Forensic Science International 232 (2013) 180–190

Contents lists available at ScienceDirect

Forensic Science International

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

The systematic profiling of false identity documents: Method validation and performance evaluation using seizures known to originate from common and different sources

Simon Baechler [a,c,*], Vincent Terrasse [b], Jean-Philippe Pujol [b], Thibaud Fritz [b], Olivier Ribaux [a], Pierre Margot [a]

[a] Institut de Police Scientifique, Ecole des Sciences Criminelles, University of Lausanne, 1015 Lausanne, Switzerland
[b] Département documents, Institut de Recherche Criminelle de la Gendarmerie Nationale, 1 Boulevard Théophile Sueur, 93110 Rosny-sous-Bois, France
[c] Service forensique, Police neuchâteloise, Rue des poudrières 14, 2006 Neuchâtel, Switzerland

ARTICLE INFO

Article history:
Received 19 February 2013
Received in revised form 16 July 2013
Accepted 29 July 2013
Available online 8 August 2013

Key words:
Forensic intelligence
Metric
Classification
Likelihood ratio
Counterfeit
Forgery

ABSTRACT

False identity documents constitute a potential powerful source of forensic intelligence because they are essential elements of transnational crime and provide cover for organized crime. In previous work, a systematic profiling method using false documents' visual features has been built within a forensic intelligence model. In the current study, the comparison process and metrics lying at the heart of this profiling method are described and evaluated. This evaluation takes advantage of 347 false identity documents of four different types seized in two countries whose sources were known to be common or different (following police investigations and dismantling of counterfeit factories). Intra-source and inter-sources variations were evaluated through the computation of more than 7500 similarity scores. The profiling method could thus be validated and its performance assessed using two complementary approaches to measuring type I and type II error rates: a binary classification and the computation of likelihood ratios. Very low error rates were measured across the four document types, demonstrating the validity and robustness of the method to link documents to a common source or to differentiate them. These results pave the way for an operational implementation of a systematic profiling process integrated in a developed forensic intelligence model.

1. Introduction

The use of false identity documents constitutes a pervasive crime that is often connected to organized crime, including that from terrorist organizations [1–6]. As such, the view that forensic intelligence could play a significant role in understanding and reducing such crimes and also threats to national security has been recognized and expressed by international experts and government agencies [7–10]. The operation of turning forensic observations detected on fraudulent identity and travel documents into actionable intelligence remains a challenge. In a previous publication [1], we proposed an intelligence model using forensic science that relies on the systematic profiling and management of false identity documents. This model encourages a paradigm shift from a case-by-case reactive approach toward a methodical, comprehensive and proactive forensic intelligence approach that uncovers, via an elementary function, links among seizures as well as patterns and trends in the data. In this perspective, the forensic profile of an entity is defined as a set of identified and measured material (visual, physical and/or chemical) features that is representative, specific and reliable enough to be of relevant use to analyzing crime or national security problems [11]. The profiling method is designed to support broad problem-solving and policing issues based on a formal memory and sustained analysis. The model relies on a theoretical and intuitive syllogism leading to the assertion that "observing similarities or differences between false identity documents' material features (i.e., their forensic profiles) helps to support the inference that these documents have been produced by the same or a different *modus operandi*, and ultimately that they have been produced by a same or a different source" [1]. Despite empirical data and examples provided illustrating what kind of intelligence can be generated through such a forensic intelligence process, this pivotal syllogism remains to be tested against practical and controlled data. The purpose of the present study is to validate the forensic profiling method and evaluate its performance using false identity documents known to

* Corresponding author. Tel.: +41 21 692 46 11.
E-mail address: Simon.Baechler@unil.ch (S. Baechler).

0379-0738/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.forsciint.2013.07.022

originate from common sources or, alternatively, from different sources, thus allowing an assessment of intra-source and inter-sources variations.

