Article

Multi-Layer Feature Based Shoeprint Verification Algorithm for Camera Sensor Images

Xinnian Wang [1,*] **, Yanjun Wu** [1] **and Tao Zhang** [2]

1 School of Information Science and Technology, Dalian Maritime University, Dalian 116026, China; wuyanjun@dlmu.edu.cn
2 School of Physics and Electronics Technology, Liaoning Normal University, Dalian 116026, China; lnnuzt@lnnu.edu.cn
* Correspondence: wxn@dlmu.edu.cn

Abstract: As a kind of forensic evidence, shoeprints have been treated as important as fingerprint and DNA evidence in forensic investigations. Shoeprint verification is used to determine whether two shoeprints could, or could not, have been made by the same shoe. Successful shoeprint verification has tremendous evidentiary value, and the result can link a suspect to a crime, or even link crime scenes to each other. In forensic practice, shoeprint verification is manually performed by forensic experts; however, it is too dependent on experts' experience. This is a meaningful and challenging problem, and there are few attempts to tackle it in the literatures. In this paper, we propose a multi-layer feature-based method to conduct shoeprint verification automatically. Firstly, we extracted multi-layer features; and then conducted multi-layer feature matching and calculated the total similarity score. Finally, we drew a verification conclusion according to the total similarity score. We conducted extensive experiments to evaluate the effectiveness of the proposed method on two shoeprint datasets. Experimental results showed that the proposed method achieved good performance with an equal error rate (EER) of 3.2% on the MUES-SV1KR2R dataset and an EER of 10.9% on the MUES-SV2HS2S dataset.

Keywords: shoeprint verification; shoeprint partition model; multi-layer feature; forensic sciences; shoeprint matching strategy

1. Introduction

With the popularity of crime investigation dramas and novels, more and more people know what investigators are looking for while solving a case. This kind of knowledge has affected the investigation and solving of cases, because some criminals may conceal their crime traces after they commit crimes. In fact, if there are no other forms of criminal evidence, as the only remaining evidence at a crime scene, shoeprint may very well reveal some available clues to a particular case.

Shoeprint conveys many important human characteristics, such as walking habits and identity, and it may play a vital role in crime investigation. The typical uses of shoeprints are to link cases, to testify whether the shoeprints are left by the same shoe, then identify the suspect. It is clear that some typical forensic evidence (fingerprints and palmprints) can be successful in linking evidence to a suspect for the evidence's uniqueness. Similar with the evidence stated previously, shoeprint evidence has its own unique identifying characteristics. The identifying characteristics can be classified into three categories: class characteristics, wear characteristics, and individual identifying characteristics [1]. Class characteristics are shoe patterns that are usually created in the manufacturing process, and different shoes may have different shoe patterns. Class characteristics are usually used to recognize or retrieve shoeprints with similar shoe patterns. Wear characteristics are those changes on the

outsoles that are caused by natural erosion or wear of the shoe which can reflect personal walking habits [1]. Wear characteristics of shoes worn by different persons are different. Individual identifying characteristics refer to random tears, cuts, and punctures on the outsoles, which are produced in the process of human wearing. Individual identifying characteristics may make shoes with the same class characteristics different. These characteristics are usually applied for shoeprint verification, which is to determine whether two shoeprints could, or could not, have been made by the same shoe or person.

Most of the researches on shoeprint tend to concentrate on shoeprint classification and retrieval by using class characteristics. Rida et al. [2] provide an overview of the research carried out in forensic shoeprint retrieval. According to the scope of representation, shoeprint classification and retrieval methods may roughly fall into three categories: global features-based methods, region features-based methods, and interest point-based methods. Methods using the global features usually take into consideration the whole shoeprint for extracting features, for example, the moments invariant methods [3,4], frequency domain methods [5–10], and convolutional neural network-based method [11,12]. Methods using the region feature usually divide shoeprints into different regions and extract features from these regions, for example, the maximally stable extreme regions (MSER)-based methods [13,14], the periodic pattern-based methods [15], the compositional active basis model-based methods [16,17], and the sparse representation methods [18]. Interest point-based methods detect interest points and use these points to recognize shoeprints (e.g., [19–21]). Luostarinen et al. [22] evaluated different shoeprint retrieval methods and compared their performance under shoeprints with different quality conditions. Their results show that methods based on local interest points with RANSAC and Fourier–Mellin transform can achieve good retrieval performance under different quality conditions. Although the above methods can achieve a good performance on shoeprint retrieval, they may fail to handle shoeprint verification. Class characteristics of a shoe can provide informative information to recognize shoe patterns but are not unique enough to verify two shoeprints.

In forensic practice, shoeprint verification is manually performed by forensic experts. Forensic experts first manually label the individual identifying characteristics, then compare them in their own opinions; and finally draw a conclusion as to whether the two shoeprints could have come from the same shoe according to their own opinions. This kind of comparison method is too dependent on experts' experience, and the conclusions of different experts are not always similar.

Yekutieli et al. [23] proposed a system to compare individual identifying characteristics with a physical matching method. The individual identifying characteristics used in their method are first marked manually, and then the contour of each individual identifying characteristic is extracted for comparison. This method cannot perform shoeprint verification in an automatic manner. To the best of our knowledge, there are no works focusing on automatic shoeprint verification. Therefore, in this paper, we focus on an automatic shoeprint verification algorithm.

The main contributions of the proposed method are summarized as follows:

(i)　we propose a multi-layer feature-based shoeprint verification algorithm that can be used in forensic practice;
(ii)　we introduce a shoeprint partition model (SPM) to analyze shoeprints, which considers foot anatomy and the relationship between shoe and foot, and facilitates analyzing shoeprints accurately in practice;
(iii)　we propose an individual identifying characteristics detection method to perform characteristics detection automatically;
(iv)　we propose a shoeprint image matching strategy. The shoeprint is divided into nineteen sections. Similarities of each section are computed respectively, and the total similarity between the two images is a weighted sum.

The rest of this paper is organized as follows: Section 2 describes the shoeprint acquisitions and datasets, Section 3 details the proposed method, Section 4 provides the performance experiments and the detailed evaluation of the proposed method and Section 5 concludes the paper.

2. Shoeprint Acquisitions and Datasets

To testify the effectiveness of the proposed methods, we build two kinds of shoeprint verification datasets named MUES-SV1KR2R and MUES-SV2HS2S. Samples from the two shoeprint image datasets are shown in Figure 1.
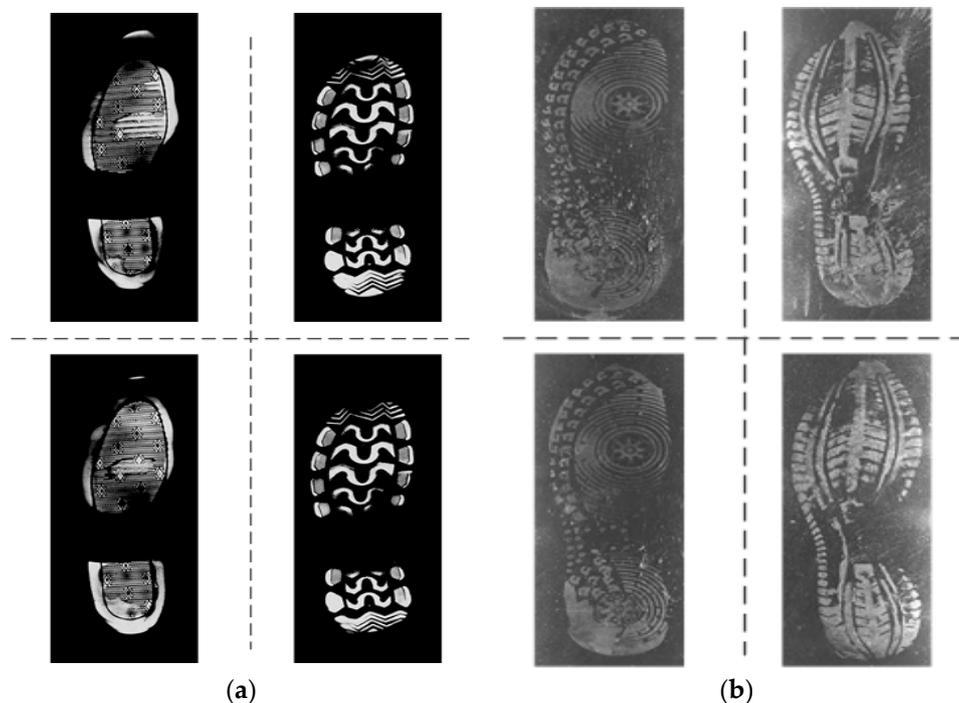

Figure 1. Sample of shoeprints in the datasets. (**a**) Samples of shoeprints in the MUES-SV1KR2R dataset; (**b**) Samples of shoeprints in the MUES-SV2HS2S dataset.

