased up to 10 minutiae to verify if this factor affects the scores of within finger variability (Fig. 2).

Finally, another configuration of 6 minutiae has been chosen (Fig. 3), in another region of the fingerprint (towards the centre), and similarly, the number of minutiae has been increased up to 10 (Fig. 4).

To model the variability of the score when a mark is compared to non-matching sources (denominator of the likelihood ratio), a sample of approximately 10,000 ten-print cards has also been introduced.

All images used are at a resolution of 500 dpi and in greyscale, in tiff format.

A Sagem DMA AFIS has been used for the computation of scores. Minutiae have been placed automatically for marks as well as for prints, using manual correction where needed, as in normal operational use. In other words, all matching minutiae used in this study have been checked manually.

3. Results

An optimized sample size of 66 images from the 704 livescans has been determined using progressive subsampling. Indeed 66 images, compared to one inked print, correspond to the minimal number of images per sample, where two independent samples still show the same shape and estimators.

A model has been estimated from the scores obtained using these images, and a Weibull distribution has been fitted (Fig. 5).

The tails of the distribution do not fit well. However, when comparing repeated samples of livescans in the same way, the departure is at least as evident, even though they are issued from one distribution. This shows that these effects in the tails of the distribution are only due to sample size.

Using the marks instead of livescan images, and the same 6 minutiae, results show that these scores can also be modelled using a Weibull distribution (Fig. 6). These scores are obtained by comparing the marks to 15 inked impressions of the same finger.

The mode of the distributions from livescan images coincides with the mode of the distribution obtained for
marks. However, a greater variance is observed for the scores from livescans, which in turn are based on much fewer observations than the marks, as well as being influenced by extreme distortion.

A box plot by visualisation method shows no significant differences between visualisation methods and livescans, except the higher variability, which is to be explained by acquisition technique (see Fig. 7), a livescan with a donor attempting to distort his finger. Furthermore, higher variability may be very much appropriate for the estimation of likelihood ratios in casework, since the support of the mark is not necessarily flat, nor is there necessarily no distortion present in the mark itself.

Friedman’s test has been used in order to see if there are significant differences between marks visualised using different detection techniques. This hypothesis has not been rejected. It has also been used to check if there is a difference between the scores obtained for the different marks when compared to the different inked prints. Here again, the hypothesis that there is a difference has not been rejected.

In conclusion, these data provide supporting evidence for using livescans in order to model within finger variability. For the study, all detection methods as well as multiple ten-print cards will be retained.

In a next step, the minutia number has been increased: 7, 8, 9, and finally 10 minutiae have been noted on the marks, using an incremental process starting from the 6 original minutiae. Minutiae were chosen according to their proximity to the initial configuration.

Fig. 8 shows how within finger variability evolves with these changes for the first peripheral configuration of minutiae.

As expected, the scores increase as more minutiae are considered. Another trend in the data is that the variance of these scores increases with each added minuita. Since more
error in the positioning (even manual) of the minutiae is introduced, this is not surprising.

When the group of minutiae is changed, scores also increase with the number of minutiae (Fig. 9).

The fit to the Weibull distribution, with each added minutiae, is well supported by the data (see Fig. 10 for 7 minutiae in the second, central, region).

For 6, 8, 9 and 10 minutiae, results are similar. In the peripheral region, the fit to a Weibull distribution observed is less optimal. The reason for this is that minutiae were often less similarly designated between the mark and the print.

The difference between minutiae configurations of a given number of minutiae in different regions of the fingerprint has been investigated; formal tests (Kolmogorov–Smirnoff test for equal distribution between two samples) have rejected the hypothesis of an absence of differences for all except the two 6-minutiae configurations. A difference between distributions of scores issued from comparisons between minutiae configurations, which have the same number of minutiae, is therefore a reality. However, since there are some doubts about the origin of some deviations in the peripheral minutiae configuration, this will be investigated further once this origin is known.

Tippett plots [7] have been used in order to check if modelling choices are admissible on the original minutia configuration. The between-finger variability is modelled using a lognormal distribution [8] that is fitted to the scores of the mark in question (the latent evaluated for a given likelihood ratio) when compared with the 100,000 fingerprints from unrelated sources.

The Tippett plot for 7 minutiae in the peripheral region is shown in Fig. 11.

Very few LRs above 1 are observed (0.25% of comparisons between different fingers), on this database when truly the print is not at the source of the mark. In the majority of cases, the evidence provides strong to very strong support for the alternative when truly this alternative is established. Conversely, when the proposition is true (in other words when the mark truly originates from the finger at the source of the print), the evidence provides almost always evidence for that proposition. The proportion of cases where the evidence is supporting the alternative when the matching proposition is true is limited (1.90%). This allows to conclude that the system is quite reliable in the given context: within finger variability, in particular, has been modelled using repeated pseudo-marks from the same finger as the latent-print comparison being evaluated.

4. Case example

Although our results constitute only a first step towards an operational system, we felt it was important to illustrate the operational use of such a model using a case example. It will stress two important points: first, once the two density functions available, the computation is straightforward and can be undertaken in the background and second, the computation process will not heavily distract fingerprint examiners from their usual practice. For each comparison a full ACE-V protocol
will have to be followed forcing the examiner to consider the concordant features and potential discrepancies. It is only when at that stage that the hypothesis of common source cannot be ruled out that a computation of a LR based on the considered features can be undertaken. For example, let us assume the following mark to be compared to a print (Fig. 12).

The comparison process has highlighted 10 minutiae in agreement, and no discordance which would lead to an exclusion.

Let us assume for the sake of the argument that an individualisation cannot be achieved due to the limited information available in terms of quantity and quality (otherwise, the statistical argument is irrelevant).

In the AFIS system used, the first finger on the list is that precise left index, with a score of 4192.

The numerator density (reflecting on the within finger variability) is obtained from the 10 minutiae configuration in the center discussed above, assuming that the above data offer a generalisation capability across fingers and individuals. Within finger variability is a Weibull with parameters (4573.6, 9.8).

The denominator density (reflecting between fingers variability) is obtained by considering all other scores in the database (from non-matching prints), and is modelled using a lognormal distribution with parameters (7.3562, 0.1634). The parameters have been directly estimated from the scores obtained by comparing the configuration of minutiae on the mark with all prints on the database. See Fig. 13 for an illustration of the fit of the lognormal distribution to the scores obtained for this mark, and Fig. 14 for an illustration showing

![Fig. 12. The mark (a) and the print (b) considered in the case example.](image)

![Fig. 13. Lognormal probability plot for between finger variability on 7 minutiae in the peripheral region.](image)
within and between finger variability as well as the score obtained in this comparison.

The likelihood ratio is

\[
LR = \frac{p(s|H,I)}{p(s|\overline{H},I)} = \frac{6.498 \times 10^{-4}}{7.5927 \times 10^{-12}} = 8.56 \times 10^7
\]

which means that the evidence provides very strong support for the view that mark and print come from the same source.

5. Discussion and conclusion

The Weibull distribution used for the modelling of within-finger variability had been mentioned previously in the literature and discarded in favour of a Gamma distribution [8]. However, in the present case, it is the model that fits the best the data, and has therefore been retained.

It has been shown that

(1) Within-finger variability, in the sense used here, i.e., the variability observed when the same minutiae are used, can be modelled using a Weibull distribution.

(2) Marks can be replaced by livescans. This has strong operational impacts. Indeed livescan images are much easier to obtain that casework realistic marks.

(3) An increase of the minutia number lead to an increase of the scores, as well as the variance of the numerator distributions.

(4) A change of the minutia configuration triggers a change of the scores distribution for a given number of minutiae.

(5) The addition of a minutia does not cause a continuous increase of the scores of the within finger variability; it depends on the minutiae present before and the one added.

An evaluation of the likelihood ratios obtained using the adopted modelling strategy, using a Tippett plot, indicates a very good performance of the system. This warrants further research about variations observed when minutiae are chosen in the delta area, when the finger, the general pattern and the donor are changed, as well as an investigation of between-finger variability. Ideally we would hope that the within finger distribution has identified and known characteristics across fingers and areas. Should that not be the case, the promising results obtained using a livescan device testify to the feasibility of modelling within finger variability on a case-by-case basis.

Finally, the case example demonstrates a way of applying this evaluative tool to an actual fingerprint comparison.

References


Nicephor[e]: A web-based solution for teaching forensic and scientific photography


Ecole des Sciences Criminelles, Université de Lausanne, Bâtiment Batochine, Lausanne CH-1015, Switzerland

Received 9 June 2006; accepted 14 June 2006
Available online 18 July 2006

Abstract

Nicephor[e] is a project funded by “Swiss Virtual Campus” and aims at creating a distant or mixed web-based learning system in forensic and scientific photography and microscopy. The practical goal is to organize series of on-line modular courses corresponding to the educational requirements of undergraduate academic programs. Additionally, this program could be used in the context of continuing educational programs. The architecture of the project is designed to guarantee a high level of knowledge in forensic and scientific photographic techniques, and to have an easy content production and the ability to create a number of different courses sharing the same content. The e-learning system Nicephor[e] consists of three different parts. The first one is a repository of learning objects that gathers all theoretical subject matter of the project such as texts, animations, images, and films. This repository is a web content management system (Typo3) that permits creating, publishing, and administrating dynamic content via a web browser as well as storing it into a database. The flexibility of the system’s architecture allows for an easy updating of the content to follow the development of photographic technology. The instructor of a course can decide which modular contents need to be included in the course, and in which order they will be accessed by students.

All the modular courses are developed in a learning management system (WebCT or Moodle) that can deal with complex learning scenarios, content distribution, students, tests, and interaction with instructor. Each course has its own learning scenario based on the goals of the course and the student’s profile. The content of each course is taken from the content management system. It is then structured in the learning management system according to the pedagogical goals defined by the instructor. The modular courses are created in a highly interactive setting and offer autoevaluating tests to the students.

The last part of the system is a digital assets management system (Extensis Portfolio). The practical portion of each course is to produce images of different marks or objects. The collection of all this material produced, indexed by the students and corrected by the instructor is essential to the development of a knowledge base of photographic techniques applied to a specific forensic subject. It represents also an extensible collection of different marks from known sources obtained under various conditions. It allows to reuse these images for creating image-based case files.

© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Forensic photography; Imaging; Training; e-Learning; Knowledge management

1. Introduction

The School of Criminal Sciences (ESC) of the University of Lausanne, the Federal School of Technology of Lausanne (EPFL), and the University of Basel develop an e-Learning framework about scientific and forensic photography named Nicephor[e]. This sits within the framework of the federal initiative Swiss Virtual Campus (http://www.virtualcampus.ch/), which is dedicated to promote high-quality e-learning to Swiss universities based on innovative information and communication technology. This paper presents the lessons learned in creating e-learning courses applied to forensic and scientific photography with theory and practical work, the solution developed, and the results obtained after 2 years of development.

The main objective of the project is to propose appropriate modular courses in scientific and forensic photography corresponding to the educational requirements of undergraduate academic programmes and continuing education. Flexibility and sustainability of the system are key conditions to achieve this goal. For this purpose, hypotheses and global strategies along four different axes have been defined [1].

* Corresponding author. Tel.: +41 21 692 46 00; fax: +41 21 692 46 05;
E-mail address: romain.voisard@unil.ch (R. Voisard).
1.1. The pedagogical axis

The pedagogical concept is based on blended learning, where the on-line course material serves as a platform to reinforce, expand, and use the theoretical ex-cathedra knowledge. For specific needs, such as with continuing education programmes, the blended learning scenario can be adapted to a full distant learning experience depending on students’ characteristics.

The web learning platform is currently integrated at the ESC in all general photographic or microscopic courses. It also covers all the photographic aspects of courses specific to a given forensic topic (scene of crime, fingerprints, microtraces, etc.). A course on digital photography has been also deployed at EPFL. This extended integration of the on-line initiative in various teaching settings delivered by a wide range of teachers promotes a culture of acceptance of e-learning activities within the institutions.

1.2. The organisational axis

This initiative had significant impact on the organisations, in particular at the ESC. Beyond the material investments (cameras, image management system, and IT) to support the project, Nicephor[e] acts as a catalyst for important changes in the way forensic images are handled in a training environment (currently in the area of crime scene investigation, fingerprints, and footwear marks). Our aim is to develop such a strategy to other fields or even across institutions. It is a move toward a system based on shared knowledge by taking advantage of training material produced by staff and from students.

1.3. The technological axis

The architecture of the system guarantees ease of use and maintenance, flexibility, and interoperability. Technological choices fulfill these criteria and benefits of stability over time by the use of SCROM/IMS specifications dedicated specifically to e-learning content. The specific options will be describe below.

1.4. The economical axis

The modularity of the course material allows for building up training programmes for varied and wider potential audiences (BSc, MSc, and continuing education). For example, the ESC has a long tradition of offering continuing education programmes to scientists working for forensic services in Switzerland and abroad. It is planned that by 2007, the ESC will be able to offer a full workshop in forensic photography using Nicephor[e], mostly in a distant learning setting.

2. Methods and technology

2.1. Strategy of development

The development team is made up of six researchers supervised by a steering committee composed of the professors involved. Based on the course material of the professors, the development team has to create the on-line material, such as texts, illustrations, animations, films, and e-learning solutions. To organise the work of each researcher, a global strategy of development have been defined (Fig. 1).

This strategy is divided in five steps. The first one is carried out only once. This crucial step determines the technical and educational directions that will be followed during the entire length of the project specifying the requirements, limitations, and objectives. Steps 2–4 are iterative procedures of content production and validation, and course implementation and evaluation. Unlike the first step, this process of production is applied to several courses in different institutions. As technology changes extremely rapidly in digital photography, the content production and technical development are planned to be extensible to keep the courses up to date. For each course, pedagogical workshops are used to explore new teaching and pedagogical strategy and implement it in a learning management system (LMS). Finally, the last step consists in the integration of the courses into the curricula of the respective institutions.

2.2. Architecture of the system

At the beginning of the project several solutions were considered to build an e-learning system. It was clear that taking advantage of an existing learning management system featuring specialized educational tools was essential. Unfortunately, those systems are not sufficient to create or truly manage content.

Fig. 1. Strategy of courses production in five main steps.
and are rather incompatible between each other. Building our own system was not an appropriate solution mostly for development time availability. Therefore, the adopted answer to our requirements had to be a framework of interconnected systems, each one dedicated to some specialized functions. Actually, the e-learning system Nicephor involves the three main following systems:

- A content management system (CMS) as repository of learning objects (LOs).
- A learning management system (LMS) as organizer and distributor of LOs in learning modules that can be used in different learning contexts.
- A digital assets management system (DAM) as extensible images collection and knowledge base of forensic photographic techniques.

2.3. The content management system

The learning objects’ repository of Nicephor requires features linked with content creation/edition/publication as well as with editors’ administration/collaboration. The most important functions are as follows:

- Easy-of-use edition, publication, and administration of content.
- Separation of content, structure, and design (dynamic web content).
- Advanced security and user management.
- Support of multi-user and multi-language content.
- Management of workflow and validation process.
- Direct exportation of content and structure into HTML pages or SCORM content packages.

The tool used in this project is Typo3. It is a free Open Source CMS that offers full flexibility and extendability while featuring an accomplished set of ready-made interfaces, functions, and modules. Typo3 is based on a core built on PHP scripts that stores data into a MySQL base. A large range of third-party extensions can be installed or created to adapt the functionality of the system to the specific project requirements.

Some specific extensions had to be developed to satisfy the needs for validation process and exportation of content and structure toward the LMS. The other requirements were mostly already fulfilled by Typo3.

2.4. The learning management system

Nicephor is compatible with most learning management systems. This interoperability is achieved by using an external CMS to store and distribute content and by complying with the SCORM/IMS specifications (Fig. 2). For instance, WebCT-Vista and Moodle are currently used as LMS for the project. Those LMS can implement complex learning scenarios, manage students and their work, and integrates number of other specific communication and e-learning tools.

A course is built into an LMS according to a specific learning scenario that defines the content, the exercises, the tests, and so on, depending on the objectives and the students. The interaction between the repository and the LMS allows for the creation of as many different courses as required, all sharing the same base of content. Another advantage of this interaction CMS–LMS is its ease of maintenance: an update of one learning object takes effect in every course that uses this particular object.

2.5. The digital assets management

Courses implemented into a learning management system have not only a theoretical portion but also a practical one. This practical work often consists of case studies or photographic exercises applied to a given forensic topic. On the one hand images are produced during the photographic exercises, and on the other the creation of the case studies requires a searchable collection of images. The aim of the Digital Assets Management system is to use staff and students’ own production to develop an extensive collection of images of different marks or objects. To achieve this, all practical photographic work is indexed with relevant metadata and maintained through the DAM system of the project: Extensis Portfolio.

Indexation of images produced by staff and students is based on Adobe’s Extensible Metadata Platform (XMP). XMP is a labelling technology that allows you to embed data about a file, known as metadata, into the file itself. The main advantage of XMP is its ability to accommodate specific metadata schemes. For example, a specific metadata schema can be created for fingerprint images to describe the general pattern, the method of detection, the quality, the comparison conclusions, etc. The same kind of XMP schema can be created and applied to other marks, objects, or scene.

Thus, for each image, producers fill an XMP schema with relevant information according to the subject and submit it to an instructor for correction. During the submission process images are grouped together by subject and the XMP data are dynamically catalogued by Extensis Portfolio (Fig. 3).

Extensis Portfolio allows to index the specific XMP schema created for each subject of image. The imported XMP values can be easily edited within the DAM and then the updated values may be embedded back into the images.

![Fig. 2. Process of content integration from the CMS into different LMS.](image-url)
Instructors use this opportunity to correct the students work and to fill specific XMP fields dedicated to the correction process.

3. Results and discussion

Nicephor[e] proposes a basis of structured resources to create on-line courses according to each present or future need of different institutions. The development challenge was to fit the system’s architecture to the creation of number of different courses given by different institutions (Fig. 4). The result of this development is a system where each institution can:

- Create content into a CMS in the desired language. This content can be shared with others institutions.
- Choose learning scenarios (blended learning or full distant learning).
- Create courses or adapt existing ones to its needs with the appropriate LMS for different level of assistance: Bachelor, Master, or continuing education program.

As mentioned before, the project has significant impact on the organisation of the ESC in the manner forensic images are handled in a training environment. The capitalisation by a
digital assets management system of all the photographic material produced, leads to a knowledge base of photographic techniques applied to forensic issues. Currently, more than 2000 images are produced per year during courses or exams. The accumulation and classification of this material in a database creates also an extensive collection of images. The main interest of this collection is the opportunity to reuse images in different contexts. Several such scenarios, for example the production of image-based case files, have already shown their efficiency using the DAM system. It is clear that the ability to reuse images is intimately linked to the searching possibilities of the collection. The search potential depends directly on the XMP schema defined for each image subject (fingerprints, shoeprints, etc.) and on the quality of the collected information. The correction procedure, executed by instructors after image integration, is precisely dedicated to maintaining the quality and the exactness of this information (Fig. 5).

In 2006, three courses using the Nicephor[e] e-learning platform have been integrated in the cursus of three different institutions. The on-line course of microscopy at BSc level that takes place at ESC, has been assessed using a protocol combining structured interviews, on-line questionnaires, and student tracking capabilities integrated to both CMS and LMS.

Overall based on the feedback from 26 students, the evaluation provides very good evidence of how well the course is integrated within the ex-cathedra teaching, enhancing their levels of understanding by offering flexible, interactive, and attractive means of approaching complex subject matter. The on-line course has been heavily used and offers an answer to the students’ needs to consolidate their knowledge outside the university campus. This evaluation framework will be extended to all the future on-line courses deployed in this programme.

4. Conclusion

Nicephor[e], a Swiss Virtual Campus funded project, proposes an organized series of on-line modular courses about scientific and forensic photography and microscopy. The architecture of this e-learning framework is based on three main systems: a content management system (Typo3), a learning management system (WebCT, Moodle) and a digital assets management system (Extensis Portfolio). This overall system fulfills the flexibility and sustainability requirements defined for the project. Different institutions can create content and use a basis of structured resources allowing to build on-line courses according to each present or future needs. Learning scenarios and objectives are highly dependant on the course. They determine the number of ECTS credits associates to the course in each institution.

The creation of a structured database of images produced by staff and students during their studies using the digital assets management system represents an extensible and searchable collection of different marks from known sources in various conditions. It gives the opportunity to reuse images in different educational contexts: the creation of image-based case files and proficiency tests or for research purpose.

The first course implementation has provided a very positive feedback from students and professors involved. It offers an attracting and stimulating solution to strengthen their knowledge.

