What can a single minutia tell about gender?

Philipp Terhörst\textsuperscript{1}, Naser Damer\textsuperscript{1}, Andreas Braun\textsuperscript{1}, Arjan Kuijper\textsuperscript{2}

\textsuperscript{1}Fraunhofer Institute for Computer Graphics Research IGD, Darmstadt, Germany
\textsuperscript{2}Mathematical and Applied Visual Computing Group, TU Darmstadt, Darmstadt, Germany

Email: \{philipp.terhoerst, naser.damer, andreas.braun, arjan.kuijper\}@igd.fraunhofer.de

Abstract—Since fingerprints are one of the most widely deployed biometrics, several applications can benefit from an accurate fingerprint gender estimation. Previous work mainly tackled the task of gender estimation based on complete fingerprints. However, partial fingerprint captures are frequently occurring in many applications including forensics and consumer electronics, with the considered ratio of the fingerprint is variable. Therefore, this work investigates gender estimation on a small, detectable, and well-defined partition of a fingerprint. It investigates gender estimation on the level of a single minutia. Working on this level, we propose a feature extraction process that is able to deal with the rotation and translation invariance problems of fingerprints. This is evaluated on a publicly available database and with five different binary classifiers. As a result, the information of a single minutia achieves a comparable accuracy on the gender classification task as previous work using quarters of aligned fingerprints with an average of more than 25 minutiae.

I. INTRODUCTION

Gender determination is a fundamental task of human beings due to the fact that many social interactions are based on gender [4]. Therefore, it is not surprising that gender is also one of the most widely deployed soft-biometrics [23]. In addition, fingerprints are the biggest commonly utilized biometric modality [18] and a well-operating fingerprint gender classifier could be used for a wide range of applications. It can contribute to context-based indexing [5], [6], surveillance [27], human-computer interactions [16] or simply to establish a person’s identity with a high degree of reliability [10]. Furthermore, it can affect forensic investigations and the rapidly growing field of consumer electronics like mobile devices. Concerning gender estimation from fingerprints, most state-of-the-art approaches focus on a ridge analysis in the spatial domain using the whole fingerprint image [20]. Consequently, these methods suffer when it comes to partial fingerprints as recent work has shown [28]. However, evaluation of partial fingerprints is a frequently occurring problem. In forensic investigations [11], latent fingerprints typically contain ridge information from partial fingerprints of always different sizes. Additionally, a number of consumer electronic devices, such as smartphones, are beginning to incorporate fingerprint sensors on only a partial area of the fingerprint due to the limited size of their sensors [8], [29]. In order to narrow the gap in research and enable these applications to use gender information from partial fingerprints, this work analyses how well the gender of a fingerprint can be estimated on well-defined level of a single minutia. For this task, a feature extraction process is proposed. This process captures the pixel values in a square-shaped region around each minutia and transforms them into feature vectors. The regions are chosen according to the location and direction of each minutia and therefore, the obtained features are robust against the rotation and translation invariance problems of fingerprints. The proposed approach was evaluated on the public available NIST Special Database 4 [9]. Moreover, five different binary classifiers were analysed on different region sizes and minutia qualities. The evaluation results point out that the proposed flexible methodology, based on a single minutia, can achieve more than 62% accuracy on the correct gender classification decision. This performance is comparable to the 54.0% - 62.9% accuracy that previous work [28] reported on quarter of aligned fingerprints consisting of an average of more than 25 minutiae.

II. RELATED WORK

Due to its importance in e.g. criminal investigations, in which fingerprint gender estimation may minimize the list of subjects, extensive research has been done. This research took place in the area of biometrics, and related fields such as anthropology, forensics, and medicine.

Most of the studies assumed that ridge configurations are controlled by genetics. In [2], the influence of genes on dermatoglyphic development was studied and a correlation between the total ridges per finger and the sex chromosomes was detected. The same correlation was found by Jantz et al. [12], [13], which lead to their assumption that the Y-chromosome may play a role in dermal ridge development. Several further studies observed a difference in ridge density, which is defined as the number of ridges that occurs in a certain area [25]. They reported a higher ridge density for females compared to males, due to finer epidermal details. This hypothesis was also supported by Gungadin et al. [30]. They found a limit of $\leq 13$ ridges/25mm\(^2\). Ridge densities below this value are more likely to have a male origin while for higher values it is more probable to come from a female origin. Similar results were confirmed from several other studies on different population groups [17], [21], [24].