Images in MUES-SV1KR2R dataset are pressure images acquired by using the shoeprint scanners, and the imaging process is shown in Figure 2. When one steps on the reflective tape (the height of the reflective tape is about 1 mm), different height downward bulges are formed on the surface of the reflective tape because of the uneven pressures. The higher the pressure is, the higher the downward bulges are. The lights of the light source shine through the transparent glass and on the surface of these downward bulges. The lights then produce shadowing between the different heights area of those downward bulges and the lights are reflected to the camera by the mirror. Different amounts of light are captured by the camera for the different heights of bulges, and a shoeprint image can be scanned at 300 dpi. Then the obtained shoeprint images are stored in the MUES-SV1KR2R dataset. Some samples in the MUES-SV1KR2R dataset are shown in Figure 1a. The MUES-SV1KR2R dataset consists of more than 1200 pairs of shoeprint images, and each pair of shoeprints has the same shoe patterns. Of 1200 shoeprint pairs, 600 pairs of shoeprints are made by the same shoes. We refer this kind of shoeprint as a reference or suspect shoeprint. Images in the dataset are registered and have the same spatial resolution. Most images in the database are clear and full. Shoeprint images are available as 3600 by 1800 pixels.
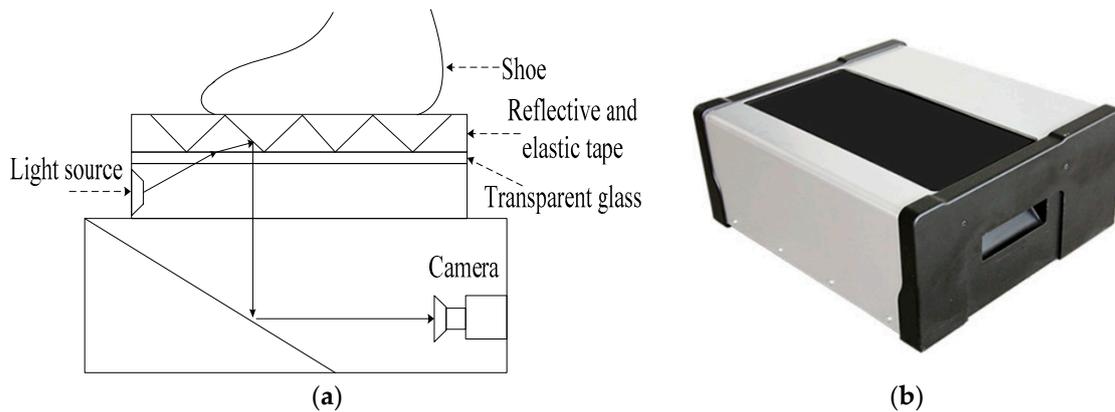
Figure 2. Illustration of the reference shoeprint scanner. (**a**) Structural diagram of the reference shoeprint scanner; (**b**) the real system of the reference shoeprint scanner.

Images in MUES-SV2HS2S are crime scene shoeprints collected by using the crime scene shoeprint scanners, and the imaging process is shown in Figure 3. For an impression left on the ground, the oblique lights from the light source shine to create shadows between the low and high areas of the impression and provide a great amount of contrast between those areas. Then the contours and contents within the impression can be revealed and captured by the camera. For impressions in sand, soil, or snow, the oblique light from the light source 1 can provide the best contrast, and for those impressions in dust or residue, the oblique light from light source 2 is most effective at a very low angle. The distance from light source 1 to the shoeprint is about 5 cm. A shoeprint image is scanned at 300 dpi according to the above process, and the shoeprint images are stored in the MUES-SV2HS2S dataset. Some samples in the MUES-SV2HS2S dataset are shown in Figure 1b. The MUES-SV2HS2S dataset consists of more than 256 pairs of shoeprints with same shoe patterns. Of 256 shoeprint pairs, 128 pairs of shoeprints are made by the same shoes. We refer this kind of shoeprint as a crime scene shoeprint. Images in this dataset are registered and have the same spatial resolution. Shoeprint images are available as 3600 by 1800 pixels.
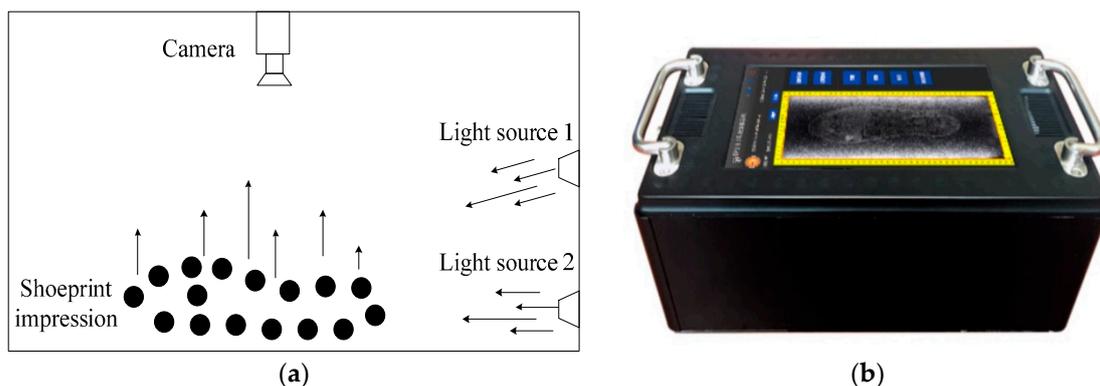
Figure 3. Illustration of the crime scene shoeprint scanner. (**a**) Structural diagram of the crime scene shoeprint scanner; (**b**) the real system of the crime scene shoeprint scanner.

3. Methods

The proposed shoeprint verification scheme consists of four basic modules: shoeprint preprocessing module, multi-layer feature extraction module, multi-layer feature matching module, and decision-making module, as shown in Figure 4. Each of these modules plays a vital role in the shoeprint verification system. The functionalities of these modules are stated as that: (1) preprocessing module aims at separating the valid shoeprints from the complex backgrounds, and conducting shoeprint registration and partition; (2) multi-layer feature extraction module aims at representing the characteristics of shoeprints on global layer, partial layer and individual identifying layer; (3) multi-layer

feature matching module compares the multi-layer features; and (4) decision-making module calculates the total similarity score, and then makes the decision that the two shoeprints are left by the same shoe or not.

| Test Shoeprint Pairs | → | Image Preprocessing | → | Multi-layer Feature Extraction | → | Feature Matching | → | Decision-making |

Figure 4. Framework of the proposed shoeprint verification system.

3.1. Shoeprint Preprocessing

This stage aimed to separate the valid shoeprints from the complex backgrounds and conducting shoeprint registration and partition. Image preprocessing has the following steps:

(1)　Shoeprint extraction: An image segmentation technique [24] is used to extract the shoeprint images from the complex backgrounds.

(2)　Image registration: In shoeprint verification applications, accurate shoeprint alignment has a determinative effect, and a FFT-based registration algorithm [25] is used to align the shoeprint images.

(3)　Shoeprint partition: A shoe partition model (SPM) is proposed to divide shoeprint image into different sections according to the structure of the foot and the relationship between shoe and foot. The SPM is to divide a shoeprint into different sections with a set of landmarks. Firstly, the contour of a shoeprint is represented by an average shape that is trained by using enough shoeprint images with various shapes. Secondly, three points (e.g., the front most point, rearmost point, and leftmost point) are marked interactively. Thirdly the other points of the contour, which are denoted with green dots as shown in Figure 5b, are estimated by using the interpolation method with the average shape and three points. Finally, the subsections are divided according to the predefined model. Each shoeprint is divided into toe section, sole section, instep section, heel section and back of heel section, and each section is further divided into several non-overlapped subsections for further analysis. The total number of subsections is 19.

(a)　　　　　　　　　　　　　(b)
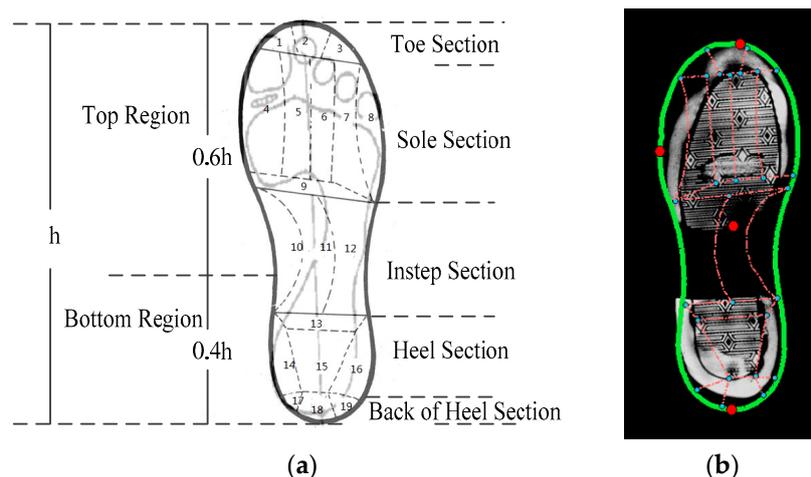
Figure 5. Illustration of the shoeprint partition model. (**a**) The layout of 19 subsections of a shoeprint; (**b**) Example of the Shoeprint Partition Model (SPM) for a full shoeprint.

3.2. Multi-Layer Feature Extraction

Most forensic experts use class, wear, and individual identifying characteristics to analyze shoeprints [1]. Class characteristics are usually caused through the manufacturing process and differ on shoe patterns. Wear characteristics are usually caused by natural erosion, and they reflect personal walking habits. Wear characteristics of shoes worn by different persons are different, even though

the shoes are of the same patterns. Individual identifying characteristics are usually attributed to something randomly added to or taken away from the outsole, and they make the shoeprints unique. We propose a multi-layer feature extraction method to represent these characteristics of shoeprints on global layer, partial layer and individual identifying layer.

3.2.1. Global Layer Feature Extraction

The global layer features that represent class characteristics are used to determine whether two shoeprints are of the same shoe patterns. Based on thousands of crime scene shoeprints, we found that patterns of the sole sections and the heel sections may be collected frequently in crime scenes, and they may determine the retrieval results in most cases. Thus, firstly, the shoeprint is divided into the bottom region and top region as shown in Figure 5a. Here, the ratio of the top region height to the bottom region height is set as 3 to 2, which is learned from training examples. Then the wavelet-Fourier–Mellin (WFM) transform coefficients of both the top and bottom regions are used to describe the class characteristics of the shoeprint [5]. For a given shoeprint image $\mathbf{u}_i$, its global layer feature extraction process has two main steps as follows.