2. Materials and methods

2.1. Data collection, visual examination and profiling

Collaboration between the Institut de Recherche Criminelle de la Gendarmerie Nationale in France (IRCGN), the Ecole des Sciences Criminelles of the University of Lausanne in Switzerland (ESC) and ten Swiss cantonal police departments allowed collecting a cross-border dataset of 347 false identity documents seized in France and Switzerland between 2000 and 2012. The dataset is distributed as follows:

- 170 counterfeit Portuguese identity cards, among them 16 documents originating from 3 known sources (one group of 8, one group of 5, and one group of 3 documents).
- 129 counterfeit French identity cards, among them 39 documents originating from 4 known sources (one group of 16, one group of 11, one group of 10 and one group of 2 documents).
- 25 counterfeit French passports, among them 4 documents originating from 1 known source.
- 23 forged British passports, among them 9 documents originating from 1 known source.

Overall, out of the 347 documents, 68 were known to originate from 9 different sources. These sources and their production were identified through police investigations and dismantling of false documents factories. All these sources are totally unrelated to one another except two of them that were connected to some extent as parts of a same criminal network, namely the groups of 10 and 2 counterfeit French ID cards mentioned above.

The 347 documents were examined with basic standard equipment [8], namely the eye, a stereomicroscope and a UV light source, in order to profile (observe and codify) a range of visual features for each document. Most of these features are also used to authenticate identity and travel documents, therefore this is sensitive information and a comprehensive list of these features cannot be given here. The profile codifies for instance what are the printing processes used to manufacture the false document, if and how security elements are imitated (UV features, watermarks, security threads, optically variable devices, embossing stamp, perforations, etc.), the serial number, what are the fonts used to write different text zones and what errors are in the machine readable zone. The number of visual features studied ranges from 25 to 49 depending on the document type. Overall, the profiling task requires three to ten minutes per document.

The analysis of false identity documents was limited to observing and codifying merely visual features because, as discussed in [1], using such routine and easy-to-acquire features offers several advantages over chemical analysis or advanced physical features. The observation of visual features is much less resource intensive (in terms of time, equipment, training and costs) and does not require working in a laboratory environment. As these features are mastered and examined routinely by diverse organizations, the codification of visual features is little prone to variability across different operators. Above all, the profiling of visual features proved already to be sufficient and efficient alone in a forensic intelligence perspective [1,12]. Another study is currently underway to specify and discuss the potential of chemical and advanced physical features for complementing visual features in specific forensic intelligence forms of inferences.

The profile of documents is composed of a combination of visual features in the method tested. This avoids the sole reliance on one specific feature or particular measurement, such as the serial number, a typical error (e.g. a misspelling) or a specific printing or reproduction defect. Even if the profile combines routine and sometimes frequent features, a multi-features (or multidimensional) approach has at least six advantages over a single-feature scheme [13]. First, a combination of routine visual features is easy to codify and any fraudulent document can be profiled within minutes. Second, it is able to cope with intra-source variations, evolving *modus operandi* or with the absence of some features (due to limited observations capabilities, low quality images or damaged documents for instance). Third, while the detection and recognition of markers of production series is a necessary prerequisite to the implementation of a single-feature scheme, this task is not always self-evident and may require a lot of cases. Profiling multiple routine features is not bound to such a prerequisite and can be implemented immediately. Forth, a higher discriminating power is achieved with a set of not-so-frequent features than with a sole very rare characteristic according to the formula [14]: $DP = 1 - \sum_i f_1^2$ (with DP the discriminating power and f the frequency of each of the i features). A very rare characteristic or specific marker of a series may be useful to identify documents belonging to this particular series, but it will be completely useless to distinguish and classify all the others documents and series where this marker is absent. Fifth, a set of features allows the investigation of the diverse levels of manufacture of false documents (production of the support, printing, forgery of security features, and personalization) whereas a single feature limits the perspective to only one level. Finally, a multidimensional approach allows processing and managing much larger datasets on longer periods of time and in an adaptive fashion.

