Automatic Shoeprint Retrieval Algorithm for Real Crime Scenes

Xinnian Wang, Huihui Sun, Qing Yu, and Chi Zhang

Dalian Maritime University, Dalian, China

Abstract. This study is to propose a fully automatic crime scene shoeprint retrieval algorithm that can be used to link scenes of crime or determine the brand of a shoe. A shoeprint contour model is proposed to roughly correct the geometry distortions. To simulate the character of the forensic experts, a region priority match and similarity estimation strategy is also proposed. The shoeprint is divided into two semantic regions, and their confidence values are computed based on the priority in the forensic practice and the quantity of reliable information. Similarities of each region are computed respectively, and the matching score between the reference image and an image in the database is the weighted sum. For regions with higher confidence value, the similarities are computed based on the proposed coarse-to-fine global invariant descriptors, which are based on Wavelet-Fourier transform and are invariant under slight geometry distortions and interference such as breaks and small holes, etc. For regions with lower confidence value, similarities are estimated based on computed similarities of regions with higher confidence value. Parameters of the proposed algorithm have learned from huge quantity of crime scene shoeprints and standard shoeprints which can cover most practical cases, and the algorithm can have better performance with minimum user intervention. The proposed algorithm has been tested on the crime scene shoeprint database composed of 210,000 shoeprints provided by the third party, and the cumulative matching score of the top 2 percent is 90.87.

1 Introduction

It is generally understood that marks left by an offender's shoeprint at a crime scene may be helpful in the subsequent investigation of the crime [5]. According to statistics, 35 percent of crime scenes had footwear prints valuable in forensic science [12], and 30 percent of all burglaries provide valuable shoeprints [11]. Shoeprints are distinctive patterns that are often found at crime scenes and have been obtaining increasing importance in forensic investigations. The most challenging task for a forensic examiner is to work with highly degraded footwear marks and matching them to the most similar shoeprint available in the database.

Some semiautomatic shoeprint retrieval methods based on various geometric patterns are reported in [11], [21], [3], and a series of patterns are chosen by human experts to classify the shoeprints. An automatic pattern classification method is proposed in [22], but it doesn't work well with debris and shadows.

Bouridane et al. [4], [1] utilize fractals to represent the shoeprints and use a mean squared noise error as the similarity measure. Accuracy of the match is 88% in classifying 145 images. This system does not attempt to answer the questions of partial, rotation or scale invariance.

Z. Geradts et al. [10] used the two-dimensional Fourier transform to classify those geometric shapes. Match was achieved with a neural network processing the Fourier transform coefficients and the positions of geometric shapes.

P. De Chazal et al. [9] use the power spectral density (PSD) to characterize the images for translational invariance, and the 2D correlation coefficient is used as the similarity measure. Results show that shoeprints are correctly matched in the top 5% of the sorted DB patterns with an 85% score. However, noisy images are not considered.

Fourier transforms modified phase only correlation (MPOC) is used in [13]. The reference DB consists of 100 different shoes available on the market and four sets of synthetic versions. The experimental result demonstrates a 100% first rank recognition rate, but the system is not invariant under translation or rotation.

Hu's seven moments are employed in [2] in order to have translation, rotation and scale invariance. Hu's moments are used on a reference DB containing 500 shoeprints and their noisy rotated versions. Results show a sharp drop of accuracy to 5.4% when the Gaussian noise variance is up to 0.2.

Gabor and Radon transform are used in [16] to extract multiresolution invariant features, and the first rank recognition rate can reach 91%.

Maximally Stable Extremal Region (MSER) feature is used in [17] to identify the features of the shoeprint and the Scale Invariant Feature Transform (SIFT) descriptors are employed to describe them. The reference DB is made of 374 shoeprints. Each pattern class consists of two images, a reference set image containing a whole left and right print, while the test set is made of an image of either a complete left or right print. They reported a 94% classification rate if viewing only 5% of the database, but no tests are performed on noisy images.

An image retrieval algorithm combing the information of the phase and the power spectral density of the Fourier transform calculated on their Mahalanobis map is employed in [8] and [7]. The reference DB consists of 35 shoeprints and the system is tested on synthetic as well as on real shoeprints coming from crime scenes. They reported 91% of the real case shoeprints found in the top 6.

Most of the above mentioned retrieval algorithms work well only with clear prints or synthetic shoeprints, but fail with crime scene shoeprints. The possible reasons are that they use features that are hard to be captured from the crime scene shoeprints, and the crime scene shoeprints are highly degraded and randomly partial. For example, (i) Real scene shoeprint images are always binarized to be separated from the backgrounds, but local invariant descriptors such as SIFT, MSER or SURF don't have good performance for binary images. (ii) There are many random extrusions, intrusions or breaks on the edges of patterns, and patterns are always randomly bridged. Fractal patterns and local invariant descriptors can be falsely extracted because of these interferences. (iii)

Fourier based methods have better performance than fractal patterns and local invariant descriptors, but they are not well correlated with the human visual system.

2 Aim

In this paper, we propose a simple but efficient low quality shoeprint images retrieval algorithm, and the test images and shoeprint images in the database all come from real crime scenes without any synthetic shoeprints or generated partials. What the proposed algorithm differs from the other existing algorithms are on the capacity of the real crime scene database, Shoeprint Contour Model used for geometry correction, hybrid Wavelet-Fourier based global invariant feature descriptors and the robust matching strategy which has better correlation with the forensic experts.