Step 1: The two corresponding regions $\mathbf{S}_{\text{top}}(i)$ and $\mathbf{S}_{\text{bottom}}(i)$ are decomposed by using Haar wavelet at a specified number of levels. We can acquire one approximation subband and three details at each level. The coefficients can be represented as:

$$\mathbf{FW}\big(\mathbf{S}_{\text{top}}(i)\big) = \left\{ \mathbf{F}_{\text{W}}^{(\mathbf{S}_{\text{top}}(i))}(l,h,v) \middle| 0 \leq l \leq L, h,v = 0,1 \right\} \tag{1}$$

$$\mathbf{FW}(\mathbf{S}_{\text{bottom}}(i)) = \left\{ \mathbf{F}_{\text{W}}^{(\mathbf{S}_{\text{bottom}}(i))}(l,h,v) \middle| 0 \leq l \leq L, h,v = 0,1 \right\} \tag{2}$$

where L denotes the maximum level. To avoid merging the useful neighbor patterns, L should meet the criterion: $2^{L-1} \leq D_{\min}$, where $D_{\min}$ represents the minimum distance between two neighbor patterns which can be specified interactively.

Step 2: Perform Fourier–Mellin transform on each band of wavelet coefficients, and then a band passed filter proposed in [25] is used to weaken the effect of connections between patterns and noises such as small holes and broken patterns. We use $\mathbf{FMW}\big(\mathbf{S}_{\text{top}}(i)\big)$ and $\mathbf{FMW}(\mathbf{S}_{\text{bottom}}(i))$ to denote the global layer feature of the shoeprint $\mathbf{u}_i$.

3.2.2. Partial Layer Feature Extraction

The partial layer features represent wear characteristics are used to determine whether two shoeprints have the same natural erosion caused by personal walking habits. Partial layer features are extracted in different subsections. The process of partial layer feature extraction are as follows, firstly the shoeprint is divided into nineteen subsections; secondly, the minimum enclosing rectangle region of each subsection is extracted and divided into non-overlapping patches; finally, the intensity distribution of each patch is used as the partial layer feature, which is represented as $\mathbf{F}_{\text{P}}$. Figure 6 shows the process of partial layer feature extraction.
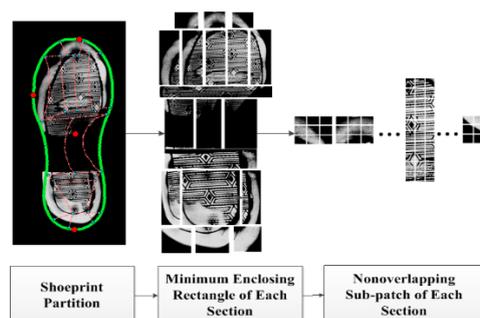

**Figure 6.** Illustration of the partial layer feature extraction method.

According to human walking habits, the importance of subsections are different, so we defined subsection weights as shown in Table 1, and subsections that likely occur wear characteristics have larger weights.

Table 1. Weight order of each subsection in a shoeprint.

Subsection Index	Section of Shoeprint	Weight Order
1, 2, 3	Toe	3
4, 5, 6, 7, 8, 9	Sole	1
13, 14, 15, 16	Heel	2
17, 18, 19	Back of Heel	4
10, 11, 12	Instep	5

3.2.3. Individual Identifying Layer Feature Extraction

Individual identifying layer feature mainly detects the individual identifying characteristics caused by the material removed from the outsole or objects stuck on the outsole. An individual identifying layer feature may be represented as an interest point, line, or region on a shoeprint image. The word 'individual' here means that an individual feature should possess some very desirable criteria including nonrepeatability, that is, the individual feature cannot be detected in a shoeprint twice or two shoeprints left by different shoes. The individual identifying characteristics on a shoeprint mainly have two kinds of structures: the corner structure and the blob-like structure. Some samples of the two structures are shown in Figure 7.
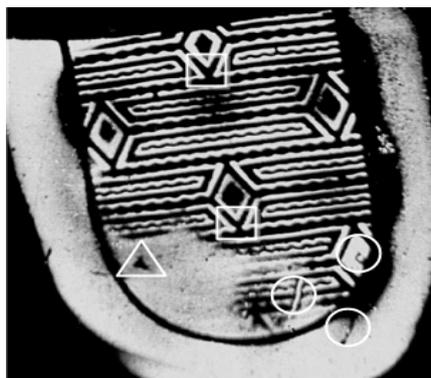
Figure 7. Examples of individual identifying characteristics. Areas labeled with circles are individual identifying characteristic with corner structure, and that labeled with triangles are individual identifying characteristics with blob-like structure. Areas labeled with rectangles are shoe patterns.

To extract individual identifying layer features, the proposed method is based on two classes of interest point detectors. The two classes of feature points are described as follows:

The first class of feature points used in the proposed method is based on Harris corners. This class of feature points is used to detect the individual identifying layer features that have a corner structure. The individual identifying characteristic with a corner structure consists of corner points that can be the intersections of two edges, and these corners usually have conventional structures, such as L-corners, T-junctions, and Y-junctions. The Harris corner detector has a strong response to the intersections of two edges, and it detects corner points by using first order derivations of an image $\mathbf{f}$ as follows:

$$\mathbf{A} = \sum_x \sum_y w(x,y) \begin{bmatrix} \mathbf{f}_x^2(x,y) & \mathbf{f}_x(x,y)\mathbf{f}_y(x,y) \\ \mathbf{f}_x(x,y)\mathbf{f}_y(x,y) & \mathbf{f}_y^2(x,y) \end{bmatrix} \tag{3}$$

where $\mathbf{f}_x$ and $\mathbf{f}_y$ denote the first order derivative of $\mathbf{f}$ with respect to the x and y direction. Let $w(x,y)$ represent the weighting function, and it is usually replaced by a Gaussian function $g(x,y,\sigma)$ of a

standard deviation σ. To detect the conventional corner structures described above, we perform the Harris corner detector on multiple scales, which detects Harris corners at different scales. A scaled image **L** with the scale parameter σ can be represented as:

$$\mathbf{L}(x, y, \sigma) = g(x, y, \sigma) \times \mathbf{f}(x, y) \tag{4}$$

The multiple scales Harris detector can be performed based on:

$$\mathbf{A}(x, y, \sigma_i, \sigma_d) = \sigma_d^2 g(x, y, \sigma_i) \begin{bmatrix} \mathbf{L}_x^2(x, y, \sigma_d) & \mathbf{L}_x(x, y, \sigma_d)\mathbf{L}_y(x, y, \sigma_d) \\ \mathbf{L}_x(x, y, \sigma_d)\mathbf{L}_y(x, y, \sigma_d) & \mathbf{L}_y^2(x, y, \sigma_d) \end{bmatrix} \tag{5}$$

where σ_i and σ_d are the integration and differentiation scale, respectively [26,27].

We can compute the corner response of a point according to the advices in [28] as follows:

$$R_1 = \lambda_1 \lambda_2 - (\lambda_1 + \lambda_2)^2 = \det(\mathbf{A}(x, y, \sigma_i, \sigma_d)) - k\text{trace}^2(\mathbf{A}(x, y, \sigma_i, \sigma_d)) \tag{6}$$

where λ_1 and λ_2 represent the eigenvalues of $\mathbf{A}(x, y, \sigma_i, \sigma_d)$. k is a parameter that can be empirically set as 0.04.

As shown in Figure 7, individual identifying characteristics with the corner structure are always described by more than one corner point. We cluster these feature points in the position space, and choose center points of each cluster as the first class of candidate individual identifying characteristic points, and these points can be denoted as $\mathbf{B}_1 = \{(x_i, y_i) | i = 1, 2, \ldots, N_1\}$ where (x_i, y_i) denotes the position and N_1 is the number of the first class of candidate individual identifying characteristic points.

The second class of feature points in the proposed method is based on Hessian detector, and a Hessian detector can be performed by using the Hessian Matrix [27]. We use the Hessian detector to detect the individual identifying layer features with the blob-like structure, because the Hessian detector has a strong response to blobs. The Hessian Matrix can be represented as follows:

$$\mathbf{A}(x, y, \sigma_d) = \begin{bmatrix} \mathbf{L}_{xx}(x, y, \sigma_d) & \mathbf{L}_{xy}(x, y, \sigma_d) \\ \mathbf{L}_{xy}(x, y, \sigma_d) & \mathbf{L}_{yy}(x, y, \sigma_d) \end{bmatrix} \tag{7}$$

We can compute the Hessian corner response of a point by using the following formula:

$$R_2 = \det(\mathbf{A}(x, y, \sigma_d)) = \mathbf{L}_{xx}\mathbf{L}_{yy} - \mathbf{L}_{xy}^2 \tag{8}$$

The second class of candidate individual identifying characteristic points are denoted as $\mathbf{B}_2 = \{(x_i, y_i) | i = 1, \ldots, N_2\}$, where (x_i, y_i) denotes the position and N_2 is the number of the second class of candidate individual identifying characteristic points.

