
An AFIS Candidate List Centric Fingerprint Likelihood Ratio Model Based on Morphometric and Spatial Analyses (MSA)

Joshua Abraham, Paul Kwan, Christophe Champod,
Chris Lennard and Claude Roux

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/51184>

1. Introduction

The use of fingerprints for identification purposes boasts worldwide adoption for a large variety of applications, from governance centric applications such as border control to personalised uses such as electronic device authentication. In addition to being an inexpensive and widely used form of biometric for authentication systems, fingerprints are also recognised as an invaluable biometric for forensic identification purposes such as law enforcement and disaster victim identification. Since the very first forensic applications, fingerprints have been utilised as one of the most commonly used form of forensic evidence worldwide.

Applications of fingerprint identification are founded on the intrinsic characteristics of the friction ridge arrangement present at the fingertips, which can be generally classified at different levels or resolutions of detail (Figure 1). Generally speaking, fingerprint patterns can be described as numerous curved lines alternated as ridges and valleys that are largely regular in terms orientation and flow, with relatively few key locations being of exception (singularities). A closer examination reveals a more detail rich feature set allowing for greater discriminatory analysis. In addition, analysis of local textural detail such as ridge shape, orientation, and frequency, have been used successfully in fingerprint matching algorithms as primary features [1] [2] or in conjunction with other landmark-based features [3].

Both biometric and forensic fingerprint identification applications rely on premises that such fingerprint characteristics are highly discriminatory and immutable amongst the general population. However, the collectability of such fingerprint characteristics from biometric scanners, ink rolled impressions, and especially, latent marks, are susceptible to adverse factors such as partiality of contact, variation in detail location and appearance due to skin elasticity (specifically for level 2 and 3 features) and applied force, environmental noises such

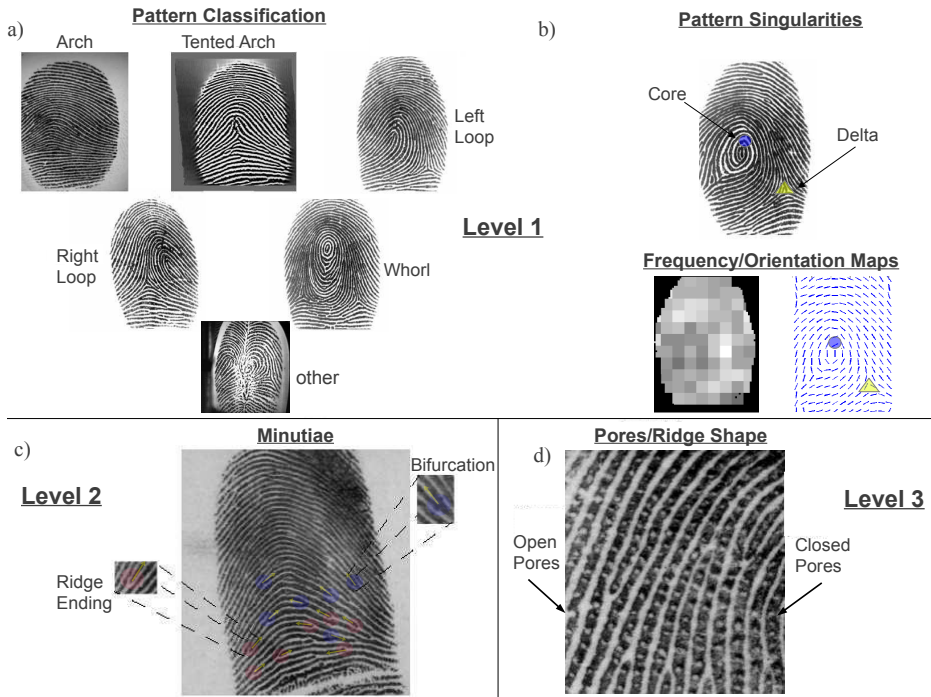


Figure 1. Level 1 features include features such as pattern class (a), singularity points and ridge frequency (b). Level 2 features (c) include minutiae with primitive types ridge endings and bifurcations. Level 3 features (d) include pores (open/closed) and ridge shape. These fingerprints were sourced from the FVC2002 [47], NIST4 [46], and NIST24 [48] databases

as moisture, dirt, slippage, and skin conditions such as dryness, scarring, warts, creases, and general ageing. Such influences generally act as a hindrance for identification, reducing both the quality and confidence of assessing matching features between impressions (Figure 2).

In this chapter, we will firstly discuss the current state of forensic fingerprint identification and how models play an important role for the future, followed by a brief introduction and review into relevant statistical models. Next, we will introduce a Likelihood Ratio (LR) model based on Support Vector Machines (SVMs) trained with features discovered via the morphometric and other spatial analyses of matching minutiae for both genuine and close imposter (or match and close non-match) populations typically recovered from Automated Fingerprint Identification System (AFIS) candidate lists. Lastly, experimentation performed on a set of over 60,000 publicly available fingerprint images (mostly sourced from NIST and FVC databases) and a distortion set of 6,000 images will be presented, illustrating that the proposed LR model is reliably guiding towards the right proposition in the identification assessment for both genuine and high ranking imposter populations, based on the discovered distortion characteristic differences of each population.

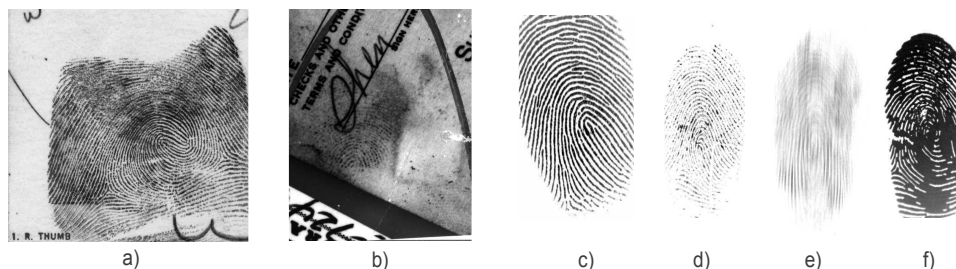


Figure 2. Examples of different fingerprint impressions, including an ink rolled print (a), latent mark (b), scanned fingerprint flats of ideal quality (c), dry skin (d), slippage (e), and over saturation (f). Fingerprints are sourced from the NIST 27 [48], FVC2004 [51], and our own databases.

1.1. Forensic fingerprint identification

Historically, the forensic identification of fingerprints has had near unanimous acceptance as a gold standard of forensic evidence, where the scientific foundations of such testimonies were rarely challenged in court proceedings. In addition, fingerprint experts have generally been regarded as expert witnesses with adequate training, scientific knowledge, relevant experience, and following a methodical process for identification, ultimately giving credibility to their expert witness testimonies.

Fingerprint experts largely follow a friction ridge identification process called ACE-V (Analysis, Comparison, Evaluation, and Verification) [5] to compare an unknown fingermark with known fingerprint exemplars. The ACE-V acronym also details the ordering of the identification process (Figure 3). In the analysis stage, all visible ridge characteristics (level 1, 2, and 3) are noted and assessed for reliability, while taking into account variations caused by pressure, distortion, contact medium, and development techniques used in the laboratory. The comparison stage involves comparing features between the latent mark and either the top n fingerprint exemplars return from an AFIS search, or specific pre-selected exemplars. If a positive identification is declared, all corresponding features are charted, along with any differences considered to be caused by environmental influence. The Evaluation stage consists of an expert making an inferential decision based on the comparison stage observations. The possible outcomes [6] are:

- exclusion: a discrepancy of features are discovered so it precludes the possibility of a common source,
- identification: a significant correspondence of features are discovered that is considered to be sufficient in itself to conclude to a common source, and
- inconclusive: not enough evidence is found for either an exclusion or identification.

The Verification stage consists of a peer review of the prior stages. Any discrepancies in evaluations are handled by a conflict resolution procedure.

Identification evaluation conclusions [7] made by fingerprint experts have historical influence from Edmond Locard's *tripartite rule* [8]. The tripartite rule is defined as follows:

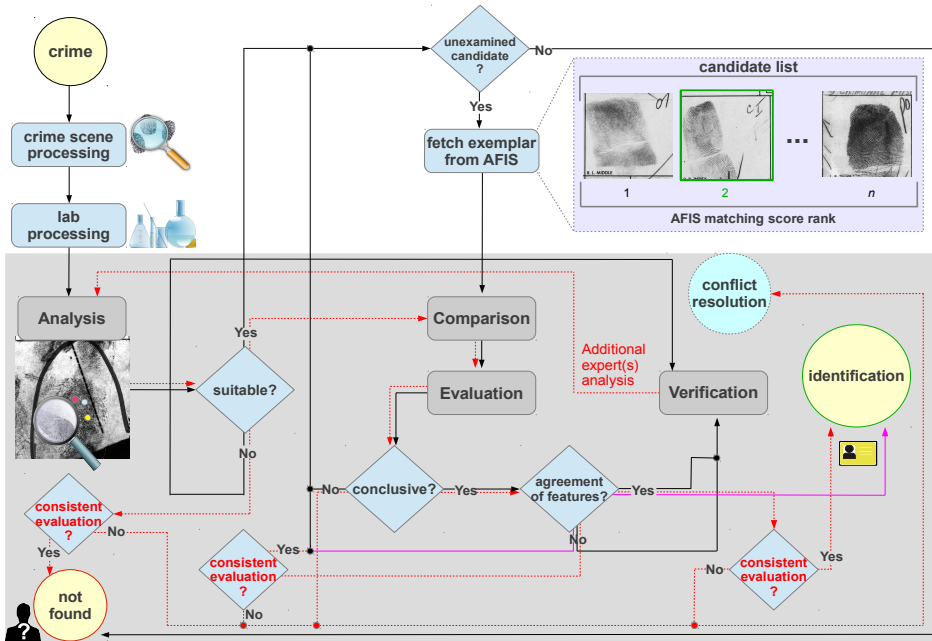


Figure 3. Flowchart of modern ACE-V process used in conjunction with AFIS. The iterative comparison of each exemplar fingerprint in the AFIS candidate list is performed until identification occurs or no more exemplars are left. The red flow lines indicate the process for the verification stage analysis. The purple flow line from the agreement of features test shows the ACE process that skips the evaluation stage.

