Nibouche, O, Bouridane, A, Gueham, M., and Laadjel, M. (2009) 'Rotation invariant matching of partial shoeprints'. Proceedings of the 13th international machine vision and image processing conference, Monte Carlo Resort, Las Vegas, Nevada, USA , July 13-16, IEEE Xplore, pp. 94-98.

ROTATION INVARIANT MATCHING OF PARTIAL SHOEPRINTS

O Nibouche, A Bouridane, D Crookes, M Gueham and M Laadjel
ECIT Centre, Queen's University of Belfast, Queen's Rd, Belfast BT9 3DT, UK
{o.nibouche, a.bouridane, d.crookes, mgueham02 ,m.laadjel01} @qub.ac.uk

Abstract

In this paper, we propose a solution for the problem of rotated partial shoeprints retrieval based on the combined use of local points of interest and SIFT descriptor. Once the generated features are encoded using SIFT descriptor, matching is carried out using RANSAC to estimate a transformation model and establish the number of its inliers which is then multiplied by the sum of point-to-point Euclidean distances below a hard threshold. We demonstrate that such combination can overcome the issue of retrieval of partial prints in the presence of rotation and noise distortions. Conducted experiments have shown that the proposed solution achieves very good matching results and outperforms similar work in the literature both in terms of performances and complexity.

1. Introduction

As a form of physical evidence, a shoemark, which is a mark made when the sole of a shoe comes into contact with a surface, can provide an important link between the criminals and the place where the crime occurred [1-3]. A shoeprint lifted from a Scene of Crime (SoC) can be checked against a database that includes the prints of shoes in the market to determine its model. It can also be matched against other SoC prints and shoeprints taken from the crime suspects so that a given shoeprint can be identified as being made by a specific shoe. Several techniques and algorithms have been reported in the literature for automatic classification, recognition, indexing and retrieval of shoe prints in the presence of rotation and noise distortions. *Chazal et al* [4] proposed a system for automatically sorting a database of shoeprints based on the outsole patterns in the Fourier domain in response to a reference shoeprint. As shown in [4], the Power Spectral Density (PSD) coefficients of the image are calculated using the Fourier Transform and used as features. A correlation function of the PSD coefficient from a reference database and a query image is used as a similarity metric [4]. To achieve invariance to rotation, matching is also carried out with rotated versions of the query image. It was suggested that the query image should be rotated in the range of [-30 30] degree with a rotation step of one degree to achieve rotation invariance within that range. That leads to matching an additional 60 copies of the same query image to the reference database in a brute force approach that attempts to better the matching score. Such drawback is overcome with the use of correlation filters [5]. Although rotated copies of the images are still used, only the reference images are rotated to generate a unique correlation filter. Still, the designed filter is only robust to rotation within the adopted training range. To achieve a high accuracy, the rotation angle to which the filter is robust in [5] is narrower than in the case of the PSD method [4]. As such multiple filters are required if robustness to a wider angle is to be attained. Multi resolution based techniques have been used in [6], where the radon transform is used to estimate the shoe print rotation angle. A print is divided into none overlapping 16×16-pixel blocks and convolved with an eight-direction Gabor filter bank. The average variance in each block across all Gabor-filtered images is used as a feature map. To insure robustness to partial prints, eight different partial prints are also processed and included in the reference database to create a 9-print class of the same shoe. A similar technique was used in [7] based on the use of directional filter banks. However, in [7], it is the energy within the filtered blocks which is used to build a feature vector. It is not clear if its energy-based features will perform well on a partial print that was not present in the training phase of the techniques in [6-7].

Following their successful use in image retrieval from large databases, model based recognition, object retrieval and texture recognition [8-10], techniques for shoe print image retrieval and classification based on extracting local features were suggested in [11]. *Pavlou et al* presented an efficient automated system for identifying shoe models based on using Maximally Stable Extremal Region (MSER) features which are transformed using SIFT descriptor [10]. Although the SIFT descriptor is rotation invariant, the experiments did not show the performances of the systems against rotation distortions.