Uniqueness is the most visible trait that is used to distinguish individual identifying characteristics from shoe patterns. For a candidate individual identifying characteristic $(x_i, y_i) \in \mathbf{B}$ and the patch centered on it, where $\mathbf{B} = \mathbf{B}_1 \cup \mathbf{B}_2 = \{(x_i, y_i) | i = 1, 2, \ldots, N\}$, and $N = N_1 + N_2$. We use the intensity distribution of each patch centered on each point of $\mathbf{B}$ to describe these points. Individual identifying characteristics are usually produced on the outsole randomly, which means that the individual identifying characteristics must be unique in a local region. In the process of individual identifying characteristic detection, periodic patterns are usually mistaken for individual identifying characteristic, as shown in Figure 8. We tell whether a point (x_i, y_i) has the same pattern as point (x_j, y_j) by calculating the normalized cross-correlation of the Fourier feature of the regions centering at these points:

$$s\left(f(\mathbf{R}_i), f(\mathbf{R}_j)\right) = \frac{\left(f(\mathbf{R}_i) - \mu_{f(\mathbf{R}_i)}\right)\left(f(\mathbf{R}_j) - \mu_{f(\mathbf{R}_j)}\right)}{\sigma_{f(\mathbf{R}_i)}\sigma_{f(\mathbf{R}_j)}} \tag{9}$$

where $f(\mathbf{R}_i)$ denote the Fourier feature of the patch centered at point (x_i, y_i). Then we delete the similar points which have the same patterns. By this way, we can get a refined individual identifying characteristic point set $\mathbf{B}_\text{p}$, and we use the intensity distribution of each patch centered on each point of $\mathbf{B}_\text{p}$ as individual identifying layer features, which is represented as $\mathbf{F}_\text{I}$.
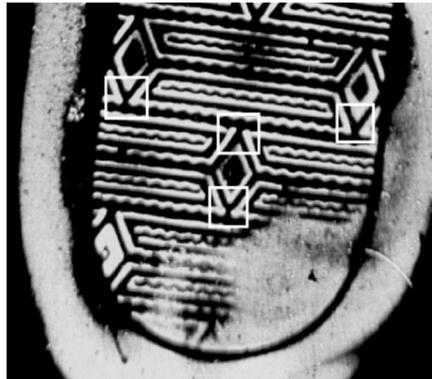

Figure 8. Periodic patterns that are labeled with white rectangles, and they are usually mistaken for individual identifying characteristic.

3.3. Multi-Layer Feature Matching

Two matching techniques are introduced in the work presented in this paper, which are the wavelet-Fourier–Mellin transform and similarity estimation (WFSE) method [5] and Ciratefi [29]. Matching using WFSE method computes the wavelet-Fourier–Mellin feature distances between the probe shoeprint and questioned shoeprint. The second matching technique matches a patch in the probe shoeprint against patches in the questioned shoeprint and computes the matching score with the best matching patch.

For the global layer feature of a pair of shoeprints $\mathbf{U}$, similarities of the top region and the bottom region are computed according to the 2D correlation coefficient, respectively, and the similarity value of class characteristics $P_\text{G}(\mathbf{U})$ is computed by their correlation coefficient of global layer features. It is a weighted sum of correlation coefficients of both $\mathbf{FMW}\big(\mathbf{S}_\text{top}(i)\big)$ and $\mathbf{FMW}(\mathbf{S}_\text{bottom}(i))$. Please refer to [5] for details.

With regard to the partial layer feature and individual identifying layer feature, we used a template matching method named Ciratefi [29] to find its best matching patch within a searching window of the questioned shoeprint. Figure 9 shows an example of the template matching method. An individual identifying layer feature $\mathbf{F}$ is extracted and encircled by the white dotted box as shown in Figure 9a, and patch $\mathbf{A}$ is an extension patch and locates at the same position in the questioned shoeprint as shown in Figure 9b, we aimed at finding the most similar patch with feature $\mathbf{F}$ in patch $\mathbf{A}$, and assigned the maximum similarity value as the matching score $S(\mathbf{F})$, which represents the similarity value of two shoeprints on feature $\mathbf{F}$. The method compares feature $\mathbf{F}$ with every patch in a sliding window manner from patch $\mathbf{A}$. The template matching method integrates three cascaded filters to yield a satisfactory result. The three filters are circular sampling filter, radial sampling filter, and template matching filter. The former two filters are used to eliminate candidate patches that have no opportunity to match with the probe template $\mathbf{F}$, and the last filter is utilized to find the best matching patch and compute the matching score.
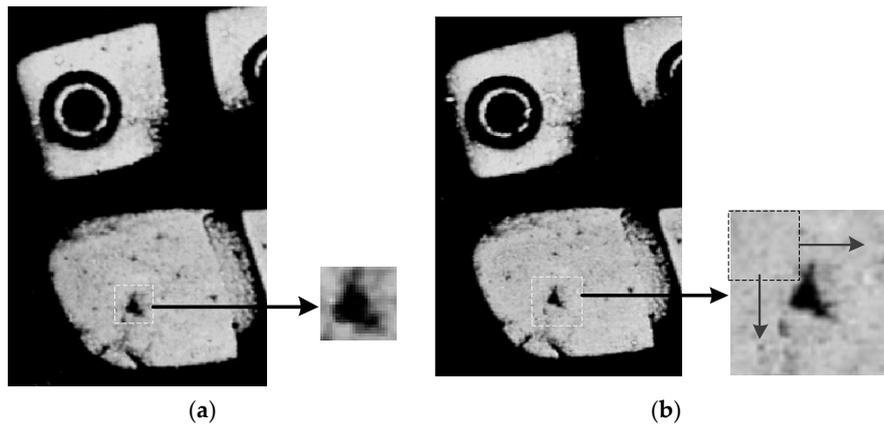
Figure 9. Illustration of template matching method. (**a**) The probe patch; (**b**) The extension patch in the questioned shoeprint.

The process of the template matching method is detailed as follows. A circular sampling filter projects the images **A** and **F** on a set of circular rings, and uses these projections to detect the first level candidate patches and their corresponding matching scales. Given the template **F**, a scaled image $\mathbf{F}_i$ with the scale parameter $\sigma_i (i = 1, 2, \ldots, n)$ can be represented by using Equation (4). Then, each scaled image $\mathbf{F}_i$ is circularly sampled in a set of circle rings situated at radius $r_k (k = 1, 2, \ldots, l)$ from the center pixel (x, y):

$$\mathbf{C}_{\mathrm{F}}(i, k) = \frac{1}{2\pi r_k} \sum_{\theta=0}^{2\pi} \mathbf{F}_i (x + r_k \cos\theta, y + r_k \sin\theta) \tag{10}$$

With regard to the image **A**, the circular projections for each (x_0, y_0) in **A** can be denoted as:

$$\mathbf{C}_{\mathrm{A}}(x_0, y_0, k) = \frac{1}{2\pi r_k} \sum_{\theta=0}^{2\pi} \mathbf{A}(x_0 + r_k \cos\theta, y_0 + r_k \sin\theta), (x_0, y_0) \in \mathbf{A} \tag{11}$$

The correlation coefficient of $\mathbf{C}_{\mathrm{F}}$ and $\mathbf{C}_{\mathrm{A}}$ is used to detect the best matching scale for each pixel (x_0, y_0) in **A**, which can be represented as:

$$X(x_0, y_0) = \max \left(\frac{(\mathbf{C}_{\mathrm{F}}(i) - \overline{\mathbf{C}_{\mathrm{F}}(i)})(\mathbf{C}_{\mathrm{A}}(x_0, y_0) - \overline{\mathbf{C}_{\mathrm{A}}(x_0, y_0)})}{\|\mathbf{C}_{\mathrm{F}}(i) - \overline{\mathbf{C}_{\mathrm{F}}(i)}\| \|(\mathbf{C}_{\mathrm{A}}(x_0, y_0) - \overline{\mathbf{C}_{\mathrm{A}}(x_0, y_0)})\|} \right) \tag{12}$$

A candidate pixel (x_0, y_0) in **A** can be detected as the first grade candidate pixel if $X(x_0, y_0) > t_1$, where t_1 denotes the threshold. The corresponding best matching scale of the first-grade candidate pixel (x_0, y_0) can be obtained as follows:

$$G_1(x_0, y_0) = \underset{i}{\mathrm{argmax}} \left(\frac{(\mathbf{C}_{\mathrm{F}}(i) - \overline{\mathbf{C}_{\mathrm{F}}(i)})(\mathbf{C}_{\mathrm{A}}(x_0, y_0) - \overline{\mathbf{C}_{\mathrm{A}}(x_0, y_0)})}{\|\mathbf{C}_{\mathrm{F}}(i) - \overline{\mathbf{C}_{\mathrm{F}}(i)}\| \|(\mathbf{C}_{\mathrm{A}}(x_0, y_0) - \overline{\mathbf{C}_{\mathrm{A}}(x_0, y_0)})\|} \right) \tag{13}$$

A radial sampling filter projects the images **A** and $\mathbf{F}_i$ on a set of radial lines, and uses these projections to eliminate the first level candidate that have no opportunity to match the probe template $\mathbf{F}_i$. Radial sampling filter can obtain the second-grade candidate pixels by eliminating some of the first grade candidate pixels; meanwhile, it can also obtain their corresponding matching orientations. Given a first-grade candidate pixel (x_0, y_0) and its corresponding matching scale $\sigma_i (i = G_1(x_0, y_0))$, the scaled template image $\mathbf{F}_i$ with the scale parameter s_i can be represented by using Equation (4). The scaled image $\mathbf{F}_i$ is radially sampled on the radial line situated at radius r_l from the center pixel (x, y) and inclination θ_j:

$$\mathbf{R}_\mathrm{F}(j) = \frac{1}{r_l} \sum_{t=0}^{r_l} \mathbf{F}_i(x + r_l \cos \theta_j, y + r_l \sin \theta_j), 0 \le j < m \tag{14}$$

For each first-grade candidate pixel (x_0, y_0), the radial projections can be denoted as:

$$\mathbf{R}_\mathrm{A}(x_0, y_0, j) = \frac{1}{r_l} \sum_{t=0}^{r_l} \mathbf{A}(x_0 + r_l \cos \theta_j, y_0 + r_l \sin \theta_j), 0 \le j < m \tag{15}$$