3 Shoeprint Database

Two databases of shoeprint images are formed by more than 4,000 and 200,000 shoeprints provided by Dalian Everspry SCI & TECH CO., LTD, China. The first database consists of clear and full 4950 shoeprints created by taking impressions of footwear outsoles provided by footwear vendors, and in this paper,we refer this kind of shoeprints as standard shoeprints. The second databases is derived from the real crime scene and composed of variable quality left or right prints. Images from the second database possibly differ on position, orientation, scale, quantity of reliable information, quality and imaging conditions. Besides of some clear full prints, most images in the database are misaligned, incomplete and degraded prints interfered with debris, shadows or other artifacts. Some typical examples of both databases are shown in Fig. 1.

4 Methods

The proposed algorithm has two phases: on-line retrieval and off-line feature extraction. In the off-line feature extraction phase, every image in the constructed

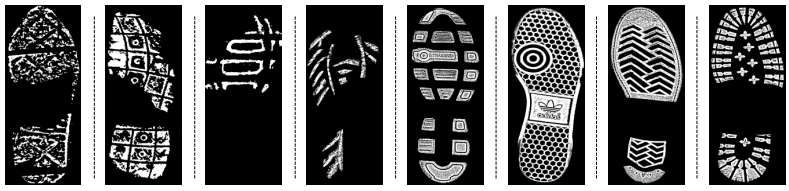

Fig. 1. Typical examples of shoeprints in the databases. The left four shoeprints are from real crime scenes, and the others are standard shoeprints.

shoeprint database is firstly preprocessed to separate the shoeprint from backgrounds, and then Wavelet-Fourier based global invariant features of each part are extracted, finally the features of each image are pooled into the shoeprint feature database prepared for print retrieval. In the on-line retrieval phase, confidence value and features of each part of every input image are computed, and then similarity measures of high confidence parts between the input image and an image in the database are computed, and similarity measures of low confidence parts are estimated based on computed similarity measures, and the similarity score between the input image and the image in the database is defined as the weighted sum of similarity measures of two parts, and finally outputs the ranked list of images based on the similarity score sorting. The flow diagram of the proposed algorithm is shown in Fig. 2.

4.1 Image Preprocessing

The goal of this stage is to separate the shoeprints from backgrounds and normalize the extracted shoeprints. Image preprocessing includes the following steps: 1) Shoeprint extraction: A local adaptive thresholding technique is used to extract the shoeprint images from backgrounds. We firstly split the image into a grid of cells and then apply a simple thresholding method (e.g. Otsu's method) on each cell to extract sub prints, and morphological operations are finally used to fill little holes and smooth edges. 2) Resolution normalization: The picture of the print is taken with a forensic scale near to the print and it is rescaled to a predefined dpi. 3) Orientation normalization: A Shoeprint Contour Model (SPCM) is proposed to normalize the shoeprint image.

The SPCM is to represent the shape of a shoeprint with a set of landmarks. Firstly enough shoeprint images with various shapes are collected to be as the training set. Secondly a set of points are labeled to annotate shoeprint contour, and finally dimensionality reduction technique are used to extract the average shoeprint contour model. In Fig. 3(a) and Fig. 3(b), given a full shoeprint image, the average SPCM model is used to estimate the initial positions, and a morpho-

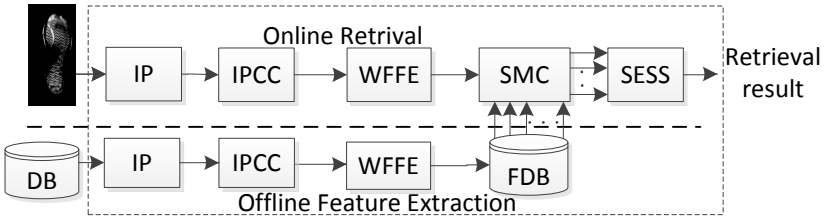

Fig. 2. The flow diagram of the proposed algorithm. IP, IPCC, WFFE, SMC, SESS, DB and FDB are abbreviations for Image Preprocessing, Image Partion and Confidence Computation, Wavelet-Fourier Based Feature Extraction, Similarity Measure Computation, Similarity Estimation and Score Sorting, Shoeprint Database and Feature Database, respectively.

logical close operation with larger size of structure element and an active contour method (e.g. Snake [23]) are used to find the best matching position between the model and the data in the input image. In Fig. 3(c), for a partial image, three points (front most point, rearmost point, and leftmost point) are marked interactively, and the shoeprint contour is estimated by the average SPCM model and the landmarks. Once the contour of the input image is estimated, the shoeprint image can be aligned to the predefined orientation and position, and the scales can be refined. In practical applications, the region of foot arch usually can't be acquired from crime scenes, and the contour of foot arch just needs to be estimated coarsely. For accurate estimation of the foot arch, another two points besides of three points marked shown in Fig. 3(c) are needed to be labeled, and the two points are close to the maximum or minimum curvature points on the left or right boundary of the foot arch region respectively, which are shown in Fig. 3(d).

4.2 Image Partion and Confidence Computation

When an experienced forensic expert compares two shoeprints, he or she may divide a full shoeprint into toe section, sole section, instep (arch) section, heel section and back of heel section, and each section has a classification priority. The sole section has the highest priority, and the heel section has the second highest priority, and the arch section has the lowest priority. In the proposed algorithm, to simulate the character of the forensic experts, we define a confidence value which is biased toward those parts which: (i) have higher priority in the forensic practice and (ii) have much more reliable information.