Acknowledgements


References

How do forensic scientists learn to become competent in casework reporting in practice: A theoretical and empirical approach

Stephen Doak a, *, Dimitris Assimakopoulos b

a Forensic Science Laboratory, Department of Justice, Equality & Law Reform, Garda HQ, Phoenix Park, Dublin 8, Ireland
b Grenoble Ecole de Management, Europole, 12 rue Pierre Semard, BP127, 38003 Grenoble, France

Received 9 June 2006; accepted 14 June 2006
Available online 1 August 2006

Abstract

In their day-to-day work, carrying out complex tasks, forensic scientists use a combination of explicit, codified standard operating procedures and tacit knowledge developed through their ongoing practice. We show that tacit knowledge is an integral part of the activities of expert forensic science practitioners who continually add to their knowledge repertoire by engaging other scientists through communities of practice. We wish to shed fresh light on the gaining of tacit knowledge by forensic scientists during their apprentice formative years, termed as legitimate peripheral participation.

In quantifying tacit knowledge exchanges, we use social network analysis, a methodology for the analysis of social structures, to map relational knowledge flows between forensic scientists within communities of practice at the Forensic Science Laboratory, Ireland. This paper sheds light on the importance of tacit knowledge within the training regime of forensic scientists and its recognition as equal to the part played by explicit knowledge.

© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Tacit knowledge; Legitimate peripheral participation; Communities of practice

1. Introduction

In their day-to-day work, carrying out complex tasks, forensic scientists use a combination of explicit, codified standard operating procedures (SOP) and tacit knowledge developed through their ongoing practice. There is a concern in the literature that the importance of explicit knowledge, present in documents or databases, is being over emphasized at the expense of inarticulate tacit knowledge, residing in people [1,2]. Indeed the management of tacit knowledge at an organizational level is relatively unexplored.

The literature catalogues tacit knowledge as the counterpart to explicit knowledge within the knowledge dichotomy or as a component within the knowledge continuum [3,4]. Tacit knowledge is expressed or carried on without words or speech [5]. The word ‘tacit’ derives from the Latin verb tacere, to be silent.

Our interest is in the development of those tacit knowledge structures that enable continuous learning and allow valuable personal knowledge to be gained by trainee forensic scientists, as they become competent experts. Scientists at different levels of a hierarchy of expertise have different tacit knowledge capacities. It has been found that the difference between experts and novices is related to their inventory of tacit knowledge [6: p. 122] where measures of tacit knowledge have the potential to explain individual differences in job performance [7].

Tacit knowledge is important to the development of professional practice and can be a source of highly effective performance in the workplace [8]. In research on large-scale engineering projects, individuals rather than turning to databases and procedure manuals to obtain information, seek knowledge in a tacit form from trusted and capable colleagues [9]. Tacit knowledge is a personal knowledge grounded in experience which because of its intricacies cannot be fully expressed [10]. Tacit knowledge consists of embodied expertise: ‘a deep understanding of complex interdependent systems that enables dynamic responses to context-specific problems [11: p. 9]’ and may be seen as a thread woven through the development of expertise [12]. Professional performance of
an expert involves sequences of routinised action punctuated by rapid intuitive decisions based on tacit understanding of the situation [13]. Expertise is based on past experiences where ‘the expert seems to remember holistic images from earlier experiences, matches and compares them and finds through the perception of diffuse signals that something in this situation is different from the memorized ones’. The expert does not have to depend ‘on time consuming sequential-analytical interpretation of information’, unlike the novice and thereby ‘is able to act in a critical situation [14: p. 690]’.

While explicit elements of practice are taught formally, tacit elements are usually learned during practice and observations while carrying out that practice [15]. In anaesthetic practice, an explicit knowledge base is insufficient for the expert and the clinical apprenticeship model of learning endures in order to pass on the other necessary form of knowledge that of tacit knowledge [16].

We show that tacit knowledge is an integral part of the activities of expert forensic science practitioners who continually add to their knowledge repertoire by engaging other scientists through communities of practice (CoP)—relatively tight-knit groups of individuals engaged in a shared practice who know each other well and work together directly. Communication within is often through face-to-face interaction, enabling members to transfer a high degree of implicit knowledge. Strong interpersonal ties create norms of direct reciprocity within a small community. Through their close quarters contact the communication reach of the community is bounded [17,18].

We wish to shed fresh light on the gaining of tacit knowledge by forensic scientists during their apprentice formative years, termed as legitimate peripheral participation (LPP), where they as peripheral members begin to become integrated into such a knowledge intensive community, because of their development of tacit knowledge competencies [19: p.100]. Learning, from the viewpoint of legitimate peripheral participation, essentially involves becoming an “insider” acquiring that particular community’s subjective viewpoint and learn to speak its language [20]. The participation of apprentices within a community is at first legitimately peripheral but over a period of time increases gradually in engagement and complexity [21: p. 95]. Newcomers (apprentices) working in social contexts with more experienced old-timers become their students through a mentoring process [22]. Once ‘newcomers have moved on from the role of peripheral participants to the status of fully legitimate members of the community, the learning they have acquired, together with its pattern and implicit complex logic, becomes part of their tacit knowledge [23: p. 291]’.

Our case study on a forensic science community shows the intricacies of tacit knowledge exchange, with an aim to allow readers to understand from a quantitative viewpoint what happens between knowledge workers during their training and their daily work practice. We use quantitative social network analysis [24] to create network maps of advice connections over time, amongst the scientists within their specialties, shedding light on the legitimate peripheral participation and knowledge dynamics of various scientists at the Forensic Science Laboratory (FSL), Ireland.

When processing a complex case, trainees may tend to rely too heavily on SOPs, as opposed to the more experienced scientists who know when to pull from their tacit knowledge reserves in order to balance their findings [25]. We use a questionnaire to look at the interplay of tacit and explicit knowledge in the training process and compare the developing tacit knowledge expertise of trainee forensic scientists with their use of explicit protocols.

This paper sheds light on the importance of tacit knowledge within the knowledge management paradigm and its recognition as equal to the part played by explicit knowledge in the training of professionals such as forensic scientists.

2. Methods

Social networks graphs were used to portray the diffusion of tacit knowledge and were visualized by utilizing Pajek [26] and Ucinet [27] software programs. Both programs represent network data in the form of a sociomatrix, i.e., a set of actors (vertices) linked with one (or more) relation(s)/ arcs. Such a one-mode matrix was used to record the advice relations given by the full complement of forensic scientist staff over a 3-day period at the FSL in early 2005. The “advice seeking” relation was used to try to best capture the tacit element of knowledge being exchanged. The respondents were told not to record those instances where only the mere exchange of information or functional business had occurred. The knowledge transferred was as a result of advice in response to a case problem where a solution was not immediately apparent. Seventy percent of the scientists in the laboratory replied. In order to try to capture all the scientists, an additional sociomatrix of “advice received” was transposed. As a result all forty-three scientists were recorded as having given some form of “advice to”. In order not to complicate the graph, this large sociomatrix was ‘binarized’ to reflect who simply gave advice to whom.

To discover those scientists who had relatively large tacit knowledge repositories in order that they could be compared to trainee scientists, both popularity and prestige, two classes of prominence, were used as a proxy measure of tacit knowledge. Popularity was measured by calculating a forensic scientist’s indegree, the number of requests for advice a scientist received. Prestige was measured by calculating the asymmetry of their advice giving over receiving.

To examine more fully the phenomenon of LPP and the concomitant exchange of tacit knowledge during training, the Biology CoP sociomatrix was extracted from the main laboratory advice sociomatrix. The frequency of communication that occurred between Biology scientists over the 3-day period was preserved in order that we could examine more fully the knowledge exchange between the scientists (i.e., the matrix was valued and asymmetrical, rather than binary).

Network maps were produced using Ucinet as the sociomatrix generator and Pajek as the visualization tool. The authors use below the Pajek definitions in describing a graph (network) as a set of vertices and arcs. The graph represents the structure of the CoP(s) network, where sets of scientists (vertices) are joined up by directional ‘advice’ relations (arcs) which point from a scientist giving advice (sender) to a scientist receiving advice (receiver). Line values in those extracted graphs are used to indicate the strength of a relation, that is the scientist giving advice a multiple of times [28: p. 6–7].

In looking at the interplay of tacit and explicit knowledge in the training process, a structured questionnaire was completed by a sample of forensic scientists from two CoPs at the FSL, one qualitative and the other quantitative in the types of case reports that they outputted.

3. Results and discussion

3.1. CoPs exchange tacit knowledge

Using social network graphs to portray the diffusion of tacit knowledge, we find based on the tacit knowledge shared between forensic scientists during their practice that the
working structure of the laboratory falls into four communities of practice (Fig. 1).

3.2. A marker for tacit knowledge levels in scientists

In order to examine the phenomenon of LPP, we needed to design a proxy to measure the levels of tacit knowledge in individual forensic scientists. We propose to use the prominence of forensic scientists as a marker to the size of their tacit knowledge repositories. In looking at a network, an actor is deemed prominent if the ties of the actor make the actor particularly visible to other actors in the network [29: p. 173]. However, some scientists may have an abundance of practical tacit knowledge built up over years of experience, but may be still peripheral in the network on account of their lacking communication skills or on their disinclination to share knowledge. The simplest measure of structural prominence is popularity, which we measure by calculating the number of requests for advice a scientist receives indegree (Fig. 2).

A scientist with a large indegree, becomes close to all other actors and is in direct contact or is adjacent to many other scientists. As a corollary, those scientists with a low indegree are peripheral in the network and seldom get asked to give any advice. We also used prestige as a more austere measure of...
prominence, where the net amount of advice a forensic scientist gave was measured. Sixteen of the forty-three forensic scientists were recorded as being prestigious within the laboratory network, with regard to the net tacit knowledge that they gave (Fig. 3). These two measures of prominence allow us to ascertain those scientists who have become central cogs in the tacit knowledge exchange network. It is evident on examining the density and prestige network graphs (Figs. 2 and 3) that the less experienced forensic scientists seek advice in casework problems from these prominent forensic scientists who have major sources of evident tacit knowledge.

3.3. Legitimate peripheral participation

The gaining of tacit knowledge during the apprentice formative years has been termed legitimate peripheral participation [21]. In mapping through the use of prominence measures those scientists with a relatively large or small measure of tacit knowledge, we can empirically observe the phenomenon of peripheral participation within a community of practice. We extracted from the laboratory advice data set, full details of advice exchange within the bounded Biology community of practice.

At the time of this study, four forensic scientists were actively being trained, two had completed their formal training and the remaining four were ensconced as fully trained and competent casework reporting forensic scientists within Biology. Length of service was used as an indicator for the scientist’s current status of training. In trying to capture whether a scientist mostly gave or received advice, the net advices/prestige ranking of each Biology member was calculated (Table 1). We found that the trainee forensic scientists (1–2 years of service) received the most advice. Those recently trained scientist (3–4 years) still received advice but to a lesser extent than the trainees. The four experienced scientists gave on balance advice (10–20 years) to their less experienced members of the biology community of practice (Fig. 4). This would be expected within the legitimate peripheral participation theory.

In using the prestige measure, it is evident that the more experienced forensic scientists have large tacit knowledge reserves when compared to the trainee/recently trained scientists (Fig. 5). Four forensic scientists within the Biology community of practice who have served between 10 and 20 years disperse their tacit knowledge to six other less experienced colleagues, all of whom have, except for one scientist, have served for less than 4 years.

The scientist with exception has served 22 years and is the longest serving member in the group (Table 1**). This forensic scientist received the most advices, which at first observation would seem to be contrary to the LPP theory. However, 1 year before this study, this forensic scientist through promotion as head of section (HOS) gained entry into the Biology CoP becoming the newest member of the group. This finding adds support to the claim in the literature of becoming newly

<p>| Table 1 |
| Advice league—Biology CoP |</p>
<table>
<thead>
<tr>
<th>Rank</th>
<th>Scientist</th>
<th>Net advices/prestige</th>
<th>Service (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FS II</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>FS I</td>
<td>12</td>
<td>19</td>
</tr>
<tr>
<td>3</td>
<td>FS II</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>FS II</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>FS II</td>
<td>−3</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>FS II</td>
<td>−4</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>FS III</td>
<td>−6</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>FS III</td>
<td>−6</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>FS III</td>
<td>−11</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>HOS FS I*</td>
<td>−11</td>
<td>22</td>
</tr>
</tbody>
</table>

Fig. 3. Prestige, of a scientist proportional to diameter of vertices—sixteen scientists within laboratory network give net advice.
peripheral, when a highly experienced individual moves to a new discipline, where there is a new CoP structure to be embedded and learning curve to climb [11].

3.4. Use of SOPs in training

We looked at the interplay of tacit and explicit knowledge in the training process, from our analysis of a completed structured questionnaire. With the questionnaire, we compared the utilization of explicit knowledge SOPs, to the use of tacit knowledge the scientists had gained through practice.

The questionnaire was completed by forensic scientists from two CoPs; one comprising the Biology community where the casework and reporting is more qualitative to the other, the Drugs community where the casework and report is very much of a quantitative output.

In their training to become competent forensic scientists, sixty five percent of those surveyed had the use of SOPs (Fig. 6). The remainder entered the laboratory before the ISO 17025
accreditation was in place. Of those scientists using SOPs, a third needed to seek extra assistance in carrying out procedures, even though SOPs existed for the full complement of procedures. This extra assistance was through advice from colleagues one hundred percent of the time. After having achieved competencies in carrying out all steps in their SOPs, none of the scientists felt sufficiently qualified to report casework.

In finding a problem in their casework process, all scientists surveyed would go to somebody else to get first hand experience. However, all found SOPs to give a beneficial baseline of knowledge and would use them as a refresher in their procedural knowledge, but three quarters of the same scientists found that the SOPs did not guide them in their day-to-day work. All only felt comfortable working on their own through continuous practice over years of service. One hundred percent of those surveyed found that the use of SOPs were an addition to the integrity and quality of the laboratory’s work.

We have shown with the developing expertise of the trainee forensic scientists that they transit from a reliance on explicit knowledge to one with a tacit knowledge framework. Explicit knowledge has qualified foundations in the first steps of a forensic scientist’s training, but is soon taken over by the tacit knowledge required to become a competent reporting caseworker.

4. Conclusion

On establishing the distinction between tacit and explicit knowledge, we have come to understand the complex relationship between explicit and tacit knowledge. We have shown that tacit knowledge is an integral part of the activities of expert forensic science practitioners. We have shone light on the process of gaining such tacit knowledge during the apprentice formative years of forensic scientists. Our case study on a forensic science community has shown the intricacies of tacit knowledge exchange. We have allowed readers to understand from a quantitative knowledge management viewpoint what happens to knowledge workers during their training and their daily work practice.

Acknowledgement

The authors wish to thank those participating staff of the Forensic Science Laboratory in the above case study.

References

A forensic image processing environment for investigation of surveillance video

M. Jeriana, S. Paolino, F. Cervelli, S. Carrato, A. Mattei, L. Garofano

DEEL, University of Trieste, v. Valerio 10, 34100 Trieste, Italy
Raggruppamento Carabinieri Investigazioni Scientifiche, Strada Fonderie 10, 43100 Parma, Italy

Received 9 June 2006; accepted 14 June 2006
Available online 25 July 2006

Abstract
We present an image processing software suite, based on the Matlab environment, specifically designed to be used as a forensic tool by law enforcement laboratories in the analysis of crime scene videos and images. Our aim is to overcome some drawbacks which normally appear when using standard image processing tools for this application, i.e. mainly the lack of full control and documentation on the operations which have been performed on the images, and the absence of new, more sophisticated algorithms which can provide improved performances and “make the difference” in critical cases.

1. Introduction
Video-surveillance systems are one of the main source of information during investigations, thanks to their wide-spread and increasing presence in our countries. However, the adopted closed-circuit devices are often affected by poor quality mainly because of economical and practical problems. Although this fact let us reflect if they can be considered more a deterrent for criminal actions rather than a valid identification system, in many cases also a low quality image can give useful information both during the first phase of the investigation and in courtrooms.

As a consequence, the images and sequences coming from video-surveillance systems need to be digitalized in order to be processed by dedicated software to enhance features useful for crime analysis. Generally, this is done either to reduce the different kinds of corruptions that have been introduced in the acquisition, conversion, and storage processes of the data or to overcome the limits of the overall system.

The characteristic problems to deal with are:
- low resolution of the images, which often implies the need to increase the size of the interesting details;
- lack of contrast;
- different types of noise or disturbances;
- blurring caused by motion or lack of focus;
- jitter or misalignment of lines due to the wear of video cassette recorder (VCR) heads;
- geometric distortions, thus severely limiting the reconstruction of the dimensions of the objects inside the image (e.g. the numerical estimation of the biometric features of subjects).

Each operation performed must be logged and certified; moreover, a complete knowledge of each step of the applied algorithms is needed to obtain full objectivity and to guarantee that the same result can be obtained following that exact procedure by anyone anywhere. Consequently, in theory the use of commercial software may not be suited from a legal point of view, due to the fact that its source code is usually not available for inspection. Open-source software, such as GIMP [1], meets this need, since each single part of the process is totally visible and open to scrutiny. However, open-source programs may not be equipped with the latest operators which have been appeared in the literature for common tasks (e.g. zooming, image noise...
reduction), or may lack tools to treat very specific problems such as recovery of a video recorded by a defective VCR.

In this paper, we present what we have called the Modular Image Processing Environment (MIPE). This software is the result of the collaboration between the developers and the end users to face the aforementioned problems. After a brief review on the state-of-the-art systems (Section 2), we describe the adopted software architecture and the operators that have been already implemented in the system (Section 3), showing some examples of their application (Section 4). We conclude presenting the current limitation of the system and describing the future of MIPE, which is, in our aim, the prototype of a more ambitious project, namely a new forensic image processing suite.

2. State-of-the-art

Presently, several products exist on the market which are dedicated to the analysis of image-based information for forensic science applications; we can recall, in no particular order and with no intention to be complete, dTective by Avid and Ocean Systems [2], Impress by Imix [3], StarWitness Video by Signalscape [4], Video Analyst by Intergraph [5], and Video Investigator by Cognitech [6].

Because of the strong correlation between the evidence source and the processing software, forensic image analysis products often come as a complete software–hardware solution. These systems perform, as a first step, the acquisition of a secure digital copy of the evidence video or images in order to prevent any possible damage of the original. The second step is dedicated to processing, and the vendors normally offer different solutions to satisfy the forensic image professional needs to analyze, enhance and edit all major image and video formats that may constitute the evidence source.

Commercial systems offer different software capabilities (e.g. available filters, algorithms and proprietary operations, courtroom oriented functions) and hardware characteristics (e.g. installed PC RAM, graphics and acquisition cards, available media readers, additional equipments such as VCRs and printers). Here we are interested mainly in the processing software part of an ideal forensic image analysis product, and the aim of MIPE is to become an affordable application for image restoration and enhancement in forensics, using validated, bleeding edge, widely accepted and open to scrutiny image processing algorithms to extract the relevant information from the recorded sequences.

All the above mentioned programs can perform the basic image editing operations: contrast and brightness adjustment, histogram equalization and editing, zooming, mirroring and rotation of the images under analysis. Some of them offer the ability to customize filters kernel and offer proprietary filters (such as JPEG dedicated deblockers). Another example are motion deblurring filters, which allow the operator to restore the details of a moving object. Inter-frame operations can also be performed, for example, employing time information recorded by the sequence in order to gain what would be otherwise lost in just a single frame. Demultiplexing algorithms are also available, so that different camera views can be converted into video clips (see Ref. [7] for a review about video contents indexing) thus allowing the isolation of the crime scene, and deinterlacing. Finally, each case can be fully organized thanks to archive and back-up facilities, and each program creates an audit log for each case, in order to make it suitable for presentations in courtrooms.

These analysis systems offer the ability to record each step of the image processing operations on a log file, however they do not guarantee complete access to the applied algorithms. In fact, although the performed operations are recorded together with their parameters, the employed algorithms may be not public.

3. The proposed system

MIPE is the system born as an answer to the above mentioned problems. Its development follows some basic guidelines that can be resumed as follows:

- complete control and knowledge on the processing which is done on each image;
- employment of state-of-the-art algorithms;
- development in strict collaboration with the final user;
- high modularity, in order to grant easy customization of the software.

For our purpose, probably the first point is the most important: in order to grant a transparent and objective result during legal procedures, the user (and the court) needs to know all the details on the processing applied to an image. As expressed in Ref. [8] it is not possible to make a clear distinction between what is called enhancement and the manipulation of the image. The best we can do is to use techniques which are widely accepted by the scientific community and give all the details on their implementation in order to make the process completely repeatable.

A complete transparency is achieved by these means:

- all the algorithms applied to an image are listed in the right sequence on a log file automatically saved with the image;
- for each employed algorithm all the involved parameters are listed;
- for each algorithm (if not a very standard one) the reference article is provided;
- for each algorithm the source code is provided.

Using bleeding edge techniques and at the same time giving the implementation details leads not only to objective results, but also better performances if compared to traditional systems.