The sets of visual features were chosen for each of the four document types according to six factors: features should always be present so that they can be codified, their possible values should be likely to vary across different sources while remaining reproducible within the production of a same source, features should be fit for an objective and reliable codification, they should be quick and easy to observe, and they should also be used to authenticate documents in order to be known and mastered by those who check and control documents. Previous cases and international intelligence alerts helped us evaluate and optimize these factors while selecting visual features.

Following observation of each document, the codification of visual features composing each profile was recorded in a dedicated computerized database called *ProfID* that we developed. This database allows the acquisition, management and comparison of documents' profiles in a short time (see infra).

2.2. Metrics and comparison process

A complete description of the comparison process framework and its integration within the forensic intelligence model has been previously published [1]. In practice, the profile of each document is compared to every other profile stored in the database according to each document type. Indeed, as a first approach, we limit ourselves to comparing profiles of documents of the same model and country in order to maximize comparability. The degree of similarity between two profiles is described by a score measuring the extent of correspondence afforded by respective features of two profiles. The computation of the score depends on the selection of a comparison metric which takes into account parameters and enables their optimization, such as the relative weight of each feature within the profile and the way these features should be combined in the similarity calculation. The comparison metric must also fit with the type and scale of data that will be compared. In the case of visual features descriptions, data is mostly qualitative, discrete, uncentered and unordered (nominal scale) – a typical example is the printing process of the false document background that may either be codified as 'inkjet', 'toner', 'offset' or 'other'. This type of data is in no way a problem or a limitation, but comparison metrics have to be selected and defined accordingly.

In order to compute a similarity score between each pair of profiles, five different comparison metrics inspired mainly by references on drug profiling [15–18] have been implemented. These metrics are mathematically described below because of the difference in data type compared to drug analysis. All metrics except Hamming may be optimized by assigning value coefficients to each feature. Coefficients were set on an empirical basis according to how each feature is understood to be source-reproducible and source-specific (based on experience and general knowledge about the manufacture of false documents). Independence or dependence between features is also taken into account. Coefficients are ranging from 1 to 3 giving the feature a different weight in the computation of similarity scores. For instance, the printing process of the document background would have a coefficient of 3 since it is an essential trait of the manufacturing process of the source, whereas the description of the cutting of the edges of the holder photograph would have a coefficient of 1 because this feature is much less source-reproducible. The length of the machine readable zone would be given a coefficient of 2 since forgers who are acquainted to the norms governing these strings of characters would repeatedly produce correct machine readable zones whereas forgers not acquainted to these norms could produce either too long, correct or too short strings of characters. In the Hamming metric, all coefficients are set to 1 by default.

Hamming: $\text{score} = \sum_{i=1}^{n} C_i$ with C taking the value of 0 or 1 depending on the correspondence of the feature; if features match within both profiles $C = 1$, else $C = 0$; n = number of features.

Manhattan: $\text{score} = \sum_{i=1}^{n} C_i$ with C taking the value of 0 or the feature coefficient depending on the correspondence of the feature; if features match within both profiles C = feature coefficient, else $C = 0$; n = number of features.

Euclid: $\text{score} = \sqrt{\sum_{i=1}^{n} C_i^2}$ with C taking the value of 0 or the feature coefficient depending on the correspondence of the feature; if features match within both profiles C = feature coefficient, else $C = 0$; n = number of features.

LnProduct: $\text{score} = \text{Ln}\left(\prod_{i=1}^{n} C_i\right)$ with C taking the value of 0 or the feature coefficient depending on the correspondence of the feature; if features match within both profiles C = feature coefficient, else $C = 0$; n = number of features.

Squared cosine correlation: $\text{score} = 100 \times \frac{(C_1 D_1 + \cdots + C_n D_n)^2}{(C_1^2 + \cdots + C_n^2) \times (D_1^2 + \cdots + D_n^2)}$ with C taking the value of 0 or the feature coefficient depending on the correspondence of the feature, and D always taking the value of the feature coefficient in order to generate a distance proportional to this coefficient; if features match within both profiles C = feature coefficient, else $C = 0$; D = feature coefficient; n = number of features.