Forensic Science International 263 (2016) 67–73

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

Image processing of false identity documents for forensic intelligence

Benjamin Talbot-Wright [a], Simon Baechler [b,c], Marie Morelato [a], Olivier Ribaux [b], Claude Roux [a,*]

[a] Centre for Forensic Science, University of Technology Sydney, P.O. Box 123, Broadway, NSW 2007, Australia
[b] Ecole des Sciences Criminelles, University of Lausanne, Batochime, 1015 Lausanne, Switzerland
[c] Service forensique, Police neuchâteloise, Rue des Poudrières 14, 2006 Neuchâtel, Switzerland

ARTICLE INFO

Article history:
Received 25 January 2016
Received in revised form 29 March 2016
Accepted 30 March 2016
Available online 6 April 2016

Keywords:
Counterfeits
Comparisons
Profiles
Transversal model
Security
Intelligence

ABSTRACT

Forensic intelligence has recently gathered increasing attention as a potential expansion of forensic science that may contribute in a wider policing and security context. Whilst the new avenue is certainly promising, relatively few attempts to incorporate models, methods and techniques into practical projects are reported. This work reports a practical application of a generalised and transversal framework for developing forensic intelligence processes referred to here as the *Transversal model* adapted from previous work. Visual features present in the images of four datasets of false identity documents were systematically profiled and compared using image processing for the detection of a series of modus operandi (M.O.) actions. The nature of these series and their relation to the notion of common source was evaluated with respect to alternative known information and inferences drawn regarding respective crime systems. 439 documents seized by police and border guard authorities across 10 jurisdictions in Switzerland with known and unknown source level links formed the datasets for this study. Training sets were developed based on both known source level data, and visually supported relationships. Performance was evaluated through the use of intra-variability and inter-variability scores drawn from over 48,000 comparisons. The optimised method exhibited significant sensitivity combined with strong specificity and demonstrates its ability to support forensic intelligence efforts.

1. Introduction

Counterfeit and falsified identity documents (CFFIDs) are a widespread crime problem and necessary accessory to a vast proportion of transnational organised crime as well as transnational collaborative crime [1–4]. False identity documents are frequently involved in human trafficking, ranging in purpose from facilitating illicit sex trade to migrant smuggling. Such documents are also held by terrorists, fugitives, drug trafficking rings, organised cross border burglars, members of counterfeit goods distribution networks, and many other individuals who may wish to pass borders undetected, or defraud police or administrative checks. Unhindered movement of criminals and terrorists across borders and control points poses a current and often understated security threat to both national and regional jurisdictions [1].

Investigating the manufacturing techniques of CFFIDS and establishing links between them may be able to connect different cases, highlight crime phenomena and provide intelligence and solutions to address multiple crime problems more efficiently [1,3]. In this context, the *Transversal model*, a generalised and transversal framework adapted from previous work [5,6], is a description of a family of processes derived from forensic science, problem solving approaches and the intelligence cycle. It is an architecture that provides conceptual and stable boundaries for forensic intelligence to operate within, and helps to define it.

The digital inputs in all instances to the model are profiles. Profiles describe the object in terms of one or more selected features, discussed in the next section. Integration of new profiles into a formalised and shared memory allows us to link them with previous profiles, and in collaboration with alternative information (investigative and police data such as arrest and interview reports, criminological data, etc.), alter and develop our hypotheses regarding the structure or behaviour of a criminal group, or assert that newly connected profiles form a new series and may be isolated. As this memory is not static, but constantly updated with new links and series, the framework is also sensitive to the time dependent evolution of criminal activity.

This enquiry and discovery of links is proposed to contribute to a better understanding of crime problems and to the consequent formulation of strategy for more effective response to crime. It is

* Corresponding author. Tel.: +61 2 95141718.
E-mail address: claude.roux@uts.edu.au (C. Roux).

http://dx.doi.org/10.1016/j.forsciint.2016.03.054
0379-0738/

important to bear in mind that method development and practical application of processes must seek to address the common transversal objectives. The terms detection, object, problem, profile, and memory are defined in the *Transversal model* [5,6]. It is indeed most important to treat the model as a template of well-defined fundamental components that may be appropriated and tuned to handle different problems, as this is the nature of the intelligence based approach, and indeed a commonality of stable and established endeavours [7,8].