The biological and anthropological studies inspired further experiments in the field of biometrics. In 1999, Acree demonstrated this trend by manually counting ridges in a well-defined area [1]. The mean epidermal ridge breadth (MRB) [15] was used to identify the gender and results in males having a 9% higher MRB value than females. In [3], [22], the gender of fingerprints was estimated on basis of the counting features, manually extracted as proposed by Acree. The manual extraction of five ridge feature (ridge count, ridge
thickness to valley thickness ratio, white lines count, pattern type concordance and ridge count asymmetry) was also used by Badawi et al. [3] who used a neural network approach on a set of 1100 female and 1100 male fingerprint images and reported an accuracy of 88.8%.

More recently, the problem of manual extracted features was tackled, and approaches on automated feature extraction were proposed. Most of these approaches followed ridge analysis in the spatial domain [14]. A method based on discrete wavelet transformation (DWT) and singular value decomposition (SVD) has been proposed by Gnanasivam et al. [7]. By decomposing different frequency ranges, a gender classification rate of 88.28% was reported. Another approach based on DWT was proposed in [26]. Here, a contrast limited adaptive histogram equalization was applied as preprocessing before training a neural network. This lead to an overall accuracy of 96%. However, 7.5% of the images were kept out and test/train split was not performed on identity level. Using frequency and spatial domain filtering, Marasco et al. [19] proposed an approach on the basis of Local Binary Pattern (LBP) and Local Phase Quantization (LPQ) operators. They reported an accuracy of 88.7% in estimating gender of fingerprints. Later on, however, Rattani et al. demonstrated in [28] that these texture descriptors (LBP & LPQ) struggle when it comes to partial fingerprints. More precisely, they observed a significant drop in accuracy down to 54.5% (LBP) and 62.9% (LPQ) when just a quarter of the fingerprint image is used. They also reported 70.5% accuracy using Binary Statistical Image Features (BSIF) on the same partial fingerprint images. However, this performance was only reported in terms of an overall classification rate. In their joint investigation of multiple fingerprints with BSIF and support vector machines, they reported a 22% higher accuracy for fingerprints of males origin than from female origins. In their gender classification experiments across different fingers, this gap grows up to 40%. Since they tested on a dataset with five times more males than female captures, this metric might be biased.

Previous approaches have not considered a gender estimation on variable portions of fingerprints, which builds the main motivation behind this work.

### III. Methodology

In order to set up gender classification algorithms which can intrinsically deal with partial fingerprints, this work offers a process to extract features from a small area around the location of a single minutia. Utilizing the area around a single minutia allows to use aligned features and therefore, tackle the rotation and translation invariance problems, which most fingerprint algorithms suffer.

The process pipeline consists of four steps: first, the minutia information is captured from the fingerprint image. Second, this information will be used to extract minutia region features. Third, these features will be normalised and finally used to make a gender decision obtained by a binary classifier.

**Minutiae extraction** - For the minutiae extraction, the NIST Fingerprint Image Software (NFIS) MINDTCT [31] was utilized, because it provides all the minutia information needed for the feature extraction process. More precisely, it determines the location $\vec{r}_m$, direction $\theta_m$, reliability $rel_m$ and type of each minutia $m$ in the fingerprint image $X$. In the case of the minutia type, it distinguishes between ridge endings (RIG) and bifurcations (BIF). The reliability $rel_m$ is based on the location of the minutia point in a quality map and the pixel intensity statistics within the immediate neighbourhood of the minutia location [31]. It describes the quality of the minutia $m$ and is defined on a range between 0 (low quality) and 1 (high quality).

**Feature extraction** - In the feature extraction process, the extracted minutiae information is used to extract the minutia region features from the grey-scale fingerprint image $X$.