The correlation coefficient of $\mathbf{R}_\mathrm{F}$ and $\mathbf{R}_\mathrm{A}$ is used to detect the best matching scale for each pixel (x_0, y_0) in $\mathbf{A}$, which can be represented as:

$$Y(x_0, y_0) = \max \left(\frac{(\mathbf{R}_\mathrm{F}(i) - \overline{\mathrm{shift}_j \mathbf{R}_\mathrm{F}(i)})(\mathbf{R}_\mathrm{A}(x_0, y_0) - \overline{\mathrm{shift}_j \mathbf{R}_\mathrm{A}(x_0, y_0)})}{\|\mathbf{R}_\mathrm{F}(i) - \overline{\mathrm{shift}_j \mathbf{R}_\mathrm{F}(i)}\| \|(\mathbf{R}_\mathrm{A}(x_0, y_0) - \overline{\mathrm{shift}_j \mathbf{R}_\mathrm{A}(x_0, y_0)})\|} \right) \tag{16}$$

where "shift" means circular shifting j positions of the argument vector. The first grade candidate pixel (x_0, y_0) in $\mathbf{A}$ can be reserved and updated as the second grade candidate pixel if $Y(x_0, y_0) > t_2$, and t_2 denotes the threshold. The corresponding best matching orientation of the second-grade candidate pixel (x_0, y_0) can be obtained as follows:

$$G_2(x_0, y_0) = \underset{i}{\mathrm{argmax}} \left(\frac{(\mathbf{R}_\mathrm{F}(i) - \overline{\mathrm{shift}_j \mathbf{R}_\mathrm{F}(i)})(\mathbf{R}_\mathrm{A}(x_0, y_0) - \overline{\mathrm{shift}_j \mathbf{R}_\mathrm{A}(x_0, y_0)})}{\|\mathbf{R}_\mathrm{F}(i) - \overline{\mathrm{shift}_j \mathbf{R}_\mathrm{F}(i)}\| \|(\mathbf{R}_\mathrm{A}(x_0, y_0) - \overline{\mathrm{shift}_j \mathbf{R}_\mathrm{A}(x_0, y_0)})\|} \right) \tag{17}$$

The template matching filter computes the similarity between the patch $\mathbf{X}$ centered at each second grade candidate pixel (x_0, y_0) and the scaled and rotated template $\mathbf{F}_i$ that use the best matching scale $\sigma_i(i = G_1(x_0, y_0))$ and the best matching angle $\theta_j(j = G_2(x_0, y_0))$, and the structural similarity (SSIM) method proposed in [30] was used to evaluate the similarity as follows:

$$SSIM(\mathbf{X}, \mathbf{F}_i) = \left[\frac{2\mu_X \mu_F + C_1}{\mu_X^2 + \mu_F^2 + C_1} \right]^\alpha \left[\frac{2\sigma_X \sigma_F + C_2}{\sigma_X^2 + \sigma_F^2 + C_2} \right]^\beta \left[\frac{\sigma_{XF} + C_3}{\sigma_X \sigma_F + C_3} \right]^\gamma \tag{18}$$

where μ_X denotes the mean of $\mathbf{X}$, σ_X denotes the standard deviation of $\mathbf{X}$, and σ_{XF} denotes the covariance of $\mathbf{X}$ and $\mathbf{F}_i$. The constants C_1, C_2, and C_3 are used to avoid the denominators being close to zeroes. α, β, and γ are parameters which are used to adjust the importance of the corresponding components.

The second grade candidate pixel (x_0, y_0) with the highest structural similarity is chosen, and the template is considered to be found at pixel (x_0, y_0), and the highest structural similarity is set as the feature similarity of this patch.

The similarity values of partial and individual identifying layer features between two shoeprints are calculated by using the Ciratefi method. The similarity values of two-layer features in subsection s_i are computed as follows:

$$P_\mathrm{P}(\mathbf{s}_i) = \sum_{\mathbf{F}_{\mathrm{P},m} \in \mathbf{s}_i} S(\mathbf{F}_{\mathrm{P},m}) / N_\mathrm{P}(\mathbf{s}_i)(i = 1, 2, \dots, 19) \tag{19}$$

$$P_\mathrm{I}(\mathbf{s}_i) = \sum_{\mathbf{F}_{\mathrm{I},n} \in \mathbf{s}_i} S(\mathbf{F}_{\mathrm{I},n}) / N_\mathrm{I}(\mathbf{s}_i)(i = 1, 2, \dots, 19) \tag{20}$$

where $\mathbf{F}_{\mathrm{P},m}$ and $\mathbf{F}_{\mathrm{I},n}$ denote the mth partial layer patch and nth individual identifying layer patch in the subsection s_i, respectively, and $S(\mathbf{F}_{\mathrm{P},m})$ and $S(\mathbf{F}_{\mathrm{I},n})$ represent the similarity value of two shoeprints on feature $\mathbf{F}_{\mathrm{P},m}$ and $\mathbf{F}_{\mathrm{I},n}$. $N_\mathrm{P}(\mathbf{s}_i)$ and $N_\mathrm{I}(\mathbf{s}_i)$ denote the number of partial layer features and individual identifying layer features in s_i, respectively. $P_\mathrm{P}(\mathbf{U}) = \left\{ P_\mathrm{P}(\mathbf{s}_i) | i = 1, 2, \cdots, 19 \right\}$ and

$P_{\rm I}(\mathbf{U}) = \left\{P_{\rm I}(\mathbf{s}_i) \middle| i = 1, 2, \cdots, 19\right\}$ represent partial and individual identifying characteristics similarity value set of two shoeprints in nineteen subsections.

3.4. Decision-Making

For fusing the three layer features at the score level, the total similarity score $p \in \mathbb{R}_+$ is used to denote the possibility that two shoeprints could have been made by the same shoe. For a pair of shoeprints $\mathbf{U}$, their total similarity score p is defined as a function of the similarity values of global, partial, and individual identifying layer features, which can be expressed as

$$p = f(P_{\rm G}, P_{\rm P}, P_{\rm I}, P_{\rm W}) = \text{sgn}(P_{\rm G} - T_{\rm G})P_{\rm PI} \tag{21}$$

where $\text{sgn}(\bullet)$ is a sign function. $T_{\rm G}$ denotes a threshold that is used to determine whether the two shoeprints have the same class characteristics. $P_{\rm PI}$ represents the combined feature score of two shoeprints on both partial and individual identifying layer features, which is mathematically defined as follows:

$$P_{\rm PI} = \frac{1}{19}\sum_{i=1}^{19} P_{\rm W}(\mathbf{s}_i)P_{\rm P}(\mathbf{s}_i)P_{\rm I}(\mathbf{s}_i) \tag{22}$$

where $P_{\rm W}(\mathbf{s}_i)$ is the weight value of subsection $\mathbf{s}_i$, and $P_{\rm W}(\mathbf{s}_i) = 1 - R(\mathbf{s}_i)/\sum_i R(\mathbf{s}_i)$, where $R(\mathbf{s}_i)$ represents the weight order, which are listed in Table 1.

For a pair of shoeprints, if $p \leq 0$, then we can make a conclusion that the two shoeprints may have different patterns; if p is greater than a positive predefined threshold $T_{\rm T}$, then we conclude that the two shoeprints could be left by the same shoe, otherwise they could not be left by the same shoe. The decision rule can be described as follows:

$$V(p) = \begin{cases} 1 & p > T_{\rm T} \\ 0 & 0 \leq p \leq T_{\rm T} \\ 0 & p < 0 \end{cases} \tag{23}$$

where $V(p) = 1$ indicates the two shoeprints are from the same shoe, and $V(p) = 0$ denotes the two shoeprints are from different shoes.

Accord to Equations (21) and (23), two thresholds $T_{\rm G}$ and $T_{\rm T}$ are needed to be pre-trained to tell whether two shoeprint are from the same shoe. We used the decision cost function (DCF) [31] to choose these two thresholds, that is to say

$$\left[\hat{T}_{\rm G}, \hat{T}_{\rm T}\right] = \arg\min_{T_{\rm G}, T_{\rm T}}\left(FAR(T_{\rm G}, T_{\rm T}) + FRR(T_{\rm G}, T_{\rm T})\right) \tag{24}$$

where $FAR(T_{\rm G}, T_{\rm T})$ and $FRR(T_{\rm G}, T_{\rm T})$ denote the false acceptance rate and the false rejection rate at the threshold $T_{\rm G}$ and $T_{\rm T}$, which can be calculated with Equations (25) and (26).

The proposed multi-layer-based shoeprint verification method is summarized in Algorithm 1.

Algorithm 1. Shoeprint Image Verification Algorithm

Input: A pair of shoeprints **U**.
Output: The total similarity score p and the verification result.
1. Image preprocessing.
2. Feature extraction. Extract partial layer feature, and individual identifying layer feature, respectively.
3. Feature matching.
4. Calculate similarity of global layer feature.
5. For r = 1, 2, ... , 19
6. Calculate similarity of partial layer feature with Equation (19).
7. Calculate similarity of individual identifying layer feature with Equation (20).
8. end For
9. Calculate total similarity score with Equation (21).
10. Judgment. Output verification result, identical or non-identical.

---

4. Experiments

In this section, we evaluate the proposed method by conducting shoeprint verification experiments on the real crime scene shoeprint image database and the suspect shoeprint image database. The false rejection rate (FRR), the false acceptance rate (FAR), and the equal error rate (EER) are used to evaluate the proposed method. The following subsections detail the experiment configuration, performance evaluation, analysis and discussion.

4.1. Experiment Configuration

In the experiments, we automatically aligned each pair of shoeprints. Thus, the main differences between the shoeprints are of a structural nature (i.e., the class characteristics, wear characteristics, and individual identifying characteristics). The registration make it possible to compare our method with some robust methods that are not invariant to rigid transformations, e.g., the histogram of oriented gradients (HOG) method [32] and the normalized cross correlation method (NCC). We also compared some RST invariant methods, e.g., the SURF algorithm [33] with the proposed method.