Given a region $\mathbf{s}$, we define its confidence value $C(\mathbf{s})$ as the product of two terms:

$$C(\mathbf{s})=P(\mathbf{s})H(\mathbf{s}) \tag{1}$$

We call $P(\mathbf{s})$ the priority term and $H(\mathbf{s})$ the information term. The confidence value of each region is used to be the weights of pooling region matching scores

(a) Average SPCM (b) Refined SPCM (c) Estimated SPCM (d) SPCM with arch contour

Fig. 3. SPCM of the shoeprint images (White circles denote the landmarks)

to be the total score. $C(\mathbf{s})$ is also used to judge whether the region can be used to retrieve shoeprints.

Based on thousands of crime scene shoeprints, we have found that patterns of the sole sections and the heel sections determine the retrieval results in most cases. Thus, the shoeprint is roughly divided into the top region and the bottom region. The top region mainly includes the toe section, sole section and parts of the arch section, and the bottom region mainly includes the other parts of the arch section, the heel section and the back of heel section. Each region is assigned a predefined priority value. The ratio of the top region height to the bottom region height is 3 to 2, which is learned from training samples. Details are shown in Fig. 4. The priority value is defined as:

$$P(\mathbf{s}) = 1 - \frac{\sum\limits_{j \in \mathbf{s}} R(j)}{\sum\limits_{i} R(i)} \tag{2}$$

where $\mathbf{s}$ represents the top region or the bottom region, $P(\mathbf{s})$ represents the priority value of region $\mathbf{s}$, $R(i)$ represents the priority order of each section. The priority orders of all sections are listed in Tab. 1.

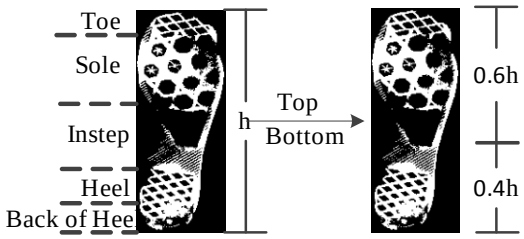

Fig. 4. Shoeprint partion

Table 1. Priority order of each section

Section number	Section of shoeprint	Rank order
1	Toe	3
2	Sole	1
3	Heel	2
4	Black of Heel	4
5	Instep	5

Information value is used to measure the amount of reliable information of the two regions. For a region with all black pixels, the information value is set

to 0. For a region with all white pixels, the information value is set to 1. The information value is defined as:

$$H(\mathbf{s}) = \frac{\sum\limits_{\mathbf{p}\in\mathbf{s}} \mathbf{I}(\mathbf{p})}{|\mathbf{s}|} \tag{3}$$

where $\mathbf{s}$ denotes the section, $|\mathbf{s}|$ is the area of $\mathbf{s}$, $\mathbf{p}$ is the pixel point, $\mathbf{I}$ represents the shoeprint image.

4.3 Wavelet-Fourier Feature Extraction

Three kinds of features which include fractal patterns, 2D Fourier Transforms or Fourier-Mellin Transform and local invariant descriptors have been commonly used to retrieve shoeprints in the literature. We have tested these features on more than 4 thousand kinds of standard shoeprints and 210 thousand real crime scene prints, and found that the three kinds of features work very well for standard shoeprints, synthetic prints or very clear and complete realcrime scene prints but failed for most of real crime scene prints.

Our perception of the universe uses different scales: Each category of observations is done in a proper scale. Using a larger scale, we can observe more details. Using a small scale, we can observe only macroscopic details of shoeprint patterns without seeing small holes, breaks, extrusions and intrusions. By changing the scale, we can observe or represent the object from coarse-to-fine. For these reasons, we use Wavelet transform [15] which represents both the spatial and frequency domain simultaneously to extract features of shoeprints across different scales. There is much redundant or irrelevant information contained in wavelet coefficients which are sensitive to translation, rotation and scaling, and Fourier-Mellin [19] transform is employed to extract discriminative invariant features in one or several special spatial-frequency subbands. We call this method Wavelet-Fourier transformation based global invariant descriptor. Since Fourier-Mellin transform can capture global invariant features, the proposed descriptor is not only global invariant under translation and rotation on each scale, but also has a multi-resolution matching ability.

The proposed feature extraction method has three steps. The first step is to transform a specified region of the input shoeprint $\mathbf{I}(\mathbf{s})$ to its wavelet domain, and $\mathbf{W}(l, h, v)$ is used to represent the wavelet coefficients where l denotes the level, h and v indicate the sub-bands of wavelet coefficients. The second step is to perform Fourier-Mellin transform on each band of wavelet coefficients and compute the power spectral density of the coefficients of Fourier-Mellin transform and filter out unnecessary coefficients. $\mathbf{M}(l, h, v)$ is used to represent the PSD of each band. The third step is to choose which bands of coefficients to be features. The flow diagram of the descriptor exaction is shown in Fig. 5.