In this paper, the issue of automatically classifying shoemarks is addressed. A critical issue that has to be overcome in order to achieve such a goal is the fact that one may have no control over the quality of the shoemarks collected form a SoC or from suspects in police custody [2]. As shown in Figure 1, frequent distortions that a SoC may encompass include partial occlusion, illumination variation, rotation, noise and affine distortions also termed foreshortening caused by nonperpendicular photography [2]. The proposed solution in this paper tackles the issues of rotation and noise distortions in partial prints. The local features are the Harris detector corners. Typically, in a shoe print up to a thousand corners are found using the detector in []. The number of detected points is reduced by creating a 4-level pyramid where a detected point is only taken into account if its Laplacian response is a local maxima in a 3×3×3 neighbourhood. Once the points are selected, the SIFT descriptor provides a rotation invariant representation of shoe prints [10]. Matching is carried out iteratively using RANSAC. Once a transformation model is found, the number of inliers is weighted by the sum of Euclidean distances below a hard threshold.

Figure 1. Left. SOC partial print with scale, rotation and illumination variation. Right. Correct match.

2. Multi-scale Harris detector and SIFT descriptor

A Harris point is any point image where the signal value changes significantly in two dimensions. Conventional “corners”, such as L-corners, T-junctions and Y-junctions, which are all intersections of two edges, satisfy this definition. However, with such a definition, a corner can also be an isolated point or an end of line. A Harris corner can be computed over a local neighbourhood *(x,y)* as a weighted sum of first order derivatives products defined as [12]:

Where the subscript indices x and y indicate a derivative of the image f with respect to the variable x and y, respectively. In practice, the weighting function is Gaussian function of standard deviation which will be noted in the remainder of the paper as . Furthermore, a scale-space representation of the image with scale parameter is defined as [8-9]:

Thus, the normalised *Laplacian of Gaussian (LoG)* of (2) can be expressed as [8-10]:

Mikolajczyk et al have extended the Harris corner detector in [12] to a multi-scale form, which can detect the corners at different scales [8-9]. It takes account of feature detection with automatic characteristic scale selection as shown in [13], where *LoG* has been demonstrated to be successful in scale selection. The multi scale detector, termed Harris-Laplace detector, exploits the high accuracy of location of a Harris corner detector and the robust scale selection of the *LoG* detector [8-10, 13]. It was shown that such detector’s points possess a better repeatability than the SIFT algorithm points while they are more abundant in images than the MSER features [8]. The scale adaptive Harris detector is based on an extension of the matrix A in (1), where, and are the integration and differentiation scales, respectively [8-9]:

The eigenvalues λ_1 and λ_2 of $A(x, y, \sigma_i, \sigma_d)$ characterises the *cornerness* of a given image neighbourhood, which makes Harris points invariant to rotation. The case where λ_1 and λ_2 are both large indicates the presence of a corner. As suggested in [8,12], rather than computing the eigenvalues of $A(x, y, \sigma_i, \sigma_d)$, one can computes the *cornerness* of a point by computing :

$$\mathrm{Cor} = \lambda_1\lambda_2 - \varkappa(\lambda_1 + \lambda_2)^2$$

$$= \det\big(A(x, y, \sigma_i, \sigma_d)\big) - \varkappa\mathrm{trace}^2\big(A(x, y, \sigma_i, \sigma_d)\big) \quad (5)$$

The computation of (5) is simpler than the computation of the eigenvalues of $A(x, y, \sigma_i, \sigma_d)$, where the value of constant $\varkappa$, which is a tunable sensitivity parameter, can be empirically set. Such a multi scale Harris point detector may detect all the corners types of points described previously. The Harris-Laplace detector computes the scale-adapted Harris formula in (5) and selects the points for which the *LoG* in (3) attains a local maximum over scale. It builds a scale-space representation and only selects points which are scale adapted Harris corners and coincide with a *LoG* local maxima at the scale. Although such approach may lead to designing a scale invariant technique, a pyramid with very few levels is considered in this paper. The aim is to reduce the number of detected points by selecting only those which are local maxima.

3. Points detector and descriptor implementation

Based on equations (4) and (5), a 4-level scale-space representation using Harris function has been built. The initial scale and the interval between two successive was set to 1.2 and 1.5, respectively. With such a large interval between two successive levels and few levels built, it is not expected to achieve scale invariance. However, selecting only the points which are local maxima reduces dramatically the number of corners selected to only a few hundreds. The constant $\varkappa$ in equation (5) was set to 0.04. The ratio of the differentiation σ_d scale to integration scaleσ_i was set to 0.7. Harris points of interest are 2-D local maxima; that is a point is selected if it is a maxima in a 3×3 neighbourhood. To remove weak and instable maxima points, only maxima points that are at least 15% of the value of the level absolute maximum are taken into account. The selected Harris points are then checked whether or not their *LoG* response achieves local maxima over scales; that is a *LoG* response of a given Harris point is more important than its adjacent pixels in a 3-D search over a 3×3×3 neighbourhood.