4.1.1. Dataset

The dataset consisted of one testing set and one training set. The definitions and descriptions of the two kinds of datasets are detailed as follows.

(i) Testing set: A testing set is a collection of shoeprint images that need to be verified. The testing set contained 1200 pairs of reference shoeprints in the MUES-SV1KR2R dataset and 256 pairs of crime scene shoeprints in the MUES-SV2HS2S dataset.

(ii) Training set: The training set is a collection of shoeprint images used to train the thresholds. The training set consisted of 300 pairs of reference shoeprint images and 100 pairs of crime scene shoeprint images. One hundred pairs of reference shoeprint images were from the same shoes. Twenty-five pairs of crime scene shoeprints are from the same shoes, and fifty pairs of shoe prints were not of the same class characteristics. Then accord to Equation (24), the optimal T_G and T_T can be achieved by operating shoeprint verification on the training set.

4.1.2. Evaluation Metric

In our experiments, we used the false acceptance rate (FAR), the false rejection rate (FRR), and the equal error rate (EER) to evaluate and characterize the proposed shoeprint verification system. These evaluation methods are always used to evaluate the performance of biometric verification, such as the finger vein verification [34,35] and palm vein verification [36].

The false acceptance rate means that deciding a claimed identity is a legitimate one while in reality it is an imposter. The FAR can be expressed as follows:

$$FAR = \frac{N_{\text{FA}}}{N_{\text{IA}}} \times 100\% \tag{25}$$

where N_{FA} and N_{IA} denote the number of false acceptances and the number of imposter attempts, respectively.

The false rejection rate means that deciding that a claimed identity is a not legitimate one while in reality it is the genuine one. The FRR can be expressed as follows:

$$FRR = \frac{N_{\text{FR}}}{N_{\text{TA}}} \times 100\% \tag{26}$$

where N_{FR} and N_{TA} denote the number of failed rejections and the number of target access, respectively.

The equal error rate (EER) represents the point where the false acceptance rate and the false rejection rate are equal, and a good system should keep this value as small as possible.

4.2. Performance Evaluation

In this section, we conducted two kinds of experiments to evaluate the performance of the proposed method. The first kind of experiment was to verify the performance of the proposed shoeprint verification method. The second kind of experiments was to compare the proposed method with some conventional methods.

### 4.2.1. Performance Evaluation of the Proposed Method

We evaluated the performance of the proposed method on two shoeprint datasets, and the datasets consisted of genuine and imposter pairs that included intra-class and inter-class matches. In this work, intra-class matching was defined as matching between a pair of shoeprint samples that were made by the same shoe, and inter-class matching was defined as matching between a pair of shoeprint samples that were left by different shoes. The experimental results are shown as follows, and Figure 10 shows the histograms of the similarity value for genuine and imposter shoeprint samples in MUES-SV1KR2R and MUES-SV2HS2S dataset, respectively. The left blue lines show the frequency distribution for inter-class similarity value, while the right red lines show the frequency distribution for intra-class similarity value. The results show that the frequency distribution of similarity value of inter-class are significantly lower than that of intra-class, which implies that the similarity value of the multiple-layer features for shoeprint samples left by the same shoe is significantly higher than the similarity value for shoeprint samples left by different shoe. The results also show that it is possible to tell whether two shoeprints are left by the same shoe.
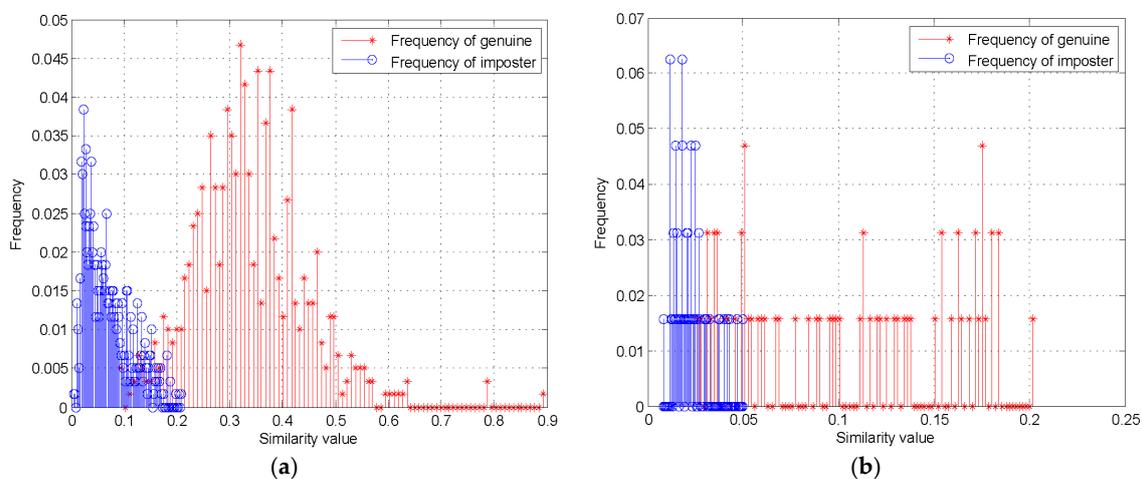
Figure 10. Genuine and imposter distributions of the proposed method on the datasets. (**a**) Results on the MUES-SV1KR2R dataset; (**b**) results on the MUES-SV2HS2S dataset.

4.2.2. Comparison and Discussion

We also used the equal error rate to evaluate the performance of the methods on two datasets by making trade-offs between the false acceptance rate and the false rejection rate. The performance of the methods is shown in the Figure 11, where the equal error rates of the proposed method were 3.2% and 10.9% on the MUES-SV1KR2R and MUES-SV2HS2S, respectively.
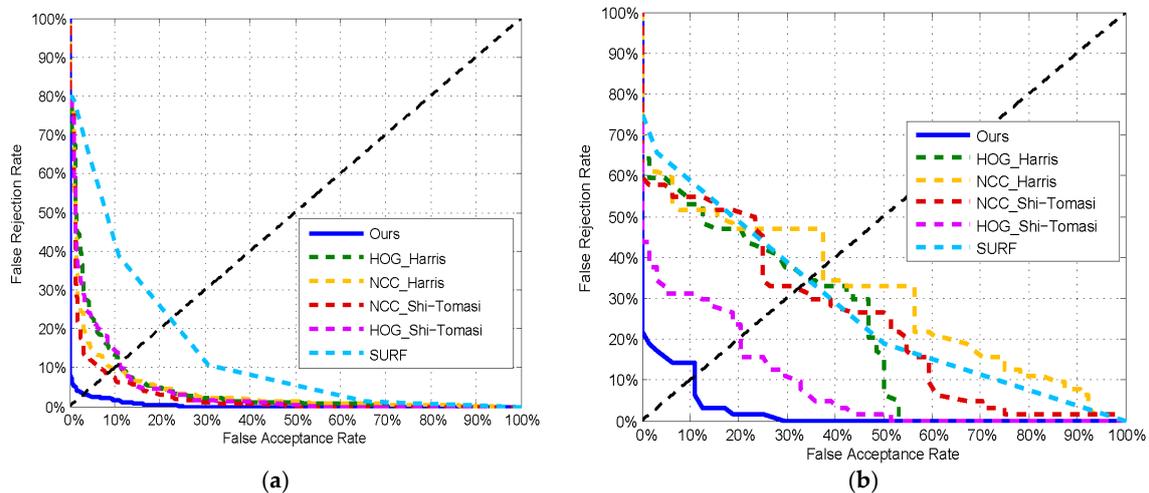
Figure 11. The receiver operating characteristics (ROC) curves of different methods on the datasets. (**a**) ROC curves on the MUES-SV1KR2R dataset; (**b**) ROC curves on the MUES-SV2HS2S dataset.

To our best knowledge, there is no automatic shoeprint verification algorithm reported up to now, so we could not compare the proposed algorithm with the state-of-the-art methods. To further test the effectiveness of the proposed method, we conducted comparisons with some existing conventional interest point-based methods. To conduct these comparisons, we implemented the conventional methods based on the corresponding literatures. In the experiments, we mainly compared our method with some interest point detections, such as the Harris [28] and Shi-Tomasi [37] corner detectors. We integrated the normalized cross correlation method (NCC) and the histograms of oriented gradients method (HOG) with the interest point detections. The patch centered on each interest point was used to describe the point, and HOG and NCC methods were used to describe the patch and match shoeprints. We also compared some RST invariant methods with the proposed method, e.g., the SURF algorithm [33].

The compared methods were all implemented with MATLAB codes. The patch size was fixed in all experiments. To ensure a fair comparison, all of the shoeprints in the datasets were aligned. The receiver operating characteristics (ROC) curve on the MUES-SV1KR2R dataset are illustrated in Figure 11a, and the corresponding results are listed in Table 2. The results show that the NCC method and HOG method perform nearly equally for all interest points, respectively, and the NCC method can perform better than the HOG method on average. The reason may be that shoeprint registration is performed by using a FFT-based registration algorithm [25], and there are variations on transformation between shoeprints. NCC methods can be robust to transformational deviations. If there are not large variations on transformation between shoeprint, the HOG method may surpass the NCC method, because HOG methods take into consideration local structural information and can be robust to small transformational deviations by pooling local structural information into histograms. The results in general show that all of the methods perform well on MUES-SV1KR2R dataset. Our approach performs significantly better compared to the compared methods.

Table 2. Summary of the experiments in terms of shoeprint verification performance on the MUES-SV1KR2R dataset measured by the equal error rate (EER).