The detailed steps of the feature extraction algorithm are as follows:

Step 1: Input the specified region $\mathbf{I}(\mathbf{s})$ with the confidence value greater than the predefined value. $\mathbf{I}(\mathbf{s})$ is decomposed using Haar Wavelet to a specified

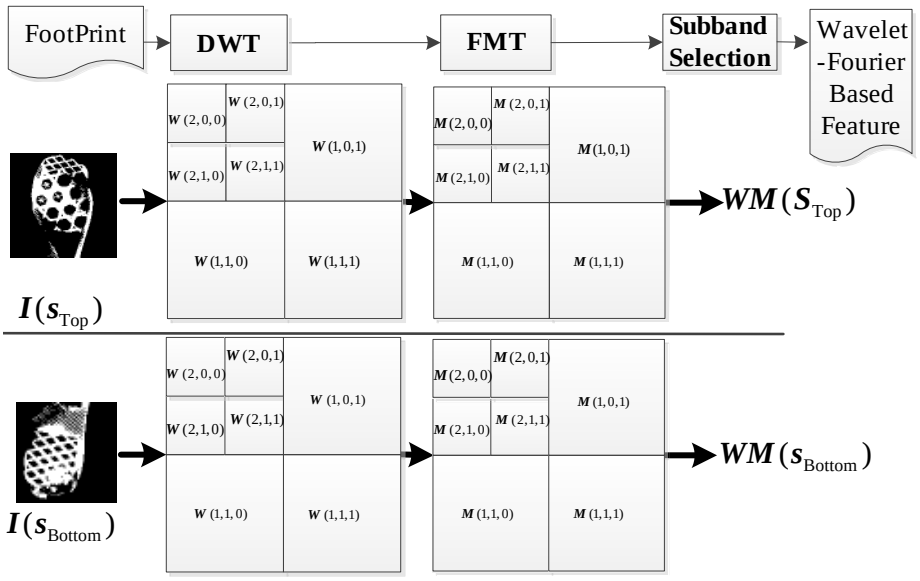

Fig. 5. Flow diagram of feature descriptor extraction

number of levels. At each level we will have one approximation subband and three details. The wavelet coefficients of $\mathbf{I}(\mathbf{s})$ can be described as:

$$\mathbf{W}(\mathbf{s}) = \{\mathbf{W}(\mathbf{l}, \mathbf{h}, \mathbf{v}) | \mathbf{0} \leq \mathbf{l} \leq \mathbf{L}, \mathbf{h}, \mathbf{v} = \mathbf{0}, \mathbf{1}\} \tag{4}$$

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

Step 2: For each band of wavelet coefficients $\mathbf{W}(l, h, v)$, the Fourier-Mellin transform is applied, and the PSD of each band denoted as $\mathbf{M}(l, h, v)$ is computed. The processes are as follows:

Step 2.1: For a given band coefficients $\mathbf{W}(l, h, v)$, the PSD of its Fourier-Mellin transform is computed, and it is denoted as $\mathbf{P}(l, h, v)$ which is invariant under translation, rotation and scaling.

Step 2.2: In a shoeprint, large connected bridges between patterns appear in the PSD as very low-frequency components, and a high-pass filter $H(\xi, \eta)$ proposed in [19] is firstly used to weaken the effects of these components. Noises such as small holes, intrusions, extrusions and broken patterns appear in the PSD as very high-frequency components, and an ideal lowpass filter whose cut-off frequency is taken for 0.8 times of the highest frequency is then used to remove them. A band-pass filtered version of $\mathbf{M}(l, h, v)$ denoted as $\mathbf{P}(l, h, v)$ is finally obtained by previous two filters.

Step 3: This step is to determine which bands of coefficients to be the retrieval features. Each subband has different contribution to the retrieval results.

For a highly degraded image, the approximation band is very important whereas details maybe interferences. For a shoeprint image with better quality, details can improve the accuracy. Subbands with rich information should be chosen intuitively. We use the standard deviation of $\mathbf{M}(l,h,v)$ to measure the information quantity, and choose the top k subbands to be the retrieval features:

$$\mathbf{WM}(\mathbf{s})=\{\hat{\mathbf{M}}^1(l,h,v)\cdots\hat{\mathbf{M}}^k(l,h,v)\} \tag{5}$$

where $\mathbf{WM}(\mathbf{s})$ represents the global invariance descriptor of region $\mathbf{s}$, $\hat{\mathbf{M}}^i(l,h,v)$ represents the ith subband ordered by the information quantity from the largest to the smallest value. For a highly degraded shoeprint, details are less important, and the approximation band of the highest level can directly be the features:

$$\mathbf{WM}(\mathbf{s})=\mathbf{M}(L,0,0) \tag{6}$$

The global invariance descriptors of the top region and the bottom region of the shoeprint with higher confidence can be computed, and they are denoted as $\mathbf{WM}(\mathbf{s}_\mathrm{T})$ and $\mathbf{WM}(\mathbf{s}_\mathrm{B})$ respectively. These features are used to measure the similarities between two shoeprints. For shoeprints in the database, they are captured from different crime scenes with different quality, features of every suband are extracted for future use when constructing the feature database. For a shoeprint to be retrieved, the features are selected according to Eq. (5) and Eq. (6).

4.4 Similarity Measure Computation

In order to compare a reference image with a database image, a measure of similarity between the images is required. The larger the measure of similarity between the reference image and the database image, the more similar the two images are. A reference image is compared to all images in the database and the similarity measure calculated for each comparison is used to rank the images in the database in a most similar to least similar order.