- Positive identifications are possible when there are more than 12 minutiae within sharp quality fingermarks.
- If 8 to 12 minutiae are involved, then the case is borderline. Certainty of identity will depend on additional information such as finger mark quality, rarity of pattern, presence of the core, delta(s), and pores, and ridge shape characteristics, along with agreement by at least 2 experts.
- If a limited number of minutiae are present, the fingermarks cannot provide certainty for an identification, but only a presumption of strength proportional to the number of minutiae.

Holistically, the tripartite rule can be viewed as a probabilistic framework, where the successful applications of the first and second rules are analogous to a statement with 100% certainty that the mark and the print share the same source, whereas the third rule covers the probability range between 0% to 100%. While some jurisdictions only apply the first rule to set a numerical standard within the ACE-V framework, other jurisdictions (such as Australia, UK, and USA [9]) adopt a holistic approach, where no strict numerical standard or feature combination is prescribed. Nevertheless, current fingerprint expert testimony is largely restricted to conclusions that convey a statement of certainty, ignoring the third rule’s probabilistic outcome.

1.2. Daubert and criticisms

Recently, there has been a number of voiced criticisms on the scientific validity of forensic fingerprint identification [10] [11] [12] [13] [14] [15]. Questions with regards to the scientific validity of forensic fingerprint identification began shortly after the *Daubert* case [17]. In the 1993 case of *Daubert v. Merrell Dow Pharmaceuticals* [18] the US Supreme Court outlined criteria concerning the admissibility of scientific expert testimony. The criteria for a valid scientific method given were as follows:

- must be based on testable and falsifiable theories/techniques,
- must be subjected to peer-review and publication,
- must have known or predictable error rates,
- must have standards and controls concerning its applications, and
- must be generally accepted by a relevant scientific community.

The objections which followed [13] [14] [15] from a number of academic and legal commentators were:

- the contextual bias of experts for decisions made within the ACE-V (Analysis, Comparison, Evaluation, and Verification) framework used in fingerprint identification
- the unfounded and unfalsifiable theoretical foundations of fingerprint feature discriminability, and
- the ‘unscientific’ absolute conclusions of identification in testimonies (i.e., either match, non-match, or inconclusive).

There have been a number of studies [16] over the last 5 years concerning contextual bias and the associated error rates of ACE-V evaluations in practice. The experiments reported by [19] led to conclusions that experts appear more susceptible to bias assessments of ‘inconclusive’ and ‘exclusion’, while false positive rates are reasonably low within simulation of the ACE-V framework. It has also been suggested from results in [20] and [21] that not all stages of ACE-V are equally vulnerable to contextual bias, with primary effects occurring in the analysis stage, with proposals on how to mediate such variability found in [22]. While contextual bias is solely concerned with the influence of the expert, the remaining criticisms can be summarised as the non-existence of a scientifically sound probabilistic framework for fingerprint evidential assessment, that has the consensual approval from the forensic science community.

The theoretical foundations of fingerprint identification primarily rest on rudimentary observational science, where a high discriminability of feature characteristics exists. However, there is a lack of consensus regarding quantifiable error rates for a given pair of ‘corresponding’ feature configurations [23]. Some critics have invoked a more traditional interpretation for discriminability [24] [25], claiming that an assumption of ‘uniqueness’ is used. This clearly violates the falsifiable requirement of *Daubert*. However, it has been argued that modern day experts do not necessarily associate discriminability with uniqueness [26]. Nevertheless, a consensus framework for calculating accurate error rates for corresponding fingerprint features needs to be established.

1.3. Role of statistical models

While a probabilistic framework for fingerprint comparisons has not been historically popular and was even previously banned by professional bodies [8], a more favourable treatment within the forensic community is given in recent times. For example, the IAI have recently rescinded their ban on reporting possible, probable, or likely conclusions [27] and support the future use of valid statistical models (provided that they are accepted as valid by the scientific community) to aid the practitioner in identification assessments. It has also been suggested in [28] that a probabilistic framework is based on strong scientific principles unlike the traditional numerical standards.

Statistical models for fingerprint identification provide a probabilistic framework that can be applied to forensic fingerprint identification to create a framework for evaluations, that do not account for the inherent uncertainties of fingerprint evidence. Moreover, the use of such statistical models as an identification framework helps answer the criticisms of scientific reliability and error rate knowledge raised by some commentators. For instance, statistical models can be used to describe the discriminatory power of a given fingerprint feature configuration, which in hand can be used to predict and estimate error rates associated with the identification of specific fingerprint features found in any given latent mark.

Statistical models could potentially act as a tool for fingerprint practitioners with evaluations made within the ACE-V framework, specifically when the confidence in identification or exclusion is not overtly clear. However, such applications require statistical models to be accurate and robust to real work scenarios.

2. Likelihood Ratio models

A *likelihood ratio* (LR) is a simple yet powerful statistic when applied to a variety of forensic science applications, including inference of identity of source for evidences such as DNA [29], ear-prints [30], glass fragments [31], and fingerprints [32] [33] [34] [35]. An LR is defined as the ratio of two likelihoods of a specific *event* occurring, each of which follow a different prior hypothesis, and thus, empirical distribution. In the forensic identification context, an event, E , may represent the recovered evidence in question, while the prior hypotheses considered for calculating the two likelihoods of E occurring are:

- H_0 : E comes from a specific known source, P , and
- H_A : E has an alternative origin to P .

Noting any additional relevant prior information collected from the crime scene as I_{cs} , the LR can be expressed as

$$LR = \frac{P(E|H_0, I_{cs})}{P(E|H_A, I_{cs})} \quad (1)$$

where $P(E|H_0, I_{cs})$ is the likelihood of the observations on the mark given that the mark was produced by the same finger as the print P , while $P(E|H_A, I_{cs})$ is the likelihood of the observations on the mark given that the mark was *not* produced by the same finger as P . The LR value can be interpreted as follows:

- $LR < 1$: the evidence has more support for hypothesis H_A ,

- $LR = 1$: the evidence has equal support from both hypotheses, and
- $LR > 1$: the evidence has more support for hypothesis H_0 .

The general LR form of equation (1) can be restated specifically for fingerprint identification evaluations. Given an unknown query impression, y , (e.g., unknown latent mark) with m' marked features (denoted as $y^{(m')}$), and a known impression, x , (e.g., known AFIS candidate or latent mark) with m marked features (denoted as $x^{(m)}$), the LR is defined as

$$LR_{finger} = \frac{P(y^{(m')}|x^{(m)}, H_0, I_{cs})}{P(y^{(m')}|x^{(m)}, H_A, I_{cs})} \quad (2)$$

where the value $P(y^{(m')}|x^{(m)}, H_0, I_{cs})$ represents the probability that impressions x and y agree given that the marks were produced by the same finger, while $P(y^{(m')}|x^{(m)}, H_A, I_{cs})$ is the probability that x and y agree given that the marks were *not* produced by the same finger, using the *closest* q corresponding features between $x^{(m)}$ and $y^{(m')}$ with $q \leq \min(m, m')$. Thus, hypotheses used to calculate the LR numerator and denominator probabilities are defined as:

- H_0 : x and y were produced by the same finger, and
- H_A : x and y were produced by different fingers.

The addendum crime scene information, I_{cs} , may include detail of surrounding fingermarks, surficial characteristics of the contacted medium, or a latent mark quality/confidence assessment. In order to measure the within-finger and between-finger variability of landmark based feature configurations required to derive values for $P(y^{(m')}|x^{(m)}, H_0, I_{cs})$ and $P(y^{(m')}|x^{(m)}, H_A, I_{cs})$, models either use statistical distributions of dissimilarity metrics (used as a proxy for direct assessment) derived from either the analysis of spatial properties [33] [34] [35], or analysis of similarity score distributions produced by the AFIS [36] [37] [38].

2.1. AFIS score based LR models

AFIS score based LR models use estimates of the genuine and imposter similarity score distributions from fingerprint matching algorithm(s) within AFIS, in order to derive a LR measure. In a practical application, a given mark and exemplar may have an AFIS similarity score of s , from which the conditional probability of the score can be calculated (Figure 4) to give an LR of

$$LR = \frac{P(s|H_0)}{P(s|H_A)} \quad (3)$$

2.1.1. Parametric Based Models

In order to estimate the score distributions used in equation (3), the authors of [36] proposed using the Weibull $W(\lambda, \beta)$ and Log-Normal $\ln \mathcal{N}(\mu, \sigma^2)$ distributions with scale/shape parameters tuned to estimate the genuine and imposter AFIS score distributions, respectively. Given query and template fingermarks with an AFIS similarity score, s , the LR is

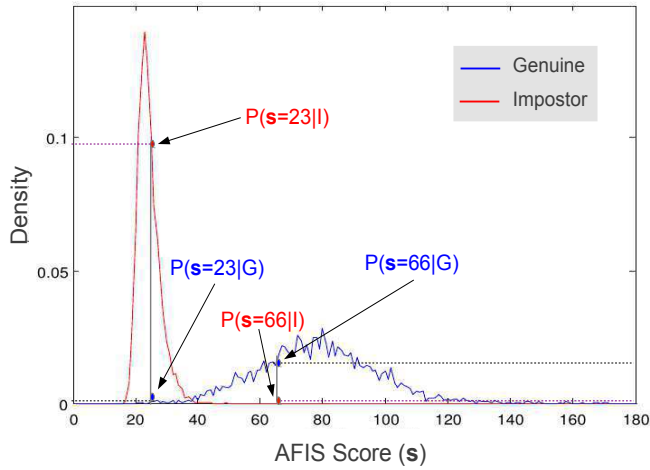


Figure 4. Typical AFIS imposter and genuine score distributions. The LR can be directly calculated for a given similarity score using the densities from these distributions.

$$LR = \frac{f_W(s|\lambda, \beta)}{f_{ln \mathcal{N}}(s|\mu, \sigma^2)} \tag{4}$$

using the proposed probability density functions of the estimated AFIS genuine and imposter score distributions.