In order to meet the above expressed requirements, Mathworks Matlab 7 [9] has been chosen as development environment. First of all, the style of programming is very clear (almost like a pseudo-code), allowing people with even very limited programming skills to understand and manage the code. All its functions are very well documented and any customization of the code is straightforward; moreover, there is a wide number of ready-to-use functions, either built-in in the basic environment or available in one of its various toolboxes. A very useful feature is that almost all its functions (except the very basic and computationally critical ones) are visible.
Another interesting characteristic of the Matlab environment is the portability of the code, which can be used with few (or none) modifications on Microsoft Windows, Apple Mac OS X and Linux, being it interpreted on run-time; this latter feature also allows very fast debugging (the source can be also compiled for a specific platform, but in this way we lose the possibility to modify it on the fly). Consequently, the time needed for the development and the customization of the system is very short if compared to lower level languages. Moreover, Matlab also provides easy and powerful tools for graphical user interface (GUI) creation.

In order to give the user the possibility of customizing the software, a common processing interface has been created, shared by most of the processing functions. Following some very simple guidelines it is possible for the user to easily interface a custom-created function to the rest of the system, both from a lower level and with respect to the integration in the existing GUI. In the present version, the fast integration is allowed for functions which accept up to three numerical parameters and one optional string that can be configured for various purposes through a choice-box. For each needed parameter it is necessary to set the minimum, the maximum and the default value, while for the options box it is required to enumerate possible values for the string. Depending on the number of used parameters, the main program will draw the right number of elements (i.e. scrollbars and textboxes) in the dialog box, together with the controls for the preview image. The function will look then perfectly integrated, as if it was belonging to the original system.\footnote{Actually, another minor operation is needed to make the function to appear in the desired menu in the main window of the program. This can be easily done by Guide, the GUI design tool provided with Matlab.}

Of course, also the user-created functions will automatically add their details to the application.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Main features implemented in MIPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic editing</td>
<td>Crop</td>
</tr>
<tr>
<td></td>
<td>Luminance and contrast</td>
</tr>
<tr>
<td></td>
<td>Grayscale conversion</td>
</tr>
<tr>
<td></td>
<td>Flip/rotate by 90°</td>
</tr>
<tr>
<td>Interpolation</td>
<td>Pixel replication</td>
</tr>
<tr>
<td></td>
<td>Bilinear</td>
</tr>
<tr>
<td></td>
<td>Bicubic</td>
</tr>
<tr>
<td></td>
<td>Rational [13]</td>
</tr>
<tr>
<td></td>
<td>LoWaD2</td>
</tr>
<tr>
<td>Image restoration</td>
<td>Shifted lines correction</td>
</tr>
<tr>
<td></td>
<td>Lens distortion</td>
</tr>
<tr>
<td></td>
<td>Perspective correction</td>
</tr>
<tr>
<td></td>
<td>Gaussian blur</td>
</tr>
<tr>
<td></td>
<td>Motion blur</td>
</tr>
<tr>
<td>Filters</td>
<td>Median</td>
</tr>
<tr>
<td></td>
<td>Usharp masking</td>
</tr>
<tr>
<td></td>
<td>High-pass</td>
</tr>
<tr>
<td></td>
<td>Rational sharpening [14,15]</td>
</tr>
<tr>
<td></td>
<td>Averaging</td>
</tr>
<tr>
<td></td>
<td>Low-pass</td>
</tr>
<tr>
<td></td>
<td>Rational noise smoothing [14]</td>
</tr>
<tr>
<td></td>
<td>Blocking artifacts reduction [10]</td>
</tr>
<tr>
<td></td>
<td>Add noise (for test purpose)</td>
</tr>
<tr>
<td>Multiple images</td>
<td>Median</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td>Simple registration</td>
</tr>
<tr>
<td></td>
<td>Projective registration</td>
</tr>
<tr>
<td>Video</td>
<td>Import frame</td>
</tr>
<tr>
<td></td>
<td>Save all frames</td>
</tr>
<tr>
<td></td>
<td>Demultiplexing</td>
</tr>
<tr>
<td></td>
<td>Save demultiplexed sequences</td>
</tr>
<tr>
<td>Deinterlacing</td>
<td>Bilinear</td>
</tr>
<tr>
<td></td>
<td>Bicubic</td>
</tr>
<tr>
<td></td>
<td>Rational [13]</td>
</tr>
<tr>
<td></td>
<td>Advanced [16]</td>
</tr>
</tbody>
</table>
on the log file. In Fig. 1 a screenshot of the proposed Modular Image Processing Environment is presented as an example.

In Table 1, most of the processing functions implemented in MIPE are presented. In some cases, different techniques are provided for performing the same tasks, although some of them may have not been deeply tested or are present just in order to compare the performances of various algorithms. Some support functions, such as the Undo/Redo functions, the possibility of easily compare two images and to display image information (size, bit depth, signal-to-noise ratio, . . .) are not mentioned in the table.

4. Experimental results and discussion

Some of the various filters and tools available in MIPE are presented here as examples, to show the potentialities of the system.

Some special algorithms have been specifically designed to cope with the rather frequent problem of the shift of subsequent lines in images due to the wear of VCR heads. They have been tested for two particular cases (one is presented in Fig. 2); however, it must be noted that it is very difficult to solve this problem in general, since the characteristics of the disturbance may vary widely from case to case.

Fig. 2. Correction of shifted lines: (top) original and (bottom) processed image.

Fig. 3. Projective registration: four images from the original set.
Another useful tool realized for MIPE let us perform frame averaging of an object which is present in different frames and captured under different point of view. Commonly available techniques allow to frame-average objects translated parallel to the image plane in consecutive frames, while more advanced algorithms also allow image rotation and rescaling. Our approach is more general and the projective registration tool lets us frame-average the object of interest (for example a car license plate) even if subject to different perspectives. The example in Fig. 3 shows how the combination of only few frames can lead to good results. Indeed, in Fig. 4 the license plate characters are easily readable.

A more common feature, that sometimes offers really good results also in the worst cases, is the implementation of motion deblurring through one of the deconvolution functions of the Matlab Image Processing Toolbox. It is easy to note in Fig. 5 that the license plate has been recovered from an apparently useless image.

Very frequently, images and sequences are stored in a format that implies a lossy compression, which causes loss of details and the introduction of artifacts on the image. MIPE offers a filter, based on Ref. [10], which is able to efficiently reduce the blocking artifacts caused by many common block-based compression tools, e.g. JPEG. An example is presented in Fig. 6.

In order to apply geometric transformations to an image with the minimum possible loss of detail, some new interpolators have been studied. In Fig. 7 our algorithm, applied to an $8 \times$ enlargement, is compared to the classic bicubic interpolation. The employed technique is an evolution of the WaDi algorithm presented in Ref. [11]. Other examples taken from the features of MIPE are presented in Ref. [12].
5. Conclusions

In this paper, a general introduction on forensic image processing and its main issues has been exposed, and a new environment for the enhancement of images coming from video-surveillance devices has been proposed.

The principles which have guided our work and their practical applications in the attempt to meet, as close as possible, the forensic image professional requirements have been depicted. Matlab has been chosen as the development platform thanks to its advanced features, the ease of programming and the possibility to work directly on interpreted functions, making the code visible and easily modifiable, thus obtaining an open to scrutiny software which can better fit legal needs. Some general features of the Modular Image Processing Environment, the software born from this work, have been described and some parts of it have been presented as examples.

Being developed in Matlab, MIPE suffers from the drawbacks of this environment, i.e., mainly the speed of execution and possible problems with memory when working with large images. In order to build a better system for the end user, we have recently started to develop a novel image processing tool written as a native application, and not as a Matlab interpreted software, while using the Matlab-based MIPE as a prototyping and testing environment. It will offer the already available algorithms a wide range of new features with respect to traditional image processing software: full control on editing the processing history, seamless integration of different kind of data (still images, groups of images, video sequences), and automatic report creation.

References


All trademarks and copyrights are the property of their respective owners.
ENFSI collaborative testing programme for ignitable liquid analysis: A review

Jeanet Hendrikse*

Netherlands Forensic Institute, Laan van Ypenburg 6, 2497 GB The Hague, The Netherlands

Received 1 June 2006; accepted 14 June 2006
Available online 24 July 2006

Abstract

The Fire and Explosion Investigation Working Group of the European Network of Forensic Science Institutes (ENFSI) is the organiser of a collaborative testing programme for ignitable liquid analysis. The testing programme was initiated in 1998. Initially to inventory the analytical methods used in this field of analysis, but with the ultimate goal to establish a European testing programme for fire debris analysts. As of today, five tests have been conducted.

This article will provide an overview of the first five ENFSI collaborative tests for ignitable liquid analysis. The background, objectives and characteristics of the testing programme are summarised, followed by an overview of the sample composition employed, the participants' performance, the difficulties and the lessons learned in each test.

© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Collaborative testing programme; Ignitable liquid analysis

1. Introduction

In 1998, the Fire and Explosion Investigation Working Group of the European Network of Forensic Science Institutes (ENFSI) initiated a collaborative testing programme for ignitable liquid analysis as one of the activities to meet the objectives of ENFSI, which are: sharing knowledge, exchanging experiences and coming to mutual agreements in the field of forensic science.1

The testing programme was established stepwise, with the ultimate objective being to test the participants’ skills in the analysis of various debris samples and in the interpretation of the analytical results obtained.

Between 1998 and 2005, five collaborative tests have been conducted. Test I was used to provide an inventory of the analytical techniques employed by the different analysts and laboratories. In Test II, the detection limits of these techniques were determined. As of Test III, artificial debris samples were distributed for testing purposes.

The tests are summarised in Table 1. The ENFSI collaborative testing programme for ignitable liquid analysis is characterised by:

1. Qualitative analysis; qualitative analysis that confirm the presence or prove the absence of ignitable liquids is sufficient in fire investigations.
2. Blind testing; the participants know it is a test sample but do not know the sample composition.
3. No prescription of methods; the participants should perform the analyses using their standard laboratory procedures.
4. Detailed reporting of results; the analysis results must be supported with both a method description and the analytical data obtained.
5. Generous timeline for analysis; the participants usually have about three months to complete the analysis and to submit the results for evaluation.

2. Sample composition

The sample composition employed in the first five ENFSI collaborative tests for ignitable liquid analysis is summarised in Table 2.

---

* Tel.: +31 70 888 6317; fax: +31 70 888 6554.
E-mail address: j.hendrikse@nfi.minjus.nl.

3. Test participants

The ENFSI collaborative testing programme for ignitable liquid analysis is organised among (mainly) forensic laboratories in Europe. The testing programme is however not limited to European laboratories and one laboratory from outside Europe has been a regular participant from the beginning.

3.1. Number of participants

The interest among laboratories in the collaborative testing programme has grown steadily since the first test. This is shown in the increasing number of participants who have become involved from Test III onwards (Fig. 1). For unknown reasons, participation in Test II was considerably lower than the other tests.

3.2. Performance of participants

In Tests III–V, the participants received artificial debris samples with the objective to test their analysis and interpretation skills. An overview of the participants’ performance in these tests is presented in Fig. 2. The results of the blank samples (III-C, IV-D and V-C) are not included in this figure, as these blanks were provided to the participants for reference purposes only.

Each test posed different challenges to the participants; these challenges are discussed in more detail in the test overview.

4. Test overview

4.1. Test I

Test I, conducted in 1998, had the objective of providing an inventory of the analytical techniques employed by fire debris analysts in the participating laboratories. For this test, a liquid sample containing a mixture of gasoline and diesel oil was distributed for analysis. The results indicated that gas chromatography was the method of choice amongst the laboratories either with flame ionisation detection (GC–FID) (64%), mass spectrometry (GC–MS) (27%), or both (9%).

Most participants analysed the liquid sample directly or as a dilution. Only one participant analysed the headspace. Headspace analysis has the disadvantage of discriminating the higher boiling compounds against the lower boiling compounds. In particular for a sample containing diesel oil this may easily lead to misinterpretation of the chromatographic data and, as a consequence, result in misclassifying this product as for example kerosene.

4.2. Test II

Test II was organised in 2000 with the objective of determining the detection limits of the analytical techniques employed by the participating laboratories and to check whether the performance of their instrumentation could be compared. The participants were requested to procure the RESTEK reference ‘Fire Debris mixture: E1387-95 Column Resolution Check Mix’, containing five aromatic and eight aliphatic hydrocarbons, 2000 µg/ml each in methylene chloride.

All participants were able to detect the lowest concentration level of 1 µg hydrocarbon/mL in methylene chloride. All participants were able to detect the lowest concentration level of 1 µg/mL RESTEK reference. This concentration level was best reached via a direct liquid injection (injection volume
of 1 μL) in splitless mode using a non-polar column (1 or 5% diphenyl methyl siloxane).

4.3. Test III

Test III was organised in 2002 with the objective to test the analysis and interpretation skills of the participants. The first ‘real’ test where the participants’ skills were challenged in terms of identifying both a polar and a non-polar ignitable liquid.

The test comprised three sand samples, one spiked with a mixture of gasoline and diesel oil (sample III-A), one spiked with denaturated spirits (sample III-B), and one blank (sample III-C). All samples were packed in vials and closed with screw caps. These screw caps, however, appeared not to be gastight for long. This resulted in false negative identifications for the highly volatile denaturated spirits in sample III-B, for those participants that had not analysed the samples directly upon receipt. No stability studies had been performed prior to conducting the test.

In sample III-A, the gasoline was identified by all participants. Seventy percent of the participants also identified the diesel oil. This heavy petroleum distillate was misclassified as, for example, kerosene or technical product by 15% of the participants, the other 15% did not identify this ignitable liquid. Most of the participants from these latter two groups had based their conclusion on headspace analysis only, whereas the majority of those that had found and correctly identified the diesel oil had used, in addition to headspace analysis, solvent extraction as a recovery method.

4.4. Test IV

Test IV was organised in 2004, the second test with the objective to test the analysis and interpretation skills of the participants. In this test the participants were asked to identify and compare liquid diesel oils, as well as to differentiate pyrolysis products of polyethylene from an ignitable liquid. Four test samples were distributed:

- two liquid diesel oils (sample IV-A and IV-B, respectively) packed in ampoules;
- two pieces of charred pinewood, one spiked with pyrolysis products of polyethylene (sample IV-C) and one blank (sample IV-D), both packed in a three-layer sealed bag.

Sample IV-A and IV-B had different origins; sample IV-A was a standard Shell diesel oil, whereas sample IV-B was a special issue Shell V-power diesel oil. They were distinguishable by:

a. Colour; diesel oil IV-A was yellow, while diesel oil IV-B was almost colourless.

b. GC-pattern (n-alkane distribution, see Fig. 3); diesel oil IV-A contained a higher proportion of lower boiling hydrocarbons than diesel oil IV-B.

Table 3
Composition RESTEK reference ‘Fire Debris mixture: E1387-95 Column Resolution Check Mix’

<table>
<thead>
<tr>
<th>No.</th>
<th>Hydrocarbon</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>n-Hexane</td>
</tr>
<tr>
<td>2</td>
<td>n-Octane</td>
</tr>
<tr>
<td>3</td>
<td>n-Decane</td>
</tr>
<tr>
<td>4</td>
<td>n-Dodecane</td>
</tr>
<tr>
<td>5</td>
<td>n-Tetradecane</td>
</tr>
<tr>
<td>6</td>
<td>n-Octadecane</td>
</tr>
<tr>
<td>7</td>
<td>n-Eicosane</td>
</tr>
<tr>
<td>8</td>
<td>2-Ethyltoluene</td>
</tr>
<tr>
<td>9</td>
<td>3-Ethyltoluene</td>
</tr>
<tr>
<td>10</td>
<td>Toluene</td>
</tr>
<tr>
<td>11</td>
<td>1,2,4-Trimethyltoluene</td>
</tr>
<tr>
<td>12</td>
<td>p-Xylene</td>
</tr>
</tbody>
</table>

2000 μg/mL each in methylene chloride, 1 mL/ampoule.
c. Isoprenoide content (ratio \(n\)-C\(_{17}\)/pristane and ratio \(n\)-C\(_{18}\)/phytane, see Fig. 3); the \(n\)-C\(_{17}\)/pristane and \(n\)-C\(_{18}\)/phytane ratio was approximately 2:1 for diesel oil IV-A and approximately 1:1 for diesel oil IV-B.

d. Aromatic content; diesel oil IV-A contained a higher concentration of aromatic hydrocarbons (factor 2–3) than diesel oil IV-B.

About 80% of the participants correctly identified the liquid samples as diesel oils. The others identified them incorrectly, mainly due to misclassification of the analytical results. Some participants, for example, misclassified liquid IV-A as being a mixture of diesel oil and gasoline due to its relatively high aromatic content. Other participants analysed the headspace of the liquids rather than the liquids themselves and as a consequence obtained \(n\)-alkane distributions that are characteristic for kerosene (due to discrimination of the higher boiling compounds), and misclassified the liquids as such.

A relatively high number of the participants (27%) left the answer to the question on common origin undecided. One participant incorrectly reported that both diesel oils could have the same origin, as this participant believed that the observed differences could be the result of weathering. While weathering may change the early eluting part of the GC-pattern due to the evaporation of the most volatile compounds, it will not affect the ratios between the less volatile compounds \(n\)-C\(_{17}\)/pristane and \(n\)-C\(_{18}\)/phytane.

In the evaluation report\(^2\) [1] for this test, it was commented that microbial degradation could not be an explanation for the observed differences. While microbial degradation may affect the \(n\)-C\(_{17}\)/pristane and \(n\)-C\(_{18}\)/phytane ratios, it would also have a marked effect on the \(n\)-alkane concentrations compared to the iso- and cycloalkane concentrations; the \(n\)-alkanes would degrade much faster than the iso- and cycloalkanes resulting in an altered chromatogram in which the intensities of the iso- and cycloalkanes are relatively enhanced towards the intensities of the \(n\)-alkanes. This was not the case in these samples.

The chromatogram of sample IV-C showed a series of \(n\)-alkene/\(n\)-alkane doublets in the range of C\(_5\) to approx. C\(_{25}\) (see Fig. 4), characteristic for pyrolysis products of polyethylene [2]. In addition, compounds such as \(\beta\)-pinene, 3-carine and other terpenes could be observed, originating from the charred

---

\(^2\) Test IV was the first test with a detailed evaluation report.
pinewood matrix itself. These would also be evident in the blank sample IV-D.

Most participants correctly identified the chemical nature of these compounds. Only 60% of the participants, however, were able to recognize them as pyrolysis products of a polymer such as polyethylene and subsequently distinguish them from an ignitable liquid. Of the remaining participants, 20% misclassified these compounds as an ignitable liquid varying from a light petroleum distillate to a weathered diesel oil and 20% left the classification undecided.

In the evaluation report, it was commented that, due to the absence, or perhaps very low presence of aromatic compounds, the observed pattern of compounds could not be assigned to a petroleum distillate, nor could it be a mixture of pyrolysed polymer and a petroleum distillate.

4.5. Test V

Test V was organised in 2005. This test had two objectives:

1. to test the interpretation skills of the participants by challenging them to differentiate pyrolysis products of rubber from an ignitable liquid;
2. to inventory the criteria used by the participants to positively identify gasoline in the absence of low boiling compounds (weathered gasoline).

Test V was the first test where quality aspects such as repeatability and stability studies were introduced, and where a test scenario and reporting templates were implemented for further professionalism of the testing programme.

The repeatability studies were performed before sample dispatch. They were performed for sample Type V-B only, as the reproducibility in preparing these test samples was considered to be most critical due to the evaporation step involved. Six test samples of V-B were randomly selected and analysed. In all six samples, the volatiles of the weathered gasoline were detected and the variation in sample composition observed was considered acceptable.

The stability of all three test sample types was monitored during the test period by randomly selecting and analysing test samples of V-A, V-B and V-C at different intervals of the test. The results indicated that, four months after sample dispatch, the sample composition of all test samples could still be recovered and detected. The variation in sample composition observed for each sample type was comparable to or even less than the variation observed in the repeatability studies for sample V-B. Based on these study results, it was considered that all test samples were stable during the test period.

The test scenario was used to contextualise the test samples. The scenario was as follows:

A fire had occurred inside a house where, due to the fire, the front door, the hall way and part of the living room was burnt. As no technical cause for the fire could be identified, a hydrocarbon detection dog was called. The dog did not indicate to any debris, but gave two positive indications on pieces of unburnt carpet, one being located underneath a flowerpot in the hallway and the other being located underneath the partially burnt sofa in the living room. From both locations a sample was collected and was labelled as sample V-A and V-B, respectively. In the living room, a piece of unburnt carpet was collected as control sample and was labelled as sample V-C. The samples were packed in three-layer sealed bags and were sent to the laboratory (i.e. participant) for analysis.