The similarity measure adapted in this paper is the 2D correlation coefficient [20]. For features $\mathbf{WM}_1$ and $\mathbf{WM}_2$, the correlation coefficient r is calculated using

$$\begin{aligned}
&\hat{\mathbf{W}}_\mathbf{1}(\mathbf{s})=\mathbf{WM}_1'(\mathbf{s})-\mathbf{W}\bar{\mathbf{M}}_1'(\mathbf{s})\\
&\hat{\mathbf{W}}_2(\mathbf{s})=\mathbf{WM}_2'(\mathbf{s})-\mathbf{W}\bar{\mathbf{M}}_2'(\mathbf{s})\\
&r(\mathbf{s})=\frac{\hat{\mathbf{W}}_\mathbf{1}(\mathbf{s})\hat{\mathbf{W}}_2(\mathbf{s})}{|\hat{\mathbf{W}}_\mathbf{1}(\mathbf{s})||\hat{\mathbf{W}}_2(\mathbf{s})|}
\end{aligned} \tag{7}$$

where $\mathbf{WM}'(\mathbf{s})$ is the 1D vector representation of $\mathbf{WM}(\mathbf{s})$, $\mathbf{W}\bar{\mathbf{M}}'(\mathbf{s})$ is the mean of $\mathbf{WM}(\mathbf{s})$, $\hat{\mathbf{W}}_\mathbf{1}(\mathbf{s})$ and $\hat{\mathbf{W}}_2(\mathbf{s})$ represent features of region $\mathbf{s}$ from different shoeprints.

4.5 Similarity Estimation and Scores Computation

The shoeprint image is divided into the top and bottom regions. Similarities of each region are computed respectively, and the total similarity between the reference image and an image in the database is the weighted sum of the similarity measures of the two regions. For an input image, the similarity between its mirror version and an image in the database is also computed. The final score is the greater one, which is insensitive to the left print or the right print. The flow diagram of the matching score computation is shown in Fig. 6.

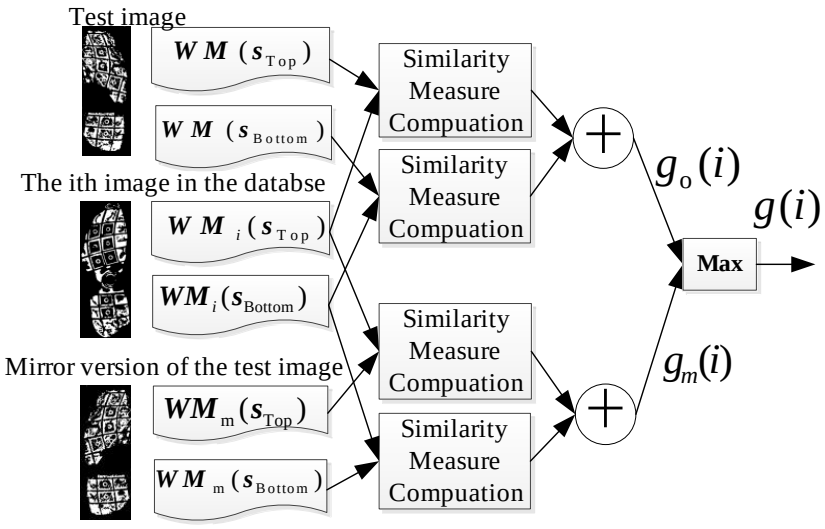

Fig. 6. Flow diagram of the matching score computation

The matching score between the test shoeprint and the ith image in the database $g(i)$ is computed according to Eq. (8). For regions of higher confidence value, the similarity measure is computed directly according to Eq. (7). To weaken the effect of missing regions and regions of lower confidence value, the similarity measure of those regions can't be computed directly according to Eq. (7). If the confidence value of the specified region of the test shoeprint image is greater than the predefined threshold, and the confidence value of the image in the database is lower, the similarity measure is estimated from computed ones of images with higher confidence values, according to Eq. (17). If the confidence value of the specified region of the test image is also lower, the similarity measure is set to a predefined value. In order to let the full shoeprints of the same pattern in the database lie in front of the partial ones, the predefined value is not simply set to 0, and it is obtained by means of trial and errors.

$$g(i)=\begin{cases} g_{\mathrm{o}}(i) & g_{\mathrm{o}}(i) > g_{\mathrm{m}}(i) \\ g_{\mathrm{m}}(i) & else \end{cases} \tag{8}$$

$$g_{\mathrm{o}}(i)=T \bullet r_{\mathrm{o}}(\mathbf{s}_{\mathrm{Top}}) + B \bullet r_{\mathrm{o}}(\mathbf{s}_{\mathrm{Bottom}}) \tag{9}$$

$$g_{\mathrm{m}}(i)=T \bullet r_{\mathrm{m}}(\mathbf{s}_{\mathrm{Top}}) + B \bullet r_{\mathrm{m}}(\mathbf{s}_{\mathrm{Bottom}}) \tag{10}$$

$$T=\frac{C(\mathbf{s}_{\mathrm{Top}})}{C(\mathbf{s}_{\mathrm{Top}}) + C(\mathbf{s}_{\mathrm{Bottom}})} \tag{11}$$

$$B=\frac{C(\mathbf{s}_{\mathrm{Bottom}})}{C(\mathbf{s}_{\mathrm{Top}}) + C(\mathbf{s}_{\mathrm{Bottom}})} \tag{12}$$