An updated variant can be found in [37], where imposter and genuine score distributions are modelled per minutiae configuration. This allows the rarity of the configuration to be accounted for.

2.1.2. Non-Match Probability Based Model

The authors of [38] proposed a model based on AFIS score distributions, using LR and Non-Match Probability (NMP) calculations. The NMP can be written mathematically as

$$NMP = P(H_A|s) = \frac{P(s|H_A)P(H_A)}{P(s|H_A)P(H_A) + P(s|H_0)P(H_0)}, \tag{5}$$

which is simply the complement of the probability that the null hypothesis (i.e., x and y come from the same known source) is true, given prior conditions x , y , and I_{cs} (i.e., background information).

Three main methods for modelling the AFIS score distributions were tested, being (i) histogram based, (ii) Gaussian kernel density based, and (iii) parametric density based estimation using the proposed distributions found in [36]. Given an AFIS score, s , the NMP and LR were calculated by setting $P(H_A) = P(H_0)$, while estimating both $P(s|H_A)$ and

$P(s|H_0)$ either by normalised bin (method (i)) or probability density (methods (ii) and (iii)) values for respective distributions. Experimentation revealed that the parametric method was biased. In addition, the authors suggest that the kernel density method is the most ideal, as it does not suffer from bias while it can be used to extrapolate NMP scores where no match has been observed, unlike the histogram based representation.

2.1.3. Practicality of AFIS based LR Models

AFIS score based LR models provide a framework that is both practically based and simple to implement in conjunction with the AFIS architecture. However, model performance is dependent on the matching algorithm of the AFIS. In fact, LR models presented will usually reflect the exact information contained in a candidate list of an AFIS query. A more complex construction, for instance, multiple AFIS matching algorithms with a mixture-of-experts statistical model would be more ideal and avoid LR values that are strictly algorithm dependent.

The scores produced from matching algorithms in AFIS detail pairwise similarity between two impressions (i.e., mark and exemplar). However, the methods used in [36] [38], which generalise the distributions for all minutiae configurations, do not allow evidential aspects such as the rarity of a given configuration to be considered. A more sound approach would be to base LR calculations on methods that do not have primary focus on only pairwise similarities, but consider statistical characteristics of features within a given population. For instance, the LR for a rare minutiae configuration should be weighted to reflect its significance. This is achieved in the method described in [37] by focusing distribution estimates of scores for each minutiae configuration.

2.2. Feature Vector based LR models

Feature Vector (FV) based LR models are based on FVs constructed from landmark (i.e., minutiae) feature analyses. A dissimilarity metric is defined that is based on the resulting FV. The distributions of such vector dissimilarity metrics are then analysed for both genuine and imposter comparisons, from which an LR is derived.

2.2.1. Delauney Triangulation FV Model

The first FV based LR model proposed in the literature can be found in [33]. FVs are derived from Delaunay triangulation (Figure 5 left) for different regions of the fingerprint. Each FV was constructed as follows:

$$x = [GP_x, R_x, Nt_x, \{A_{1x}, L_{1x-2x}\}, \{A_{2x}, L_{2x-3x}\}, \{A_{3x}, L_{3x-1x}\}] \quad (6)$$

where GP_x is the pattern of the mark, R_x is the region of the fingerprint, Nt_x is the number of minutiae that are ridge endings in the triangle (with $Nt_x \in \{0, 1, 2, 3\}$), A_{ix} is the angle of the i^{th} minutia, and $L_{ix-((i+1) \bmod 3)x}$ is the length in pixels between the i^{th} and the $((i + 1) \bmod 3)^{th}$ minutiae, for a given query fingerprint. Likewise, these structures are created for candidate fingerprint(s):

$$y = [GP_y, R_y, Nt_y, \{A_{1y}, L_{1y-2y}\}, \{A_{2y}, L_{2y-3y}\}, \{A_{3y}, L_{3y-1y}\}] \quad (7)$$

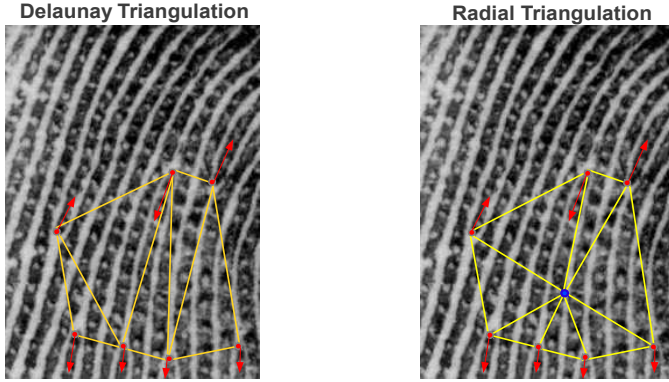


Figure 5. Delaunay triangulation (left) and radial triangulation (right) differences for a configuration of 7 minutiae. The blue point for the radial triangulation illustration represents the centroid (i.e., arithmetic mean of minutiae x-y coordinates).

The FVs can be decomposed into *continuous* and *discrete* components, representing the measurement based and count/categorical features, respectively. Thus, the likelihood ratio is rewritten as:

$$LR = \underbrace{\frac{P(x_c, y_c | x_d, y_d, H_0, I_{cs})}{P(x_c, y_c | x_d, y_d, H_A, I_{cs})}}_{LR_{c|d}} \cdot \underbrace{\frac{P(x_d, y_d | H_0, I_{cs})}{P(x_d, y_d | H_A, I_{cs})}}_{LR_d} = LR_{c|d} \cdot LR_d \tag{8}$$

where LR_d is formed as a prior likelihood ratio with discrete FVs $x_d = [GP_x, R_x, Nt_x]$ and $y_d = [GP_y, R_y, Nt_y]$, while continuous FVs x_c and y_c contain then remaining features in x and y , respectively. The discrete likelihood numerator takes the value of 1, while the denominator was calculated using frequencies for general patterns multiplied by region and minutia-type combination probabilities observed from large datasets.

A dissimilarity metric, $d(x_c, y_c)$, was created for comparing the continuous FV defined as:

$$d(x_c, y_c) = \Delta^2 A_1 + \Delta^2 L_{1-2} + \Delta^2 A_2 + \Delta^2 L_{2-3} + \Delta^2 A_3 + \Delta^2 L_{3-1} \tag{9}$$

with Δ^2 as the squared difference of corresponding variables from x_c and y_c . This was used to calculate the continuous likelihood value, with:

$$LR_{c|d} = \frac{P(d(x_c, y_c) | x_d, y_d, H_0, I_{cs})}{P(d(x_c, y_c) | x_d, y_d, H_A, I_{cs})}. \tag{10}$$

Finger/Region	LR True < 1	LR False > 1
Index/All	2.94 %	1.99 %
Middle/All	1.99 %	1.84 %
Thumbs/All	3.27 %	3.24 %
Index/Core	4.19 %	1.36 %
Middle/Core	3.65 %	1.37 %
Thumbs/Core	3.74 %	2.43 %
Index/Delta	1.95 %	2.62 %
Middle/Delta	2.96 %	2.58 %
Thumbs/Delta	2.39 %	5.20 %

Table 1. Some likelihood ratio error rate results for different finger/region combinations.

Density functions of both $P(d(x_c, y_c)|x_d, y_d, H_0, I_{cs})$ and $P(d(x_c, y_c)|x_d, y_d, H_A, I_{cs})$ were estimated using a kernel smoothing method. All LR numerator and denominator likelihood calculations were derived from these distribution estimates.

Two experiments were configured in order to evaluate within-finger (i.e., genuine) and between-finger (i.e., imposter) LRs. Ideally, LRs for within-finger comparisons should be larger than all between-finger ratios. The within-finger experiment used 216 fingerprints from 4 different fingers under various different distortion levels. The between-finger datasets included the same 818 fingerprints used in the minutia-type probability calculations. Delaunay triangulation had to be manually adjusted in some cases due to different triangulation results occurring under high distortion levels. Error rates for LRs greater than 1 for false comparisons (i.e., between-finger) and LRs less than 1 for true comparisons (i.e., within-finger) for index, middle, and thumbs, are given in Table 1. These errors rates indicate the power that 3 minutiae (in each triangle) have in creating an LR value dichotomy between within and between finger comparisons.

2.2.2. Radial Triangulation FV Model: I

Although the triangular structures of [33] performed reasonably well in producing higher LRs for within-finger comparisons against between-finger comparisons, there are issues with the proposed FV structure’s robustness towards distortion. In addition, LRs could potentially have increased dichotomy between imposter and genuine comparisons by including more minutiae in the FV structures, rather than restricting each FV to only have three minutiae.

The authors of [34] defined *radial triangulation* FVs based on n minutiae $x = [GP_x, x_s]$ with:

$$x^{(n)} = [\{T_{x,1}, RA_{x,1}, R_{x,1}, L_{x,1,2}, S_{x,1}\}, \{T_{x,2}, RA_{x,2}, R_{x,2}, L_{x,2,3}, S_{x,2}\}, \dots, \{T_{x,n}, RA_{x,n}, R_{x,n}, L_{x,n,1}, S_{x,n}\}], \tag{11}$$

(and similarly for y and $y^{(n)}$), where GP denotes the general pattern, T_k is the minutia type, RA_k is the direction of minutia k relative to the image, R_k is the radius from the k th minutia to the centroid (Figure 5 right), $L_{k,k+1}$ is the length of the polygon side from minutia k to $k + 1$, and S_k is the area of the triangle defined by minutia k , $(k + 1) \bmod n$, and the centroid.