The chromatogram of sample V-A (see Fig. 5) showed remarkable similarities with gasoline, but was distinguishable from gasoline by the distortions in pattern (i.e. the aliphatic hydrocarbon and C$_3$-alkylated benzene pattern) and the presence$^3$ of interfering products such as alkenes, limonene and cumene. Nevertheless, almost 80% of the participants misclassified it as an ignitable liquid, showing how easily a false positive identification is made.

Some participants reported a positive result without any comments. Some others misclassified the volatile compounds in sample V-A on the basis that the volatiles detected in the sample were absent in control sample V-C. In the evaluation

$^3$ For the identification of these interfering products, mass spectrometry detection was needed.
report [3], it was commented that these participants may have misunderstood the use of the control sample: as the matrix from a control sample and a (test) sample are never exactly the same (both are taken from different locations), it can never be excluded that the (test) sample is contaminated with substances other than an ignitable liquid whereas the control sample is not. A few participants did keep the possibility of contamination in mind stating that the volatiles identified in sample V-A could be from pieces of glue underneath the carpet, which were ‘coincidently’ missing in control sample V-C. However, despite this explanation, these participants incorrectly reported the volatiles in sample V-A as a positive for an ignitable liquid.

Some other participants misclassified most of the volatiles in sample V-A as gasoline, and used the scenario to explain the distortion in pattern and the presence of additional volatiles (such as the alkenes, limonene and cumene). For example, the alkenes were explained as being pyrolysis products from the flower pot, whereas limonene and cumene were considered to originate from a citrus plant in the flowerpot or from a previously used carpet cleaning solvent. This could indicate that the test scenario was somewhat misleading.

The chromatogram of sample V-B (see Fig. 6) showed a pattern that was similar to the aromatic fraction of weathered gasoline. The low boiling compounds (including methyl tert-butylether), characteristic for gasoline, were absent; trace amounts of these compounds could only be detected with more detailed analysis using Solid Phase MicroExtraction (Carboxen/ Polydimethylsiloxane) GC–MS. Most participants classified the volatile compounds in this sample indeed as (weathered) gasoline, and based the classification solely on the aromatic fraction observed. In the evaluation report, the evaluators debated whether this classification was correct as the aromatic fraction could just as well originate from an aromatic solvent.

For some participants (12%), the concentration level of the weathered gasoline was too low to detect (below their limit of detection). Others (9%) did identify the aromatic fraction, but considered the concentration level as being too low (below their limit of quantitation) to positively report it as an ignitable
liquid. No plausible explanation for this could be found in the recovery and analysis methods employed, as the different participants used different methods, varying from static to dynamic headspace. The explanation, more likely, should be found in the packing of the sample. As the packing of all samples in this test did not contain much room for vapour, some participants had commented that they had to re-pack the samples prior to analysing them with their normal laboratory procedures. Re-packing may have posed the risk of losing volatiles (and perhaps of introducing contamination), and subsequently may have contributed in a negative identification.

5. Lessons learned and future recommendation

The most important lessons that have been learned from the first five ENFSI collaborative tests are, that:

- Some pyrolysis products can be remarkably similar to an ignitable liquid and, as a consequence, may easily lead to a false positive identification. A good example in this respect was the pyrolysis products of rubber in test sample V-A where almost 80% of the participants had misclassified them as an ignitable liquid. As false positive identifications certainly are a reason for concern in real case work, it is strongly recommended to provide more training on pyrolysis products in future tests.
- GC-FID is good for screening, but GC–MS is considered necessary for identification, in particular when interfering products are to be distinguished from an ignitable liquid. In time, the number of MS-users has significantly increased; in Test V only 6% of the participants still relied their analysis on FID only.
- A liquid sample should be analysed directly or as a dilution, analysing the headspace instead may easily lead to misclassification due to the discrimination of the higher against the lower boiling compounds.
- The packing material of the test samples must be gastight during the test period; for this reason stability studies have been introduced in Test V, it is recommended to perform stability studies prior to conducting future tests.

To further improve the overall testing programme, a test scenario and reporting templates have been implemented in Test V. The reason for implementing a test scenario was to make the test more realistic and possibly provide the participants with information that could be of use to them during the analysis and interpretation phase. The reporting templates were designed to assist the participants in reporting their results in a more uniform manner. At the same time however, they simplified the evaluation of the test results. It is therefore recommended to continue using a test scenario and reporting templates in future tests, as long as the reporting templates are practical and the test scenarios are not misleading.

6. Conclusions

One of the goals of the ENFSI Fire and Explosion Investigation Working Group has been met; a European collaborative testing programme for ignitable liquid analysis has been established. It is the intention of the Working Group to conduct this testing programme annually. The interest in participation is still growing.

Acknowledgements

The ENFSI Fire and Explosion Investigation Working Group gratefully acknowledges Dr. Peter van Bebber from the Forensic Science Institute of the Bundeskriminalamt in Wiesbaden, Germany, for initiating the ENFSI collaborative testing programme for ignitable liquid analysis.

In addition, an acknowledgement is in place for those institutes that have prepared the test samples in one or more of these tests; these institutes are the Forensic Science Institute of the Bundeskriminalamt in Wiesbaden, Germany, the Laboratoire Central de la Prefecture de Police in Paris, France, and the Forensic Science Institute of the Landeskriminalamt in Stuttgart, Germany. And an acknowledgement for those scientists that have contributed in the evaluation phase of one or more of these tests.

Last but not the least, the author wishes to thank Dr. Niamh Nic Daeid from the University of Strathclyde Royal College in Glasgow, Scotland, for proof-reading this article and for providing valuable editorial remarks.

References

Cocaine profiling for strategic intelligence purposes, a cross-border project between France and Switzerland
Part I. Optimisation and harmonisation of the profiling method

S. Lociceroa, P. Hayoza, P. Esseivaa,*, L. Dujourdyb, F. Besacierb, P. Margaota

a Institut de Police Scientifique, Ecole des Sciences Criminelles, Université de Lausanne, BCH, 1015 Lausanne-Dorigny, Switzerland
b Laboratoire de Police Scientifique de Lyon, 31 Avenue Franklin Roosevelt, 69134 Ecully Cedex, France

Received 29 May 2006; accepted 14 June 2006
Available online 28 July 2006

Abstract
Optimisation and harmonisation of analytical and statistical methodology have been carried out between two forensic laboratories (Lausanne, CH and Lyon, F) in order to provide drug intelligence for cross-border cocaine seizures. The aim was to improve the gas chromatographic analysis of cocaine samples for profiling. Some important validation parameters were tested to verify the developed method and demonstrate its profiling capacity: the selectivity of the method with retention time reproducibility, the choice of a derivatisation agent improving the chromatography (MSTFA, BSA, TMSI and BSTFA + TMCS 1%), the cutting agents influence (matrix effect), the influence of the sample storage conditions and the sample quantity to weigh for analyses. Eight main alkaloids, which represent the sample signature, have been selected: ecgonine methyl ester, ecgonine, tropacocaine, benzoylecgonine, norcocaine, cis- and trans-cinnamoylcocaine and 3,4,5-trimethoxycocaine. Their stability in the solvent used (CHCl3/pyridine) was demonstrated.

In order to reach the final objective, which is the comparison of samples seized and analyzed in two different laboratories, the harmonisation of the profiling method between the two laboratories had to be ensured and is the subject of ongoing research.

Keywords: Cocaine; Drug profiling; GC-FID; Derivatisation agent; Cutting agents; Alkaloid stability; Storage conditions

1. Introduction
Recent European projects on analytical harmonisation in drug intelligence have been particularly focused on synthetic drugs like XTC and phenethylamine derivatives [1–4]. Nevertheless, according to the UE Drugs Agency (EMCDDA) annual report 2005, cocaine has become a major problem in the European drug trafficking landscape. Similarly, the latest statistics about drug trafficking in Switzerland (the Federal Office of Police) show that cocaine seizures nearly doubled between 2003 and 2004 to reach quantities never seen before.

The use of profiling methods has been shown to produce powerful means to establish links between seizures, thus providing drug intelligence. Cocaine profiling methods are used by some laboratories, but there is no or little useful information exchange both concerning the data obtained and the methods used in order to improve the fight against illicit traffic. No research concerning the harmonisation of the cocaine profiling between two different laboratories have been tested so far. This project was designed to provide cross-border drug intelligence between France and Switzerland concerning the traffic of cocaine.

Two laboratories have developed this collaboration over 2 years to design a harmonised analytical and statistical methodology for cocaine seized at the border as well as inside either one of the two countries: the Forensic Science Institute of Lausanne, Switzerland (IPS) and the Forensic Science Laboratory of Lyon, France (LPS). This project, named CASTEL (Caractérisation Analytique des Saisies Transfrontalières pour l’Etablissement de Liens), was financed by a European program, Interreg IIIA (1/FU/9.3/1), which promotes cross border cooperation between adjacent regions.

The process of profiling consists in two interconnected steps. The first one is dedicated to the development of analytical
Methodologies allowing the extraction of a signature of the illicit product. The second is directly focused on the integration of the results for law enforcement applications. The innovating challenge concerns the possibility of highlighting existing links between cocaine sample seizures analysed on two different analytical instruments or at two different times. Such an approach is fraught with difficulties that have stopped potential initiatives in the past. A second important point is to evaluate the potential of this approach for law enforcement needs and intelligence.

The project unfolds in three main parts which are themselves divided in several tasks.

1. Analytical and statistical methodology:
   (1) Study of the samples behaviour and optimisation of the analytical method in order to choose interesting compounds to profile with regards to: retention time reproducibility, potential overlap with current cutting agents, derivatising agents and cutting agents influence, storage influence, sample quantity to weight, alkaloids stability in the solvent used.
   (2) Adjustment of cocaine samples comparison method in the two laboratories, validation and interpretation of the results.
   (3) Harmonisation of statistical methodology within and between laboratories (linked and unlinked samples analysed in both laboratories, 10 pre-treatments, 6 correlation coefficients and distance measurements tested).

2. Exchange of information on a network:
   (1) Construction of a common protected database.
   (2) Development of a data-processing platform via Internet to distribute the information.

3. Training/teaching:
   (1) Establishment of continuing education dedicated to law enforcement forces.

2. Target compounds for profiling and choice of the analytical method

The usual issue is focused on the compounds of choice to characterize and compare the chemical profile of cocaine samples. We notice that many studies have shown that cocaine samples contain several compounds many of which are alkaloids. Some of them come from the coca leaf and are co-extracted with cocaine; others result from a hydrolysis, an oxidation or a thermo-degradation of the cocaine during its analysis [5]. Fig. 1 shows a diagram of the formation and decomposition for some alkaloids and principal compounds of interest.

Since the first cocaine sample chromatography in 1962 by Brochmann-Hanssen [6] with a non polar stationary phase, many columns and chromatography techniques were tested in order to analyse cocaine in plant material or in powder samples to understand its production [7,8]. The literature describes different classes of compounds that could be used for the determination of chemical profiles. Organic compounds, solvents [9–12], inorganic compounds [13] or isotopic ratio [14] have been investigated for this purpose.

Moore [15,16] was the first to detect and describe alkaloids other than cocaine by GC at the beginning of the seventies. Moore and Casale have reviewed the analysis of cocaine and its precursors [7]. An important new point is the use of a
derivatisation agent for the sample preparation. Since that time, the knowledge of the cocaine production having become a priority for the fight against drug trafficking in the 1980s, researchers have been very prolific concerning the study or the discovery of compounds contained in cocaine samples. But all these compounds are unfortunately not detected using a unique method and a decision has to be taken concerning the choice of the compounds which will represent the cocaine samples’ signature in order to establish links between samples. Many authors have carried out cocaine profiling but two GC methods stand out, those of Janzen [17] and Casale [18] which are at the origin of the method used in this research.

The profiling method has to comply with different criteria. First, results have to be obtained quickly in order to provide, in quasi real time, useful information to the police forces and justice (purity). Moreover, it is important to highlight a chemical signature of cocaine samples which is reliable and discriminating when dealing with profiling. Guéniat and Esseiva PhD researches [19,20] have focused on a method which allows in one run the quantification of cocaine, the determination of major cutting agents as well as the separation of the major cocaine alkaloids. They obtained a cocaine sample profile with GC-FID reducing the number of compounds of interest to seven: ecgonine methyl ester (Eme), ecgonine, tropacocaine, benzoylecgonine (BenzyloEc), nor-cocaine, cis- and trans-cinnamoylcocaine. A eighth compound was added to the profile: the 3,4,5-trimethoxycocaine (Tmc) [21] which was already used in Lyon and detected with the optimized method.

The following parameters was tested in order to verify and validate the new method developed and demonstrate its capacity to be used for profiling: the selectivity of the method (retention time study), the choice of a derivatisation agent improving the chromatography, the cutting agents influence (matrix effect), the influence of the sample storage conditions, the sample quantity to weigh for analyses and the stability of the target alkaloids according to the sample purity.

A further study will be to select the best statistical method to compare the analytical results obtained in the two different laboratories. Intra and inter laboratories variability will be studied in order to evaluate threshold values to be used in the decision process concerning the presence of a link or not between different samples (to be submitted as part II).

3. Materials

3.1. Samples

The following street cocaine samples, stored at IPS and seized between 2000 and 2004, were analysed:

(1) Sample A belongs to a seizure containing three samples whose purity was at 85%; it was used for retention time study and for derivatisation and cutting agents’ tests.

(2) Sample B belongs to a seizure containing three samples whose purity was at 74%; it was used to determine the influence of storage conditions.

(3) Sample C belongs to a seizure containing four samples whose purity was at 54%; it was used for the sample mass test.

(4) Sample D belongs to a seizure containing five samples whose purity was at 84%; it was used for the stability study in a matrix at 20% and 84%.

3.2. Standards and chemicals

Several standards were necessary for the analysis but those of tropacocaine, cis- and trans-cinnamoylcocaine and Tmc were not available. We obtained tropacocaine by extraction from cocaine seizures.

Table 1 describes the different standards, solvents and derivatisation agents used for the analyses.

3.3. Tropacocaine extraction

The extraction of tropacocaine was achieved by thin layer chromatography (TLC) using the procedure proposed by Mari et al. [22]. A cocaine sample of known composition and purity containing tropacocaine was dissolved in ethanol and applied to a chromatographic plate activated beforehand by a solution of potassium hydroxide (0.1 M) in methanol. The visual inspection under ultraviolet (254 nm) allowed the detection of several spots, one of which being tropacocaine (confirmed by GC-MS analysis). Tropacocaine was extracted from the plate by scraping off the silica followed by dissolution in ethanol, filtering and evaporation under nitrogen [23].

3.4. Sample preparation

Trimethylsilylation with MSTFA was selected according to the technique developed by Moore [24] and after testing several derivatisation agents (see Section 4). Each of the cocaine samples was dissolved in 500 μl of chloroform/ pyridine (5:1) containing 1 mg/ml of heneicosane as internal standard (IS). Hundred microlitres of MSTFA were added and the vial was shaken vigorously and placed in an ultrasonic bath during 15 min. Finally, it was placed in the oven for 60 min at 80 °C.

3.5. Instrumentation

A Perkin-Elmer Autosystem XL gas chromatograph interfaced with a flame ionisation detector (FID) and equipped with a split/splitless injection system was used at the IPS. An Agilent Technologies 6850 gas chromatograph interfaced with an FID and equipped with a split/splitless injection system was used at the LPS. The following conditions were selected: Column: DB-1.

Table 1

<table>
<thead>
<tr>
<th>Standards, solvents and derivatisation agents used for the analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Product</strong></td>
</tr>
<tr>
<td>Heneicosane</td>
</tr>
<tr>
<td>Chloroform</td>
</tr>
<tr>
<td>Pyridine</td>
</tr>
<tr>
<td>Cocaine</td>
</tr>
<tr>
<td>Ecgonine methyl ester (Eme)</td>
</tr>
<tr>
<td>Ecgonine</td>
</tr>
<tr>
<td>Benzyloecgonine</td>
</tr>
<tr>
<td>Norcocaine</td>
</tr>
<tr>
<td>MSTFA</td>
</tr>
<tr>
<td>BSA</td>
</tr>
<tr>
<td>TMSI</td>
</tr>
<tr>
<td>BSTFA + TMCS 1%</td>
</tr>
<tr>
<td>Phenacetin</td>
</tr>
<tr>
<td>Lidocaine</td>
</tr>
<tr>
<td>Caffeine</td>
</tr>
<tr>
<td>Procaiene</td>
</tr>
<tr>
<td>Glucose</td>
</tr>
<tr>
<td>Mannitol</td>
</tr>
<tr>
<td>Lactose</td>
</tr>
<tr>
<td>Inositol</td>
</tr>
</tbody>
</table>
30 m x 0.25 mm i.d. x 0.25 µm film thickness; carrier gas: Helium, 22 psi; injector temperature at 230 °C; oven programme: isothermal for 1 min at 180 °C, 4 °C/min to 275 °C and hold for 2.25 min; injection volume: 1 µl (Lausanne) and 2 µl (Lyon), split at 20:1.

4. Experimental

4.1. Samples comparison

Samples were compared using the methodology developed and routinely used at the IPS for the illicit cocaine and heroin seizures [25]. First, sample areas are normalised by the cocaine area in order to avoid or decrease the influence of the dilution and other analytical factors such as variations in volume of samples injected. Then, comparison between samples is done by means of the cosine function which allows calculating the correlation coefficient (C) between two samples [26].

For each sequence, the same street sample was analysed several times in order to control the quality of the experiment and evaluate the inter variation of the sample. These different replicates were compared using the cosine function. The mean correlation value (C = 99.89) and the standard deviation (S.D. = 0.10) were calculated. These values were used as criteria to evaluate if two samples compared are linked or not. If the correlation coefficients were under the determined limits, then samples were considered as unlinked.

4.2. Chromatographic method

Cocaine profiling consists in selecting the target compounds of the complex matrix in order to extract a chemical profile for comparison purposes.

Fig. 2. (a) Optimized chromatogram of a cocaine sample analysed in Lausanne (CH). (b) Optimized chromatogram of a cocaine sample analysed in Lyon (F).
In order to verify that the eight compounds selected are suitable for sample comparison, we studied their respective chromatographic behaviour under different analytical conditions.

Fig. 2a and b represents the optimized chromatograms of cocaine samples obtained at Lausanne and Lyon. Target alkaloids are well separated with the temperature programme used on different analytical instruments.

4.3. Selectivity, retention parameters

Since cocaine samples can contain different cutting agents, it was important to compare their retention times to those of target alkaloids in order to determine potential overlaps. For example, phenacetin may not be totally derivatised and its underivatized peak does coelute with ecgonine. Consequently, the ecgonine peak area cannot be used for profiling in this case. In the same manner, non-derivatised lactose co-elutes with Tmc.

Absolute and relative retention times were used to identify the components in cocaine samples.

Table 2 shows the absolute and relative retention time (RT and RRT) of cocaine alkaloids and cutting agents in relation to the internal standard. According to the usual limits used, excellent results with RSDs largely inferior to 5% were obtained in most cases. Therefore, both the retention times of the alkaloids and cutting agents are reproducible.

4.4. Derivatisation agent and cutting agent influence

The profiling is routinely realised starting from the eight major alkaloids encountered in cocaine samples analysed by GC-FID with a MSTFA derivatisation. The purpose was to test several derivatisation agents on samples cut with current cutting agents in order to improve the chromatography of some problematic substances (especially sugars). The trimethylsilylation and acetylation of the hydroxyl groups in the alkaloids are well known methods to improve their chromatography. Following the literature, BSA, TMSI and BSTFA + TMCS 1% were chosen [18,24,27].

Sample A was selected as a reference sample for the comparisons. This sample was cut with the most common adulterants and diluents found in Europe: phenacetin, lidocaine, caffeine, procaine, lactose, inositol, mannitol and glucose. Mixtures with each were made at different levels of purities (10%, 30%, 50% and 70% of cocaine), representing 512 analyses.

Profiles thus obtained, major alkaloids were singled out in order to compare the different profiles and study which derivatising agent is the most appropriate for profiling given the presence of common adulterants. Each normalised sample was compared to the reference sample with the cosine function and the results are represented in the Fig. 3. The lower and upper limits represent one and three times the standard deviation of the reference sample.

Fig. 3 shows rapidly that the best correlations are obtained with MSTFA which was selected as the derivatisation agent of choice for cocaine profiling. Ninety three percent of the samples derivatised with MSTFA match with the reference sample whereas only 89%, 65% and 77% of the samples match when they are derivatised with BSA, TMSI and BSTFA, respectively.