$$r_{\mathrm{o}}(\mathbf{s}_{\mathrm{T}})=\begin{cases} r\left(\mathbf{s}_{\mathrm{T}}\right) & C(\mathbf{s}_{\mathrm{T}}) > C_{\mathrm{th}}(\mathbf{s}_{\mathrm{T}}), C_i(\mathbf{s}_{\mathrm{T}}) > C_{\mathrm{th}}(\mathbf{s}_{\mathrm{T}}) \\ r_{\mathrm{oest}}(\mathbf{s}_{\mathrm{T}}) & C(\mathbf{s}_{\mathrm{T}}) > C_{\mathrm{th}}(\mathbf{s}_{\mathrm{T}}), C_i(\mathbf{s}_{\mathrm{T}}) < C_{\mathrm{th}}(\mathbf{s}_{\mathrm{T}}) \\ r_{\mathrm{Top}} & else \end{cases} \tag{13}$$

$$r_{\mathrm{o}}(\mathbf{s}_{\mathrm{B}})=\begin{cases} r\left(\mathbf{s}_{\mathrm{B}}\right) & C(\mathbf{s}_{\mathrm{B}}) > C_{\mathrm{th}}(\mathbf{s}_{\mathrm{B}}), C_i(\mathbf{s}_{\mathrm{B}}) > C_{\mathrm{th}}(\mathbf{s}_{\mathrm{B}}) \\ r_{\mathrm{oest}}(\mathbf{s}_{\mathrm{B}}) & C(\mathbf{s}_{\mathrm{B}}) > C_{\mathrm{th}}(\mathbf{s}_{\mathrm{B}}), C_i(\mathbf{s}_{\mathrm{B}}) < C_{\mathrm{th}}(\mathbf{s}_{\mathrm{B}}) \\ r_{\mathrm{Bottom}} & else \end{cases} \tag{14}$$

$$r_{\mathrm{m}}(\mathbf{s}_{\mathrm{T}})=\begin{cases} r\left(\mathbf{s}_{\mathrm{T}}\right) & C(\mathbf{s}_{\mathrm{T}}) > C_{\mathrm{th}}(\mathbf{s}_{\mathrm{T}}), C_i(\mathbf{s}_{\mathrm{T}}) > C_{\mathrm{th}}(\mathbf{s}_{\mathrm{T}}) \\ r_{\mathrm{oest}}(\mathbf{s}_{\mathrm{T}}) & C(\mathbf{s}_{\mathrm{T}}) > C_{\mathrm{th}}(\mathbf{s}_{\mathrm{T}}), C_i(\mathbf{s}_{\mathrm{T}}) < C_{\mathrm{th}}(\mathbf{s}_{\mathrm{T}}) \\ r_{\mathrm{Top}} & else \end{cases} \tag{15}$$

$$r_{\mathrm{m}}(\mathbf{s}_{\mathrm{B}})=\begin{cases} r\left(\mathbf{s}_{\mathrm{B}}\right) & C(\mathbf{s}_{\mathrm{B}}) > C_{\mathrm{th}}(\mathbf{s}_{\mathrm{B}}), C_i(\mathbf{s}_{\mathrm{B}}) > C_{\mathrm{th}}(\mathbf{s}_{\mathrm{B}}) \\ r_{\mathrm{oest}}(\mathbf{s}_{\mathrm{B}}) & C(\mathbf{s}_{\mathrm{B}}) > C_{\mathrm{th}}(\mathbf{s}_{\mathrm{B}}), C_i(\mathbf{s}_{\mathrm{B}}) < C_{\mathrm{th}}(\mathbf{s}_{\mathrm{B}}) \\ r_{\mathrm{Bottom}} & else \end{cases} \tag{16}$$

where $C_{\mathrm{th}}(\mathbf{s}_{\mathrm{T}})$ and $C_{\mathrm{th}}(\mathbf{s}_{\mathrm{B}})$ are the predefined thresholds, r_{Top} and r_{Bottom} are predefined default similarity values. After the similarities of regions with higher confidence values have been computed according to Eq. (7), $r_{\mathrm{oest}}(\mathbf{s})$ is estimated from computed similarity measures:

$$r_{\mathrm{oest}}(\mathbf{s})=\frac{a(\mathbf{s}) + \varphi b(\mathbf{s})}{1+\varphi} \tag{17}$$

$$a(\mathbf{s})=\min_i(r_i(\mathbf{s})), s.t. C_i(\mathbf{s}) > C_{\mathrm{th}}(\mathbf{s}) \tag{18}$$

$$b(\mathbf{s})=\max_i(r_i(\mathbf{s})), s.t. C_i(\mathbf{s}) > C_{\mathrm{th}}(\mathbf{s}) \tag{19}$$

where $r_i(\mathbf{s})$ represents the similarity measure, φ is the golden ratio which is 0.618.

5 Experiments and Results

5.1 Performance Evaluation Measure

The proposed algorithm is designed to sort shoeprint images of the database in response to a test image and present the ranked list to the user for final evaluation. An algorithm with higher performance will present fewer nonmatching images than an algorithm with lower performance. Cumulative Matching Characteristic Curve (CMC) is used to measure the accuracy performance of a retrieval algorithm operating in the closed-set identification task [18]. Images in the gallery set are compared and ranked based on their similarity with the test (probe) images. The CMC shows how often the probe image appears in the top n matches.