The LR was then calculated as

$$LR = \frac{P(x^{(n)}, y^{(n)} | GP_x, GP_y, H_0, I_{cs})}{P(x^{(n)}, y^{(n)} | GP_x, GP_y, H_A, I_{cs})} \cdot \frac{P(GP_x, GP_y | H_0, I_{cs})}{P(GP_x, GP_y | H_A, I_{cs})} = LR_{n|g} \cdot LR_g \quad (12)$$

The component LR_g is formed as a prior likelihood with $P(GP_x, GP_y | H_0, I_{cs}) = 1$ and $P(GP_x, GP_y | H_A, I_{cs})$ equal to the FBI pattern frequency data. Noting that the centroid FVs can be arranged in n different ways (accounting for clockwise rotation):

$$y_j^{(n)} = (\{T_{y,k}, RA_{y,k}, R_{y,k}, L_{y,k,(k+1) \bmod n}, S_{y,k}\}, \\ k = j, (j + 1) \bmod n, \dots, (j - 1) \bmod n),$$

for $j = 1, 2, \dots, n$, $LR_{n|g}$ was defined as

$$LR_{n|g} = \frac{P(d(x^{(n)}, y^{(n)}) | GP_x, GP_y, H_0, I_{cs})}{P(d(x^{(n)}, y^{(n)}) | GP_x, GP_y, H_A, I_{cs})} \quad (13)$$

where the dissimilarity metric is

$$d(x^{(n)}, y^{(n)}) = \min_{i=1, \dots, n} d(x^{(n)}, y_i^{(n)}). \quad (14)$$

The calculation of each of the $d(x^{(n)}, y_i^{(n)})$ is the Euclidean distance of respective FVs which are normalised to take a similar range of values. The two conditional probability density functions of $P(d(x^{(n)}, y^{(n)}) | GP_x, GP_y, H_0, I_{cs})$ and $P(d(x^{(n)}, y^{(n)}) | GP_x, GP_y, H_A, I_{cs})$ were estimated using mixture models of normal distributions with a mixture of three and four distributions, respectfully, using the EM algorithm to estimate distributions for each finger and number of minutiae used.

This method modelled within and between finger variability more accurately in comparison to the earlier related work in [33], due to the flexibility of the centroid structures containing more than three minutiae. For example, the addition of one extra minutia halved the LR error rate for some fingerprint patterns. In addition, the prior likelihood is more flexible in real life applications as it is not dependent on identifying the specific fingerprint region (which is more robust for real life fingerprint-to-exemplar comparisons).

2.2.3. Radial Triangulation FV Model: II

The authors of [35] proposed a FV based LR model using radial triangulation structures. In addition, they tuned the model using distortion and examination influence models. The radial triangulation FVs used were based on the structures defined in [34], where five features are stored per minutia, giving

$$y_i^{(n)} = (\{\delta_j, \sigma_j, \theta_j, \alpha_j, \tau_j\}, i = j, (j + 1) \bmod n, \dots, (j - 1) \bmod n),$$

for a configuration $y^{(n)}$ starting from the i^{th} minutia, for $i = 1, 2, \dots, n$, where δ_j is the distance between the j^{th} minutia and the centroid point, σ_j is the distance between the j^{th} minutia and the next contiguous minutia (in a clockwise direction), θ_j is the angle between the direction of a minutia and the line from the centroid point, α_j is the area of the triangle constituted by the j^{th} minutia, the next contiguous minutia and the centre of the polygon, and τ_j is the type of the j^{th} minutia (ridge ending, bifurcation, unknown).

The distance between configurations $x^{(n)}$ and $y^{(n)}$, each representing n minutiae, is

$$d(x^{(n)}, y^{(n)}) = \min_{i=1, \dots, n} d_c(x^{(n)}, y_i^{(n)}) \tag{15}$$

where

$$d_c(x^{(n)}, y_i^{(n)}) = \sum_{j=1}^n \Delta_j \tag{16}$$

with

$$\begin{aligned} \Delta_j = & q_\delta \cdot (x^{(n)}(\delta_j) - y_i^{(n)}(\delta_j))^2 + q_\sigma \cdot (x^{(n)}(\sigma_j) - y_i^{(n)}(\sigma_j))^2 \\ & + q_\theta \cdot d_\theta(x^{(n)}(\theta_j), y_i^{(n)}(\theta_j))^2 + q_\alpha \cdot (x^{(n)}(\alpha_j) - y_i^{(n)}(\alpha_j))^2 \\ & + q_\tau \cdot d_T(x^{(n)}(\tau_j), y_i^{(n)}(\tau_j))^2 \end{aligned} \tag{17}$$

where $x^{(n)}(\delta_j)$ (and $y_i^{(n)}(\delta_j)$) is the normalised value for δ for the j^{th} minutiae, and likewise for all other normalised vector components σ , θ , α , and τ , while d_θ is the angular difference and d_T is the defined minutiae type difference metric. The multipliers (i.e., q_δ , q_σ , q_θ , q_α , and q_τ) are tuned via a heuristic based procedure.

The proposed LR calculation makes use of:

- distortion model: based on the Thin Plate Spline (TPS) bending energy matrices representing the non-affine differences of minutiae spatial detail trained from a dataset focused on finger variability,
- examiner influence model: created to represent the variability of examiners when labelling minutiae in fingerprint images.

Let $y^{(k)}$ be the configuration of a fingerprint, $x_{\min}^{(k)}$ the closest k configuration found, and $z_{i, \min}^{(k)}$ the closest configuration for the i^{th} member of a reference database containing N impressions. Synthetic FVs can be generated from minute modifications to minutiae locations

represented by a given FV, via Monte-Carlo simulation of both distortion and examiner influence models. A set of M synthetic FVs are created for $x_{\min}^{(k)}$ ($\{\zeta_1^{(k)}, \dots, \zeta_M^{(k)}\}$) and for each $z_{i,\min}^{(k)}$ ($\{\zeta_{i,1}^{(k)}, \dots, \zeta_{i,M}^{(k)}\}$), from which the LR is given as

$$LR = \frac{N \sum_{i=1}^M \psi \left(d(y^{(k)}, \zeta_i^{(k)}) \right)}{\sum_{i=1}^N \sum_{j=1}^M \psi \left(d(y^{(k)}, \zeta_{i,j}^{(k)}) \right)} \quad (18)$$

where ψ is defined as

$$\psi(d(y^{(k)}, \bullet)) = \exp \left(\frac{-\lambda_1 d(y^{(k)}, \bullet)}{T^{(k)}} \right) + \frac{B(d(y^{(k)}, \bullet), \lambda_2 k)}{B(d_0, \lambda_2 k)} \quad (19)$$

which is a mixture of Exponential and Beta functions with tuned parameters λ_1 and λ_2 , while d_0 is the smallest value into which distances were binned, and $T^{(k)}$ is the 95th percentile of simulated scores from the examiner influence model applied on $y^{(k)}$. Experimental results from a large validation dataset showed that the proposed LR model can generally distinguish within and between finger comparisons with high accuracy, while an increased dichotomy arose from increasing the configuration size.

2.2.4. Practicality of FV based LR Models

Generally speaking, to implement robust FV based statistical models for forensic applications, the following must be considered:

- Any quantitative measures used should be based on the data driven discovery of statistical relationships of features. Thus, a rich dataset for both within and between finger data is essential.
- Effects of skin distortion must be considered in models. Latent marks can be highly distorted from skin elasticity and applied pressure. For instance, differences in both minutiae location (relative to other features) and type (also known as type transfer) can occur when different distortion exists.
- Features used in models must be robust to noisy environmental factors, whilst maintaining a high level of discriminatory power. For instance, level 1 features such as classification may not be available due to partiality. In addition, level 2 sub-features such as ridge count between minutiae, minutiae type, and level 3 features such as pores, may not be available in a latent mark due to the material properties of the contacted medium or other environmental noise that regularly exist in latent mark occurrences.
- The model should be robust towards reasonable variations in feature markings from practitioners in the analysis phase of ACE-V. For instance, minutiae locations can vary slightly depending on where a particular practitioner marks a given minutia.

The LR models proposed in [33] and [34] use dissimilarity measures of FVs (equations (9) and (14)) which are potentially erroneous as minutiae types can change, particularly in distorted impressions. While the method in [35] has clearly improved the dissimilarity function by introducing tuned multipliers, squared differences in angle, area, and distance based measures are ultimately not probabilistically based. A joint probabilistic based metric for each FV component using distributions for both imposter and genuine populations would be more consistent with the overall LR framework.

With regards to skin distortion, the radial triangulation FV structures of [34] [35] are robust, unlike the Delaunay triangulation structure of [33]. Furthermore, the model proposed in [35] models realistic skin distortion encountered on flat surfaces by measuring the bending energy matrix for a specialised distortion set. However, this only accounts for the non-affine variation. Affine transformations such as shear and uniform compression/dilation are not accounted for. Such information can be particularly significant for comparisons of small minutiae configurations encountered in latent marks. For instance, a direct downward application of force may have prominent shear and scale variations (in addition to non-affine differences) for minutiae configurations, in comparison to the corresponding configurations of another impression from the same finger having no notable downward force applied.

3. Proposed method: Morphometric and Spatial Analyses (MSA) based Likelihood Ratio model

In this section, we present a newly formulated FV based LR model that focuses on the important sub-population of close non-matches (i.e., highly similar imposters), with intended practicality for fingermark-to-exemplar identification scenarios where only sparse minutiae triplet information may be available for comparisons. First we discuss relevant background material concerning morphometric and spatial measures to be used in the FVs of the proposed model. The proposed model is presented, which is based on a novel machine learning framework, followed by a proposed LR calculation that focuses on the candidate list population of an AFIS match query (i.e., containing close non-match exemplars and/or a matching exemplar). Finally, an experimental framework centred around the simulation of fingermark-to-exemplar close non-match discovery is introduced, followed by experimental results.

3.1. Morphometric and spatial metrics

The foundations of the morphometric and spatial analyses used in the proposed FV based LR model are presented. This includes a non-parametric multidimensional goodness-of-fit statistic, along with several other morphometrical measures that describe and contrast shape characteristics between two given configurations. In addition, a method for finding close non-match minutiae configurations is presented.