Moreover, MSTFA is the only one to make possible the profiling for samples cut with lactose or mannitol whatever the sample’s purity. This information is important because these two diluents are very often used by traffickers. Some important chromatographic problems appear with glucose, particularly when its

Table 2

| Table 2 Absolute and relative retention time (RT and RRT) of cocaine alkaloids and cutting agents in relation to the internal standard |
|--------------|-----------------|----------------|-----------------|
| Alkaloids     | RT (min)        | RSD<sub>RT</sub> (%) | RRT (min)       | RSD<sub>RRT</sub> (%) |
| Eme          | 3.62            | 0.16            | 0.33            | 0.19             |
| Phenacetin (1)| 3.76            | 0.48            | 0.34            | 0.46             |
| Phenacetin (2)| 4.35            | 0.16            | 0.39            | 0.14             |
| Ecgonine     | 4.40            | 0.08            | 0.40            | 0.10             |
| Lactocaine (1)| 5.55            | 0.48            | 0.50            | 0.43             |
| Caffeine     | 6.08            | 1.46            | 0.55            | 1.40             |
| Lactocaine (2)| 7.10            | 0.28            | 0.64            | 0.23             |
| Glucose (1)  | 7.23            | 0.14            | 0.65            | 0.08             |
| Glucose (2)  | 8.18            | 0.41            | 0.74            | 0.37             |
| Tropacocaine | 8.35            | 0.62            | 0.75            | 0.60             |
| Mannitol     | 9.18            | 0.45            | 0.83            | 0.38             |
| Glucose (3)  | 9.73            | 0.12            | 0.88            | 0.07             |
| Internal standard | 11.14       | 0.08            | 1               | –                |
| Inositol     | 12.19           | 0.43            | 1.10            | 0.40             |
| Cocaine      | 12.92           | 0.60            | 1.16            | 0.63             |
| Procaine     | 13.27           | 0.64            | 1.19            | 0.60             |
| BenzoylEc    | 14.28           | 0.05            | 1.28            | 0.10             |
| Norcocaine   | 14.70           | 0.03            | 1.32            | 0.08             |
| Cis          | 16.29           | 0.12            | 1.46            | 0.16             |
| Trans        | 18.96           | 0.07            | 1.70            | 0.12             |
| Lactose (1)  | 23.33           | 0.34            | 2.10            | 0.32             |
| Sucrose      | 23.49           | 0.27            | 2.11            | 0.29             |
| Tmc          | 24.16           | 0.02            | 2.17            | 0.04             |
| Lactose (2)  | 25.64           | 1.22            | 2.31            | 1.23             |

Note: phenacetine, lactose and lidocaine give two peaks (1) and (2), glucose gives three peaks (1) (2) and (3), RT: retention time; RRT: relative retention time; RSD: relative standard deviation; Eme: ecgonine methyl ester; BenzoylEc: benzoylecgonine; Cis: cis-cinnamoylecgonine methyl ester; Trans: trans-cinnamoylecgonine methyl ester; Tmc: 3,4,5-trimethoxyecgonine.

Fig. 3 shows rapidly that the best correlations are obtained with MSTFA which was selected as the derivatisation agent of choice for cocaine profiling. Ninety three percent of the samples derivatised with MSTFA match with the reference sample whereas only 89%, 65% and 77% of the samples match when they are derivatised with BSA, TMSI and BSTFA, respectively.
concentration is higher than 70%. Hence, results of profiling should be taken with caution when such is the case.

It was also observed that whatever the cutting agent used, bad results were obtained using BSA and BSTFA + TMCS with samples cut with agents at a concentration superior or equal to 90%.

The profiling is based on the hypothesis that samples that come from the same batch should be linked whatever the derivatisation or cutting agent used. Even if results obtained with BSA, TMSI and BSTFA + TMCS demonstrate their incapacity to chromatograph cocaine samples well derivatisation with MSTFA results in an excellent correlation coefficient for the large majority of the blends tested and was therefore selected.

4.5. Influence of the samples’ storage condition

The influence of the storage condition on the profiling was investigated because samples are being stored at different uncontrolled locations in France and Switzerland. Moreover, this is a common question regarding best practice considering sample preservation.

The street sample B used for this test was stored in a secure room at ambient temperature (closed vial). Its analysis at time zero was decided to be the reference value. The stability of its chemical profile was tested after submission to the following three different conditions: storage in a secure room at 20 °C; in a drying oven at 37 °C and in a refrigerator at 5 °C. Furthermore, for each storage condition, two samples were put inside a desiccator and two others outside a desiccator (one in an open vial, the other in a closed vial).

Two 10 mg samples, coming from the different storage conditions, were analysed each week during 3 months. A calibration curve was done for each analysis to control the procedure and quantify the cocaine sample.

The temperature and humidity variations, inside and outside the desiccator, were recorded with an ELPRO® (elproLOG, 3.2×) sensor.

Sensor 1: Temperature outside desiccator ($T_{\text{outside desiccator}}$).
Sensor 2: Humidity outside desiccator ($H_{\text{outside desiccator}}$).
Sensor 3: Temperature inside desiccator ($T_{\text{inside desiccator}}$).
Sensor 4: Humidity inside desiccator ($H_{\text{inside desiccator}}$).

Table 3 represents the averages of the temperature and humidity obtained during the experiment. Relative standard deviations are represented in subscript.

The important RSDs obtained for some measures show that temperatures and humidity were not constant within all storage conditions. The effects of these variations on the profiling were studied.

As explained earlier (samples comparison), peak areas of each analysis were normalised and compared to the reference sample with the cosine function.

Fig. 4 represents the influence of the temperature and humidity on samples stored in a vial open or closed and Fig. 5 shows the results obtained for samples stored outside and inside a desiccator. Validation limits are represented by the standard deviation.

![Fig. 4. Correlation coefficients comparisons of the reference sample stores during 3 months in an opened and closed vial.](image1)

![Fig. 5. Correlation coefficients comparisons of the reference sample stores during 3 months inside or outside a desiccator.](image2)
Figs. 4 and 5 show that all the samples were linked whether the vial was capped or not and with varying humidity. It is therefore possible to compare all the samples and the environmental factors did not influence significantly the comparison. Fig. 6 summarises the correlations obtained and shows that all the samples profiles have the same appearance; in other words, over the period of study (3 months), temperature, humidity and time have no measurable influence on the profiling. Consequently, profiling can be performed on cocaine samples stored during up to at least 3 months, at room temperature, in closed vials (to avoid contamination between samples) without any further precaution.

4.6. Sample mass for analysis

Another important factor to be tested is the sample quantity weighed for analyses. A compound which is injected in a too important quantity saturates the detector and thus influences the reliability of the analysis. But the quantity has to be sufficient to see all the selected alkaloids. Three different cocaine sample quantities were tested (sample C): 8, 10 and 12 mg (8 and 12 mg are usually used in Lausanne and in Lyon, respectively).

Three replicates by weight were analysed twice. Peak area of each alkaloid was divided by the internal standard peak area. Relative standard deviations are represented in Fig. 7.

Except for Tmc which was under the detection limits, very good results were obtained for the different alkaloids. RSDs were usually below 5% with all three injection masses. During analysis the cocaine peak saturated the detector quickly at 12 and 8 mg was below the good detection limit. Ten milligrams was selected to proceed with the research.

4.7. Stability study of the target alkaloids

Depending on the number of samples to analyse, it is possible that a sample remains several hours in solution before being injected. The stability of the alkaloids was tested during 24 h in three different conditions studying order to measure any matrix effect:

1. Individual stability in the solvent used for analysis.
2. Stability in a matrix at 84%.
3. Stability in a matrix at 20%.

4.7.1. Standards stability

The stability of Eme, ecgonine, tropacocaine, benzoylecgonine and norcocode was measured. Tropacocaine was obtained using thin layer chromatography (TLC) separation as explained previously.

3 mg × 1 mg of each standard compound were analysed every 3 h during 24 h. Fig. 8 shows the relative standard deviation (RSD in %) for the peak area of each alkaloid normalised by the internal standard peak area.

A very good stability of all alkaloids inside the matrix studied was observed (indeed RSDs are all inferior to 4%).

4.7.2. Stability in a matrix at 84%

After this first study, the behaviour of the alkaloids was examined when they are all in the same matrix and without cutting agent.

The cocaine sample D was pure at 84% and contained all the alkaloids studied. Three samples of 10 mg were weighed and analysed 15 times over 6 days. Fig. 9 shows the relative standard deviation (RSD in %) for peak area of each alkaloid normalised by the internal standard peak area.

A very good stability of all alkaloids inside the matrix studied was observed (indeed RSDs are quasi all inferior to 3%).

4.7.3. Stability in a matrix at 20%

The cocaine sample used is the same as previously (sample D). Inositol was selected to obtain a matrix at 20%. Three samples of 10 mg were weighed and analysed five times over 2 days. Fig. 10 shows the relative standard deviations (RSD in %) for the peak area of each alkaloid divided by the internal standard peak area.

Bad results were obtained for norcocode (sample 3) and Tmc. These different peaks were indeed rarely or not detected because they were under the detection limits of the instrument. Thereby their stability could not be...
measured correctly, as it was observed before with the study of the cutting agent influence.

However it was known that these different samples were linked because they were selected from the same batch. When applying the calculation method previously used to determine the correlation coefficient between cocaine samples, this coefficient was found to be 99.89% and confirmed that samples were indeed well linked. It can be concluded that samples are stable also in a matrix at 20%.

5. Conclusion

The purpose of the research was to compare cocaine chemical profiles analysed on two different GC-FIDs in two different laboratories in order to determine how results could be compared in a drug intelligence perspective. The first step was focused on the optimisation of the analytical method and its robustness. The verification of some chromatographic parameters (such as the retention time stability, detection of peak overlapping, determination of the ideal quantity to perform analyses) constituted the initial step.

Secondly, in order to correct and improve the target alkaloids chromatography, several derivatisation agents were tested and MSTFA was found to be the most useful with the method used.

The influence of cutting agents on the profiling was also tested and the limits of the profiling were determined depending on the cutting agent and cocaine concentration.

Furthermore, the profile comparison of a sample stored in different conditions showed no significant influence on the profile by all temperature and humidity variations evaluated.

Finally, stability studies of the target compounds for cocaine samples profiling were done in different matrices. Ecggonine methyl ester, ecgonine, tropacocaine, benzoylecgonine, norcocaine, cis-and trans-cinnamoylcoecgonine and 3,4,5-trimethoxy-cocaine are compounds that were shown to be convenient to profile cocaine samples.

Once the method was optimized, statistical methodologies allowing sample comparison needed to be tested for optimal profile determinations. The principal alkaloids are selected as quantitative and continuous variables for the chemical comparison process in order to compare the chemical profiles.

The pre-treatment step becomes an important part of the chemometric analysis and is part of the ongoing project. Finally, different statistical methods (distances and correlation coefficients) will be tested in order to characterize the quality of the established links between samples. Pre-treatment and treatment will be fundamental steps in the comparison of samples analysed on both side of the Franco-Swiss border and may help create the basis for international collaboration and comparison in drug intelligence.

Acknowledgments

The authors acknowledge Swiss Confederation and European Community for funding the present research project (INTERREG III A, 1/FU/9.3/1).

Nicole Egli and Daniel Alvarez are acknowledged for their technical support.

References

Sugar and fatty acid analysis in ecstasy tablets

Ines Baer*, Pierre Margot

Institut de Police Scientifique, University of Lausanne, BCH, Lausanne-Dorigny 1015, Switzerland

Received 8 June 2006; accepted 14 June 2006
Available online 4 August 2006

Abstract

Sugars and stearates (composed of fatty acids) are both frequent components used in the production of ecstasy tablets. Their analysis can therefore provide supplementary information useful for drug intelligence. Links established using these substances would be very significant as they should give us information about the manufacturer of the tablets. Two methods have been developed for the analysis of sugars and fatty acids by GC–MS and were applied to 109 ecstasy tablets. Characterisation of the samples should allow the differentiation of a certain number of them and furthermore their classification into groups.

This is obtained by analysing the raw data using chemometric methods. Several pre-treatments have been tested together with six similarity measures on a small number of ecstasy samples in order to determine which combination would best characterise one ecstasy sample and differentiate it from the others at the same time. Normalisation followed by the fourth square and applied together with the squared cosine function appeared to give the best results and has been applied to all samples. The correlation values obtained of each sample with all others express the probability of a presence of a link between two samples.

In order to verify the signification of these values, and thus of a link, all samples have been compared considering the data visually according to three selected criterions. The 109 examined samples could be divided into 67 groups, with 43 of them containing only one sample. Considering the distribution of their correlation values, sample pairs showing a value below 0.23 can be considered as linked. As the excipients are necessarily related to the blending, which also includes the active substance, and variation in the excipient content has been proven by the grouping of the samples, a low similarity value does indicate a link with regard to the producer. In conclusion, it appears that the result obtained with the excipients is certainly very valuable, but all other available information has to be taken into account as well before making any conclusions.

© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Ecstasy; Sugar; Fatty acids; Chemometrics; Profiling

1. Introduction

Contrary to the other “common” narcotics, such as cocaine or heroine, ecstasy is by definition a tablet and therefore has to be treated differently. Similarly to the others, the main substance of ecstasy is mixed with other components. The difference in the case of tablets is that these components are not defined as cutting agents, but as excipients. The term is of pharmatechnological origin and designates substances necessary to produce solid forms of drugs, such as tablets and capsules. The important point for investigation is that once excipients are added to the active substance and the tablet is pressed, its composition hardly ever changes (eventually decomposition or very slight changes in quantities—depending on the quality of the tablet). Therefore, a seized tablet presents the same composition than at the initial production.

The fabrication of ecstasy can be divided into three steps: (1) synthesis of the active substance, (2) addition of excipients (and eventually adulterants), and (3) the compression. These steps can be carried out by the same laboratory or by separate facilities. Consequently, tablets of the same visual appearance might not present the same composition and vice versa. This fact must be taken into account when trying to compare and group ecstasy seizures.

The comparison of ecstasy tablets can be done at different levels:

(1) By trace analysis of the active substance which is characterised by the way of synthesis. This was investigated for amphetamine in the SMT Project [1], and is currently examined for MDMA in another project (CHAMP).
(2) By visual and physical description, such as logo, colour, diameter, etc. [2].

(3) Finally, it is possible to consider substances such as excipients. Dyes have been examined by Goldmann [3]. Herein, carbohydrates (used as diluents and binders) and fatty acids (components of stearates which are used as lubricants) are the focus of the research.

The aim of comparison at any level is to establish eventual links between different samples. By characterising the tablets, it is possible to differentiate a certain number and furthermore classify them into groups which allow an identification of the source of production. This information can be used to estimate for how long a laboratory was in production, as well as to know how the illicit drug was distributed and to what extent. This form of intelligence allows a strategic vision of the illicit drug market and is usually resumed under the name of Drug Profiling.

The purpose of the presented research project is to evaluate the potential of the two types of excipients chosen - carbohydrates and fatty acids - for providing useful intelligence in the context of ecstasy profiling.

2. Experimental

2.1. Sample preparations

One hundred and nine ecstasy seizures have been selected. One tablet per seizure was ground and homogenised in an agate mortar. Three weigh-ins were taken for each type of analysis and each weigh-in was analysed twice. Thus, six measures were obtained per sample and method. The sample preparation for carbohydrate analysis has been taken from Sweeney et al. [4]. About 2 mg of ecstasy powder are dissolved in 1 ml pyridine. Then, 200 µl of hexamethyldisilazane (HMDS) and 100 µl of trimethylchlorosilane (TMCS) are added and the mixture is vortexed for 30 s. The solution can be analysed after having rested up to 180 min.

As for fatty acid analysis, the procedure is based on Iverson et al. [5], but has been modified. About 25 mg of the ecstasy sample is mixed with 1.5 ml nitrogen. The vials are closed tightly and put into the oven for 1 h at 175 °C. The oven temperature program was: 140 °C for 2 min, followed by heating up to 230 °C with a rate of 5 °C/min and held for 2 min. The injector and GC–MS interface temperatures were set to 250 and 280 °C, respectively.

3. Results and discussion

3.1. Determination of a comparison method

An adequate comparison method had to be determined for the characterisation of linked samples and a good separation from non-linked samples. For the linked sample population, the replicas of each tablet have been used. The non-linked sample group was an arbitrary choice based on the information available, such as administrative data, physical and chemical characteristics. We would like to emphasise that the analysed sample size allows only an exploratory analysis whose results have to be considered with caution. Consequently, the determined comparison method does not represent a final result, and another method might appear more appropriate with a larger sample set.

Table 1 shows the variables which have been used for sample comparison.

3.2. Choice of a pre-treatment

For this first part, only a small sample set of 15 tablets has been used. The variables in Table 1 reveal that missing peaks had to be expected as only very few samples present all these peaks. The choice of a pre-treatment included the question how to handle these missing peaks. The replacement by 0, 200 and an additional higher limit value were tested [1]. The latter was determined in function of the minima area integrated by both analytical methods. Finally, two minimal and maximal values (with regard to noise) were chosen and tested as replacement: 0, 200, 1E4 and 6E4.

Six combinations of pre-treatments have been tested. Taking into account the replacement of the missing peaks, this procedure resulted in the application of twenty pre-treatments. The latter have been evaluated by considering the false positive, the estimate of discrimination, and the confirmation by PCA analysis [1,6]. The combinations which appeared to be the most appropriated are represented in Table 2.

3.3. Choice of a comparison method

The chosen pre-treatments were applied together with six similarity measures (i.e. Pearson correlation [1], Squared cosine function [7,8], Similarity Index [9,10], Canberra Index...
of the histograms, especially the overlapping zone. The latter should be small, or completely absent, and resemble a regular triangular shape. As for numerical values—the maximum value of the linked sample group (MaxL), the standard deviation of the linked sample group (STDEV_L) and the false positives (%FP) and negatives (%FN) had to be small. In contrast to this, the minimum value of the non-linked sample group (MinNL), the estimate of discrimination (D), and the percentage of linked samples below the minima value of the non-linked sample group (%L < NL) had to be high.

The selected method was the Squared cosine function applied after N4R (6E4) pre-treatment. A general view of the histogram and a zoom of the overlapping zone are shown in Fig. 2.

3.4. Comparison of all ecstasy samples

The selected comparison method has been applied to the mean values of the peak areas of all 109 ecstasy samples, thus giving correlation values for each possible sample pair (altogether 5886). The correlation value expresses the probability of a presence of a link between two samples. In case of a small value, the presence of a link becomes very probable. Whereas if it is high, the absence of a link is more likely.

In order to verify the signification of these values, the data of all samples have been visually compared according to the following three criterions. (1) Qualitative—samples were grouped according to their active substance, the type of sugar and the fatty acids. For the latter, three distinct types appeared, which could be clearly differentiated: C16 < C18, C16 ≈ C18 and C16 > C18. (2) The proportions of the two or three main peaks, i.e. the principal sugar peak and the palmitic and stearic acid. (3) The distribution of the minor fatty acids. The histogram in Fig. 3 reveals that the profile does not change for the six replicas of the same ecstasy sample, but can vary between different samples.

The 109 examined samples could be divided into 67 groups, with 43 of them containing only one sample. The distribution of the correlation values below 1 is represented in Fig. 4.

![Fig. 1. Schema of a histogram with some comparison criterions.](image1)

![Fig. 2. Squared cosine function after N4R (6E4). Histogram overview (left) and extract (right).](image2)
threshold of 0.23 (MinNL) determined in the previous chapter, seemed to be appropriate, since more than 90% of the group values and only two non-grouped sample pairs were below this value. Therefore, sample pairs showing a correlation below 0.23 can be considered as linked.

However, the two non-grouped sample pairs presenting a very high correlation (a very low value) have to be explained. The lower value (Fig. 4(1)) was obtained with two samples containing a different illicit substance and almost no excipients. Consequently, they are identical with respect to the excipients, as there is no data to show any variation. Consequently, it has to be considered that when no excipient can be detected, much care has to be taken for the comparison of the results. Nonetheless, the absence of excipients can also be considered as a significant characteristic as it is rare to find nothing except the illicit substance. However, all other available information (visual and physical characteristics) has to be taken into account before drawing any conclusions. This is particularly significant in the second case (Fig. 4(2)), where the sample pair presents indeed a similar excipient content, but shows differences in the active substance and physical characteristics.

Finally, we evaluated the significance of the link produced by excipient analysis. When considering the basic production steps as presented in Fig. 5, it appears that the excipients are necessarily related to the blend before compression.

The most general assumption which can be made is that a high correlation indicates that the blends used for the two corresponding tablets contained the same excipient, except in the previously mentioned case of lacking excipients. This correspondence can be very specific with regard to the fatty acids which are the ingredients of metallic stearates used as lubricant. Despite the name, stearates also contain palmitates

---

**Fig. 3.** Fatty acid distribution for six replica of ecstasy sample 1301 (left)—and for sample 1301 and 1507A (right).

**Fig. 4.** Distribution of correlation values according to the groups.
and other fatty acid salts. The US pharmacopoeia states that stearate and palmitate together should account for not less than 90% of the fatty acid content. This still leaves a considerable range of materials to be supplied as stearate and it can be expected to observe variations in the fatty acid content. This was confirmed by the various number and proportions of fatty acids observed in the analysed ecstasy samples. Furthermore, it is recommended to use a single supplier for a given formulation as stearate from a new supplier might have other effects on compaction and dissolution.

