Forensic Science International 260 (2016) 74–84

Contents lists available at ScienceDirect

Forensic Science International

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

Development of a systematic computer vision-based method to analyse and compare images of false identity documents for forensic intelligence purposes–Part I: Acquisition, calibration and validation issues

Marie Auberson [a], Simon Baechler [a,b,*], Michaël Zasso [a,c], Thibault Genessay [a], Luc Patiny [c], Pierre Esseiva [a]

[a] Ecole des Sciences Criminelles, Université de Lausanne, 1015 Lausanne, Switzerland
[b] Service forensique, Police neuchâteloise, Rue des poudrières 14, 2006 Neuchâtel, Switzerland
[c] Institute of Chemical Sciences and Engineering, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 18 July 2015
Received in revised form 20 November 2015
Accepted 13 January 2016
Available online 25 January 2016

Keywords:
Image analysis and comparison
Histograms
Profiling
Profiles
Filters
Scanner
Counterfeit
Forgery

ABSTRACT

Following their detection and seizure by police and border guard authorities, false identity and travel documents are usually scanned, producing digital images. This research investigates the potential of these images to classify false identity documents, highlight links between documents produced by a same *modus operandi* or same source, and thus support forensic intelligence efforts. Inspired by previous research work about digital images of Ecstasy tablets, a systematic and complete method has been developed to acquire, collect, process and compare images of false identity documents.

This first part of the article highlights the critical steps of the method and the development of a prototype that processes regions of interest extracted from images. Acquisition conditions have been fine-tuned in order to optimise reproducibility and comparability of images. Different filters and comparison metrics have been evaluated and the performance of the method has been assessed using two calibration and validation sets of documents, made up of 101 Italian driving licenses and 96 Portuguese passports seized in Switzerland, among which some were known to come from common sources. Results indicate that the use of *Hue* and *Edge* filters or their combination to extract profiles from images, and then the comparison of profiles with a Canberra distance-based metric provides the most accurate classification of documents.

The method appears also to be quick, efficient and inexpensive. It can be easily operated from remote locations and shared amongst different organisations, which makes it very convenient for future operational applications. The method could serve as a first fast triage method that may help target more resource-intensive profiling methods (based on a visual, physical or chemical examination of documents for instance). Its contribution to forensic intelligence and its application to several sets of false identity documents seized by police and border guards will be developed in a forthcoming article (part II).

1. Introduction

Identity and travel document fraud is connected to organised crime and terrorist groups who produce, distribute and use false documents to support their various criminal activities. Means are however lacking to trace back production and distribution networks, and the task of distinguishing organised crime implications from anecdotic frauds in the daily mass of seizures is a solid challenge [1]. The systematic processing of false identity documents material features within a forensic intelligence framework may provide a part of the solution, and a model has been proposed to formalise and foster such an approach [2–5]. The rationale of this model assumes that based on the observation of a similarity between the forensic profiles of two false documents, one can infer that these documents were produced according to a common *modus operandi* (manufacturing method), and ultimately that they may originate from a common source (forger, organisation or factory). The model was originally articulated on a

[*] Corresponding author.
E-mail address: simon.baechler@unil.ch (S. Baechler).

http://dx.doi.org/10.1016/j.forsciint.2016.01.016
0379-0738/© 2016 Elsevier Ireland Ltd. All rights reserved.

comparison method based on the profiling of a combination of visual and physical features of fraudulent documents (printing techniques, security elements imitations, errors in the machine readable zone, etc.), which achieved convincing results in detecting and monitoring linkages between seizures, crime patterns and trends ([6,2,3,23]) . The validity and interest of the forensic intelligence model are however not bound to that specific comparison method. Alternative profiling methods can be imagined, as substitutes or as additional bricks, increasing the simplicity, flexibility, rapidity, robustness, cost-effectiveness and/or the added value in terms of intelligence.

In this perspective, inspired by a successful method developed to manage digital images of Ecstasy tablets for intelligence purposes [7] and following a preliminary feasibility study with promising results [8], a computer vision-based method has been developed to process systematically images of false identity and travel documents. Compared with the visual and physical profiling of false documents, a digital image analysis and comparison method would have the key advantage of facilitating the exchange of forensic case data (i.e. dematerialised transfer of images) between the different forensic and policing stakeholders–organisations, jurisdictions, countries–thus fostering combined and timely forensic intelligence efforts, reducing linkage blindness and bringing out crime patterns. This article (Part I) presents the critical steps of the method, the development of a prototype and how the method was calibrated and validated to meet scientific and operational criteria. A forthcoming article (Part II) will detail the method contribution to forensic intelligence and its application to several sets of false identity documents seized by police and border guard authorities in Switzerland.