3.1.1. Multidimensional Kolmogorov-Smirnov Statistic for Landmarks

A general multidimensional Kolmogorov-Smirnov (KS) statistic for two empirical distributions has been proposed in [39] with properties of high efficiency, high statistical power, and distributional freeness. Like the classic one dimensional KS test, the multidimensional variant looks for the largest absolute difference between the empirical

and cumulative distribution functions, as a measure of fit. Without losing generality, let two sets with m and n points in \mathbb{R}^3 be denoted as $X = \{(x_1, y_1, z_1), \dots, (x_m, y_m, z_m)\}$ and $Y = \{(x'_1, y'_1, z'_1), \dots, (x'_n, y'_n, z'_n)\}$, respectively. For each point $(x_i, y_i, z_i) \in X$ we can divide the plane into eight defined regions

$$\begin{aligned} q_{i,1} &= \{(x, y, z) | x < x_i, y < y_i, z < z_i\}, \\ q_{i,2} &= \{(x, y, z) | x < x_i, y < y_i, z > z_i\}, \\ &\vdots \\ q_{i,8} &= \{(x, y, z) | x \geq x_i, y \geq y_i, z \geq z_i\}, \end{aligned}$$

and similarly for each $(x'_j, y'_j, z'_j) \in Y$,

$$\begin{aligned} q'_{j,1} &= \{(x, y, z) | x < x'_j, y < y'_j, z < z'_j\}, \\ q'_{j,2} &= \{(x, y, z) | x < x'_j, y < y'_j, z > z'_j\}, \\ &\vdots \\ q'_{j,8} &= \{(x, y, z) | x \geq x'_j, y \geq y'_j, z \geq z'_j\}. \end{aligned}$$

Further defining

$$D_m = \max_{\substack{i=1, \dots, m \\ s=1, \dots, 8}} | |X \cap q_{i,s}| - |Y \cap q_{i,s}| | \tag{20}$$

which is the maximum pairwise difference of point tallies for X and Y within each of the eight defined regions centred and evaluated at each point in X , and likewise,

$$D_n = \max_{\substack{j=1, \dots, n \\ s=1, \dots, 8}} | |X \cap q'_{j,s}| - |Y \cap q'_{j,s}| | \tag{21}$$

which is the maximum pairwise difference of point tallies for the eight defined regions centred and evaluated at each point in Y , the three dimensional KS statistic is

$$Z_{m,n,3D} = \sqrt{n.m / (n + m)} \cdot \left(\frac{D_m + D_n}{2} \right). \tag{22}$$

The three dimensional KS statistic can be specific to the minutiae triplet space where each minutia spatial and directional detail is represented as a three dimensional point, (x, y, θ) . Given $m = n$ matching minutiae correspondences from two configurations X and Y , alignment is performed prior to calculating the statistic, in order to ensure that minutiae correspondences are close together both spatially and directionally. However, direction has a circular nature that must be handled differently from the spatial detail. Instead of raw angular values, we use the orientation difference defined as

$$z = z(\theta, \theta_0) = \frac{\pi}{2} - \min(2\pi - |\theta - \theta_0|, |\theta - \theta_0|) \tag{23}$$

where $z \in [-\frac{\pi}{2}, \frac{\pi}{2}]$. Each minutia, (x, y, θ) , is then transformed to $(x, y, z(\theta, \theta_0))$ if the centred minutia used to create the eight regions has a direction of θ_0 , while region borders are defined in the third dimension by $z \geq 0$ and $z < 0$.

3.1.2. Thin Plate Spline and Derived Measures

The Thin Plate Spline (TPS) [40] is based on the algebraic expression of physical bending energy of an infinitely thin metal plate on point constraints after finding the optimal affine transformations for the accurate modelling of surfaces that undergo *natural warping* (i.e., where a diffeomorphism exists). Two sets of landmarks from each surface are paired in order to provide an interpolation map on $\mathbb{R}^2 \rightarrow \mathbb{R}^2$. TPS decomposes the interpolation into an affine transform that can be considered as the transformation that expresses the global geometric dependence of the point sets, and a non-affine transform that fine tunes the interpolation of the point sets. The inclusion of the affine transform component allows TPS to be invariant under both rotation and scale.

Given n control points

$$\{\mathbf{p}_1 = (x_1, y_1), \mathbf{p}_2 = (x_2, y_2), \dots, \mathbf{p}_n = (x_n, y_n)\}$$

from an input image in \mathbb{R}^2 and control points

$$\{\mathbf{p}'_1 = (x'_1, y'_1), \mathbf{p}'_2 = (x'_2, y'_2), \dots, \mathbf{p}'_n = (x'_n, y'_n)\}$$

from a target image \mathbb{R}^2 , the following matrices are defined in TPS:

$$\mathbf{K} = \begin{bmatrix} 0 & u(r_{12}) & \dots & u(r_{1n}) \\ u(r_{21}) & 0 & \dots & u(r_{2n}) \\ \dots & \dots & \dots & \dots \\ u(r_{n1}) & u(r_{n2}) & \dots & 0 \end{bmatrix},$$

where $u(r) = r^2 \log r^2$ with r as the Euclidean distance, $r_{ij} = \|p_i - p_j\|$,

$$\mathbf{P} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \dots & \dots & \dots \\ 1 & x_n & y_n \end{bmatrix}, \mathbf{V} = \begin{bmatrix} x'_1 & x'_2 & \dots & x'_n \\ y'_1 & y'_2 & \dots & y'_n \end{bmatrix}, \mathbf{Y} = [\mathbf{V} \mathbf{0}_{2 \times 3}]^T, \mathbf{L} = \begin{bmatrix} \mathbf{K} & \mathbf{P} \\ \mathbf{P}^T & \mathbf{0}_{3 \times 3} \end{bmatrix},$$

where \mathbf{K} , \mathbf{P} , \mathbf{V} , \mathbf{Y} , \mathbf{L} have dimensions $n \times n$, $3 \times n$, $2 \times n$, $(n + 3) \times 2$, and $(n + 3) \times (n + 3)$, respectively. The vector $W = (w_1, w_2, \dots, w_n)$ and the coefficients a_1, a_x, a_y , can be calculated by the equation

$$\mathbf{L}^{-1}\mathbf{Y} = (\mathbf{W} | a_1 \ a_x \ a_y)^T. \tag{24}$$

The elements of $\mathbf{L}^{-1}\mathbf{Y}$ are used to define the TPS interpolation function

$$f(x, y) = [f_x(x, y), f_y(x, y)], \tag{25}$$

with the coordinates compiled from the first column of $\mathbf{L}^{-1}\mathbf{Y}$ giving

$$f_x(x, y) = a_{1,x} + a_{x,x}x + a_{y,x}y + \sum_{i=1}^n w_{i,x}U(\|\mathbf{p}_i - (x, y)\|) \tag{26}$$

where $[a_{1,x} \ a_{x,x} \ a_{y,x}]^T$ is the affine transform component for x , and likewise for the second column, where

$$f_y(x, y) = a_{1,y} + a_{x,y}x + a_{y,y}y + \sum_{i=1}^n w_{i,y}U(\|\mathbf{p}_i - (x, y)\|) \tag{27}$$

with $[a_{1,y} \ a_{x,y} \ a_{y,y}]^T$ as the affine component for y . Each point (or minutia location in our application) can now be updated as

$$(x_{new}, y_{new}) = (f_x(x, y), f_y(x, y)). \tag{28}$$

It can be shown that the function $f(x, y)$ is the interpolation that minimises

$$I_f \propto \mathbf{W}\mathbf{K}\mathbf{W}^T = \mathbf{V}(\mathbf{L}_n^{-1}\mathbf{K}\mathbf{L}_n^{-1})\mathbf{V}^T, \tag{29}$$

where I_f is the *bending energy* measure

$$I_f = \int \int_{\mathbb{R}^2} \left(\frac{\partial^2 z}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 z}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 z}{\partial y^2} \right)^2 dx dy \tag{30}$$

and \mathbf{L}_n is the $n \times n$ sub-matrix of \mathbf{L} . Affine transform based metrics relating to shear, rotation, and scale (i.e., compression and dilation) can be calculated straight from Singular Value Decomposition (SVD) of the affine matrix

$$\mathbf{U}\mathbf{S}\mathbf{V}^T = SVD \left(\begin{bmatrix} a_{x,x} & a_{x,y} \\ a_{y,x} & a_{y,y} \end{bmatrix} \right). \tag{31}$$

From this decomposition, we define an angle cost

$$d_\theta = \min(\theta, 2\pi - \theta) \tag{32}$$

with $\theta = |(\arctan(\mathbf{V}_{1,2}, \mathbf{V}_{1,1}) - \arctan(\mathbf{U}_{1,2}, \mathbf{U}_{1,1}))|$, a shear cost

$$d_{shear} = \log(\mathbf{S}_{1,1}/\mathbf{S}_{2,2}), \tag{33}$$

and a scale cost

$$d_{scale} = \log \left(\max \left(\mathbf{S}_{1,1}, \mathbf{S}_{2,2}, \frac{1}{\mathbf{S}_{1,1}}, \frac{1}{\mathbf{S}_{2,2}} \right) \right). \tag{34}$$

3.1.3. Shape Size and Difference Measures

Shape size measures are useful metrics for comparing general shape characteristics. Given a matrix \mathbf{X} of dimensions $k \times m$, representing a set of k m -dimensional points, the *centroid size* [41] is defined as

$$S(\mathbf{X}) = \sqrt{\sum_{i=1}^k \|(\mathbf{X})_i - \bar{\mathbf{X}}\|^2}, \tag{35}$$

where $(\mathbf{X})_i$ is the i^{th} row of \mathbf{X} and $\bar{\mathbf{X}}$ is the arithmetic mean of the points in \mathbf{X} (i.e., centroid point). Given a second landmark configuration \mathbf{Y} also with k m -dimensional points, we define the shape size difference as

$$d_S = |S(\mathbf{X}) - S(\mathbf{Y})|. \tag{36}$$

Another useful shape metric is derived from the partial Procrustes method [41], which finds the optimal superimposition of one set of landmarks, \mathbf{X} , onto another, \mathbf{Y} , using translation and rotation affine operators:

$$\min_{\Gamma, \gamma} \|\mathbf{Y} - \mathbf{X}\Gamma - \mathbf{1}_k \gamma^T\|^2 \tag{37}$$

where $\mathbf{1}_k$ is a $(k \times 1)$ vector of ones, Γ is a $m \times m$ rotation matrix and γ is the $(m \times 1)$ translation offset vector. Using centred landmarks, $\mathbf{X}_c = \mathbf{C}\mathbf{X}$ and $\mathbf{Y}_c = \mathbf{C}\mathbf{Y}$ where $\mathbf{C} = \mathbf{I}_k - \frac{1}{k} \mathbf{1}_k \mathbf{1}_k^T$, the ordinary partial Procrustes sum of squares is

$$OSS_p(\mathbf{X}_c, \mathbf{Y}_c) = \text{trace}(\mathbf{X}_c^T \mathbf{X}_c) + \text{trace}(\mathbf{Y}_c^T \mathbf{Y}_c) - 2\|\mathbf{X}_c\| \|\mathbf{Y}_c\| \cos \rho(\mathbf{X}_c, \mathbf{Y}_c) \tag{38}$$

with $\rho(\mathbf{X}_c, \mathbf{Y}_c)$ as the Procrustes distance defined as

$$\rho(\mathbf{X}_c, \mathbf{Y}_c) = \arccos \left(\sum_{i=1}^m \lambda_i \right) \tag{39}$$

where $\lambda_1, \dots, \lambda_m$ are the square roots of the eigenvalues of $Z_X^T Z_Y Z_Y^T Z_X$ with $Z_X = \mathbf{H}\mathbf{X}/\|\mathbf{H}\mathbf{X}\|$ and $Z_Y = \mathbf{H}\mathbf{Y}/\|\mathbf{H}\mathbf{Y}\|$ for the Helmert sub-matrix, \mathbf{H} , with dimension $k \times k$.