Another interesting point concerning the presence of stearates is its hydrophobic character. It has been shown that it might reduce dissolution rate and bioavailability of several acids observed in the analysed ecstasy samples. Furthermore, it was confirmed by the various number and proportions of fatty acids from a new supplier might have other effects on compaction and dissolution.

Another interesting point concerning the presence of stearates is its hydrophobic character. It has been shown that it might reduce dissolution rate and bioavailability of several acids observed in the analysed ecstasy samples. Furthermore, it was confirmed by the various number and proportions of fatty acids from a new supplier might have other effects on compaction and dissolution.

Although focussing on the chemical composition of the ecstasy tablets, their physical characteristics need to be considered as well, since they are also related to the producer. All determined groups are composed of tablets presenting identical physical characteristics. Only one exception has been found where two different logos were observed (diamond + double lightning). The corresponding tablets are however linked on the basis of the physical characteristics [2], i.e. presenting the same particular tablet shape and strong colour.

The physical correspondence was therefore confirmed by the excipient composition. Nonetheless, there were several tablets showing identical physical properties and as a result were sometimes attributed to the same class, but were differentiated by their excipients. In these cases care should be taken before making any classes as the significance of the various characteristics must be discussed.

4. Conclusions

We showed variation in excipient content leading to the classification of the samples into groups. In particular, fatty acids appeared to give specific profiles which were useful for sample characterisation. The variation observed between the groups was confirmed by the similarity values we obtained. The evaluation of these data gave rise to the conclusion that a high correlation in excipient content (corresponding to a low similarity value) indicates a link towards the tablet producer.

However, we furthermore showed that it is important to consider all available information about the remaining characteristics of the tablets. This is even imperative when dealing with tablets where no excipient has been detected by GC–MS analysis.

References

[13] M. Perkal, Y.L. Ng, J.R. Pearson, Impurity profiling of methampheta-
NIR analysis of cellulose and lactose—Application to ecstasy tablet analysis

Ines Baer a,*, Robert Gurny b, Pierre Margot a

a Institut de Police Scientifique, University of Lausanne, BCH, 1015 Lausanne-Dorigny, Switzerland
b School of Pharmacy, Department of Pharmaceutics and Biopharmaceutics, 1211 Geneva, Switzerland

Received 6 June 2006; accepted 14 June 2006
Available online 24 July 2006

Abstract

Cellulose and lactose are the most frequently used excipients in illicit ecstasy production. The aim of this project was to use near infrared reflectance spectroscopy (NIRS) for the determination of the different chemical forms of these two substances, as well as for the differentiation of their origin (producer). It was possible to distinguish between the different chemical forms of both compounds, as well as between their origins (producers), although within limits. Furthermore, the possibilities to apply NIRS for the analysis of substances such as found in illicit tablets were studied. First, a few cellulose and lactose samples were chosen to make mixtures with amphetamine at three degrees of purity (5, 10 and 15%), in order to study the resulting changes in the spectra as well as to simultaneously quantify amphetamine and identify the excipient. A PLS2 model could be build to predict concentrations and excipient. Secondarily, the technique was to be applied to real ecstasy tablets. About 40 ecstasy seizures were analysed with the aim to determine the excipient and to check them against each other. Identification of the excipients was not always obvious, especially when more than one excipient were present. However, a comparison between tablets appeared to give groups of similar samples. NIR analysis results in spectra representing the tablet blend as a whole taking into account all absorbing compounds. Although NIRS seems to be an appropriate method for ecstasy profiling, little is known about intra- and intervariability of compression batches.

# 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Cellulose; Lactose; Determination of origin; NIR; Ecstasy; Excipients

1. Introduction

Near infrared reflectance spectroscopy (NIRS) has gained great importance in pharmaceutical technology. Its applications are diverse starting with on-line measurement in pharmaceutical manufacturing [1,2], moisture measurement [2,3], verification of quality and identity of products [4,5], measurement of polymorphism or degree of crystallinity [6,7], study of physical properties due to manufacturing (e.g., particle size) [8], identification and comparison of excipients [9–11], and many more [12–14]. In the forensic context the technique has been applied for the identification and quantification of the illicit substance in ecstasy tablets [15–17]. The great advantage of this technique lies in the fast analysis (<1 min) and the simple sample preparation.

In this project, NIR was applied to the study of lactose and cellulose in order to determine the different chemical forms of these two substances, as well as the differentiation of their origin (producer). This could be useful in a forensic context, e.g., for the highlighting of a production net for example, and also for cases of counterfeit legal drugs.

Cellulose could not be analysed before in our laboratory and it was suspected to be present in ecstasy tablets where no or only a small amount of sugars were detected. Therefore, the technique was to be applied to real ecstasy tablets with the aim to determine the excipient and to check them against each other. Additionally, mixtures of amphetamine and chosen cellulose and lactose standards were prepared in order to test quantification and simultaneous identification of the excipients. This was tested in view of a possible application in routine drug analysis.
2. Experimental

2.1. Materials

2.1.1. Cellulose/lactose standards

Twenty-five cellulose (15 suppliers) and 23 lactose (10 suppliers) samples were used in this study. In the beginning, all samples were analysed as powders directly in their containers. Then, all samples were compressed and measured partly as a whole tablet (10 measures on every side), partly as crushed powder (10 measures).

2.1.2. Amphetamine blends

Six cellulose (four microcrystalline and two methylcellulose) and seven lactose standards (four lactose monohydrate and three anhydrous lactose) were chosen to make mixtures with amphetamine at three purities (5, 10 and 15%). Amphetamine was chosen due to its high availability. All the mixtures were homogenised in an agate mortar before analysis.

2.1.3. Ecstasy

Thirty-nine ecstasy tablets from 35 seizures have been analysed. The configuration of the instrument did not allow an analysis in the solid phase, due to the fact that the surface of the tablets was smaller than the surface of the probe. Thus, the samples were homogenised in an agate mortar before analysis. Five measurements were recorded for every tablet, taking special care that the powder was regularly shaken to vary particle disposition. In addition to lactose and cellulose standards, maize starch, potassium carbonate, talc, silica, sodium bicarbonate, maltodextrine, maltose, sorbitol, and Mg stearate (Sigma–Aldrich, Switzerland) have been analysed.

2.2. Instrumentation

Analyses were realized on a FOSS SmartProbe™ Analyzer 6500, equipped with a monochromator for 400–2500 nm coverage and a SmartProbe module with integrated detector and a 3 m long optic interactance fibre bundle. The spectra were recorded in diffuse reflection mode in the full wavelength range and at 2 nm intervals. Individual spectra were determined as averages of 32 scans. A ceramic standard served as reference.

2.3. Data treatment

2.3.1. Software

The software used for the acquisition of the spectra and first observation after having applied pre-treatments was Vision®, specifically designed for use with the FOSS NIRS systems. The raw spectra displayed after acquisition are reflectance spectra transformed into absorbance spectra by a logarithmic function (log 1/R) based on the Lambert–Beer law. Chemometric treatments and multivariate analysis were performed with the Unscrambler®, a software developed by Camo, Norway. The data could be directly imported from Vision software by using NSAS files.

2.3.2. Treatments

Several pre-processing methods and combinations were tested to obtain optimal classification or prediction models such as 1/X transformation, standard normal variate (SNV), multiplicative scatter correction (MSC), and derivatives. Ammonium nitrate (1, 8, 9, 12, 18, 19, 20) The 1/X transformation transforms the data points from wavelength units into wave number units, resulting in a more realistic presentation. SNV and MSC have been developed to compensate scatter-induced baseline offsets in reflectance spectra. They consist in a standardisation of the spectra. For the former, every spectrum is treated separately and a normal distribution is not necessary, in contrast to MSC. The latter requires a large set of data and is therefore more likely to produce a true mean spectrum. However, SNV and MSC give similar results. Derivatives are taken to improve the resolution and to reduce baseline offsets. However, this procedure also amplifies spectral noise, and is therefore often combined with a Savitzky–Golay algorithm, which includes a smoothing function.

2.3.3. Multivariate analysis

Principal component analysis (PCA) was used since it allows for the detection of sample patterns such as particular groupings, and may distinguish useful information from noise and meaningless variation. Partial least square (PLS) regression has been performed as well. The purpose of the method is: (1) to describe the relationship between a set of predictors (X-matrix) and a set of responses (Y-matrix) (the mathematical form of this relationship is called a model), and (2) to predict new values for which only the X-values are known. As in the case of the PCA, the PLS regression is based on projection principles [13, 18, 19, 21, 22].

3. Results

3.1. Cellulose/lactose standards

3.1.1. Differentiation of the chemical form

In the case of cellulose, seven different forms have been analysed by NIRS. The raw spectra are shown in Fig. 1. While the curves show varying shapes, similarities are visible. First, a
data pre-processing method has been determined in order to produce a PCA model separating the seven types of cellulose. Wavelength selection was necessary to avoid the influence of water, which has strong overtones in the 1450 and 1930 nm region \[2,4,14,18\].

It appeared that complete separation could be achieved, however, with CMC and microcrystalline cellulose (Cell) being close to each other (Fig. 2a). But as it seemed possible to differentiate the seven groups, the pre-processing method was used to create a PLS1 model in order to predict the type of cellulose. Only four of the seven cellulose forms were used for this work (i.e., Cell, CMC, HPMC and MC) as only a few samples were available in the case of the other three. The regression plot of the created model is shown in Fig. 2(b). The model was tested with 21 randomly chosen spectra, all of which have been correctly identified.

For the lactose samples, only three types had to be differentiated: lactose monohydrate (LacMh), anhydrous lactose (AnhLac) and blends (Ludipress, LudipressLCE, Cellactose, MicroceLac, StarLac). The spectra could be easily distinguished and have also been treated for PCA analysis. Lactose monohydrate and anhydrous lactose were already separated by using the raw spectra. However, to differentiate the blends from the lactose monohydrate, pre-processing of the data [1st Savitsky–Golay derivative (11pt average, 3rd polynomial order) and mean normalisation] and wavelength selection was necessary. As expected, the blends which present variable lactose monohydrate content were a little dispersed while those with the highest lactose monohydrate content were very close to the pure standards. A PLS1 model using the same pre-processing method allowed for a correct prediction of 20 randomly chosen lactose samples (no blends).

3.1.2. Differentiation of producers

This part was carried out with spectra from powdered and microcrystalline cellulose (formerly designated as Cell), as this is the chemical form most frequently used as excipient and the most likely to be found in ecstasy tablets. The results shown in Fig. 3 reveal that the samples are not visually distinguishable apart from the two powdered cellulose standards (CFFSanacel and CFFQualicel).

![Fig. 2. (a) PCA after log(1/X) transformation, SNV and 2nd Savitsky–Golay derivative (5pt average, 2nd polynomial order) of the seven cellulose groups; (b) predicted vs. measured regression plot for the prediction of the cellulose types.](image-url)
After PCA, spectra of the same standards appeared to form little subgroups. One strongly influencing factor was probably the compression. For the same substance the measures after compression were differentiated from those before compression. Hence, we decided to consider the compressed and non-compressed spectra separately and found that different producers could indeed be distinguished, but only by using the measures of the powdered samples.

As this was an exploratory research in the context of ecstasy analysis, differentiation after compression was judged more important. Considering the PCA of the spectra taken from tablets, the subgroups were still present and could be explained by the different way of measuring. Measurements of the same standard carried out on the tablet surface, the crushed tablet or on a second tablet would be differentiated by PCA, resulting in subgroups, which sometimes superpose with those of other standards. Thus, an important conclusion is the importance of coherence in sample measuring and processing. When PCA was carried out but with spectra produced in exactly the same conditions, the producers could be differentiated except for Ceolus and Vivapur (designated by the arrow in Fig. 4). Only measures of crushed tablets have been taken for this PCA. The samples that still form subgroups are those where measures have been made on two crushed tablets.

Similar observations were made with the lactose monohydrate standards where we also tried to separate the different origins. A test with measures taken under the same conditions did not show a neat separation as in the case of cellulose. However, the different groups were not overlapping.

To summarize, we showed that differentiation of the producers is possible, however, under specific conditions. We would like to remind that the work was carried out with standard samples directly obtained from the producer—a situation far from “ecstasy reality”. Ecstasies are compressed blends, maybe not very complex blends, but still complex enough not to allow such a specific differentiation. For the determination of the chemical form of the excipient alone it would be necessary to proceed to an extraction. But the results are still very interesting as they demonstrate that already very small differences are detected.
3.2. Amphetamine blends

Before starting any data analysis, the spectra have been compared in order to see if a regular change in concentration could be detected and in what wavelength region it would be observed. In Fig. 5, the three blends of amphetamine with CFFSanacel are shown in the wavelength section where the changes were the most obvious. The three groups can be easily distinguished. Changes due to the varying concentration could also be observed in other wavelength regions (around 1140 and 2200 nm), but were less visible.

The question of a prediction of the illicit substance content in ecstasy has been thoroughly studied by Sondermann and Schneider [15–17], who have shown that good prediction is possible. The method was therefore tested on our amphetamine blends. In addition, simultaneous identification of the excipient was attempted.

Good separation of the excipients and of the three concentrations could be achieved after wavelength selection and pre-processing using log(1/X), SNV and the 2nd Savitsky–Golay derivative (5pt average, 2nd polynomial order). The same pre-processing method and wavelength selection have been tested for the creation of a PLS2 model. In Fig. 6, the predicted values for the amphetamine concentration and the identity of the excipient are plotted against the measured values. As slope and correlation were near one for both Y-variables, it was decided to test the model on a set of selected spectra representing the whole data set used for the construction of the model. The result of the prediction of amphetamine concentration and excipient identity in the samples is represented in Fig. 7.

All data were correctly attributed to the corresponding sample. However, the average deviation for amphetamine concentration is higher than for the prediction of the excipient. Thus, to obtain a more reliable prediction model, blends with more components and a wider range of concentrations should be used. Additionally, data processing could be optimised as only a few possibilities have been tested here.

3.3. Ecstasy

3.3.1. Excipient identification

The identification of the excipient was rather difficult. The principle of NIR analysis implies that the spectrum reflects the

![Fig. 5. Amphetamine blends with CFFSanacel in the region between 1400 and 1800 nm (after SNV treatment).](image)

![Fig. 6. PLS2 after log(1/X), SNV and the 2nd Savitsky–Golay derivative (5pt average, 2nd polynomial order). Predicted vs. measured regression plots for the amphetamine concentration (left) and the excipient (right).](image)
absorbance depending on the various functional groups. This means that substances with similar chemical structure will absorb in similar wavelength regions. Blends will therefore present spectra of superposing absorptions and compounds might only be recognised if they have a particular functional group. This fact explains why MDMA is generally recognised in our tablets. On the other hand, the chemical structures of cellulose and sugars are the same. Since the tablets have all been analysed by GC/MS in the framework of another project we knew if sugars were present, which was very useful in the evaluation of the spectra. The example shown in Fig. 8 was a simple case, where cellulose is easily recognised by the close correspondence of the spectrum to the cellulose standard.

More complicated cases consisted in samples with probably more than one excipient. The shape of the curve did not correspond to the sugar detected by earlier GC analysis. Mixtures were difficult to evaluate, and no tests could be made to verify how a spectrum evolves as a function of the excipient content (cellulose alone, cellulose/lactose, lactose alone). However, in 24 of 39 tablets cellulose was detected, and 3 of them presented lactose monohydrate as well.

Finally, there were six cases where no excipient could be determined apart from a likely presence of talc. The MDMA content of the concerned tablets is either 40 or 60%. Therefore, another substance must be present next to talc, as its absorbance does not seem high enough to explain the 40–60% content found in the tablets (Fig. 9). Unfortunately, an identification of the compound could not be realised.

3.3.2. Comparison of ecstasy samples

Simultaneously to the first part described above, the spectra of the different tablets were compared with each other in order to detect similarities. Those spectra were noted that were identical or very similar after SNV treatment. An example of
two tablets presenting different logos, but an identical spectrum, is shown in Fig. 10.

However, such a visual comparison is longsome and difficult due to the high amount of information present to compare. Furthermore, additional mathematical treatments are necessary to unveil latent information, but which make the spectra more complex. Therefore, the data have been imported to The Unscrambler® software in order to test PCA analysis and to verify the groups formed after having applied pre-processing methods.

The groups formed by PCA are similar to those observed by visual comparison. However, an evaluation of the significance of these similarities is difficult, as measures have been taken on punctual ecstasy seizures and not on one or larger seizures. The latter would have been necessary for a proper interpretation of the group formation. Nonetheless, it was interesting to observe correspondences mostly in samples presenting the same physical and general chemical characteristics. Unfortunately, no further conclusion about any tendency should be drawn from that due to the small sample size.

4. Conclusions

We showed in the cellulose and lactose standard evaluation that already small differences indicate different origins as well. In the case of the cellulose standards, not only the 7 chemical forms could be distinguished, but also 8 of 10 producers of microcrystalline cellulose. Furthermore, a prediction of amphetamine concentration and excipient type was demonstrated which is very useful in routine analysis. However, ecstasy tablets are far from being standards. We suppose that the NIR spectra for tablets coming from one compression batch are rather variable. Due to the lack of information about this intravariability, it is difficult to judge how much spectral variance is acceptable to consider samples as linked or not. NIR analysis results in spectra representing the tablet blend as a
whole taking into account all absorbing compounds. From this point of view, NIR analysis seems to be an appropriate method for ecstasy profiling. However, the least is known about intra- and intervariability of compression batches for this method. We would like to remind that profiling was not the original purpose of this project. The possibility of profiling using NIR analysis only appeared during the comparison of the ecstasy spectra and the detection of similarities to results obtained by GC–MS analysis. Furthermore, this technique represents an interesting issue in the profiling problematic.

Acknowledgements

Many thanks to Prof. Robert Gurny for having this project made possible and providing workspace at the School of Pharmacy in Geneva. The authors are grateful to Tradall (Meyrin, CH) for having lent their FOSS NIR instrument during 6 months.

References

Pattern detection in forensic case data using graph theory: Application to heroin cutting agents

Anne-Laure Terrettaz-Zufferey a,1,*, Frédéric Ratle b,2, Olivier Ribaix a,1, Pierre Esseiva a,1, Mikhail Kanevski b,2

a Institut de Police Scientifique et de Criminologie, Ecole des Sciences Criminelles, Université de Lausanne, Batochime, CH-1015, Switzerland
b Institut de Géomatique et d’Analyse du Risque, Faculté des Géosciences et de l’Environnement, Université de Lausanne, Amphipôle, CH-1015, Switzerland

Received 2 June 2006; accepted 14 June 2006
Available online 1 August 2006

Abstract

Pattern recognition techniques can be very useful in forensic sciences to point out to relevant sets of events and potentially encourage an intelligence-led style of policing. In this study, these techniques have been applied to categorical data corresponding to cutting agents found in heroin seizures. An application of graph theoretic methods has been performed, in order to highlight the possible relationships between the location of seizures and co-occurrences of particular heroin cutting agents. An analysis of the co-occurrences to establish several main combinations has been done. Results illustrate the practical potential of mathematical models in forensic data analysis.

#2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Heroin seizures; Profiling; Adjacency matrix; Co-occurrences; Swiss cantons

1. Introduction

Pattern recognition techniques have been applied successfully in forensic sciences mainly to the problem of image processing [1,2]. Such methods are also used in the field of drug profiling, as described in ref. [3]. They provide information about links between seizures through the processing of the chemical profile of illicit drug seizures using statistical models of the data [4,5]. Other pattern recognition methods have been applied to illicit drug data using clustering and classification algorithms in refs. [6–8]. Alternative methods of clustering data with graphs is described in refs. [9,10]. An application of graph theory for modelling time intervals of real-case palaeontological data was made in ref. [12]. It is reasonable to suppose that such methods could be adapted and used in the field of drug profiling.

This study focuses on cutting agents in heroin seizures. They represent semi-quantitative data, collected in the same experimental analysis as the major constituents of heroin [4]. Their distribution in the four regions of investigation (canton of Vaud, Geneva, Neuchâtel and Tessin) is not homogeneous. Some products are present in all regions and other tend to be endemic. This fact encourages the study of the co-occurrences of cutting agents. This investigation aims to assess the performance of a combinatorial method to process cutting agents data. The application of mathematical models related to graph theory [11] can potentially help through visualisation the detection of regularities or patterns in the co-occurrences of cutting agents. On this basis, hypotheses can be developed about particular aspects of drug processing and distribution processes.