Associated with Harris points detector is a descriptor which provide a hash signature of the neighbourhood of a given point. The SIFT descriptor computes a weighted gradient magnitude histogram of gradient location and orientation in a region surrounding the detected point of interest [10]. First, to assign an orientation to a given point of interest, at the level in the scale-space representation in which the point was detected, a 36-bin gradient histogram covering the 360 degree range of orientations is computed. In the resulting histogram, the absolute peak and any local peak within 80% of its value are selected as orientation angles. This approach together with the subsequent interpolation suggested in [10], lead to creating multiple points in the same space location and scale, though with a different orientation. Thus far, to each point is assigned a spatial location *(x,y)*, a scale σ and an orientation θ. To build a SIFT descriptor; a circular patch centred at a point of interest is selected. The selected neighbourhood is mean and standard deviation normalised. The gradient magnitudes and orientations are sampled around the key point location to a 16 × 16 pixels neighbourhood which is the size of the descriptor window [10]. Such window grid is formed of 4 × 4 blocks each of 4 × 4 pixels. The gradient angle associated with every block is quantised into 8 directions using the gradient magnitude. The resulting 3-D histogram is a 128 dimensional feature vector.

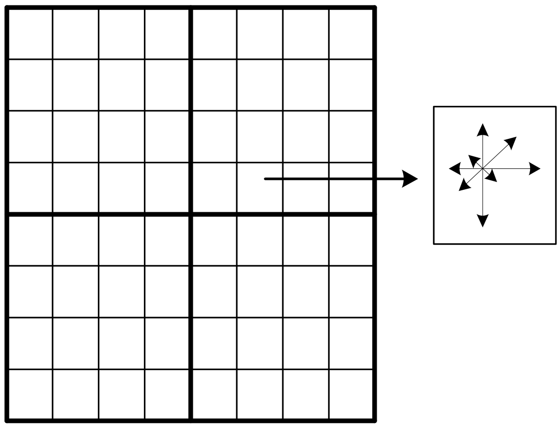

Figure 2. SIFT Descriptor in [10]

Matching is carried out in two steps. First, inliers that belong to a rotation transformation are found. The score from this step is the number of computed inliers that belong to the estimated rotation transformation; that is the number of points in the query image that match other points in a reference image on a point-to-point basis. In the second step, one computes a matrix of point-to-point distances between the reference and query images. Such strategy sums up all distances below a threshold, set in the presented experiments to 0.005. The distance used to build the point-to-point distance matrix is the Euclidean distance of any two points' normalised descriptors. As with RANSAC voting, the highest is the score, the better is the matching. Let be the number of detected points in a query image using the multi-scale Harris detector in equations (5). Similarly, let be the number of detected points in a reference image. A matching score based on the points extracted using the Harris detector can be obtained from a matrix formed from the Euclidean distance d_{ij} elements below a threshold:

(6)

Finally, the matching score is the result of multiplying the number of inliers by the score computed by (6).

4. Experiments and results

Experiments were run on a reference database of 300 shoe prints from Foster & Freeman [3]. To simulate scene of crime prints, degraded images from the reference database were created. Divided into three query databases, the degradations include:

- Rotation distortions
- Noise distortions
- rotation distortions with Gaussian noise perturbations

To simulate partial prints in SoCs, random quarter prints were selected to build the above four query databases. As such, a shoe is divided into its toe and heel parts, which are then divided into a left and right part. Each of the above four test databases was built separately; that is it was not required that all databases should be built from the same partial prints. The selected quarter print is then rotated and/or Gaussian noise added to constitute the above three databases. Figure 3 shows three query images with different amount of added Gaussian noise and their correct match. Each query databases is formed of 300 prints which are matched against the 300 prints in the reference database. Although such approach is not conventional as data is not divided into training data and test data, it is common in local image features literature [10]. It circumvents some very strict data protection regulations in force in the UK. Furthermore, when the proposed solution is compared with similar techniques in the literature, the same test constraints are applied to all of them, making the comparison as fair and extensive as possible. Other ways of building training and test databases can be carried out by asking supposed suspects to provide multiple prints of their shoes, from which few will be selected for training and the remaining prints, which may be further in lab-processed, are used for test. Carried out in a controlled environment, the way in which the prints are collectedly implies good quality prints and does not reveal the performances of the algorithms in the presence of shoe print degradations. Cumulative Matching Characteristic (CMC) curve is used for comparison. Our results are compared against the work in [4] which is a PSD feature based shoeprint matching algorithm. Such technique achieves rotation invariance within a given range. In a brute force matching style, it uses rotated copies of the query print for matching from which the best result of correlations between the query rotated copies and the reference image is taken into account.