3.1.4. Close Non-Match Discovery and Alignment

In order to reproduce the process of an examiner querying a minutiae configuration marked on fingerprint with an AFIS, a method for finding close configurations was developed. To find close non-matches for a particular minutiae configuration, we employed a simple search algorithm based solely on minutiae triplet features, in order to maintain robustness towards such fingerprint-to-exemplar match scenarios. The minutiae triplet features are extracted in a fully automated manner using the NIST mindtct tool [49] without particular attention to spurious results, besides minimum quality requirements as rated by the mindtct algorithm.

Algorithm 1 *findCloseTripletConfigs*: Find all close triplet configurations to \mathbf{X}

Require: A minutiae triplet set \mathbf{X} and a dataset of exemplars D .

```

candidateList = null
for all minutiae configurations  $\mathbf{Y} \in D$  with  $|\mathbf{X}| = |\mathbf{Y}|$  do
  for all minutiae  $(x_Y, y_Y, \theta_Y) \in \mathbf{Y}$  do
    found  $\leftarrow$  false
    for all minutiae  $(x_X, y_X, \theta_X) \in \mathbf{X}$  do
       $\mathbf{Y}' \leftarrow \mathbf{Y}$ 
      rotate  $\mathbf{Y}'$  by  $(\theta_X - \theta_Y)$  {This includes rotating minutiae angles.}
      translate  $\mathbf{Y}'$  by offset  $(x_X - x_Y, y_X - y_Y)$ 
      if  $\mathbf{Y}'$  is close to  $\mathbf{X}$  then
        found = true
        goto finished:
      end if
    end for
  end for
  finished:
  if found = true then
     $\mathbf{Y}' \leftarrow$  PartialProcrustes( $\mathbf{X}, \mathbf{Y}'$ ) {Translate/Rotate  $\mathbf{Y}'$  using partial Procrustes}
    TPS( $\mathbf{X}, \mathbf{Y}'$ ) {non-affine registration by TPS}
    if  $I_f < I_{max}$  then
      add  $\mathbf{Y}'$  to candidateList {Add if bending energy < limit (equation (29))}
    end if
  end if
end for
return candidateList

```

Once feature extraction is complete, the close match search algorithm (Algorithm 1) finds all equally sized close minutiae configurations in a given dataset of exemplars to a specified minutiae set configuration (i.e., potentially marked from a latent) in an iterative manner by assessing all possible minutiae triplet pairs via a crude affine transform based alignment on configuration structures. Recorded close minutiae configurations are then re-aligned using the partial Procrustes method using the discovered minutiae pairings. Unlike the Procrustes method, the partial Procrustes method does not alter scale of either landmarks. For the application of fingerprint alignment, ignoring scale provides a more accurate comparison of landmarks since all minutiae structures are already normalised by the resolution and dimensions of the digital image. The TPS registration is then applied for a non-affine transformation. If the bending energy is higher than a defined threshold, we ignore the

potential match due to the likely *unnatural* distortion encountered. Finally, a candidate list with all close minutiae configurations is produced for analysis.

3.2. Proposed model

We now propose an LR model based on what is found in [4], developed specifically to aid AFIS candidate list assessments, using the intrinsic differences of morphometric and spatial analyses (which we label as MSA) between match and close non-match comparisons, learnt from a two-class probabilistic machine learning framework.

3.2.1. Feature Vector Definition

Given two matching configurations X and Y (discovered from the procedure described in Section 3.1.4) a FV based on the previously discussed morphometric and spatial analyses is defined as:

$$\mathbf{x}_i = \{Z_{m,n,3D}, I_f, d_\theta, d_{shear}, d_{scale}, S(\mathbf{X}), d_S, OSS_p(\mathbf{X}_c, \mathbf{Y}_c), d_{mc}\} \quad (40)$$

where $Z_{m,n,3D}$ is the three dimensional KS statistic of equation (22) using the transformed triplet points, I_f , d_θ , d_{shear} , and d_{scale} are the defined measures of equations (29) and (32-34) resulting from registering \mathbf{X} onto \mathbf{Y} via TPS, $S(\mathbf{X})$ and d_S are the shape size and difference metric of equations (35-36), $OSS_p(\mathbf{X}_c, \mathbf{Y}_c)$ is the ordinary partial Procrustes sum of squares of equation (38), and d_{mc} is the difference of the number of interior minutiae within the convex hulls of \mathbf{X} and \mathbf{Y} . The d_{mc} measure is an optional component to the FV dependent on the clarity of a fingerprint's detail within the given minutiae configuration. For the experiments presented later in this chapter, we will exclude this measure.

The compulsory measures used in the proposed feature vector rely solely on features that are robust to the adverse environmental conditions of latent marks, all of which are based on minutiae triplet detail. The FV structures are categorised by genuine/imposter (or match/close non-match) classes, number of minutiae in the matching configurations, and configuration area (categorised as small, medium, and large).

3.2.2. Machine Learning of Feature Vectors

Using the categories prescribed for the defined FVs, a probabilistic machine learning framework is applied for finding the probabilities for match and close non-match classes. The probabilistic framework employed [42] is based on Support Vector Machines (SVMs) with unthresholded output, defined as

$$f(\mathbf{x}) = h(\mathbf{x}) + b \quad (41)$$

with

$$h(\mathbf{x}) = \sum_i y_i \alpha_i k(\mathbf{x}_i, \mathbf{x}) \quad (42)$$

where $k(\bullet, \bullet)$ is the kernel function, and the target output $y_i \in \{-1, 1\}$ represents the two classes (i.e., ‘close non-match’ and ‘match’, respectively). We use the radial basis function

$$k(\mathbf{x}_i, \mathbf{x}) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}\|^2) \tag{43}$$

due to the observed non-linear relationships of the proposed FV. Training the SVM minimises the error function

$$C \sum_i (1 - y_i f(\mathbf{x}_i))_+ + \frac{1}{2} \|h\|_{\mathcal{F}} \tag{44}$$

where C is the soft margin parameter (i.e., regularisation term which provides a way to control overfitting) and \mathcal{F} is the Reproducing Kernel Hilbert Space (RKHS) induced by the kernel k . Thus, the norm of h is penalised in addition to the approximate training misclassification rate. By transforming the target values with

$$t_i = \frac{y_i + 1}{2}, \tag{45}$$

the posterior probabilities $P(y_i = 1|f(\mathbf{x}_i))$ and $P(y_i = -1|f(\mathbf{x}_i))$ which represents the probabilities that \mathbf{x}_i is of classes ‘match’ and ‘close non-match’, respectively, can now be estimated by fitting a sigmoid function after the SVM output with

$$P(\mathbf{x}_i \text{ is a match} | f(\mathbf{x}_i)) = P(y_i = 1 | f(\mathbf{x}_i)) = \frac{1}{1 + \exp(Af(\mathbf{x}_i) + B)} \tag{46}$$

and

$$P(\mathbf{x}_i \text{ is a close non-match} | f(\mathbf{x}_i)) = P(y_i = -1 | f(\mathbf{x}_i)) = 1 - \frac{1}{1 + \exp(Af(\mathbf{x}_i) + B)}. \tag{47}$$

The parameters A and B are found by minimising the negative log-likelihood of the training data:

$$\arg \min_{A,B} \left[- \left(\sum_i t_i \log \left(\frac{1}{1 + \exp(Af(\mathbf{x}_i) + B)} \right) + (1 - t_i) \log \left(1 - \frac{1}{1 + \exp(Af(\mathbf{x}_i) + B)} \right) \right) \right] \tag{48}$$

using any optimisation algorithm, such as the Levenberg-Marquardt algorithm [43].

3.2.3. Likelihood Ratio Calculation

The probability distributions of equations (47-48) are posterior probabilities. Nevertheless, for simplicity of the initial application, we assume uniform distributions for $P(f(\mathbf{x}_i)) = z$ for some constant, z , whereas $P(\mathbf{x}_i \text{ is a match}) = a$ and $P(\mathbf{x}_i \text{ is a close non-match}) = 1 - a$

where a reflects the proportion of close minutiae configuration comparisons that are ground truth matches. Thus, the LR is equivalent to the posterior ratio (PR)

$$LR = \left(\frac{1-a}{a}\right) \cdot PR = \left(\frac{1-a}{a}\right) \cdot \frac{P(\mathbf{x}_i \text{ is a match} | f(\mathbf{x}_i))}{P(\mathbf{x}_i \text{ is a close non-match} | f(\mathbf{x}_i))}. \quad (49)$$

For future consideration, the probabilities $P(\mathbf{x}_i \text{ is a match})$ and $P(\mathbf{x}_i \text{ is a close non-match})$ can be adaptively based on Cumulative Match Characteristic (CMC) curve [44] statistics of a given AFIS system or any other relevant background information.

