A TRANSLATIONAL AND ROTATIONAL INVARIANT DESCRIPTOR FOR AUTOMATIC FOOTWEAR RETRIEVAL OF REAL CASES SHOE MARKS

Federico Cervelli, Francesca Dardi, and Sergio Carrato

Dip. di Elettrotecnica, Elettronica ed Informatica, University of Trieste
Via Valerio 10, 34100, Trieste, Italy
phone: + (39) 040 558 7147, fax: + (39) 040 558 3460, email: federico.cervelli@deei.units.it
web: http://www.units.it/ipl/

ABSTRACT

Shoe marks found on the crime scene are invaluable for the identification of the culprit when no other piece of evidence is available. Thus semi-automatic and automatic systems have been recently proposed to find the make and model of the footwear that left the shoe marks. The systems proposed up to now have two main drawbacks, as they (i) are generally not based on rotation and translation invariant descriptions, and (ii) are tested on synthetic shoe marks, i.e. on shoeprints with added synthetic noise. Here we show the results of a translation and rotation invariant description based on the Fourier transform properties: the test is made on both synthetic and real shoe marks and a comparison with algorithms proposed by others is presented.

1. INTRODUCTION

During the investigation of a crime the forensic experts are in charge of a thorough and detailed analysis of the evidence, as this strongly increases the chances to identify the authors of the crime.

A key element during the investigation is the ability to search specialized archives in order to identify and qualify the evidence, such as fingerprints, DNA sequences, faces, spent cartridges and different chemicals and compounds. Pattern recognition has been employed successfully in many of the aforementioned fields, even on classical one like fingerprint recognition [15].

Recently great attention and research effort have been given to the semi-automatic and automatic retrieval of footwear from the shoe marks, i.e. from the shoeprints found on the crime scene. In particular, shoe marks allow the prosecutor to understand the crime dynamics [13] and, with no suspect or few elements available, the knowledge of the make and of the model of the shoe that left the shoe mark is a precious information for police officers to lead investigation.

In this paper we will give a review of the state of the art on automatic footwear retrieval systems (Sec. 2), then we will give a brief description of the description used and the similarity measure employed in this work (Sec. 3), and we will show the results obtained testing this as well as other systems on both synthetic and real shoe marks (Sec. 4). Finally, we will (Sec. 5) draw the conclusions and point out the future work.

2. STATE OF THE ART

Even using a fully automated retrieval system, is a general Law requirement that the forensic expert is in charge of his or her forensic analysis and, after the query to the reference database, the expert will proceed with the eye-analysis of the higher ranking results. The first systems to be reported were human based ones [12, 19, 2], i.e. human experts classified the shoe soles and shoe marks using a series of patterns chosen in a special vocabulary defined on purpose.

Here we give an overview of the automatic footwear retrieval systems that have been proposed since then.

In [11] soles and shoeprints are first photographed and binarized. Then labeling is performed and the binary pixels are grouped into shapes. The shapes are classified by their Fourier transform components and finally a single-hidden-layer feedforward neural network trained with back propagation is used as the recognition network. Regrettably the performance of the system is not reported and its development was abandoned.

Fractals and mean square noise error are employed in [3] to represent and compare shoeprints, respectively. Fractal decomposition produces a list of spatial transformations that regenerate the original image when applied recursively to the image itself, thus producing small changes when applied to look alike and causing large differences when applied to different patterns. The mean square noise error is then used as the similarity measure. The reference database (DB) is composed of 145 gray level shoeprint images. Tests are made adding Gaussian noise, rotation and translation to the DB. Reported results give 60% correct match at 9° rotation and a severe drop to 10% after a 10 pixel translation. No results are given on noise performance.

Simple Fourier transform is implemented in [8]. The power spectral density (PSD) is calculated and used to characterize the images, in order to have translational invariance, while rotational invariance is obtained brute-force through rotation of the query image in the range $\pm 30^\circ$. The 2D correlation coefficient between the PSDs is used as the similarity measure. Average Match Score and Cumulative Matching Score are used to optimize the system parameters and to assess it, respectively. Results show that shoeprints are correctly matched in the first 5% of the sorted DB patterns with an 85% score. Here noisy images are not considered.