Method for Shoeprint Verification	Performance (EER), %	Computation Time, s
Harris_HOG	11.1	2.1
Harris_NCC	9.8	3.1
Shi-Tomasi_HOG	11.2	9.7
Shi-Tomasi_NCC	8.5	10.0
SURF	22.4	14.7
Ours	**3.2**	**280.6**

The receiver operating characteristics (ROC) curve on the MUES-SV2HS2S dataset are illustrated in Figure 11b, and the corresponding results are listed in Table 3. The results show that HOG method can perform better than the NCC method. The reason may be that shoeprints in the MUES-SV2HS2S dataset are always of low quality and shoeprints made by the same shoe are not exactly the same due to the varying imaging conditions. HOG methods take into consideration local structural information and can be robust to small transformational deviations by pooling local structural information into histograms. The results also show that the EER of all of the compared methods increase a lot, compared with shoeprint verification on MUES-SV1KR2R dataset. The main reason may be that shoeprint images in the MUES-SV2HS2S dataset are of low quality and with complicated backgrounds, and it results in that the interest points always lying on the shoeprint impression boundary and background. Our approach performes significantly better compared to the other feature detector and descriptor combinations. The results are summarized in Table 1. Our method also performes better than on the Harris, Shi-Tomasi, and SURF interest point method. That is mainly because (i) the multi-layer features used in our method use global and local information to analyze shoeprints and integrate their strengths to yield more satisfactory verification results; and (ii) we analyze and estimate the similarity of the local structural information by using the SSIM method.

Table 3. Summary of the experiments in terms of shoeprint verification performance on the MUES-SV2HS2S dataset measured by the EER.

Method for Shoeprint Verification	Performance (EER), %	Computation Time, s
Harris_HOG	34.4	2.1
Harris_NCC	37.5	3.2
Shi-Tomasi_HOG	20.3	10.4
Shi-Tomasi_NCC	31.3	11.0
SURF	34.3	15.2
Ours	**10.9**	**293.3**

Figure 12 shows three samples of shoeprint in the MUES-SV2HS2S dataset. The shoeprints in Figure 12 have the same shoe pattern, and the shoeprint shown in Figure 12b is left by a different shoe from Figure 12a,c. The main individual identifying characteristics are labeled with red circles as shown in Figure 12a. Figures 13 and 14 provide a visual illustration of the shoeprint verification results of the proposed method and the compared methods. We extract individual identifying characteristics on Figure 12a and compare these extracted characteristics at the same position in Figure 12b,c. The results between Figure 12a,b are shown in Figure 13, and the results between Figure 12a,c are shown in Figure 14. In Figure 13, the unmatched points are labeled with blue '*' s. The unmatched points denote that the two shoeprints in the point location were different. The matched points are labeled with blue '*'s in Figure 14. The results show that our approach can tell not only the differences between two shoeprint images but also the characteristics in common. The results show that the proposed method outperformed the compared methods.
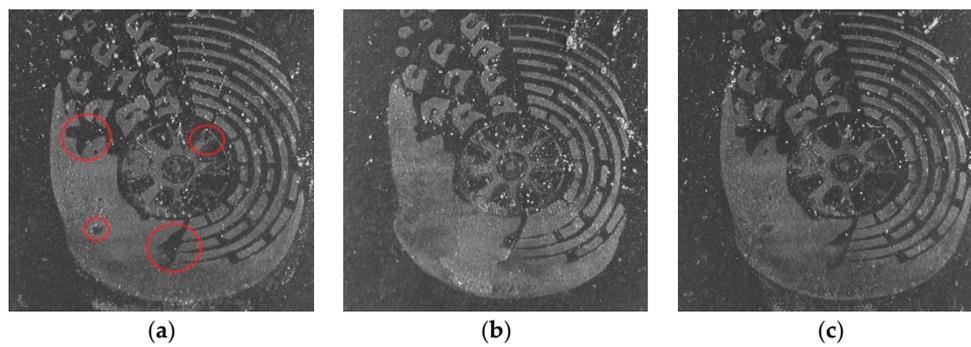
Figure 12. Samples of the probe shoeprint and two questioned shoeprints. (**a**) The probe shoeprint sample; (**b**) the questioned shoeprint sample that is made by different shoes with (**a**); (**c**) the questioned shoeprint sample that is made by the same shoe with (**a**).
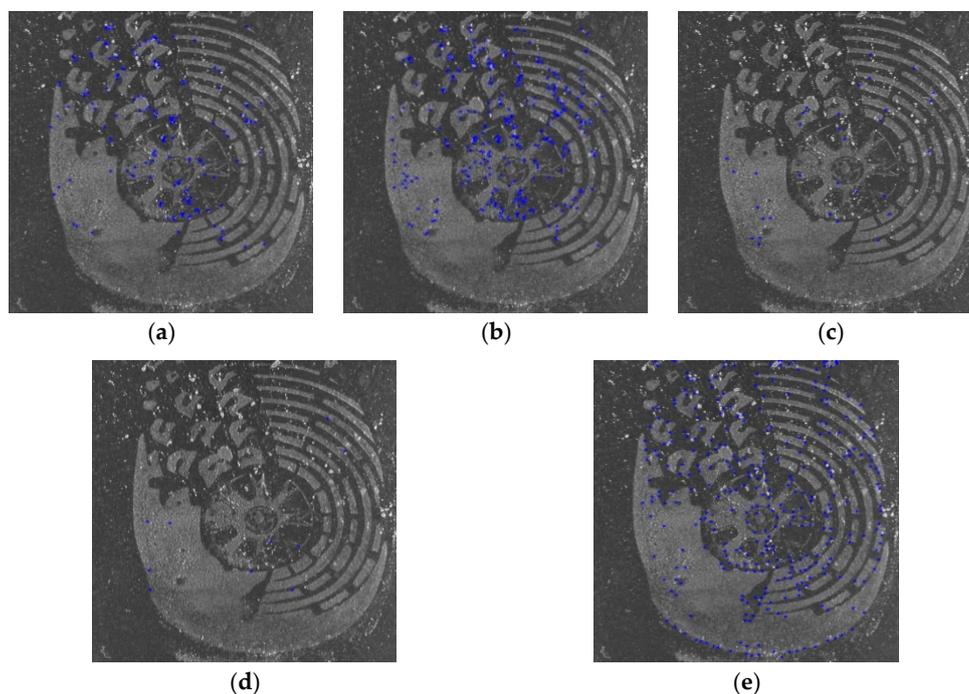
Figure 13. The results of our method and the compared methods on shoeprint made by different shoes. The unmatched points are labeled with blue '*'s. The more points are labeled, which should include red circled points in Figure 12a at least, the better the algorithm performs. (**a**) The results of Harris_HOG; (**b**) the results of Harris_NCC; (**c**) The results of Shitomasi_NCC; (**d**) the results of Shitomasi_HOG; (**e**) the results of our method.

Finally, we compare the time-consumption of the methods stated previously. This kind of experiments is conducted on a Windows 64-bit system with a 3.3-GHz CPU and 8-GB RAM. The time consumptions of each method on the two datasets are shown in Tables 2 and 3. Experimental results show that the proposed method required 280.6 s and 293.3 s of computation time for shoeprint verification on the two datasets. Although the proposed method requires more time to verify shoeprints than the compared methods, the performance of the proposed method surpasses the compared methods. Compared with the methods stated above, the proposed method matches not only the individual identifying layer feature, but also the global layer and partial layer feature, so it requires more computation. It is also for that more time is spent on the calling the Ciratefi method, because the Ciratefi method needs a lot of computations. In future works, we would like to reduce computation time by introducing parallel computing technology and optimizing the program code.
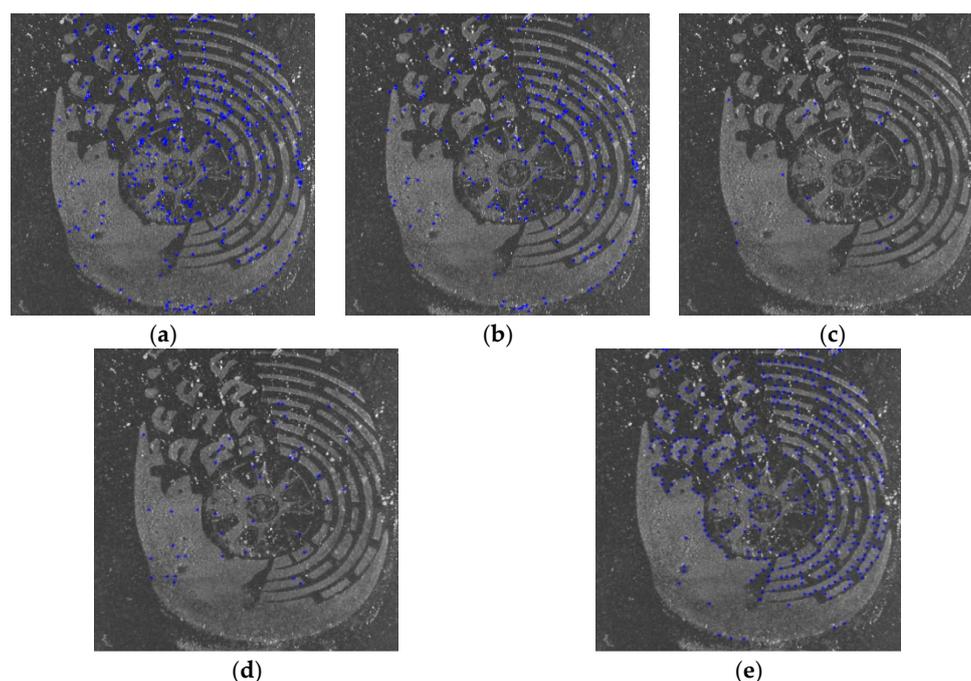
Figure 14. The results of the proposed and the compared methods on shoeprints made by the same shoe. The matched points are labeled with blue '*'s. The more points are labeled, which should include red circled points in Figure 12a at least, the better the algorithm performs. (**a**) The results of Harris_HOG; (**b**) the results of Harris_NCC; (**c**) the results of Shitomasi_NCC; (**d**) the results of Shitomasi_HOG; (**e**) the results of our method.