2. Methodology

2.1. The data

The studied area covers four Swiss cantons: Vaud, Geneva, Neuchâtel and Tessin (Fig. 1). In the framework of the profiling methodology [3], cutting agents are detected through gas chromatography analysis in the same way as major constituents. They are transformed into a binary format: the presence or absence of the agents in the samples are indicated, respectively, as 1s and 0s in a matrix. Eighteen cutting agents have been considered.
2.2. Basic definitions and notations

A graph $G = (V, E)$ is specified by its vertex set, $V$, and edge set, $E$, that represents connections between each vertex. We usually define $V$ as: $V = v_1, \ldots, v_n$. A matrix called the adjacency matrix, $A_G$, is associated to the graph, $G$, and its entries $a_{ij}$ are given by:

$$a_{ij} = \begin{cases} 1, & \text{if } \{i, j\} \in E \\ 0, & \text{otherwise} \end{cases}$$
2.3. Method

Fig. 2 illustrates, with a small subset of the data, the modelling process followed in this study:

(1) Each year of the binary data concerning 18 cutting agents extracted from the analysis seizures from one region is considered separately.

(2) \( V = \{v_1, \ldots, v_{18}\} \) represents the list of the 18 detected cutting agents and corresponds to the vertex of the graph; \( E = \{v_i, v_j\} \) represents the edges between the vertices, which appear if two cutting agents are found together at least one time during 1 year.

(3) Construction of the adjacency matrix \( A_G \) whose entries are \( a_{ij} \in \{0, 1\} \), and of the corresponding graph.

Cutting agents are in the same alphabetical order in the dataset as on the graphs. No dependencies between products are represented with the graphs.

2.4. Tools

All experiments were done using UA-Graph [13] and Matlab. The graph theory implementation was taken from free toolboxes proposed on the Mathworks website (http://www.mathworks.com).

Fig. 3. Graphs of the co-occurrences for cantons of Neuchâtel and Tessin \([n\text{ is the number of samples in the dataset}].\)
3. Results

The co-occurrence of cutting agents has been highlighted for the four studied cantons during 3 years. The adjacency matrix has been computed for years with enough relationships between products. The lack of co-occurrences between cutting agents for cantons of Neuchâtel and Tessin does not allow the construction of an adjacency matrix, and thus of a graph. Figs. 3 and 4 show the structures of the graphs by years and cantons and the combination are mentioned for cases without graph representation.

4. Discussion

Graph structures range from very simple to highly complex. It can be observed that complexity does not depend solely on the quantity of seizures. Various interpretations are possible.

It is known that the data from Tessin (see Fig. 3) represents exclusively important seizures made at the border. They are articulated around the combination “caffeine–paracetamol–griseofulvin” (c–p–g). Data collected at Geneva and Vaud contains a broader variety of seizure’s types (border, result of investigations and street seizures). They show more complex

| 1  | citric acid       | 7  | fructose       | 13 | paracetamol    |
| 2  | fatty acid        | 8  | glucose        | 14 | phenobarbital  |
| 3  | aspirin           | 9  | glycerol       | 15 | phosphate      |
| 4  | caffeine          | 10 | griseofulvin   | 16 | pimozide       |
| 5  | diazepam          | 11 | lactose        | 17 | pircaine       |
| 6  | diethylene        | 12 | manitol        | 18 | sucrose        |

Fig. 4. Graphs of the co-occurrences for cantons of Geneva and Vaud [n is the number of samples in the dataset].
structures (Fig. 4). In all the regions, mostly the same c–p–g mixture appears with other combinations of cutting agents. A plausible hypothesis is that the cutting process occurs also within the country, at a very local level, whereas a first cutting process of this nature is performed earlier in the distribution process. Consequently, those variables show a promising potential in an intelligence perspective for analysing the local distribution process of heroin.

The higher complexity of the graph structures representing data from the canton of Geneva can be explained by the strategic situation of the region, which entails an international airport, several customs and a highway bounding France and Switzerland. There is evidence that these factors influence the diversity of the illicit drug traffic that is reflected through the graphs.

The complexity of the graphs seems to evolve during the period under consideration, rather independently of the number of seizures. It could reveal a time-dependent change in the cutting process. This hypothesis is supported by the observed decrease in the global purity of the seized heroin in Switzerland (http://www.sgrm.ch/, last accessed 29/05/2006). A possible interpretation is that heroin was scarce during a particular period around 2004–2005. This could have led to a more intensive cutting, possibly earlier in the drug fabrication process, in order to fill market demand.

Differences can also be observed through graph comparisons. For instance, the presence of mannitol varies as a function of the year and the region under consideration. These types of variations can justify a spatio-temporal analysis of flows of specific cutting agents.

Graph visualisation has shown to be very promising for the detection of patterns leading to hypothesis that can be then tested. The intelligence potential of the approach is therefore validated. Perspectives for improving the method are numerous.

The number of co-occurrences of the cutting agents for the period under consideration is not represented. It could eventually be shown through the adaptation of the width of the edges. Moreover, co-occurrences of more than two products can be highlighted through edges colours. It can also be observed that some combinations of cutting agents are more frequent than others. This could lead to a reorganisation of the graphs in order to make them more readable, but they must stay comparable. Finally, the chosen period for each graph is somewhat arbitrary. Ideally, the influence of the length of the period would be worth investigating. The development of a dedicated software tool would improve the possibilities to explore and compare in a more comprehensive way those graphical structures on every aspect.

It is planned that the potential of this visualisation method will be assessed on other types of forensic case data.

5. Conclusion

Graph theory has been used to represent the co-occurrence and evolution of heroin cutting agents within specific regions. This visualisation method has helped to develop a great variety of hypotheses explaining the local cutting process.

From this preliminary study, ways for improvement have been suggested. Particularly, refining the visualisation process could be reached through the development of an exploratory software tool that would allow the detailed inspection of all aspects of the structures.

This promising method can be integrated into broader intelligence approaches and will be tested in order to detect patterns within other complex forensic datasets.

Acknowledgement

This work was supported by the Swiss National Science Foundation (grant no. 105211-107862).

References

Forensic drug Intelligence: An important tool in law enforcement

Pierre Esseiva, Sylvain Ioset, Frédéric Anglada, Laëtitia Gasté, Olivier Ribaux, Pierre Margot, Alain Gallusser, Alex Biedermann, Yves Specht, Edmond Ottinger

a The University of Lausanne, École des Sciences Criminelles, Institut de Police Scientifique, Batochime 1015 Lausanne-Dorigny, Switzerland
b Federal Office of Police, Forensic Science Unit, Nussbaumstrasse 29, 3003 Bern, Switzerland
c Federal Office of Police, Investigation Office, Av. Bergières 42, 1004 Lausanne, Switzerland
d Office of the Attorney General of Switzerland, Taubenstrasse 16, 3003 Bern, Switzerland

Received 8 June 2006; accepted 14 June 2006
Available online 26 July 2006

Abstract

Organised criminality is a great concern for national/international security. The demonstration of complex crimes is increasingly dependent on knowledge distributed within law-enforcement agencies and scientific disciplines. This separation of knowledge creates difficulties in reconstructing and prosecuting such crimes.

Basic interdisciplinary research in drug intelligence combined with crime analysis, forensic intelligence, and traditional law enforcement investigation is leading to important advances in crime investigation support. Laboratory results constitute one highly dependable source of information that is both reliable and testable. Their operational use can support investigation and even provide undetected connections or organisation of structure.

The foremost difficulties encountered by drug analysts are not principally of a chemical or analytical nature, but methodologies to extract parameters or features that are deemed to be crucial for handling and contextualising drug profiling data. An organised memory has been developed in order to provide accurate, timely, useful and meaningful information for linking spatially and temporally distinct events on a national and international level (including cross-border phenomena).

Literature has already pointed out that forensic case data are amenable for use in an intelligence perspective if data and knowledge of specialised actors are appropriately organised, shared and processed. As a particular form of forensic case data, the authors’ research focuses on parameters obtained through the systematic physical and chemical profiling of samples of illicit drugs. The procedure is used to infer and characterise links between samples that originate from the same and different seizures. The discussion will not, however, focus on how samples are actually analysed and compared as substantial literature on this topic already exists. Rather, attention is primarily drawn to an active and close collaboration between magistrates, forensic scientists, law enforcement investigators and crime analysts from different institutions with the aim of generating, using and validating relevant profiling case data as integral part of investigative and crime analysis processes.

Original advances are highlighted through experiences from criminal investigations of offences related to the unlawful importation, exportation, supply and possession of illicit drugs.

Keywords: Profiling; Illicit drugs; Forensic intelligence; Criminal investigation; Trans-institutional collaboration

1. Introduction

The criminal prosecution of organised crime constitutes a major difficulty and plays an increasingly important role in matters concerning a nation’s security. It is generally recognised that the investigation and prosecution of phenomena of serious cross-border crime require a close interaction between local, national and international actors in matters such as information, coordination and analysis. Criminal investigation proceedings in the context of drug trafficking, including its financing, are typical examples of this.

In Switzerland – a context to which the authors will refer to several times throughout this paper – the Federal Criminal Police (FCP), a main division of the Federal Office of Police (fedpol), carries out investigations under the direction of the
Office of the Attorney General in cases of serious crimes, notably organised crime, money laundering and corruption. The latter are criminal acts under federal jurisdiction according to Art. 340 and 340bis of the Swiss Criminal Code and additional legislation such as the Narcotics Act. Experience has shown that investigation of such serious forms of crime is a highly demanding task and may be fraught with various difficulties. Information obtained through traditional methods of investigation, such as telephone controls, surveillances and so on, may be incomplete, distorted and uncertain to some degree. Individuals who have been arrested may be uncooperative and provide inaccurate and/or contradictory information. Moreover, offenders are frequently affiliated to highly structured hierarchies and greatly benefit from recent advances made in the mobility of information (internet, mobile phones, etc.).

In order to obtain more accurate and consolidated views of criminal activities, a great need is often felt for more “reliable” sources of information, possibly based on scientific evidence. Literature has emphasised the potential of forensic case data for generating information that is capable of providing support and insight while aiding to overcome investigative obstacles. A term that is frequently used in this context is “intelligence” which, in a forensic perspective, can be taken to refer to the timely, accurate and usable product of logically processed forensic case data [7]. An innovative development of these ideas as a methodology for the analysis of serial crimes has been provided by Ribaux and Margot [9]. However, current uses of forensic data, such as fingerprint and DNA databases, are still rarely seen as a means to go beyond traditional identification [12].

Besides frequently used kinds of evidence, such as shoe marks, tool marks, DNA and fingerprints, recent literature has also reported methodologies for generating, organising and using data obtained through the systematic physical and chemical profiling of illicit drugs [2,3].

This paper describes the operational use of a practically implemented drug profiling methodology. Particular attention will be drawn to the collection, processing, sharing and exchange of profiling data. An emphasis is also made on the importance of having a close collaboration between all key actors involved in the investigative process, notably investigating magistrates, forensic scientists, police investigators and crime analysts. The potential of the described methodology for generating timely and relevant intelligence is illustrated by means of practical experiences gathered during a complex criminal investigation in organised crime.

2. Profiling of illicit drugs

2.1. Definitions

Some definitions need to be clarified to understand some concepts developed herein.

Drug profiling is the extraction of a drug sample’s chemical and/or physical profile, to be used in the application of policies against the illegal use of drugs (law enforcement, legislation, public health, etc.).

The profile of a drug sample is a subset of the sample’s characteristics specifically chosen with respect to the purpose of the process.

A class is a group of samples having similar profiles. It is a result of statistical methods applied to the output of the analytical process.

Profile and class can be chemical, physical or both, depending on the nature of characteristics considered.

2.2. Preliminaries and aim

Important contributions in this field have been made since the seventies introducing the notion of profiling and the interpretation of chemical links. There is a considerable amount of research that concentrates on the pertinence of chemical profiles, as well as on the notion of harmonisation of methods which currently is an extensively debated topic. Recent literature mainly focuses on the development and the improvement of new analytical techniques for eliciting chemical profiles or on the determination of mathematical methodology aiming at evaluating their similarity [4].

Currently, the most popular analytical method for characterising samples of drugs is gas chromatography, notably due to its high sensitivity. Different aspects can be investigated such as the determination of trace level impurities, occluded solvent or stable isotopes. Chemometrics and statistical issues have also been investigated in order to provide in-depth study of variables necessary for the selection of relevant impurities (stable and discriminatory) as well as statistical processing improving the classification of drugs samples. These approaches reflect the main tendencies of current developments in drug profiling. Great efforts have particularly been deployed in extracting profiles from samples and in building databases with the aim of classifying and comparing generated data. An exhaustive literature review of the aforementioned field of research can be found in ref. [5].

Aspects such as link management and interpretation as well as transmission of information to police forces do not seem, however to have been studied and explored equally well. The major aim of the present article will thus be to address some of the issues that pertain to the latter observation. An ongoing, close collaboration is described between the principal actors involved in the prosecution of illegal drug trafficking and related offences. These professionals are police investigators, magistrates, crime analysts and forensic scientists affiliated to national law enforcement or research agencies. Discussion will mainly focus on ways in which forensic case data may be used as integral part of investigative and criminal analysis processes. The specific case of systematic physical and chemical profiling of samples of illegal drugs will be considered. At times, the discussion will include a deeper insight of the notion of link through experiences made in real cases.

2.3. Drug profiling in Switzerland

Unlike countries such as Finland or the Netherlands, there is no centralised laboratory in Switzerland that analyses samples
of illegal drugs. There are currently about 10 laboratories distributed among different cantons and various linguistic regions. These laboratories operate with independent infrastructures and analyse samples on request by local prosecutors (investigating magistrates). One of the disadvantages of this situation is a deficiency in harmonisation. Currently, each laboratory uses its own drug profiling strategy (if any), implying considerable differences in sampling procedures, analytical methods, database structure, link management etc. Therefore, the comparison of samples from a seizure made in Geneva with samples from a seizure made in Bern, for instance, is impossible unless a laboratory is provided with samples from both seizures of interest. In other words, results obtained on samples analysed in a laboratory from one canton are not directly amenable for comparison with results obtained by a laboratory from another canton, leading to a kind of linkage blindness [5].

This is a fundamental difference compared with DNA evidence: here various standardised kits are available for analysis. These kits are such that when applied by different laboratories, results can be expected to be comparable. The nature of the data and the harmonisation greatly facilitated the construction and operational implementation of many of today’s national DNA databases.

Another difficulty encountered in practice - probably not only in Switzerland - relates to differences in what may be termed “philosophy”. For instance, part of the forensic science community considers the essence of drug profiling being specific comparisons of samples originating from selected cases (Fig. 1 (i)). Such a procedure is unnecessarily restricting as it confines itself to the inspection of only those samples that have been selected prior to the examination stage. Another way to look at drug profiling - which will be advocated throughout this paper - is to consider a systematic process in which each new sample is compared with existing entries of an organised memory built upon earlier cases (Fig. (ii)). This is where the two approaches are fundamentally different. They address two different levels of information in the judicial process as a whole. It may be used mainly, if not solely, in the following frameworks:

1. as evidence to be presented in court. Attention is focused on a restricted subset of samples and a single event, meaning that potential linkages beyond this level can usually not be addressed;
2. as a piece of intelligence in the course of the investigation [6].

The results in the first framework will typically be used to confirm a relation previously inferred on the basis of other information. The drug profiling methodology will thus be used to compare two samples in a specifically determined case scenario as supportive or corroborating evidence.

The second option does a more complete and complex use of the potential of the chemical and/or physical information. The aim is to provide some intelligence about a phenomenon, or about a specific instance of the phenomenon under investigation. This may take a whole new dimension, notably when there is other information, obtained through traditional investigative methods, suggesting that the seizures could be linked. The drug profiling methodology achieves this goal by providing the investigator with a list of connected cases that were not previously considered, together with other information, such as singular cutting agents.

This approach is further discussed below, in relation to operational activities.

2.4. An operational approach to systematic physical and chemical profiling

A systematic profiling of samples of illicit drugs including heroin and cocaine has to be considered as an overall methodology developed and practically implemented by the Institut de police scientifique (IPS) at the University of Lausanne [2,3]. The major distinctive feature that such a system has over those based on singular case to case comparisons is to allow searches in an organised memory (Fig. 1). This, in turn, provides for the possibility of finding previously unsuspected connections between samples seized in different cases and is essential justification for maintaining and implementing a drug profile database. It provides a solid basis for the production of a form of intelligence that focuses on finding analogies between new cases and those in the memory of the database.

Fig. 1. Schematic representation of different approaches to drug profiling: (i) punctual comparison of two specifically selected samples of drugs, (ii) systematic analysis, classification and comparison of samples of drugs.
A formal outline of the process is as follows. A memory is used as a core part of the database. Its role is to regroup samples from previously analysed seizures/cases. The memory is organised in such a way that samples whose profiles are satisfying well defined similarity criteria are grouped into classes. Characteristics chosen for classifying samples can be defined at different levels of detail. For example, a typical chemical class consists of samples with comparable proportions of impurities [1–3] (usually, a threshold is fixed using pre-existing knowledge). On the other hand, a physical class can contain pills with the same logo.

The notion of chemical and/or physical class is essentially statistical in nature as it is derived from evaluations made on databases containing samples known to originate from the same and different seizures respectively. Further details on how these criteria were derived can be found, for example, in ref. [3]. Whenever a new candidate is to be added to the memory, the sample is analysed in such a way that classification according to the specified criteria can be operated automatically. Two distinct stages underlie the process during which it is decided whether a new candidate should be assigned to a specific chemical/physical class. The first stage consists of a non-supervised strategy (currently performed by principal component analysis, a descriptive mathematical method [4]) that selects candidates with similar chemical profiles.

The second stage is a consideration of one of two possible scenarios. In one scenario, new candidates have profiles similar to that of a pre-existing chemical/physical class. In that case, a detailed comparison (e.g., through correlation measurements) is obtained between the candidate and the samples contained in the target class. If links can be confirmed (according to specified criteria), the connected cases are aggregated within the same class. Note that this operation will modify its own structure as well as that of the memory, which will need to be updated. This procedure can be compared to supervised classification methods used in mathematics. In the other scenario, new candidates may have a profile similar to that of a sample of the memory previously not associated with a specific class. Again, detailed comparisons are performed and if the criteria are met, a new class is created and the modified memory is updated. It must be noted that the memory may also be updated when additional circumstantial information, gathered through investigation, becomes available. The latter can aid, together with existing data on chemical/physical classes, to refine profile descriptors. For the remainder of this paper, we shall not outline in further detail how these comparisons are made in practice, because much specialised literature is available on this topic.

We shall rather emphasise and discuss the meaning of the highlighted links. In other words, what is the signification of these links? This question is certainly the most challenging one in the area of drug profiling and also a topic that has not yet been discussed in literature. The proposed approach uses real-case study in order to bring some provisory directions in the utilisation/interpretation of this information. This step is of major importance if we want to promote a more systematic utilisation of drug profiling by law enforcement. As long as this problematic has not been studied and formalised, it is highly probable that the information connecting to drug profiling will stay at the laboratory level.

3. Real-case study

3.1. Case description

In the course of an investigation carried out by the Swiss Federal Criminal Police, directed by the Office of the Attorney General of Switzerland, certain members affiliated to a highly structured familial clan became suspected of trafficking large quantities of heroin to central European countries, mainly heading towards Switzerland. Investigations aimed at gathering information and evidence in order to track, and, ultimately, prove the connection between criminal acts committed at temporally and geographically distinct occasions (including possible cross-border phenomena). However, many of the usual obstacles typically occurring in cases of organised crime were encountered as described earlier. Investigators and magistrates approached forensic scientists and crime analysts in order to seek collaboration and aid in finding ways in which complex criminal investigations could be supported by information derived from forensic case data. Apparently disconnected case samples were analysed and were grouped/classified according to the methodology presented in Section 2.3. This led to case connections of different types which needed further interpretation.

3.2. Results and discussion

3.2.1. Samples and sampling

Twenty-seven heroin seizures potentially connected from a law enforcement point of view were transferred for analysis in the IPS laboratory. In about half of these cases a sampling had been made by a member of the laboratory or a collaborator of the Swiss Federal Criminal Police acquainted with the specific requirements of drug sampling. Indeed the database demonstrates that a seizure can be composed of mixed heroin samples with different chemical profiles. For example, about 20% of the analysed seizures in the database show multiple profiles. This is particularly important, because multi class seizures can link chemical classes as shown in Fig. 2.

In contrast, there were several seizures from which – due to procedural obstacles – not proper samples exemplars were available. For instance, there was a seizure of about 200 kg from which only one sample was received by the laboratory. Clearly, the chemical composition of this seizure is fraught with uncertainties and not well represented. Actually, there is no chance of discovering more than one profile in this seizure (though there were probably many more than one). Important links, for example between chemical classes such as presented in Fig. 3, will be missed, resulting in a decrease of relevant investigative information. To illustrate this problem, and taking into account only the seizures sampled in a profiling perspective (representative sampling of the seizure), about 40% of the latter (three out of seven) had more than one chemical profile. This
was sufficient to create a wide net of relations (right part of Fig. 2) and not only several seizures linked to a unique chemical class (left part of Fig. 2). If the other half of the seizures had been properly sampled, much more useful information might have been extracted from the data.