As already noted, the LR formulas are based on different distributions specified per FV categories of minutiae count and the area of the given configuration. This allows the LR models to capture any spatial and morphometric relational differences between such defined categories. Unlike previous LR methods that are based on the distributions of a dissimilarity metric, the proposed method is based on class predictions based on a number of measures, some of which do not implicitly or explicitly rate or score a configuration's dissimilarity (e.g. centroid size, $S(\mathbf{X}_i)$). Instead, statistical relationships of the FV measures and classes are learnt by SVMs in a supervised manner, only for class predictions.

In its current proposed form, the LR of equation (49) is not an evidential weight for the entire population, but rather, an evidential weight specifically for a given candidate list.

3.3. Experimentation

3.3.1. Experimental Databases

Without access to large scale AFISs, a sparse number of fingerprint-to-exemplar datasets exists in the public domain (i.e., NIST27 is the only known dataset with only 258 sets). Thus, to study the within-finger characteristics, a distortion set was built.

We follow a methodology similar to that of [35] where live scanned fingerprints have eleven directions applied, eight of which are linear directions, two torsional, and central application of force. Using a readily available live scan device (Suprema Inc. Realscan-D: 500ppi with rolls, single and dual finger flats), we follow a similar methodology, described as follows:

- sixteen different linear directions of force,
- four torsion directions of force,
- central direction of force,
- all directions described above have at least three levels of force applied,
- at least five rolled acquisitions are collected,
- finally, numerous impressions with emphasis on partiality and high distortion are obtained by recording fifteen frames per second, while each finger manoeuvres about the scan area in a freestyle manner for a minimum of sixty seconds.

This gave a minimum total of 968 impressions per finger. A total of 6,000 impressions from six different fingers (from five individuals) were obtained for our within-finger dataset, most of which are partial impressions from the freestyle methodology. For the between-finger

comparisons, we use the within-finger set in addition to the public databases of NIST 14 [45] (27000×2 impressions), NIST 4 [46] (2000×2 impressions), FVC 2002 [47] ($3 \times 110 \times 8$ flat scan/swipe impressions), and the NIST 27 database [48] (258 exemplars + 258 latents), providing over 60,000 additional impressions.

3.3.2. SVM Training Procedure

A simple training/evaluation methodology was used in the experiments. After finding all FVs for similar configurations, a random selection of 50% of the FVs were used to train each respective SVM by the previously defined categories (i.e., minutiae configuration count and area). The remaining 50% of FVs were used to evaluate the LR model accuracy. The process was then repeated by swapping the training and test sets (i.e., two-fold cross-validation). Due to the large size of the within-finger database, a substantially larger number of within-finger candidates are returned. To alleviate this, we randomly sampled the within-finger candidates to be of equal number to the between-finger counterparts (i.e., $a = 0.5$ in equation (49)). All individual features within each FV were scaled to have a range of $[0, 1]$, using pre-defined maximum and minimum values specific to each feature component.

A naive approach was used to find the parameters for the SVMs. The radial basis kernel parameter, γ , and the soft learning parameter, C , of equations (43) and (44), respectively, were selected using a grid based search, using the cross-validation framework to measure the test accuracy for each parameter combination, (γ, C) . The parameter combination with the highest test accuracy was selected for each constructed SVM.

3.3.3. Experimental Results

Experiments were conducted for minutiae configurations of sizes of 6, 7, and 8 (Figure 6) from the within-finger dataset, using configurations marked manually by an iterative circular growth around a first minutiae until the desired configuration sizes were met. From the configuration sizes, a total of 12144, 4500, and 1492 candidates were used, respectively, from both the within (50%) and between (50%) finger datasets. The focus on these configuration settings were due to three reasons: firstly, the high computational overhead involved in the candidate list retrieval for the prescribed datasets, secondly, configurations of such sizes perform poorly in modern day AFIS systems [50], and finally, such configuration sizes are traditionally contentious in terms of Locard's tripartite rule, where a probabilistic approach is prescribed to be used.

The area sizes used for categorising the minutiae configurations were calculated by adding up the individual areas of triangular regions created using Delaunay triangulation. Small, medium, and large configuration area categories were defined as $0 < A < 4.2mm^2$, $4.2mm^2 \leq A < 6.25mm^2$, and $A \geq 6.25mm^2$, respectively.

The results clearly indicate a stronger dichotomy of match and close non-match populations when the number of minutiae was increased. In addition, the dichotomy was marginally stronger for larger configuration areas with six minutiae. Overall, the majority of FV's of class 'match' derive significantly large LR values.

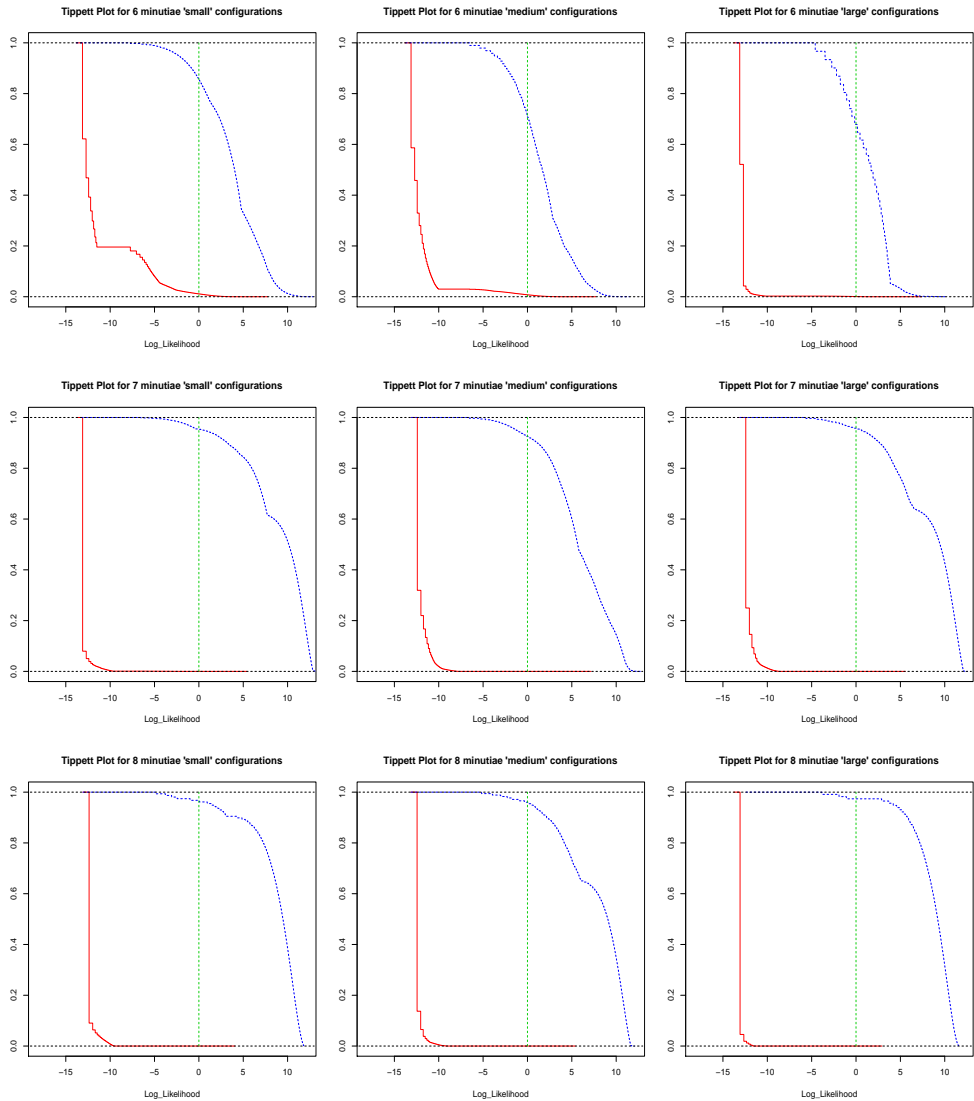


Figure 6. Tippett plots for minutiae configurations of 6 (top row), 7 (middle row), and 8 (bottom row) minutiae with small, medium, and large area categories (left to right, respectively), calculated from $P(x_i \text{ is a match} | f(x_i))$ and $P(x_i \text{ is a close non-match} | f(x_i))$ distributions. The x -axis represents the logarithm (base 2) of the LR values in equation (49) for match (blue line) and close non-match (red line) populations, while the y -axis represents proportion of such values being greater than x . The green vertical dotted line at $x = 0$ signifies a marker for $LR = 1$ (i.e., $x = \log_2 1 = 0$).

4. Summary

A new FV based LR model using morphometric and spatial analysis (MSA) with SVMs, while focusing on candidate list results of AFIS, has been proposed. This is the first LR model known to the authors that use machine learning as a core component to learn spatial feature relationships of close non-match and match populations. For robust applications for fingerprint-to-exemplar comparisons, only minutiae triplet information were used to train the SVMs. Experimental results illustrate the effectiveness of the proposed method in distinguishing match and close non-match configurations.

The proposed model is a preliminary proposal and is not focused on evidential value for judicial purposes. However, minor modifications can potentially allow the model to also be used for evidential assessments. For future research, we hope to evaluate the model with commercial AFIS environments containing a large set of exemplars.

Author details

Joshua Abraham^{1,*}, Paul Kwan², Christophe Champod³,
Chris Lennard⁴ and Claude Roux¹

* Address all correspondence to: joshua.abraham@uts.edu.au

1 Centre for Forensic Science, University of Technology, Sydney, Australia

2 School of Science and Technology, University of New England, Australia

3 Institute of Forensic Science, University of Lausanne, Switzerland

4 National Centre for Forensic Studies, University of Canberra, Australia

References

- [1] J.C. Yang (2008), D.S. Park. A Fingerprint Verification Algorithm Using Tessellated Invariant Moment Features, *Neurocomputing*, Vol. 71, No. 10-12, pages 1939-1946.
- [2] J.C. Yang (2011). Non-minutiae based fingerprint descriptor, in *Biometrics, Jucheng Yang (Ed.)*, ISBN: 978-953-307-618-8, *InTech*, pages 79-98.
- [3] J. Abraham (2011), P. Kwan, J. Gao. Fingerprint Matching using a Hybrid Shape and Orientation Descriptor, in *State of the art in Biometrics, Jucheng Yang and Loris Nanni (Eds.)*, ISBN: 978-953-307-489-4, *InTech*, pages 25-56.
- [4] J. Abraham (2012), C. Champod, C. Lennard, C. Roux. Spatial Analysis of Corresponding Fingerprint Features from Match and Close Non-Match Populations, *Forensic Science International*, DOI: 10.1016/j.forsciint.2012.10.034
- [5] C. Champod (2004), C. J. Lennard, P. Margot, M. Stoilovic. Fingerprints and Other Ridge Skin Impressions, *CRC Press*. 2004.
- [6] C. Champod (2000). Fingerprints (Dactyloscopy): Standard of Proof, in *Encyclopedia of Forensic Sciences*, J. Siegel, P. Saukko and G. Knupfer (Eds.), London: Academic Press, pages 884-890.