2. Hypotheses and premises

For the purposes of this paper, two new definitions will be proposed, applied and left open to scrutiny in the future:

Action: "The *action*" will specifically refer to the application of modus operandi (M.O.); the instance specific act, movement, behaviour and decisions of the offender. It includes activity both deliberate and unintentional, which may also produce non-M.O. specific traces. M.O. refers to the conditions and parameters, whereas the *action* refers to a particular and concrete application, and is subject to error and variation. In Aristotelian terms, M.O. is the potentiality whereas the *action* is the actuality [9]. This definition is useful in a forensic intelligence context as it allows the distinction between truly linked objects, and the traces that might be discovered and compared. It has to be highlighted that it differs from *activity* as understood in a forensic evidence evaluation perspective [10,11].

Repetition: A profile whose score with another falls within the optimised cut-off threshold[1] will constitute a repetition (i.e. a repeating profile). This is a pre-validation allocation to reflect only that significant similarity is detected, not to imply that this similarity is yet interpreted to indicate something in particular. Note that a series of *actions* is a ground truth; a repetition is a detection we hope reflects that. A cluster of repetitions in the context of classification, would be called a class. Repetitions and classes can be viewed as working hypotheses that nourish and foster the forensic intelligence process.

In the end, forensic intelligence seeks to assist decision making by providing intelligence regarding the actions or prevalence of figures and organisations relevant to a crime series or crime problem. It does this under the premise that repetitions in relevant objects likely indicate a repeating M.O. [1,3]. Three hypotheses exist inherently here which must be addressed and tested sequentially.

1. *Hypothesis of representativeness*: The process requires descriptive profiles. It is hypothesised that false identity documents contain representative features, which when encoded will produce descriptive profiles representative to the *action* of which they are an effect; This is in effect the representative hypothesis as coined by Margot [12].
2. *Hypothesis of detection*: It is hypothesised that profiling of representative features across a large number of objects will allow detection of repetitions and meaningful patterns, which could allow the detection of distinct *actions*.
3. *Hypothesis of inference*: Distinct series of repetitions are hypothesised to beget inferences regarding *actions*' relationship with each other, particularly repeating M.O. and thus in turn the nature of the crime system, or overlaps between crime systems based on alternative data.

[1] A dual approach combining a deterministic and a Bayesian approach was proposed in [5]. However, for the purpose of this study, only the deterministic approach was tested and the evaluation system covers in this article will only be based on the deterministic approach.

Neither the M.O. nor the *action* is expected to be formally described, nor is that the aim. The process at the base level is repetition detection, targeting visual effects of the *action* under the hypothesis that a certain degree of similarity, detected in repetition *implies* a repeating *action* and M.O. It is expected that distinct series of repetitions will resolve clusters of similar *actions*, which might in turn have a simple correlation to M.O. The 'repetition' does not distinguish this by definition. More potent inferences can be made in the company of complementary information and within the context of the problem at a later stage.

This paper thus iterates an approach to rapidly producing intelligence from relevant information from series of objects in the form of an experiment in the profiling (of images) of false identity documents. It illustrates a concrete application of the *Transversal model* that leverages image processing techniques in a forensic intelligence perspective.

3. Approach and methodology

This section considers what is deemed as the "*heart of the forensic intelligence framework*" [5] by presenting a method to profile and compare a collection of digital images from various CFFIDS which results are then evaluated in order to feed analysis and further steps of the *Transversal model*.

3.1. Data descriptions

The datasets were procured and acquired in their entirety in previous work [13] and supplied by various police jurisdictions throughout Switzerland, including those of Aargau, Bern, Fribourg, Geneva, Luzern, Neuchâtel, Ticino, Valais, Vaud, and Zürich. Using classical document examination techniques, specimens were pre-classified as authentic, counterfeit, forged or stolen blank by the specialists within the police jurisdictions they were obtained from. Four nationalities and three different types of false documents (counterfeits, forgeries and stolen blanks) were studied to train and test the method on diverse environments. The four sorts of documents were chosen since they represent current threats encountered in Switzerland by police and border guard authorities. The data is summarised below in Table 1

3.2. The script platform

A platform constructed at the Ecole Polytechnique Fédérale de Lausanne (EPFL), known as the Script platform, is an online platform operated using JavaScript as input for the acquisition and manipulation of various kinds of data as well as the visualisation of results [13]. All profile comparison and performance analysis was carried out using this platform. The feature extraction itself was carried out using a Region of Interest (ROI) designation tool written in Python language (The Python tool).

Table 1
Summary of data.

Summary of data					
	Total	Authentic (Au)	Counterfeit (Cf)	Forged (Fg)	Stolen blank (Sb)
Bulgarian passport	106	6	40	57	3
Bulgarian ID card	121	7	30	84	0
Portuguese passport	97	4	48	44	1
Romanian ID	115	10	105	0	0