This demonstrates that an adapted sampling process is necessary to have an accurate representation of the composition of seizures. That step is crucial in the profiling process.

3.2.2. Analysis and comparisons

Using the standardised analytical methodology [3], the 27 seizures were compared together (Fig. 4) and with the memory of the database (Fig. 5). This latter comparison highlighted 40 seizures that had chemical links to these. However, different degrees of connection have been identified and a first attempt of the structure of these hierarchical classifications is found hereafter.

In summary 67 (40 + 27) different seizures analysed within 6 years (1999–2005) have been linked through 34 chemical classes.

Analysing the obtained results, a classification of the different configurations of chemical linkages has been proposed. The simplest case was observed when two seizures had samples that belonged to the same chemical class in common. In this instance, the connection is done directly by the chemical class, there is no intermediate between the two

![Diagram](image_url)
seizures; this configuration has been defined as a chemical link of degree 1 (Fig. 3).

A chemical link of degree 2 between two seizures was obtained when two seizures were linked indirectly, through a third one. In other word and using the diagram illustrated in Fig. 3, seizure no.1 is linked with degree 2 with seizure no.3 because both of them are linked through a degree 1 linkage with seizure no. 2, even if they do not share a sample with a similar chemical profile.

Degrees 3, 4, …, n are deducted using the same logical path. It is important to point out that two seizures could be simultaneously linked in different degrees due to the fact that a seizure could contain multiple samples with different chemical profiles.

These data represent the analytical output provided by the laboratory. However, confronted with law enforcement information they could be contextualised and validated. Integrating the different types of information in the casework example, the following complex representations could be enhanced due to the transitivity of the links (Fig. 4).

3.2.3. Output

Fig. 4 represents the number of seizures potentially connected in the presented case. This was, initially, the police officer’s view resulting from the inquiry process. Various sources of information had been used by the investigator. The validity and the reliability of these sources vary considerably and had to be estimated (following existing guidelines). For instance, when two individuals are arrested simultaneously, each of whom is found in possession of illegal substances with comparable composition, then it appears reasonable to suspect that the two individuals may be organisationally related. In contrary, if the investigation link was based solely on name similarities (assuming that the persons were from the same family), the link is considered weak. The situation is similar to that encountered with other trace evidence. They have to be confronted with the police officer information in order to reconstruct the puzzle. When confronted and combined with the profiling data assumptions made by the police officer could be validated (or sometimes weakened). The use of drug profiling, in this investigation context, is a valuable step because it is based on material evidential information [10].

In this specific example, four major clusters (1–4 in Fig. 4) had been extracted by the profiling methodology. Each cluster validated the investigator hypothesis and allowed updating his perception of the traffic. Of course, seizures that were not part of a cluster could not be automatically rejected due to the complexity of the creation and interpretation of chemical links [11].

When confronted to the memory of the profiling database and extracting links in an intelligence perspective, the perception of the complex interactions could be illustrated as follows: (Fig. 5).

As a result of the comparison in the database, the four clusters observed in Fig. 4 have been grouped in two bigger entities (1 and 2 in Fig. 5), thus connecting physically clusters that might have been considered by the analyst as separate without further evidence. The confrontation with the database has helped demonstrate the need of the interaction between different types of data and the resulting benefits of having a more complete and precise view of the investigation and the traffic organisation itself.

The 40 seizures extracted from the database were chemically linked with the 27 seizures belonging to the case. The police officer in charge of the investigation had not suspected any of these connections with his case. It helped complete the overall
picture of the extent of the traffic and made visible some unsuspected links.

3.2.4. Added value for the investigation

The present case study clearly illustrates that the prime customers of the forensic intelligence information are mainly law enforcement agencies and investigators. Therefore, the process would achieve its goal by being close to these interlocutors to help them focus on complex phenomena such as drug trafficking organisations. The deciphering of this type of structures and organisations has such importance that the usual sample to sample comparison in the probative process is relegated to a relatively insignificant role.

The two main benefits for the investigation could be described as follows: (1) confirmation of a hypothesis proposed in the course of an investigation and (2) orientation towards cases that were unconnected due to the absence of any evidence of such links. In both cases, the maintenance of a database is a very important feature of the system. It is the only practical way to detect connections that remained unsuspected by investigators; this is a powerful forensic intelligence application.

**Confirmation.** The hypothesis proposed by the investigators can be assessed in several ways and the results of such assessment influence the next action to take. The result of the analysis of an illegal substance produces valuable elements in this context. The investigator can use it to strengthen his reporting about the relationship between cases or dealers. The chemical link on its own can be the starting point of further investigation about a relationship or one more element confirming it. The link can be crucial in decision making about the resources to allocate to the piece of case under study.

**Orientation.** This is the result of the extraction of specific cases already stored in the database. It highlights potentially connected cases in a retroactive way and has the potential of directing the investigation towards previously unsuspected relations.

Although this is a retrospective approach, this process can be used to detect trends and patterns, which could ultimately help to proactively draw hypothesis about the structure of drug trafficking.

4. General discussion

Chemical and physical classes are a kind of forensic case data that have the potential of describing phenomena and series [8]. However, this sort of database, due to the complexity associated with its entities, could not be used in the same way as DNA or fingerprint databases (e.g., AFIS). While the latter allow for an efficient retrieving of candidates as potential sources of, respectively, a questioned DNA profile and a questioned friction ridge mark, the relation between two samples of drugs with the same profile is of a quite different nature. With profiles of drugs, links could exist on different levels with respect to the chain of production, the trafficking route, the way of distribution and supply [11].

This situation is close to that observed and described in ref. [7] with toolmarks and shoemarks. A chemical/physical link between two drug samples (or a similarity between toolmarks, shoemark design) does not necessarily determine an association between the persons associated to each sample [11]. These series or classes have to be considered as means of analysing and reasoning on patterns that combine different sources of data rather than as a simple support in the identification of source, as highlighted by Ribaux et al. [7]. The profile described can also be influenced or tuned iteratively to take into account traditional police information. The strength of combining the information to build up intelligence is one major concept introduced in this paper.

Moreover, the chemical/physical classes have a high potential in an operational or strategic perspective, i.e. the database has not a sole use to support individual investigations as demonstrated through the case example. It has been shown that they can confirm hypotheses made by the investigating officer (relation between two offenders) or point out to other connections not suspected during the enquiries. In a strategic perspective, these classes can highlight the magnitude and the volume of a drug market, the extent of an organisation over a territory, the risk that certain illicit products produce over a consumer population, logistic and material requirements for the drug trafficking organisation. This may represent crucial information for law enforcement agencies to manage their resources to tackle a perceived specific problem.

Throughout this paper, the databases have been described as structures allowing an inference process such as described and introduced by Ribaux et al. [7] who approached the problem of classification and grouping in serial burglary investigations. A similar model has been transposed here to drug profiling. Indeed the classification through profiling is the strategy adopted to allow for grouping of similar profiles in a chemical class in order to support the investigations and gain a more complete view and understanding of a phenomenon (e.g., a traffic of illegal drugs).

Nevertheless, this approach could not be achieved by providing raw information to the investigator, the combining of the two sources of information and the resulting interpretation needs a thorough understanding of the inferential process and a good interlocutor that sees the value of sharing and combining the information. This approach is neither natural to traditional forensic scientist and to law enforcement investigators and the change of paradigm is not trivial especially since some in the legal profession insist on a complete independence of science from the investigative process. This, in our view, is a untenable misuse of science and negates its powers to determine physical, reliable and controllable data in support of investigative inferences that is the key of forensic intelligence.

5. Conclusion

The current situation in forensic drug analysis is mainly focused on the harmonisation of analytical methods as well as...
the creation of databases. These instruments provide a huge amount of untapped information. The challenge for forensic scientist and law enforcement is to extract information of a very rich nature from this source of data. This article presents concepts developed and tested in Switzerland to evaluate the complexity of the information brought by these analyses and databases, and of the inference processes that can be made in a forensic intelligence perspective. A case example of a successful complex investigation of an international drug ring based on these processes demonstrates the need to change the paradigm of forensic evidence as a pure probative exercise to a powerful investigative science.

References


Morphologic, experimental–comparative investigation as an identification of the injuring instrument method

R. Sitienë*, A. Zakaras, A. Pauliukevičius, G. Kisielius

Institute of Forensic Medicine, Mykolas Romeris University, Medicine Criminalistics Laboratory, S. Žakausko 12, LT-08234 Vilnius, Lithuania

Received 9 June 2006; accepted 14 June 2006
Available online 28 July 2006

Abstract

Aiming to identify the injuring tool characteristics and the tool itself morphologic, experimental–comparative investigations of the skin wound, rib and cartilage injuries taken during the autopsy are performed. During 1995–2004, 489 investigations were performed for this purpose. In 418 cases, knives were submitted for identification of the specific injuring tool (in total—835 knives). In 205 cases the investigation included not only skin wounds, but also the injured rib cartilages.

Identification investigations were performed by investigating both the skin wounds morphologic characteristics and dynamic traces—trails in the rib cartilage tissue left by the micro relief of the knife blade edge.

In the case of the investigated and experimental skin wounds characteristics coincidence the experimental and comparative dynamic traces investigation was performed when the traces were suitable for the tool identification purposes. In the case of the investigated and experimental skin wounds, dynamic traces coincidence, the totality of the coincided characteristics was considered individual. In those cases, the conclusion included the fact that the injury had been made by a particular knife. According to our data during 1995–2004 in 23 cases—15.9% (5.5% out of the total investigated cases), the knife identification was based on the skin wound characteristics and dynamic traces in the rib cartilage tissue. In our opinion, the dynamic traces in the rib cartilage tissue investigations supplement the identification field and are valuable in the tool identification.

In 11 (2.6% out of the total investigated cases) cases knives were identified only by the skin wounds morphologic characteristics, the ribs being not injured or dynamic traces being not suitable for the tool identification.

© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Stab-cut injury; Dynamic traces; Experimental–comparative investigation; Identification

1. Introduction

In 2001–2005, in Lithuania, violent deaths made 53.4% of the total 8812 averagely performed forensic medicine autopsies, including approximately 140 cases (1.57%) of stab-cut injury fatalities per year.

A proper investigation of such injuries is very important for legal institutions. The assessment of morphological characteristics of the skin wounds and injury canals enables to determine the characteristics of the injuring instrument. In the case of disposal of the probable injuring instruments sometimes the determination of a particular instrument is possible. The identification investigation is done, while assessing the morphology of stab-cut skin wounds, dynamic traces in the cartilage or bone tissues in the injury canal. In cases of the instrument’s penetration into the body through clothing the instrument characteristics related information may be supplemented by findings of the clothing damage investigation.

Forensic medicine literature references data emphasize the importance of the wound characteristics assessment in the identification of the injuring instrument [1–4].

Ochima (2004) especially values the information carried by the canals of injuries inflicted parallel to the body surface as reflecting the injuring instrument’s blade profile [5].

Ivanov argues that traditionally the determination of a particular injuring instrument is based on the usage of traces left by the knife blade’s micro relief in the rib cartilage. The works of Kariakin (1966), Voiler (1972) and Savičienė show that only traces left during the blade’s penetration by the knife blade edge bezel have identification value. Edelev (1990) argues that the formation of traces depends upon the blade’s lean degree during the penetration as well as on the roughness of
lateral surfaces. Following Zagriadskaja (1968), Kostiliov (1977), Ivanov’s (2000) findings lateral cartilages are injured only in 20% of the total fatal stab-cut injuries cases. Traces suitable for traces investigation are determined only in 9.8% of cases. The investigated dynamic traces allow the determination of a particular injuring instrument only in 2.2% of the total number of expertise carried out [6].

At autopsy the investigation of both morphologic characteristics of the skin wounds and dynamic traces in the cartilage and bone tissues in the injury canal is not sufficiently detailed due to the autopsy conditions (microscopy, microscopy, etc.). For this reason after the macroscopic investigation at autopsy the skin wounds as well as compound parts of the injury canal are excised and the investigation is continued under the laboratory conditions.

A prolonged investigation of such injuries is carried out in the Medicine Criminalistics laboratory. Its specialists have accumulated East European countries’ experience and have reached a high professional level in morphologic investigation of injuries and the injuring instrument identification spheres.

2. Materials and methods

Aiming to determine the injuring instrument’s characteristics or identify the instrument itself skin wounds, injuries of ribs or their cartilages are taken at autopsy and their further morphological, experimental–comparative investigation is carried out. For this purpose in 1995–2004, 489 investigations were performed, including 418 cases with knives submitted for the determination of a particular injuring instrument (the total 835 knives). Two hundred and five cases included the investigation of both the skin wounds and injured rib cartilages (including 145 cases when knives were submitted for investigation).

For a long time the attitude towards stab-cut wounds as a source of identification information was rather sceptical. During the period of the lack of reliable methods on preserving and reconstruction of the initial form of the skin wounds, their investigational value was especially low. Introduction of various solutions able to stop and eliminate post mortem changes and at the same time preserve and reconstruct the initial form of the skin wound into practice has proved that beside general (group) characteristics the details on the structure of stab-cut wounds can reflect an injuring instrument’s individual characteristics.

The laboratory work starts with the preparation of objects. The initial form of wounds is reconstructed by swelling a flap of dried skin in Ratnevski I solution (1972) [7]. After the swelling procedure it is possible to reliably assess the total morphological (micro irregularities of edges or the picture of micro details, shape of the ends) and micro metric (lengths of compound parts, course, width) data regarding the wounds. Fig. 1(a and b) show the effectiveness of the method.

The investigation is not affected by a dry condition of the rib cartilages because their proper swelling in water does not change dynamic traces and they still remain informative.

Identification investigations are performed, when assessing morphologic characteristics of the skin wounds and dynamic traces—tracks in the rib cartilage tissues have been left by the micro relief of knife’s blade edges. Exclusion of experimental and comparative investigations makes identification investigation of instruments impossible.

The solution of this issue was based on the following tactics: (1) to collect from the injury investigation data all the possible information regarding the instrument’s characteristics—the length and the biggest width of the penetrated part of the blade, the instrument’s blade, the tip zone blade, characteristics of the blunt edge of the blade (thickness, existence or absence of edges, their sharpness, asymmetry of the knife blunt back edges’ impact), characteristics of the tip. (2) To investigate the submitted knives, without experiment excluding the ones evidently inadequate to the determined instrument’s characteristics. (3) To establish marks reflecting the knife’s characteristics by modeling injuries in the experimental material with the help of knives. (4) To compare the information on the instrument’s characteristics obtained from the injury investigation data with the experimental injury investigation data. (5) To carry out an experimental–comparative investigation of clothing damages (in the case of their existence). (6) In case the injury contains dynamic traces, to carry out the traces investigation. (7) To carry out the metallization traces investigation. (8) To make the conclusion based on the total complex investigation data.

During the search for a suitable experimental material, the skin wounds and injuries were modelled on the chosen material (fresh pig skin, smoked pig skin, leather clothing, foam material, paper, etc.). In the end a colorless, transparent, 200 μm thick polythene membrane (made by applying a low pressure technology) has been chosen. Damages made on this material perfectly reflect properties of a single-edged knife’s blade. During the experimental–comparative investigation it has been established that in all cases when the skin contained marks of an asymmetric impact of the blunt edge of the knife (the end of the wound having the shape of asymmetric “M”), they were also expressed in the polythene membrane damages (asymmetric “Y”). In the cases of a rounded knife’s blunt edge the end of the wound had the “U” shape; the same shape was reflected in the polythene material damage. When the pointed zone blade was not sharp, microscopically margins of the initial part of the wound were rough or scratched, the microscopic roughness of the damage was also observed on the polythene membrane. When the knife’s blade was not sharp and the wound margins were rough, the same kind of roughness was observed on the membrane damages. However, differences have also been established. In the case of the skin wound end formed by the knife’s blunt back edge being of the “M” shape, in the membrane it is always of the “Y” shape; if the knife’s tip left intersection signs in the skin wound, on the membrane at the place of the tip’s impact the damage margins were deformed, dent or in the case of a blunt tip a round or ellipse shaped defect—of 0.1 mm diameter or 0.2 mm × 0.3 mm size with microscopically deformed margins appeared. Other morphologic differences have not been observed. Awareness of these
differences beforehand eliminated obstacles for carrying out an experimental–comparative investigation.

During the investigation of cartilage damages castings of cartilage injuries of canal walls were prepared by using the casting-material for forensic use “Mikrosil” (made in Sweden). Suitability of dynamic traces for investigation purposes was assessed.

The comparative investigation of the dynamic traces being investigated and experimental dynamic traces was carried out by applying the comparative optical system and Laboratory Universal Computer Image Analysis “Lucia Forensic”.

The importance of morphologic characteristics, experimental and comparative investigation in identification of the injuring instruments can be proved by the following case from practice.

Fig. 2(a–c) shows how successfully coinciding investigated and experimental dynamic traces supplement the identificational field.

Some pictures from practice (see Figs. 3–6) demonstrate a possibility of the instrument identification according to the coinciding characteristics of the investigated wounds and experimental injuries.

3. Results and discussion

Experimental and comparative investigation of dynamic traces was carried out in the case of coincidence of the investigated and experimental skin wounds characteristics when the traces were suitable for identification of the instrument. In the case of coincidence of the investigated...
and experimental skin wounds characteristics, investigated and experimental dynamic traces, the totality of coinciding features was considered individual. In such cases the conclusion about the injury having been inflicted by a particular knife was made. According to our findings in 1995–2004, in 23 cases (5.5% out of the total investigated cases) knives were identified on the basis of the skin wounds characteristics and dynamic traces in the cartilage tissue. We argue that the investigation of dynamic traces in the cartilage tissue supplements the identificational field and is valuable in identifying the instrument.

In 11 cases (2.6% out of the total investigated cases) knives were identified only by morphological characteristics of the skin wounds, the ribs being not injured or dynamic traces being unsuitable for identification.
In the skin wounds the totality of identificational characteristics consisted of the length of the main parts of wounds, features of formed blade tips, characteristics of the tip’s impact, traces left by the blunt back edges of the knife blade, either their symmetry or asymmetry; in the case of the butt’s asymmetry—if it is left or right, microscopic characteristics of the double-edged tips left by the butt’s edges, features of the double-edged tips, double-edge micrometry, distance between the double-edged tips, micrometry of the course of traces left by the tip zone in conjunction with the length of the main part of the wound; if the knife blade’s defect leaves a trace in the wound, its localization in conjunction with the length of the main part of the wound allows to approximately (sometimes, quite precisely) determine the smallest distance between the instrument’s blade defect and butt; micro morphologic characteristics of the wound margins reflecting the degree of instrument blade’s sharpness. These data are supplemented by the findings of metallization traces investigation. They might be supplemented by traces left on the knife’s anchor or handle. The totality of these characteristics may be individual. When assessing the characteristics it is important to differentiate between their relationship to the instrument’s features and injury making mechanism. Often it becomes obvious during the experimental and comparative investigation. For this reason before the experimental and comparative investigation the identificational value of the investigated injury characteristics cannot be considered final because during the latter investigation it can be specified more precisely.

Contrary to some authors’ arguments that the wound excision may cause morphological artifacts of wounds [2] our long-term practice has proved that the initial form of the fatality’s flap of skin dried in the room temperature may be reconstructed with the help of Ratnevski I solution even after a year’s period. During the skin wound micrometry the following fact is taken into consideration: after the excision of the wound from the fatality’s body the skin shrinks due to its elasticity, at the same time causing shrinking of the wound (11–16%, depending on the relation of the longitudinal axes to the course of Langer lines). It is not an

Fig. 5. The place of the knife blade’s impact. As the tip is broken (a) it leaves a clear trace in the wound (b) and membrane damage (c).
obstacle for determining the true micrometrical data, which existed before the excision of the wound from the body [8]. If at autopsy only the length of wounds is measured, the reconstruction of the initial form of the wound distinguishes such micro details of wounds whose micrometrical data are very useful for the experimental–comparative investigation.

So, the morphological, experimental and comparative investigation in 8.1% of cases enabled to determine a particular injuring instrument. Three hundred and eighty-one (45.62%) out of eight hundred and thirty-five knives submitted for investigation were excluded from further investigation by categorically denying their usage during the infliction of injuries. Thus, categorical conclusions regarding the instruments were made for 49.7% of the total number of instruments having been submitted for investigation. In the rest of cases it was possible to make only probability conclusions (40.3%) or the tool usage could be neither verified, nor denied (10%).

4. Conclusion

We argue that a prolonged, morphologic and experimental–comparative investigation of stab-cut injuries (skin wounds and rib cartilages) carried out under laboratory conditions is significant and valuable for identification of a particular crime instrument.

References