For each minutia $m$, represented as a tuple $\{x_m, y_m, \theta_m, rel_m\}$, the region of size $s \times s$ pixels around its location $\vec{r}_m = (x_m, y_m)^T$ is extracted as follows. First, the image $X$ is rotated counter clockwise in a right-handed Cartesian coordinate system by angle $\theta$ around the center $\bar{c}$ such that the direction of the minutia shows upwards (directional alignment). Second, the minutia location

$$\vec{r}_m = \left( x'_m(\theta) y'_m(\theta) \right) = R_{cc}(\theta) \cdot [\vec{r}_m - \bar{c}] + \bar{c}$$

(1)

in the rotated image $X'_{rot(\theta),m}$ is calculated where $R_{cc}(\theta)$ is a counter clockwise rotation matrix around angle $\theta$. Next, the grey-scale values in $X'_{rot(\theta),m}$ are extracted as features in a square-shaped minutia region. Therefore, the positions $i'_m(k, \theta)$ and $j'_m(k, \theta)$ in the rotated region are calculated

$$i'_m(k, \theta) = y_m(\theta) + \lfloor k/s \rfloor - \lfloor s/2 \rfloor$$

(2)

$$j'_m(k, \theta) = x'_m(\theta) + k \mod s - \lfloor s/2 \rfloor$$

(3)

for $k \in [0, s^2 - 1]$. Here, $\lfloor \cdot \rfloor$ and $\mod$ describes the floor function and the modulo operator. Then, a feature vector $\vec{f}_m$ for minutia $m$ is determined. Its $k^{th}$ component is given by

$$f_k^m = \left(X'_{rot(\theta),m}\right)_{i'_m(k,\theta),j'_m(k,\theta)}$$

(4)

**Normalisation** - To make the scale of the features comparable across different minutia regions, a z-score normalisation

$$z^m_k = \frac{f_k^m - \mu_k}{\sigma_k}$$

(5)

is applied on each feature of the minutia regions. The values $\mu_k$ and $\sigma_k$ represent the mean and the variance of the $k^{th}$ components of the training features.

**Classification** - For estimating the gender of the fingerprint image, a binary classifier is trained on the minutia region features $\vec{f}_m$ from the training minutiae to predict whether the fingerprint is from female or male origin.

### IV. Experimental setup

**Database** - In order to investigate the gender information of a single minutia, a fingerprint database with special properties is needed. First, the database has to contain gender information of the fingerprint images. Second, it has to resemble...
fingertips that can occur in forensics scenarios. Therefore, the public available NIST Special Database 4 [9] was chosen for the experiments. It consists of 4000 8-bit grey scale rolled fingerprint images for 10 fingers captured on left and right hands. For each of the 2000 subjects one finger was randomly chosen and captured in two sessions. The images split of in 375 pairs of female and 1625 pairs of male captures, which are roughly equally distributed over the different fingers. The resolution per 512x512 pixel image is 19.7 pixels per millimeter. On average, there were around 63 bifurcations (BIF) and 66 ridge endings (RIG) per fingerprint in the database with a reliability $\geq 0$. For a reliability $\geq 0.5$, a fingerprint contains on average 22 BIF and 22 RIG. Therefore, more than half a million minutia regions can be used for this analysis.

Evaluation metric - Previous work [28] proposed metrics for the performance of gender estimators known as correct overall classification rate (COCR), correct female classification rate (CFCR) and correct male classification rate (CMCR). While COCR describes the percentage of test samples whose gender was correctly classified, CFCR/CMCR specify the percentage of female/male captures correctly classified as female/male. Since this work has to deal with unbalanced classes (only 18.75% of the captures are female), we report the COCR of a classifier at an operating point (score threshold) in which the CFCR equals the CMCR. Otherwise, the COCR will be significantly biased from the majority class.

Investigations - In order to analyse how well a single minutia can be used to classify gender, several experiment settings have been investigated. At the beginning, the performance dependency of the minutia region size $s \times s$ (in pixels) has been analysed. This aims at understanding how much information is needed to achieve a certain gender estimation performance. Beside the amount of information, the quality of the information plays a major role for the classification process. Therefore, the importance of the reliability $rel$ is analysed. Different reliability thresholds $rel_{th}$ are defined and only minutiae with a reliability above these thresholds are integrated into the workflow. Next, the minutia region size and its reliability are investigated together, since the amount of information stored in such a region and its quality can jointly affect the performance. Alignments often allow to capture patterns more easily. Thus, the effect of the directional alignment was analysed and finally, five different binary classifiers, namely AdaBoost, logistic regression, random forest, $k$-nearest neighbours and support vector machines, are compared in two working scenarios.

Workflow details - For each of this experimental setups, the workflow described in Section III is applied. It consists of the minutiae information extraction, the minutia region extraction, the normalisation and classification process. For the experiments, each result, reported in COCR, was produced using 10 fold cross-validation using all fingers. In each fold, the minutia regions from 20 randomly chosen identities were utilized for training and the remaining for testing. To find descent hyperparameters for the binary classifiers, 50 steps of Bayesian optimization were applied for each minutia type, reliability, and region size combination.