Fourier transforms modified phase only correlation (MPOC) was employed in [14], given that the phase information is much more important than the FFT amplitude in preserving the features of image patterns. The reference DB is made of 100 elements, used to generate synthetic shoe marks. Four sets of synthetic shoe marks are created by adding white Gaussian noise or applying motion blur. Synthetic shoe marks are also obtained by pasting the shoeprints into texture images taken from the Brodatz album [4]. The test demonstrates a 100% first rank recognition rate, but the

© EURASIP, 2010 ISSN 2076-1465

system is not invariant under translation or rotation.

To address translation, rotation and scale invariance, Hu's seven moments were used in [1]. Hu's moments are employed on a reference DB containing 500 shoeprints, showing that the first and the second moments are the most discriminating. Synthetic shoe marks are generated adding zero mean Gaussian noise, and tests are made with a rotation from -90° to +90°. Results show a severe drop of accuracy, which is as low as 5.4% for a 20% Gaussian noise variance.

A Maximally Stable Extremal Region (MSER) detector is used in [17] to identify the features of the shoeprint and the Scale Invariant Feature Transform (SIFT) algorithm is employed to describe them. The combination of MSER an SIFT gives the method a good repeatability under affine transformations, together with the possibility to deal with partial or hidden data. The reference DB is made of 374 shoeprints, each 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 report a 94% classification rate if viewing only 5% of the database, but again no test is performed on noisy images.

Finally an image retrieval algorithm combining the information of the phase of the Fourier transform of the shoe mark images with the power spectral density of the Fourier transform calculated on their Mahalanobis map is employed in [6, 7]. The reference DB is made of 35 shoeprints and the system is tested on synthetic as well as on real shoe marks coming from crime scenes. A comparison with other methods is made and the results show that the performance of the system is on par, if not better than the other, with the 100% of the real shoe marks found in only 24% of the known shoes DB.

In this paper we make a comparison of the most successful methods described above (i.e. the ones in [8], [14], and [7]) with a method based on the Fourier transform properties, on both synthetic and real shoe marks coming from crime scenes[1].

3. FOURIER BASED DESCRIPTION AND SIMILARITY MEASURE

In this section, we present the needed tools to match images which are translated and rotated with respect to each other, using the translation and rotation properties of the Fourier transforms.

Let $f_1(x,y)$ and $f_2(x,y)$ be two images differing only by a displacement (x_0,y_0), so that:

$$f_2(x,y) = f_1(x-x_0, y-y_0). \tag{1}$$

Their Fourier transforms $F_1(\xi,\eta)$ and $F_2(\xi,\eta)$ are related by:

$$F_2(\xi,\eta) = e^{-j2\pi(\xi x_0+\eta y_0)} F_1(\xi,\eta) \tag{2}$$

and the cross-power spectrum $Q_{1,2}$ of $f_1(x,y)$ and $f_2(x,y)$ is given by:

$$Q_{1,2}(\xi,\eta) = \frac{F_1(\xi,\eta)F_2^*(\xi,\eta)}{\left|F_1(\xi,\eta)F_2^*(\xi,\eta)\right|} = e^{j2\pi(\xi x_0+\eta y_0)} \tag{3}$$

where $F_2^*(\xi,\eta)$ is the complex conjugate of $F_2(\xi,\eta)$.

[1]We preliminary tested [6] a SIFT matching algorithm [16], similar to the one employed in [17], with poor results.

Thus, if $f_1(x,y)$ and $f_2(x,y)$ are related only by a translation, the inverse Fourier transform of $Q_{1,2}(\xi,\eta)$ is a pulse which is zero everywhere except nearby the point of maximum (x_0,y_0), which also gives the displacement between $f_1(x,y)$ and $f_2(x,y)$.