5.2 Test Images and Gallery Set

To evaluate the performance of the proposed algorithm, twelve groups of crime scene shoeprint images provided by the third party are used to be test images. Shoeprint images of each group have the same patterns, and have been acquired from different crime scenes with different image quality. The test images cover most patterns and possible cases that can be found in the real crime scenes. Each group has different number of images, and the total number of the test images is 72. Images from each group differ on position, orientation, scale, quantity of reliable information, quality and imaging conditions. In order to verify that the proposed algorithm is insensitive to geometry distortions, each test image is synthetically transformed into different translation, rotation and scaling versions, and another 432 images are inserted into the gallery sets.

Two gallery sets are constructed to test the performance. The first gallery set denoted as $\mathbf{GS}_1$ includes 72 test images, 432 generated geometry distortion versions of the test images and another 9592 crime scene shoeprints, and $\mathbf{GS}_1$ is often used by the third party to evaluate the performance of the retrieval algorithms. The second gallery set denoted as $\mathbf{GS}_2$ consists of about 210,000 real crime scene shoeprints of China. The proposed algorithm is used to retrieve the test images on these gallery sets and is evaluated by the CMC measure. It should be noted that many other shoeprints in the database have the same patterns with test images, but they aren't labeled, so the practical cumulative matching score of the proposed algorithm should be much higher than the experimental results.

5.3 Parameters Selection

The proposed algorithm has four main parameters, which are $C_{\mathrm{th}}(\mathbf{s}_{\mathrm{T}})$, $C_{\mathrm{th}}(\mathbf{s}_{\mathrm{B}})$, r_{Top} and r_{Bottom}. This experiment is to select the optimal value for every parameters, and the experiment is conducted on the gallery set $\mathbf{GS}_1$. The experimental result shows that the cumulative matching score can reach the highest one when the confidence value threshold is 0.03 and the estimated similarity is 0.3. For more details, please refer to Section I of the supplementary material.

5.4 Performance of Anti-geometry Distortion

To verify the anti-geometry distortion ability of the proposed algorithm, 72 test images are respectively inputted into the retrieval algorithm, and statistically count the number of their synthetically generated geometry distortion versions lying in the top 7 in the ranked list of results. Since the gallery includes the test images, the first one in the sorted list is the test image itself. In theory, the transformed versions of the test image should be at the top of the sorted list, but in practice, the generated images may lose some information because of out of range, and some images of the same category may also lie in the top 7. Although the average percent is just more than 81%, the left 19% images are almost from the same classes. These results show the proposed algorithm is robust to geometry distortions. For more details, please refer to Section II of the supplementary material.

5.5 Performance Comparisons with The State of Art Algorithms

Performance experiments are conducted on the two gallery sets $\mathbf{GS}_1$ and $\mathbf{GS}_2$ respectively, and the accuracies are shown in Tab. 2. The CMC curves are shown in Fig. 7. The experimental results show that the accuracy of top 2% is more than 87.5% on the gallery set $\mathbf{GS}_1$ and 90.87% on the gallery set $\mathbf{GS}_2$. Furthermore, there are also a lot of unmarked same pattern images on the top 2% of the ranked lists, and the accuracy of correct match would be much higher than the current accuracy. For more details, please refer to Section III of the supplementary material.

As stated in introduction, most algorithms have better performance under their assumed conditions such as qualities of test images and capacity of the gallery set, etc. Due to lack of public evaluation databases and public available test softwares, comparisons just depend on what the literature reported, and the comparison results are listed in Tab. 3.

Table 2. Retrieval performance of the proposed algorithm on $\mathbf{GS}_1$ and $\mathbf{GS}_2$

Datebase	Performance	Ranks of Retrieval Results						
		0.1%	0.2%	0.3%	0.4%	0.5%	1%	2%
$\mathbf{GS}_1$	Accuracy(%)	45.2	64.1	69.4	73.8	75.8	81.8	87.5
	Cumulative Number	228	323	350	372	382	412	441
$\mathbf{GS}_2$	Accuracy(%)	68.7	75.2	77.6	80.0	80.8	84.7	90.9
	Cumulative Number	345	379	391	403	407	427	458

6 Conclusion

In this study, we proposed a simple but efficient low quality footprint images retrieval algorithm, and the proposed algorithm has been tested on the database

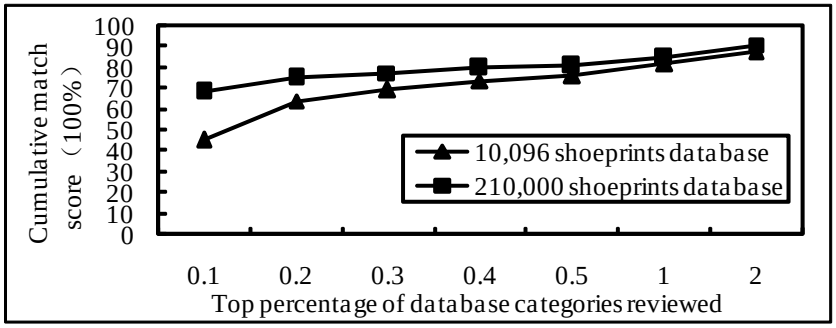

Fig. 7. CMC of the proposed algorithm on two gallery sets

Table 3. Comparison with other algorithms

Methods	Performance	Gallery set description
[4]	88%	145 full-print images with no spatial or rotational variations
[9]	85%(Top 5%)	476 complete shoeprint images belonging to 140 pattern groups
[13]	100%(the first rank)	100 shoeprint images and 6,400 generated images
[2]	99.4%(Gaussian noise with variance 0.01)	500 clear and full shoeprints from shoe manufactures
[8, 7]	91%(Top 6)	35 shoeprints
Ours	45.2%(Top 0.1%) 87.5%(Top 2%)	10,096 crime scene shoeprints
	90.9%(Top 2%) 94.1%(Top 5%)	210,000 crime scene shoeprints