5. Conclusions

In this work, we have proposed a multi-layer feature-based method to conduct shoeprint verification automatically. Multi-layer features were extracted to compute the similarity values of two shoeprints on each layer, and then a total similarity score was computed based on these similarity values and a conclusion was made by a piecewise function of total similarity score. Comparative experiments with state-of-the-art methods were conducted, and the experimental results showed that: (1) our method had a better performance than the SURF method as well as some other state-of-the-art methods; and (2) the higher the quality of the shoeprint images were, the lower the EER of the methods can be achieved. In conclusion, the EERs on the two shoeprint image databases, which were 3.2% and 10.9%, demonstrate that our method is an effective shoeprint verification method, especially for images derived by using the shoeprint scanners. Our next work is to design a much more effective classifier to improve the verification algorithm, and we will also focus on how to improve the performance on complex backgrounds.

Author Contributions: X.W. and Y.W. conceived and designed the experiments; Y.W. performed the experiments; T.Z. analyzed the data; Y.W. wrote the original draft; X.W. and T.Z. reviewed and edited the manuscript; X.W. supervised this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1.　Bodziak, W.J. *Footwear Impression Evidence: Detection, Recovery, and Examination*; CRC Press: Boca Raton, FL, USA, 2000.

2.　Rida, I.; Bakshi, S.; Proença, H.; Fei, L.; Nait-Ali, A.; Hadid, A. Forensic shoe-print identification: A brief survey. *arXiv* **2019**, arXiv:1901.01431.

3. Algarni, G.; Amiane, M. A novel technique for automatic shoeprint image retrieval. *Forensic Sci. Int.* **2008**, *181*, 10–14. [CrossRef] [PubMed]

4. Wei, C.H.; Hsin, C.; Gwo, C.Y. Alignment of Core Point for Shoeprint Analysis and Retrieval. In Proceedings of the International Conference on Information Science, Electronics and Electrical Engineering (ISEEE), Sapporo, Japan, 26–28 April 2014; pp. 1069–1072. [CrossRef]

5. Wang, X.N.; Sun, H.H.; Yu, Q.; Zhang, C. Automatic Shoeprint Retrieval Algorithm for Real Crime Scenes. In Proceedings of the Asian Conference on Computer Vision, Singapore, 1–5 November 2014; pp. 399–413. [CrossRef]

6. Wang, X.N.; Zhan, C.; Wu, Y.J.; Shu, Y.Y. A manifold ranking based method using hybrid features for crime scene shoeprint retrieval. *Multimed. Tools Appl.* **2016**, *76*, 21629–21649. [CrossRef]

7. Gueham, M.; Bouridane, A.; Crookes, D. Automatic Recognition of Partial Shoeprints Based on Phase-Only Correlation. In Proceedings of the IEEE International Conference on Image Processing, San Antonio, TX, USA, 16–19 September 2007; pp. 441–444. [CrossRef]

8. Wu, Y.J.; Wang, X.N.; Zhang, T. Crime Scene Shoeprint Retrieval Using Hybrid Features and Neighboring Images. *Information* **2019**, *10*, 45. [CrossRef]

9. de Chazal, P.; Flynn, D.; Reilly, R.B. Automated processing of shoeprint images based on the Fourier transform for use in forensic science. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 341–350. [CrossRef]

10. Deshmukh, M.P.; Patil, P.M. Automatic shoeprint matching system for crime scene investigation. *Int. J. Comput. Sci. Commun. Technol.* **2009**, *2*, 281–287.

11. Kong, B.; Supancic, J.; Ramanan, D.; Fowlkes, C. Cross-Domain Forensic Shoeprint Matching. In Proceedings of the 28th British Machine Vision Conference (BMVC), London, UK, 4–7 September 2017.

12. Kong, B.; Supancic, J.; Ramanan, D.; Fowlkes, C. Cross-Domain Image Matching with Deep Feature Maps. *Int. J. Comput. Vis.* **2019**. [CrossRef]

13. Pavlou, M.; Allinson, N.M. Automated encoding of footwear patterns for fast indexing. *Image Vis. Comput.* **2009**, *27*, 402–409. [CrossRef]

14. Pavlou, M.; Allinson, N.M. Automatic extraction and classification of footwear patterns. In Proceedings of the 7th International Conference on Intelligent Data Engineering and Automated Learning, Burgos, Spain, 20–23 September 2006; pp. 721–728. [CrossRef]

15. Kortylewski, A.; Albrecht, T.; Vetter, T. Unsupervised Footwear Impression Analysis and Retrieval from Crime Scene Data. In Proceedings of the Asian Conference on Computer Vision, Singapore, 1–5 November 2014; pp. 644–658. [CrossRef]

16. Kortylewski, A.; Vetter, T. Probabilistic Compositional Active Basis Models for Robust Pattern Recognition. In Proceedings of the 27th British Machine Vision Conference (BMVC), York, UK, 19–22 September 2016.

17. Kortylewski, A. Model-Based Image Analysis for Forensic Shoe Print Recognition. Ph.D. Thesis, University of Basel, Basel, Switzerland, 2017.

18. Alizadeh, S.; Kose, C. Automatic Retrieval of Shoeprint Images Using Blocked Sparse Representation. *Forensic Sci. Int.* **2017**, *277*, 103–114. [CrossRef]

19. Richetelli, N.; Lee, M.C.; Lasky, C.A.; Gump, M.E.; Speir, J.A. Classification of footwear outsole patterns using Fourier transform and local interest points. *Forensic Sci. Int.* **2017**, *275*, 102–109. [CrossRef]

20. Wang, H.X.; Fan, J.H.; Li, Y. Research of shoeprint image matching based on SIFT algorithm. *J. Comput. Methods Sci. Eng.* **2016**, *16*, 349–359. [CrossRef]

21. Almaadeed, S.; Bouridane, A.; Crookes, D.; Nibouche, O. Partial shoeprint retrieval using multiple point-of-interest detectors and SIFT descriptors. *Integr. Comput. Aided Eng.* **2015**, *22*, 41–58. [CrossRef]

22. Luostarinen, T.; Lehmussola, A. Measuring the Accuracy of Automatic Shoeprint Recognition Methods. *J. Forensic Sci.* **2014**, *59*, 8. [CrossRef]

23. Yekutieli, Y.; Shor, Y.; Wiesner, S.; Tsach, T. *Expert Assisting Computerized System for Evaluating the Degree of Certainty in 2D Shoeprints*; The U.S. Department of Justice: Washington, DC, USA, 2012.

24. Chunming, L.; Rui, H.; Zhaohua, D.; Gatenby, J.C.; Metaxas, D.N.; Gore, J.C. A Level Set Method for Image Segmentation in the Presence of Intensity Inhomogeneities with Application to MRI. *IEEE Trans. Image Process.* **2011**, *20*, 2007–2016. [CrossRef]

25. Reddy, B.S.; Chatterji, B.N. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Trans. Image Process.* **1996**, *5*, 1266–1271. [CrossRef]

26. Mikolajczyk, K.; Schmid, C. A Performance Evaluation of Local Descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1615–1630. [CrossRef]

27. Mikolajczyk, K.; Tuytelaars, T.; Schmid, C.; Zisserman, A.; Matas, J.; Schaffalitzky, F.; Kadir, T.; Gool, L.V. A Comparison of Affine Region Detectors. *Int. J. Comput. Vis.* **2005**, *65*, 43–72. [CrossRef]

28. Harris, C.; Stephens, M. A Combined Corner and Edge Detector. In Proceedings of the 4th Alvey Vision Conference, Manchester, UK, 31 August–2 September 1988; pp. 147–151. [CrossRef]

29. de Araújo, S.A.; Kim, H.Y. Ciratefi: An RST-invariant template matching with extension to color images. *Integr. Comput. Aided Eng.* **2011**, *18*, 75–90. [CrossRef]

30. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef]

31. el Hannani, A.; Petrovska-Delacrétaz, D.; Fauve, B. Text-independent Speaker Verification. In *Guide to Biometric Reference Systems and Performance Evaluation*; Petrovska-Delacrétaz, D., Chollet, G., Dorizzi, B., Eds.; Springer: London, UK, 2009; pp. 167–211.

32. Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, CA, USA, 20–25 June 2005; pp. 886–893. [CrossRef]

33. Bay, H.; Tuytelaars, T.; Gool, L.V. SURF: Speeded up robust features. In Proceedings of the 9th European Conference on Computer Vision (ECCV), Berlin/Heidelberg, Germany, 7–13 May 2006; pp. 404–417. [CrossRef]

34. Yang, G.; Xiao, R.; Yin, Y. Finger Vein Recognition Based on Personalized Weight Maps. *Sensors* **2013**, *13*, 12093–12112. [CrossRef]

35. Xi, X.; Yang, G.; Yin, Y. Finger Vein Recognition with Personalized Feature Selection. *Sensors* **2013**, *13*, 11243–11259. [CrossRef]

36. Ma, X.; Jing, X.; Huang, H. A Novel Palm Vein Recognition Scheme Based on an Adaptive Gabor Filter. *IET Biom.* **2016**, *6*, 325–333. [CrossRef]

37. Shi, J.; Tomasi, C. Good features to track. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 21–23 June 1994; pp. 593–600. [CrossRef]