Now, suppose $f_2(x,y)$ is a translated and rotated version of $f_1(x,y)$, such that:

$$f_2(x,y) = f_1(cos\theta_0 x + sin\theta_0 y - x_0, -sin\theta_0 x + cos\theta_0 y - y_0). \tag{4}$$

Given the Fourier transform rotation property we have that $F_1(\xi,\eta)$ and $F_2(\xi,\eta)$ are related by:

$$\begin{aligned} F_2(\xi,\eta) = & e^{-j2\pi(\xi x_0+\eta y_0)} \\ & F_1(cos\theta_0\xi + sin\theta_0\eta, -sin\theta_0\xi + cos\theta_0\eta) \end{aligned} \tag{5}$$

If we look at the magnitudes $P_1(\xi,\eta)$ and $P_2(\xi,\eta)$ of $F_1(\xi,\eta)$ and $F_2(\xi,\eta)$, respectively, the following relation holds:

$$P_2(\xi,\eta) = P_1(cos\theta_0\xi + sin\theta_0\eta, -sin\theta_0\xi + cos\theta_0\eta) \tag{6}$$

i.e. the magnitude of the second image is the rotated replica of the magnitude of the second one.

If the coordinates of the obtained magnitudes are transformed into polar coordinates $(\xi,\eta) \rightarrow (\rho,\theta)$, with $\rho = \sqrt{\xi^2+\eta^2}$ and $\theta = \arctan(\eta/\xi)$ we have:

$$P_2(\rho,\theta) = P_1(\rho,\theta-\theta_0) \tag{7}$$

and the rotation turns to be a translation θ_0 in the θ coordinate.

Following again the procedure described in the case of pure translation, we can calculate the cross-power spectrum $\tilde{Q}_{1,2}(\xi',\eta')$ between the Fourier transform of $P_1(\rho,\theta)$ and $P_2(\rho,\theta)$. Its inverse Fourier transform will be a pulse like function with its maximum located at the displacement $(0,\theta_0)$.

In an ideal case the height of the peak would be near one and would be departing from this value for different images. Therefore, in this work we consider the peak height of the inverse Fourier transform of $\tilde{Q}_{1,2}(\xi',\eta')$ as the similarity measure for image matching: if two images are similar, their inverse Fourier transform will have a distinct sharp peak, if they are not similar, the peak will drop significantly.

The steps described above give a translation and rotation invariant description and a similarity measure. The procedure could be made more general including scale invariance, performing a Fourier-Mellin transform [5]. However, the reference DB specifications are known and during the crime scene analysis all evidence is fully documented within a metrical context, thus scale should not be an issue.

All the tests described below were made following the above mentioned procedure using standard MATLAB [21] commands, without any fine tuning.

3.1 Spectral Weighting Functions

We performed some tests on the system performance multiplying the Fourier magnitudes $P_1(\xi,\eta)$ and $P_2(\xi,\eta)$ with two different filters, i.e. a high-pass emphasis filter $H(\xi,\eta)$ and a band-pass function $W_\beta(\xi,\eta)$.

The high-pass filter [20] nulls the contribution of the frequencies around $\rho = 0$ that get oversampled due to the Cartesian to polar coordinate transformation. The filter is given by:

$$X(\xi,\eta) = \cos(\pi\xi)cos(\pi\eta) \quad (8)$$
$$H(\xi,\eta) = (1 - X(\xi,\eta))(2 - X(\xi,\eta)) \quad (9)$$

with ξ and η between -0.5 and 0.5 (in unit frequency).

The band-pass weighting function [14] has the same shape as the spectrum of a Laplacian of Gaussian (LoG) function and is used to eliminate both high and low frequency components. The filter is given by:

$$W_\beta(\xi,\eta) = \frac{\xi^2+\eta^2}{\alpha} \exp\left(-\frac{\xi^2+\eta^2}{2\beta^2}\right) \quad (10)$$

where β controls the function extension and $\alpha = 4\pi\beta^4$ is used for normalization.

During our test we tried them separately and together, with different values of β, ranging from 50 to 150.

4. SETUP OF THE TEST

We have built a reference DB (*RDB*) made of known shoeprints, realized starting from the images available at the ENFSI [10] WGM website [9] and then adapted as explained in Sec. 4.1

Two different sets of shoe marks are used in our test to query the *RDB*:

- a synthetic shoe marks set (*SyntS*), realized as detailed in Sec. 4.2;
- a real shoe marks set (*RealS*), realized starting from the images available at the ENFSI WGM website [9] and then adapted as explained in Sec. 4.1.

4.1 Reference DB and real shoe marks set

The images were first converted to gray scale and then resized to 512×512 pixel size, to obtain a reference DB made of full shoeprints, Fig. 2.A).