provided by the third party. The gallery database consists of more than 210,000 real crime scene images, and most images in the database are misaligned, incomplete and degraded prints interfered with debris, shadows and other artifacts. The accuracy of the proposed algorithm is more than 90.87% within top two percent of the ranked list of images. The average retrieval time for an image on the gallery set composed of 210,000 images is about 30s on an ordinary PC with a 3.10 GHz CPU and 8GB RAMs. Parameters of the proposed algorithm have learned from huge quantity of crime scene shoeprints and standard shoeprints, and they are applicable in most cases. Therefore, the retrieval results don't depend on the skills of the operators.

Our future work is to increase both the precision rate and the recall rate of retrieving crime scene shoeprints with less reliable information.

Acknowledgement. This research has been supported by the Fundamental Research Funds for the Central Universities.

References

1. Alexander, A., Bouridane, A., Crookes, D.: Automatic classification and recognition of shoeprints. In: Proc. International Conference on Image Processing and Its Applications **2** (1999) 638–641
2. Algarni, G., Amiane,M.: A novel technique for automatic shoeprint image retrieval. Forensic Science International **181** (2008) 10–14
3. Ashley ,W. : What shoe was that? The use of computerized image database to assist in identification. Forensic Science International **82** (1996) 7–20
4. Bouridane, A., Alexander, A., Nibouche, M., Crookes, D.: Application of fractals to the detection and classification of shoeprints. In: Proc. International Conference on Image Processing **1** (2000) 474–477
5. Bouridane, A.: Imaging for forensics and security: from theory to practice. Springer US (2009)
6. Brodatz, P.: Textures: a photographic album for artists designers. Dover, New York (1996)
7. Cervelli,F., Dardi,F., Carrato,S.: An automatic footwear retrieval system for shoe marks from real crime scenes. In: Proc. International Symposium on Image and Signal Processing and Analysis (2009) 668-672
8. Cervelli, F., Dardi, F., Carrato, S.: A texture based shoe retrieval system for shoe marks of real crime scenes. In: Proc. International Conference on Image Analysis and Processing (2009) 384–393
9. Chazal, P. D., Flynn, J., Reilly, R.B.: Automated processing of shoeprint images based on the Fourier transform for use in forensic science. IEEE Transactions on Pattern Analysis and Machine Intelligence **27** (2005) 341–350
10. Geradts, Z., Keijzer, J.: The image data rEBEZO for shoeprint with developments for automatic classification of shoe outsole designs. Forensic Science International **82** (1996) 21–23
11. Girod, A.: Computer classification of the shoeprint of burglar soles. Forensic Science International **82** (1996) 59–65
12. Girod, A.: Shoeprints - coherent exploitation and management In: Proc. the European Meeting for Shoeprint Toolmark Examiners. (1997)
13. Gueham, M., Bouridane, A., Crookes, D.: Automatic recognition of partial shoeprints based on phase-only correlation. In: Proc. IEEE International Conference on Image Processing **4** (2007) 441-444
14. Gueham, M., Bouridane, A., Crookes, D., Nibouche, O.: Automatic recognition of shoeprints using Fourier-Mellin transform. In: Proc. NASA/ESA Conference on Adaptive Hardware and Systems (2008) 487,491,22-25
15. Neelamani, R.N., Hyeokho, C., Baraniuk, R.: Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems. IEEE Transactions on Signal Processing **52** (2004) 418–433
16. Patil, P.M., Kulkarni, J.V.: Rotation and intensity invariant shoeprint matching using Gabor transform with application to forensic science. Pattern Recognition **42** (2009) 1308–1317
17. Pavlou, M., Allinson, N.: Automated encoding of footwear patterns for fast indexing. Image and Vision Computing **27** (2009) 402–409
18. Phillips, P.J., Moon, H., Rizvi, S.A., Rauss, P.J.: The FERET evaluation ethodology for face-recognition algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence **22** (2000) 1090–1104

19. Reddy, B.S., Chatterji, B.N.: An FFT-based technique for translation, rotation, and scale-invariant image registration. IEEE Transactions on Image Processing **5** (1996) 1266-1271
20. Russ, J.C.: The image processing handbook, second ed. CRC Press (1995)
21. Sawyer, N.: SHOE-FIT A computerized shoe print database. In: Proc. European Convention on Security and Detection (1995) 86–89
22. Tang, Y., Srihari, S.N., Kasiviswanathan, H., Corso J.J.: Footwear print retrieval system for real crime scene marks. In: Proc. International Workshop on Computational Forensics (2011) 88–100
23. Xu, C.Y., Prince, J.L.: Snakes, shapes, and gradient vector flow. IEEE Transactions on Image Processing **7** (1998) 359–369