Forensic Science International 235 (2014) 86–101

Contents lists available at ScienceDirect

Forensic Science International

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

Evidence evaluation in fingerprint comparison and automated fingerprint identification systems—Modeling between finger variability

N.M. Egli Anthonioz *, C. Champod

Institut de Police Scientifique, School of Criminal Sciences, Building Batochime, University of Lausanne, CH-1015 Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 22 June 2012
Received in revised form 28 August 2013
Accepted 3 December 2013
Available online 18 December 2013

Keywords:
Fingerprint evaluation
Likelihood ratio
Between-finger variability

ABSTRACT

In the context of the investigation of the use of automated fingerprint identification systems (AFIS) for the evaluation of fingerprint evidence, the current study presents investigations into the variability of scores from an AFIS system when fingermarks from a known donor are compared to fingerprints that are not from the same source. The ultimate goal is to propose a model, based on likelihood ratios, which allows the evaluation of mark-to-print comparisons. In particular, this model, through its use of AFIS technology, benefits from the possibility of using a large amount of data, as well as from an already built-in proximity measure, the AFIS score. More precisely, the numerator of the LR is obtained from scores issued from comparisons between impressions from the same source and showing the same minutia configuration. The denominator of the LR is obtained by extracting scores from comparisons of the questioned mark with a database of non-matching sources. This paper focuses solely on the assignment of the denominator of the LR. We refer to it by the generic term of between-finger variability. The issues addressed in this paper in relation to between-finger variability are the required sample size, the influence of the finger number and general pattern, as well as that of the number of minutiae included and their configuration on a given finger. Results show that reliable estimation of between-finger variability is feasible with 10,000 scores. These scores should come from the appropriate finger number/general pattern combination as defined by the mark. Furthermore, strategies of obtaining between-finger variability when these elements cannot be conclusively seen on the mark (and its position with respect to other marks for finger number) have been presented. These results immediately allow case-by-case estimation of the between-finger variability in an operational setting.

1. Introduction

The use of statistical models in the context of fingerprint comparisons has been proffered to be the way forward [1], even though such models are not used presently in operational settings. The possibility for a change in these practices has been opened recently in the United States of America, since the use of valid mathematical models for the evaluation of mark-to-print comparisons is allowed by the International Association for Identification [2]. Among the strong reasons for using models is the fact that these models build on solid observations of data, and bring transparency to the decision process. The data used for the building of the model, the assumptions, the modeling steps, and the features used can be (and are) fully described, and the models' merits can be assessed. Furthermore, the models can be tested, and associated performance measures (including error rates in a decisional setting or rates of misleading evidence in a probabilistic one) can be computed, which is an advantage in the current North American judicial setting following Daubert v. Merrell Dow Pharmaceuticals, Inc. [3] and its progeny.

The 2009 National Academy of Sciences report [4] highlights the importance of the use of data-based models to underpin the conclusions in the forensic comparison sciences. A dearth of such data-based information is noted in that report [4]. Several studies concerning probability models for fingerprint evaluation have been published in recent years. These studies add to the knowledge in this domain and directly address this lack of information. Neumann et al. [5,6] have proposed a model integrating both within- and between-finger variability in a likelihood ratio, for the purpose of the evaluation of a mark-to-print comparison; the same is true of Neumann et al. [7]. Egli et al. [8] present an approach using AFIS scores, for the assessment of within-finger variability. A full model is the subject of a thesis [9]. Also, the extension of a model to compute likelihood ratios not with respect to a given finger, but where the prosecution and defense hypotheses concern

* Corresponding author. Tel.: +41 216924630; fax: +41 216924605.
E-mail addresses: Nicole.EgliAnthonioz@unil.ch (N.M. Egli Anthonioz), Christophe.Champod@unil.ch (C. Champod).

0379-0738/$ – see front matter
http://dx.doi.org/10.1016/j.forsciint.2013.12.003

the suspect (respectively, his having left the mark or not) is presented in Neumann et al. [10]. How consideration of the general pattern can be integrated into such a model is presented in Neumann et al. [11].