- [7] C. Champod (2009). Friction Ridge Examination (Fingerprints): Interpretation Of, in *Wiley Encyclopedia of Forensic Science (Vol. 3)*, A. Moenssens and A. Jamieson (Eds.), Chichester, UK: John Wiley & Sons, pages 1277-1282.
- [8] C. Champod (1995). Edmond Locard–numerical standards and “probable” identifications, *J. Forensic Ident.*, Vol. 45, pages 136-163.
- [9] J. Polski (2011), R. Smith, R. Garrett, et.al. The Report of the International Association for Identification, Standardization II Committee, Grant no. 2006-DN-BX-K249 awarded by the U.S. Department of Justice, Washington, DC, March 2011.
- [10] M. Saks (2010). Forensic identification: From a faith-based “Science” to a scientific science, *Forensic Science International Vol. 201*, pages 14-17.
- [11] S. A. Cole (2008). The ‘Opinionization’ of Fingerprint Evidence, *BioSocieties*, Vol. 3, pages 105-113.
- [12] J. J. Koehler (2010), M. J. Saks. Individualization Claims in Forensic Science: Still Unwarranted, *75 Brook. L. Rev.* 1187-1208.
- [13] L. Haber (2008), R. N. Haber. Scientific validation of fingerprint evidence under Daubert Law, *Probability and Risk Vol. 7, No. 2*, pages 87-109.
- [14] S. A. Cole (2007). Toward Evidence-Based Evidence: Supporting Forensic Knowledge Claims in the Post-Daubert Era, *Tulsa Law Review*, Vol 43, pages 263-283.
- [15] S. A. Cole (2009). Forensics without Uniqueness, Conclusions without Individualization: The New Epistemology of Forensic Identification, *Law Probability and Risk*, Vol. 8, pages 233-255.
- [16] I. E. Dror (2010), S. A. Cole. The Vision in ‘Blind’ Justice: Expert Perception, Judgement, and Visual Cognition in Forensic Pattern Recognition, *Psychonomic Bulletin & Review*, Vol. 17, pages 161-167.
- [17] M. Page (2011), J. Taylor, M. Blenkin. Forensic Identification Science Evidence Since Daubert: Part I-A Quantitative Analysis of the Exclusion of Forensic Identification Science Evidence, *Journal of Forensic Sciences*, Vol. 56, No. 5, pages 1180-1184.
- [18] *Daubert v. Merrel Dow Pharmaceuticals* (1993), 113 S. Ct. 2786.
- [19] G. Langenburg (2009), C. Champod, P. Wertheim. Testing for Potential Contextual Bias Effects During the Verification Stage of the ACE-V Methodology when Conducting Fingerprint Comparisons, *Journal of Forensic Sciences*, Vol. 54, No. 3, pages 571-582.
- [20] L. J. Hall (2008), E. Player. Will the introduction of an emotional context affect fingerprint analysis and decision-making?, *Forensic Science International*, Vol. 181, pages 36-39.

- [21] B. Schiffer (2007), C. Champod. The potential (negative) influence of observational biases at the analysis stage of fingermark individualisation, *Forensic Science International*, Vol. 167, pages 116-120.
- [22] I. E. Dror (2011), C. Champod, G. Langenburg, D. Charlton, H. Hunt. Cognitive issues in fingerprint analysis: Inter- and intra-expert consistency and the effect of a 'target' comparison, *Forensic Science International*, Vol. 208, pages 10-17.
- [23] J.J. Koehler (2008). Fingerprint Error Rates and Proficiency Tests: What They are and Why They Matter. *Hastings Law Journal*, Vol. 59, No. 5, pages 1077.
- [24] M. J. Saks (2008), J. J. Koehler. The Individualization Fallacy in Forensic Science Evidence, *Vanderbilt Law Rev.*, Vol 61, pages 199-219.
- [25] S. A. Cole (2006). Is Fingerprint Identification Valid? Rhetorics of Reliability in Fingerprint Proponents' Discourse, *Law & Policy*, pages 109-135.
- [26] D. H. Kaye (2010). Probability, Individualization, and Uniqueness in Forensic Science Evidence: Listening to the Academies. *Brooklyn Law Review*, Vol. 75, No. 4, pages 1163-1185.
- [27] IAI (2010) Resolution 2010-18, *International Association For Identification*, <http://www.theiai.org>
- [28] C. Champod (2001), I. W. Evett. A Probabilistic Approach to Fingerprint Evidence, *J. Forensic Ident.*, Vol. 51, No. 2, pages 101-122.
- [29] A. Collins (1994), N. E. Morton. Likelihood ratios for DNA identification, *Proc Natl Acad Sci U S A.*, Vol. 91, No. 13, pages 6007-6011.
- [30] C. Champod (2001), I. W. Evett, B. Kuchler. Earmarks as evidence: a critical review, *J Forensic Sci.*, Vol. 46, No. 6, pages 1275-1284.
- [31] G. Zadora (2009). Evaluation of evidence value of glass fragments by likelihood ratio and Bayesian Network approaches, *Anal Chim Acta.*, Vol. 642, No. 1-2, pages 279-290.
- [32] D. A. Stoney (1985). Quantitative Assessment of Fingerprint Individuality, *D. Crim. Dissertation*, University of California, Graduate Division of the University of California.
- [33] C. Neumann (2006), C. Champod, R. Puch-Solis, N. Egli, A. Anthonioz, D. Meuwly, A. Bromage-Griffiths. Computation of Likelihood Ratios in Fingerprint Identification for Configurations of Three Minutiae, *Journal of Forensic Sciences*, Vol. 51, No. 6, pages 1255-1266.
- [34] C. Neumann (2007), C. Champod, R. Puch-Solis, N. Egli, A. Anthonioz. Computation of likelihood ratios in fingerprint identification for configurations of any number of minutiae, *Journal of Forensic Sciences*, Vol. 52, No. 1, pages 54-64.

- [35] C. Neumann (2012), C. Champod, R. Puch-Solis, N. Egli, Quantifying the weight of evidence from a forensic fingerprint comparison: a new paradigm, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, Vol. 175, No. 2, pages 371-415.
- [36] N. M. Egli (2007), C. Champod, P. Margot. Evidence evaluation in fingerprint comparison and automated fingerprint identification systems—modelling within finger variability, *Forensic Science International*, Vol. 167, No. 2-3, pages 189-195.
- [37] N. M. Egli (2009). Interpretation of Partial Fingermarks Using an Automated Fingerprint Identification System, *PhD Thesis, University of Lausanne*
- [38] Heeseung Choi (2011), A. Nagar, A. K. Jain. On the Evidential Value of Fingerprints, in *Biometrics (IJCB)*, 2011 *International Joint Conference on*, pages 1-8.
- [39] G. Fasano (1987), A. Franceschini. A multidimensional version of the Kolmogorov-Smirnov test, *Mon. Not. R. astr. Soc.*, Vol. 225, pages 155-170.
- [40] F. Bookstein (1989). Principal Warps: Thin-Plate Splines and the Decomposition of Deformations., *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 11, No. 6, pages 567-585.
- [41] I. Dryden (1998) and K. Mardia. Statistical Shape Analysis, *John Wiley & Sons*.
- [42] J. Platt (1999). Probabilistic outputs for support vector machines and comparison to regularized likelihood methods, In: A. Smola, P. Bartlett, B. Scholkopf, and D. Schuurmans (Eds.): *Advances in Large Margin Classifiers*. Cambridge, MA, MIT Press.
- [43] C. Kelley (1999). Iterative Methods for Optimization, *SIAM Frontiers in Applied Mathematics*, No. 18.
- [44] H. Moon (2001), P. Phillips. Computational and performance aspects of pca-based face-recognition algorithms, In *Perception*, Vol. 30, No. 3, pages 303-321.
- [45] C. I. Watson (1993), NIST Special Database 14, NIST Mated Fingerprint Card Pairs 2 (MFCP2), <http://www.nist.gov/srd/nist14.cfm>
- [46] C. I. Watson (1992), NIST Special Database 4, NIST 8-bit Gray Scale Images of Fingerprint Image Groups (FIGS), <http://www.nist.gov/srd/nist4.cfm>
- [47] D. Maio (2002), D. Maltoni, R. Cappelli, J.L. Wayman, A. K. Jain. FVC2002: Second Fingerprint Verification Competition, In *Proceedings of 16th International Conference on Pattern Recognition (ICPR2002)*, pages 811-814.
- [48] M. D. Garris (2000), R. M. McCabe, NIST Special Database 27, Fingerprint Minutiae from Latent and Matching Tenprint Images, <http://www.nist.gov/srd/nist27.cfm>
- [49] NIST Biometric Image Software (2012), <http://www.nist.gov/itl/iad/ig/nbis.cfm>

- [50] M. Indovina (2012), V. Dvornychenko, R. A. Hicklin, G. I. Kiebusinski, ELFT-EFS Evaluation of Latent Fingerprint Technologies: Extended Feature Sets [Evaluation #2], *NISTIR 7859*, <http://dx.doi.org/10.6028/NIST.IR.7859>
- [51] D. Maio (2004), D. Maltoni, R. Cappelli, J. L. Wayman, A. K. Jain, FVC2004: Third Fingerprint Verification Competition, *Proc. International Conference on Biometric Authentication (ICBA)*, pages 1-7.