V. Results

In order to investigate the dependency of the minutia region size $s \times s$ and therefore, how much minutia information is needed to achieve a certain performance, Figure 1 shows the performance for two reliability thresholds and both minutia types, namely ridge ending (RIG) and bifurcation (BIF). The solid line shows the performance in terms of the COCR, while the shaded area describes its standard deviation. In all cases, it can be observed that for a region size of 2 x 2 pixels the COCR is close to a random behaviour of 50% and increases for larger size values, because a total of 4 pixels are a big information bottleneck and adding additional information helps the classifier to find a reliable decision. Around a region size of 15 x 15 pixels an saturation in accuracy (COCR) can be seen. In addition, the standard deviation is constantly the smallest for logistic regression, followed by random forest, while for AdaBoost the performance and its deviation varies a lot. The best performance can constantly be observed by the random forest classifier, which is able to robustly deal with the noise prone data. Generally, the overall performance is higher for a reliability $rel \geq 0.5$. Regardless of the reliability, the region features around ridge endings (RIG) achieve a better performance than bifurcations (BIF). Further, it can be seen that even for a small region size of 6 x 6 pixels a COCR of above 58% is achieved. This shows that a single minutia is able to estimate the gender of a fingerprint.

In Figure 2 the performance dependent on the reliability was analysed for a fixed region size of 10 x 10 pixels. As before, the same ordering of the performances and its standard deviation can be observed independent of the used reliability threshold. For RIG, it can further been seen that increasing the reliability threshold also increase the COCR by up to 2%, while for BIF none such a trend can be observed.

A joint analysis of the quantity (region size) and the quality (reliability) of the minutia information over the COCR is shown in Figure 3. Here, a blue value indicates a poor performance while a red value refers to a good one. It can clearly be seen that in all cases the region size strongly affects the overall performance. As observed before, for a very small region size of 2 x 2 pixels the COCR is close to 50% and quickly rises for increasing region sizes. For logistic regression, a higher reliability threshold positively affects the performance. Probably the simplicity of the algorithm favours minutia regions with a higher reliability due to less noise. On the other hand, random forest is able to robustly find more complex and more noise-prone pattern if enough information is available. Therefore, its performance is only weakly affected by the reliability and shows a constant high COCR above a region size bigger than 8 x 8 to 10 x 10 pixels.

To study the effect of the used directional alignment, in Figure 4 the COCR is shown over the region size with and without this alignment. Without the directional alignment the minutia region is extracted from the non-rotated fingerprint image ($\theta_{mn}$,
Fig. 1: Analysis of the region height/width $s$: COCR is plotted over the region height/width $s$ for two reliability thresholds $\text{rel}_{th}$ and both minutia types. The standard deviation is shown as the shaded area.

Fig. 2: Analysis of the reliability dependence: COCR is plotted over the different reliability thresholds $\text{rel}_{th}$ for a minutia region size 10x10 pixels and both minutia types. The standard deviation is shown as the shaded area.
(a) Logistic regression - RIG

(b) Logistic regression - BIF

(c) Random forest - RIG

(d) Random forest - BIF

Fig. 3: Plots of the parameter space: COCR is reported over different minutia region sizes and reliability thresholds $\text{rel}_\text{th}$ for two classification algorithms over RIG and BIF.

$= 0 \forall m)$. It can be seen that aligning the region according to the minutia direction moderately helps the classifiers to learn the pattern and achieve higher COCR. However, this effect is probably attenuated by the fact that MINDTCT only returns the direction of a minutia in steps of $11.25^\circ$.

For a more complete overview Table I presents the COCR and their standard deviations for some of the tested classifiers. There, the COCR is reported for two region sizes (10 x 10 and 20 x 20 pixels), two reliability thresholds ($\text{rel}_\text{th}$=0.0, 0.5) and over both minutia types (RIG and BIF). Five different binary classifier were evaluated: AdaBoost (Ada), logistic regression (LogReg), random forest (RF), k-nearest neighbour (KNN) and support vector machines (SVM). In all scenarios it can be observed that random forest outperforms all other classifiers. Furthermore, RIG generally performs significantly better than BIF. This can be explained by the fact that most gender information is stored in the ridge breadth, which is harder to determine when a ridge splits up in two as in the case of bifurcations. For BIF random forest achieved a COCR of $(61.9 \pm 0.7)$% and for RIG a COCR of $(62.8 \pm 0.5)$%.