The models described in these publications use a likelihood ratio based approach. While none of these models can completely solve issues related to the examiners' choice of characteristics to be considered, they can help underpin the examiners' conclusions, add transparency and help the assessment of the associative strength of a group of features found in agreement. Furthermore, and perhaps even more importantly, they help the assessment of whether or not the two groups of features compared (in most models minutiae) indeed correspond to what is expected if they were both left by the same finger. This last point is not necessarily addressed in all approaches that have been proposed in recent years in particular those centered mostly around fingerprint individuality [12–23]. It is the consideration of the within-finger variability that assesses this expectation. Only the models based on likelihood ratios address this question explicitly, while other models handle this variability by allowing the characteristics to vary within certain thresholds and still be considered as "matching". The same type of approach is used in other forensic areas, where so-called score-based likelihood ratios are also used. Different approaches for the computation of score-based likelihood ratios are discussed in Hepler et al. [24], and illustrated on the particular example of documents. In speaker recognition [25] the approach that uses, according to [24], a trace-anchored denominator is proposed. The same is true for Alexander et al. [26], Alexander [27], Gonzalez-Rodriguez et al. [28] and Ramos-Castro et al. [29], for speaker recognition as well as other biometric measurements. The between-finger variability as computed here uses the same, trace-anchored, approach. A previous study addressed within-finger variability [8], the present article presents one approach for the assignment of the between-finger variability and an investigation of variables that have an impact on this distribution. In particular, the required database size in order to estimate a stable distribution is assessed, and a distribution fitted. Furthermore, the dependence of the distribution on the finger number as well as the general pattern is investigated. These are the elements needed for the estimation of the between-finger distribution and therefore the denominator of the LR. The impact of the number of minutiae in a configuration and of their placement on the distribution is also presented. Knowing the impact of the number of minutiae and their configuration is not of importance in a case-by-case approach, given that in such an approach the minutiae are defined by the case at hand. Such knowledge, however, opens the possibility of estimating between variability without having to extract large numbers of scores. Indeed, if the between finger variability mainly depended on the number of minutiae, approximations of this distribution could be feasible. If so, the number of data needed for its estimation could be greatly reduced. This may be particularly important for some finger number/general pattern combinations, where the obtention of the necessary number of scores may require a huge size of the background database. For example, on right thumbs, left loops are rare (0.2% of right thumbs show left loops, which comes to 1/5000 fingers showing this combination). In order to therefore obtain a given number of prints showing this finger number/general pattern combination, and given this expectation of observing it in 1/5000 fingers, a database including overall approximately 5000 times more fingerprints than the minimum number of scores required would be needed. This problem could be solved through approaches allowing to estimate between-finger variability using far fewer between-finger scores, taking advantage of knowledge about the overall behavior of this distribution.

While general articles, discussing possible approaches to the computation of score-based likelihood ratios in general [24] as well as between-source variability in different contexts [29] have been published, no publication has investigated the different factors that impact, or not, the between-finger variability, such as the general pattern, the number of minutiae, the number of observations, and the finger number. While the results obtained for these elements are dependent on a particular system, a methodology can be derived from this approach, analyzing different factors that may impact on the distribution.

2. Materials and methods

2.1. Between-finger variability

The likelihood ratio is, here, considered as follows. The score issued from an AFIS system for a comparison between a mark and a print, s, is taken to be the evidential information. The numerator, $f(s|H)$ is the probability density of the observed score s if the latent mark and a given rolled inked print originate from the same source (H). This source, the suspect's finger, is at the origin of the within-finger variability. This variability models therefore the scores expected when different impressions from the suspect's finger are compared. The evidential score is considered to represent the comparison between a mark and an (inked) print. Therefore, the comparisons carried out in order to obtain the scores for the within-finger distribution are also those between marks and (inked) prints, unless substitutes for marks can be found. Such substitutes could be for example livescan images or inked prints, that is, more easily acquired representations of the finger surface than developed marks, if the scores obtained between such substitutes represent well the scores between marks and rolled inked prints. Details concerning within-finger distribution are described in Egli [9] and Egli et al. [8]. The denominator, $f(s|\overline{H})$ is the relative likelihood of observing the evidential score if the mark and the rolled inked print do not come from the same source. It is related to the between-finger variability. This distribution is modeled on the basis of the scores obtained when the mark is compared to a database of fingerprints unconnected to the mark and is the focus of the present contribution.

2.2. Data

2.2.1. Images

A background database of 685,245 rolled inked fingerprints (from 68,543 ten-print cards) was used. The number of fingers is not exactly ten times the number of ten-print cards since some fingers were not present on the cards. Comparisons to absent fingers yield scores of 0, and these scores have been excluded from all calculations. The ten-print cards used as a reference database are those that have been excluded from the Swiss central database mainly due to age when the repository was purged, and were received in a completely anonymous form. The prints were excluded before the year 2003; when exactly they were inserted, excluded and whether the donors present are more frequently of particular ethnic groups than the general population is unknown. The cards reflect fingerprints of individuals arrested approximately between 1940 and 1960. Due to the evolution of the Swiss population over the second half of the 20th century it is likely that the ethnic composition of the population in these cards differs from that in the current Swiss population. It is known that a very large majority of donors is male; otherwise, the ten-print cards are totally anonymous. The ten-print cards were inserted into a Sagem DMA AFIS and encoded automatically. Minutiae and general patterns were designated by the system and have not been checked by an operator. The results presented here are based on