2. Method

The proposed computer vision-based method is grounded in the above-mentioned forensic intelligence model rationale and relies on the hypothesis that digital images processing methods–such as colour, hue or texture extraction and comparison–enable the measurement of similarities and differences between false identity documents, thus supporting the inference of a common or different source/manufacturing method at the origin of the documents [2].

The process underlying the method is first described generally (Fig. 1), its steps being described further in the next sections. The method requires to digitalise false identity documents using a sensor (such as a scanner or digital camera), which is already done routinely by police and border guards who detect and seize fraudulent documents. To get rid of sensitive personal data and to target the document features that do not change according to the bearer (name, date of issue or photograph for instance), selected regions of interest of the image are automatically extracted and cropped. The cropped images are then imported in an image collection management system named *Script* that is developed by the Ecole Polytechnique Fédérale de Lausanne [7] based on open source packages (*Script*, https://github.com/cheminfo/script; *ImageJ*, https://github.com/cheminfo/imagej-plugin; *Visualizer*, https://github.com/npellet/visualizer). This system, hosted on a secure web-access server, enables operators in different locations or from different organisations to easily upload images in a common framework. From an image processing point of view, the *Script* system is able to filter images using 10 different filters (Fig. 2): Red/Green/Blue from the RGB colour space; Grey levels; Hue/Saturation/Brightness from the HSB space; Contrast; Texture; Edge [8]. It then extracts histograms representing the pixels distribution [9,10]. Profiles derived from histograms of each image can be systematically compared with those of other images using different metrics, thus returning a similarity score (or distance) for each pair of images. Scores are used to derive a classification and are finally interpreted in an intelligence perspective to assess if documents could be linked to a same origin or not (Morelato *et al.*, 2015, [5]).

2.1. Image Acquisition Conditions

For an image analysis method to be reliable and valid in comparing specimens, the sensors through which the specimens–here false identity documents–are digitalised have to produce reproducible images. Reproducibility is a fundamental requirement to ensure the comparability of images, in particular when they are produced at different times, different locations and/or by different sensors [11–13,5]. Therefore, besides the scientific issue regarding the variability caused by sensors, a real challenge is posed on an operational level considering that several stakeholders–various police and border guards authorities or forensic labs throughout a country or even internationally–wishing to contribute to the acquisition of images may possess very different sensors and use computers operating different system specifications. To account for these scientific and operational issues, four mutually exclusive hypotheses concerning sensors reproducibility and its impact on the comparability of images have to be assessed:

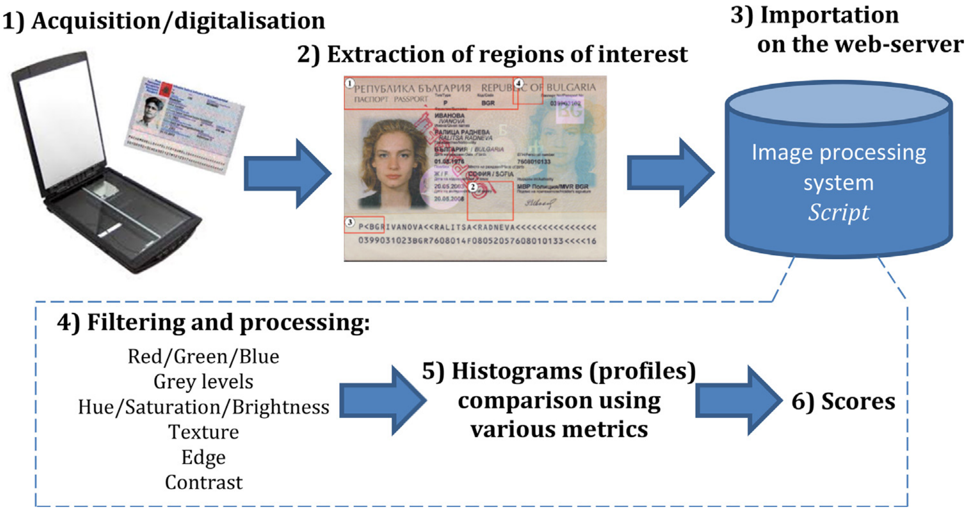
1) Acquisition/digitalisation
2) Extraction of regions of interest
3) Importation on the web-server
Image processing system *Script*
4) Filtering and processing:
Red/Green/Blue
Grey levels
Hue/Saturation/Brightness
Texture
Edge
Contrast
5) Histograms (profiles) comparison using various metrics
6) Scores
1
2
3
4
РЕПУБЛИКА БЪЛГАРИЯ REPUBLIC OF BULGARIA
P<BGRIVANOVA<<RALITSA<RADNEVA<<<<<<<<<<<<<<<
0399031023BGR7608014F08052057608010133<<<<16

Fig. 1. Process of the image analysis and comparison method.