Fig. 4: Performance with and without directional alignment for $\text{rel}_\text{th} = 0$ (RIG). The solid line shows the performance with alignment, while the dashed one presents the performance without.
was reached at a region size 20 x 20 pixels and \( rel_{th} = 0.5 \). These numbers achieved by a small minutia region are still comparable to COCR reported by other recent works, however on other databases. In [28], 65.8% to 71.7% COCR are reported on full fingerprints (containing on average more than 100 minutiae) and on one-fourth portion of the original fingerprints (containing on average more than 25 minutiae) 54.0% to 62.9% COCR are reported.

For the sake of completeness, for a region size of 30 x 30 pixels and a reliability threshold of \( rel_{th} = 0.5 \) the classifier is able to achieve a COCR of \( (63.2 \pm 0.4)\% \) on RIG. However, in this database the average distance to the nearest minutia is about 8-9 pixels (for \( rel_{th} = 0 \)) and therefore, this size can be seen to contain information beyond the minutia region.

VI. CONCLUSION

Fingerprint gender estimation can be used in a wide range of applications. However, applications in forensics or consumer electronics require dealing with partial fingerprints and this is a task which current methods can’t deal well with. This work investigates fingerprint gender estimation on the level of a single minutia. It proposes a flexible methodology to create minutia region features, which are robust to the rotation and translation invariance problems of fingerprints. The experiment results point out that ridge endings reach higher gender classification rates than bifurcations. The proposed method showed that a single minutia can achieve a correct overall gender classification rate (COCR) of above 62%, which is comparable to previous work using quarter of fingerprints containing an average of more than 25 minutiae. Future work can utilize the proposed method as a building block to create more flexible and accurate fingerprint gender estimators.

ACKNOWLEDGEMENT

This work was supported by the German Federal Ministry of Education and Research (BMBF) as well as by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within the Center for Research in Security and Privacy (CRISP).

<table>
<thead>
<tr>
<th>Minutia</th>
<th>Size s x s</th>
<th>( rel_{th} )</th>
<th>Ada</th>
<th>LogReg</th>
<th>RF</th>
<th>KNN</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIG</td>
<td>10x10</td>
<td>0.0</td>
<td>58.7 ± 2.0</td>
<td>58.1 ± 0.3</td>
<td>60.3 ± 0.5</td>
<td>58.0 ± 2.6</td>
<td>59.3 ± 0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>57.9 ± 1.0</td>
<td>59.6 ± 0.6</td>
<td>61.2 ± 0.6</td>
<td>58.6 ± 2.6</td>
<td>58.8 ± 0.6</td>
</tr>
<tr>
<td></td>
<td>20x20</td>
<td>0.0</td>
<td>57.8 ± 1.3</td>
<td>59.4 ± 0.3</td>
<td>61.8 ± 0.8</td>
<td>58.7 ± 2.9</td>
<td>60.5 ± 0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>58.7 ± 1.2</td>
<td>61.4 ± 0.6</td>
<td>62.8 ± 0.5</td>
<td>60.7 ± 2.8</td>
<td>61.9 ± 0.7</td>
</tr>
<tr>
<td>BIF</td>
<td>10x10</td>
<td>0.0</td>
<td>56.1 ± 1.9</td>
<td>57.4 ± 0.3</td>
<td>58.7 ± 0.6</td>
<td>55.2 ± 2.3</td>
<td>58.3 ± 0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>56.7 ± 1.4</td>
<td>57.8 ± 0.7</td>
<td>58.8 ± 0.6</td>
<td>56.3 ± 3.8</td>
<td>58.7 ± 0.7</td>
</tr>
<tr>
<td></td>
<td>20x20</td>
<td>0.0</td>
<td>55.8 ± 1.2</td>
<td>58.3 ± 0.3</td>
<td>61.2 ± 0.6</td>
<td>58.3 ± 1.6</td>
<td>59.8 ± 0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>57.6 ± 1.5</td>
<td>60.3 ± 0.7</td>
<td>61.9 ± 0.7</td>
<td>59.0 ± 2.1</td>
<td>60.5 ± 0.5</td>
</tr>
</tbody>
</table>

TABLE I: COCR and its standard deviation reported for different minutia types, minutia region sizes s x s in pixels, reliability thresholds \( rel_{th} \) and different algorithms.

REFERENCES