To allow the comparison with previous works [8, 14, 7], three zones of interest of 100×100 pixel size were cropped from each shoeprint and used for the test, Fig. 1. Cropping was performed also for the crime scenes shoe marks in order to increase their number for testing purposes.

Following this procedure and starting from the images available at the ENFSI WGM website, we built: a reference DB made of 25 known shoeprints, a reference DB made of 75 known shoeprints crops, and a set of real shoe marks made of 35 items (corresponding to 16 known shoeprints).

Then, we have queried each element of the *RealS* against those in the cropped shoeprints *RDB*, and against those in the full shoeprints *RDB*, using the Fourier based invariant description employed in this work and using the methods described in [8], in [14] and in [7]. Finally results were compared in the form of top-one, top-five, top-ten and top-twenty ranks.

4.2 Synthetic shoe marks set

We want to compare the method used in this paper with the ones employed in other works, but the field still lacks a standard database for comparisons; moreover, all other existing works only consider synthetic shoe marks.

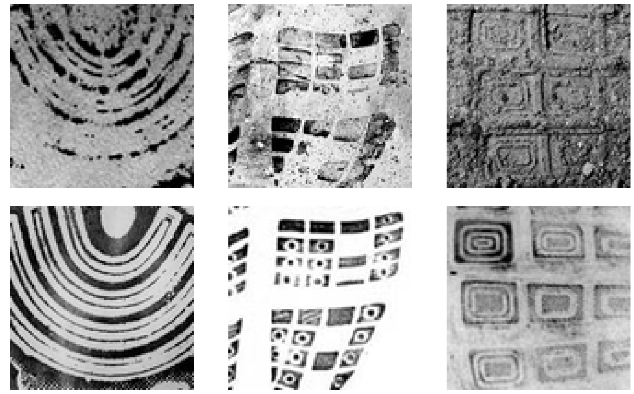

Figure 1: Examples of shoe marks (top), and matching shoeprints (bottom).

Thus, starting from the *RDB*, described in Sec. 4.1, two subsets of synthetic shoe marks were created [14]:

- Subset 1: 300 shoe marks obtained by adding to each shoeprint in the *RDB* a white gaussian noise with zero mean and variance σ^2 = 1, 5, 10 and 15 %;
- Subset 2: 375 shoe marks obtained by blurring each shoeprint in the *RDB* with a motion blur of d = 2, 5, 10, 15 and 20 pixel.

Some samples of Subset 1 are shown in Fig. 2 for gaussian noise.

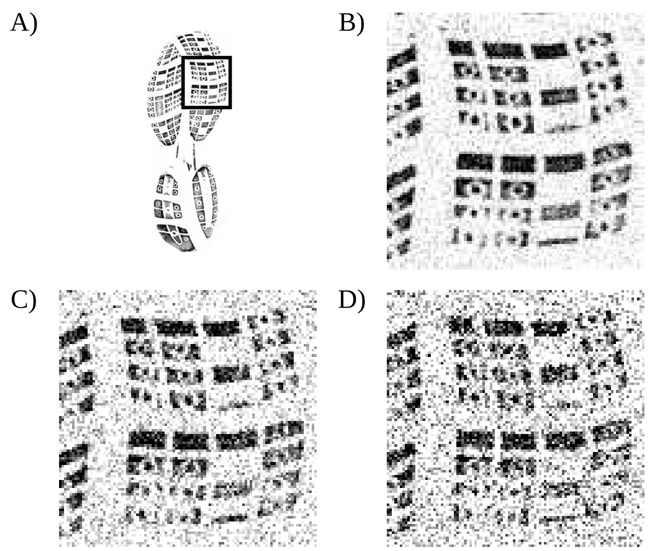

Figure 2: Reference DB sample and synthetic shoe marks obtained by adding increasing white Gaussian noise to it: A) original image, B) σ^2 = 1%, C) σ^2 = 5%, D) σ^2 = 10%.

Again, we have queried each element of the *SyntS* against those in the cropped shoeprints *RDB*, and against those in the full shoeprints *RDB*, using the Fourier based invariant described in this paper and employing the methods described in [8], in [14] and in [7]. We then compared the results in the form of top-one, top-five, top-ten and top-twenty ranks.