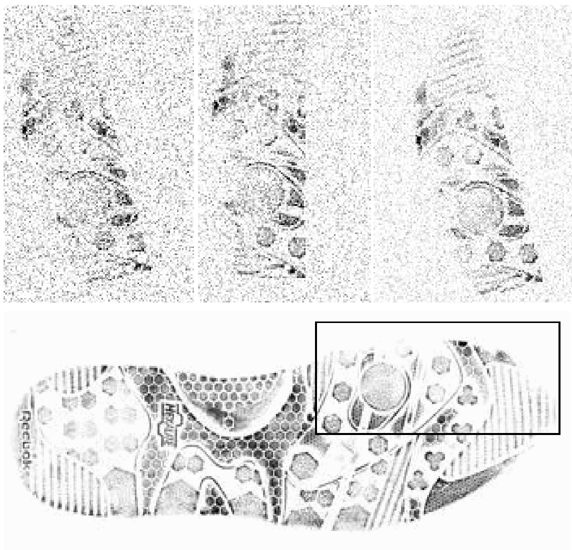

Figure 3. Query prints with different rotation angles and noise levels: Top, partial prints with a noise ratio of: Left 20%, Centre 15%, Right 10%
Bottom, correct match

The first test was carried out on noisy quarters. The Gaussian noise is expressed as the ratio of Gaussian noise variance to the power of the shoe print image. The evaluation of the performances of the proposed technique detailed in Table 1 and Figure 4. It shows that the proposed technique performs better than the PSD method in [4]. As a matter of fact, despite having its performances drop as the level of noise increases, the probability of finding the correct match within the returned top 10 matches is in the worst case about 0.9. Still, in our experiments, the proposed technique clearly outperforms results of the PSD method. This is evidenced in Figure 4 where the CMC performances of the proposed technique with the highest level of noise in our experiment are better than the PSD method wit the lowest level of noise. The goal of the second test was to evaluate the performances of proposed technique against rotation distortions. The performances are measures for a rotation angle between 0 to 30 degrees and then for a rotation angle of 45 degrees, which is outside the range of the PSD method. Even within the range of the PSD method, the proposed technique achieves much better performances. However, when the performances of the PSD drop dramatically for a rotation angle outside its range, the proposed technique retains its invariance to rotation, which is clearly demonstrated in Figure 3 and Table 2. The Third and finale test was carried out on prints that encompass both rotation and noise distortions, where rotation angle were selected randomly between 15 and 30 degrees and the noise levels were set to 10%, 15% and 20%. Once gain, the proposed technique achieves much higher performances than the PSD method as shown in Figure 6 and Table 3. Even with an additive Gaussian noise level of 20%, the CMC of the proposed technique rallies rapidly so that there is a probability of about 82% of finding the correct match within the list of the top 10 returned prints. At this level of noise, the proposed technique performs better than the PSD method at a noise level of only 10%.

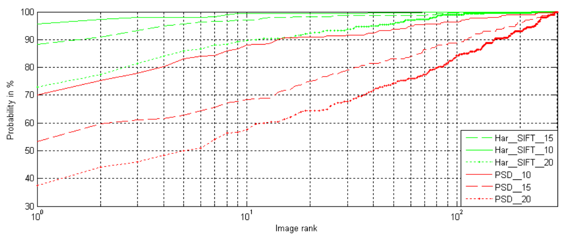

Figure 4. CMC performances for partial prints with noise perturbations

Table 1. Performances evaluation for noisy partial prints.

	1st Rank	3rd Rank	5th Rank	10th Rank
Har_SIFT (10%)	95.67	98	98	99.33
Har_SIFT (15%)	88.33	93.33	95.67	97
Har_SIFT (20%)	73	81.33	86	89.67
PSD (10%)	70	78	83	88
PSD (15%)	53.33	61	62.67	68.33
PSD (20%)	37.33	46	50	57.67

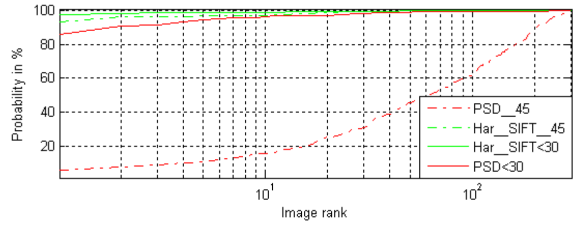

Figure 5. CMC performances for partial prints with rotation distortions.

Table 2. Performances evaluation for rotation-distorted partial prints.

	1st Rank	3rd Rank	5th Rank	10th Rank
Har_SFIT (0°≤ angle≤30°)	97.67	98.67	98.67	99
Har_SIFT (45°)	91	96.33	97	97.33
PSD (0°≤ angle≤30°)	85.67	91.67	94.33	96.67
PSD (45°)	5.67	9	10.67	15.67

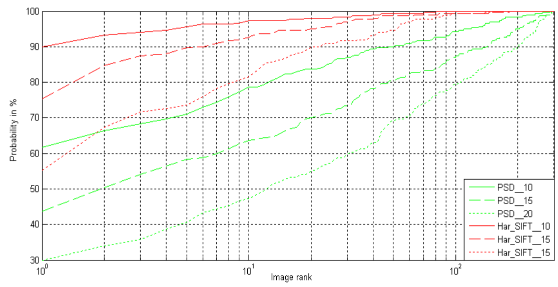

Figure 6. CMC performances for partial prints with rotation and noise distortions.

Table 3. Performances evaluation for noisy and rotated partial prints.

	1st Rank	3rd Rank	5th Rank	10th Rank
Har_SIFT (10%)	90	94	95.67	97.33
Har_SIFT (15%)	75.33	87.33	89.67	92.67
Har_SIFT (20%)	55.33	71.67	73.67	81.67
PSD (10%)	61.67	68.33	71	78.67
PSD (15%)	43.67	54	58.33	63.67
PSD (20%)	30	35.67	40.67	47.33

5. Conclusions

In this paper, we have suggested a technique for retrieval of shoe prints based on combining Harris points and SIFT descriptor. Experiments were conducted on partial synthetic images with rotation and Gaussian noise distortions. The suggested solution in this paper achieves excellent classification performances and outperforms the results of similar work in the literature. It is also faster and much simpler to implement as one no longer requires to rotate the query print to achieve a limited rotation invariance.

6. References

[1] W.J. Bodziak, “Footwear Impression Evidence – Detection, Recovery and Examination”, CRC Press LLC, 2000.

[2] T. Thompson and S. Black, “Forensic Human Identification: An Introduction”, CRC Press, 2007.

[3] Foster and Freeman Ltd, 2009, http://www.fosterfreeman.co.uk, last time visited 15 April 2009.

[4] P. D. Chazal, J. Flynn, and R.B. Reilly, “Automated processing of shoeprint images based on the Fourier transform for use in forensic science,” IEEE Trans. Pattern Analy. Machine Intell., vol. 27, no. 3, pp. 341-350, March. 2005.

[5] M. Gueham, A. Bouridane and D. Crookes, “ Automatic Recognition of Partial Shoeprints Using a Correlation Filter Classifier”, in Proc. of the 2008 Machine Vision and Image Processing International Conference, IMVIP '08, pp: 37-42.

[6] P.M. Patil and J.V. Kulkarni, “Rotation and intensity invariant shoeprint matching using Gabor transform with application to forensic science”, to appear in Pattern Recognition (in Press).

[7] L. Ghouti, A. Bouridane and D. Crookes, “Classification of Shoeprint Images Using Directional Filterbanks”, in Proceedings of the Visual Information Engineering (VIE 2006) Conference, pp. 167-173, Bangalore, India, September 2006.

[8] K. Mikolajczyk and C. Schmid. Indexing based on scale invariant interest points. In Proceedings of the 8th International Conference on Computer Vision, Vancouver, Canada, pages 525–531, 2001.

[9] K. Mikolajczyk and C. Schmid, A performance evaluation of local descriptors, IEEE PAMI 27 (10) (2005), pp. 1615–1630.

[10] David G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, 60, 2 (2004), pp. 91-110.

[11] M. Pavlou and N. M. Allinson, “ Automated encoding of footwear patterns for fast indexing”, to appear in Image and Vision Computing, Volume 27, Issue 4, 3 March 2009, Pages 402-409.

[12] C. Harris and M. Stephens. A combined corner and edge detector,. In Alvey Vision Conference, pages 147–151, 1988.

[13] T. Lindeberg. Feature detection with automatic scale selection. International Journal of Computer Vision, 30(2):79–116, 1998.