SyntS	Algorithm															
	PSD				MPOC				Mahalanobis				This Work (Crops)			
	1	5	10	20	1	5	10	20	1	5	10	20	1	5	10	20
1%	99	100	100	100	100	100	100	100	95	99	100	100	100	100	100	100
5%	82	89	93	98	100	100	100	100	78	93	99	99	92	94	95	95
10%	71	75	82	86	100	100	100	100	74	86	92	98	84	86	87	88
15%	67	72	74	76	100	100	100	100	–	–	–	–	71	73	76	80
2 pxl	98	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
5 pxl	95	100	100	100	100	100	100	100	87	98	99	100	100	100	100	100
10 pxl	60	69	75	79	100	100	100	100	39	74	79	93	100	100	100	100
15 pxl	68	82	85	89	96	100	100	100	–	–	–	–	95	100	100	100
20 pxl	34	42	60	74	100	100	100	100	8	38	52	72	98	100	100	100

Table 1: Results for synthetic shoe marks queried on the cropped shoeprints *RDB* for the different methods. Shown is the percentage of shoe marks that are correctly matched at the first place and in top-five, top-ten and top-twenty rank.

SyntS	This Work							
	Crops				Shoes			
	1	5	10	20	1	5	10	20
1%	100	100	100	100	100	100	100	100
5%	92	94	95	95	93	94	97	99
10%	84	86	87	88	79	82	87	93
15%	71	73	76	80	58	65	70	92
2 pxl	100	100	100	100	100	100	100	100
5 pxl	100	100	100	100	97	97	99	99
10 pxl	100	100	100	100	75	79	88	97
15 pxl	95	100	100	100	68	72	82	93
20 pxl	98	100	100	100	53	62	74	88

Table 2: Results for synthetic shoe marks queried on the cropped shoeprints *RDB* (left) and on the full shoeprints *RDB* (right) for the proposed invariant method. Shown is the percentage of shoe marks that are correctly matched at the first place and in top-five, top-ten and top-twenty rank.

5. RESULTS AND DISCUSSION

We first tested the performance of the system when querying the cropped *RDB* with the *SyntS*, and make a comparison with the following methods[2]: MPOC [14], PSD [8] and Mahalanobis [7]. Results are shown in Table 1.

As can be seen from the table, the invariant description employed performs better than the PSD and Mahalanobis methods, but doesn't reach the 100% score of the MPOC case, although the latter is not translation and rotation invariant.

We then tested the system when querying the full shoeprints *RDB* with the *SyntS*. The results are shown in Table 2.

As can be seen querying the full shoeprints causes a degradation of the results, especially when applying motion blur. In Fig. 3 the results in the case of a $d = 10$ pxl synthetic motion blur are shown in the form of cumulative matching characteristics curves [18]: the horizontal axis of the graph is the percentage of the reference DB reviewed and the vertical axis is the probability of a match.

We then tested the method employed in this work querying the full shoeprints *RDB* with the shoe marks coming from the crime scenes. During this test we tried several combinations of the high-pass emphasize filter and the weight function, with β values ranging from 50 to 150. The cumulative match score results are shown in Fig. 4, which shows the best results are obtained when using both the high-pass emphasize filter $H(\xi,\eta)$ and the weight function $W_{\beta=150}(\xi,\eta)$.

[2]Both MPOC and PSD were implemented by us on purpose.

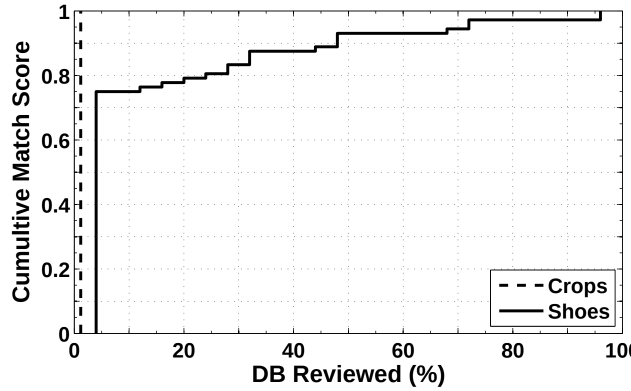

Figure 3: Results for synthetic shoe marks in the case of d = 10 pxl motion blur, queried on the cropped *RDB* (- -) and on the full shoeprints *RDB* (—).

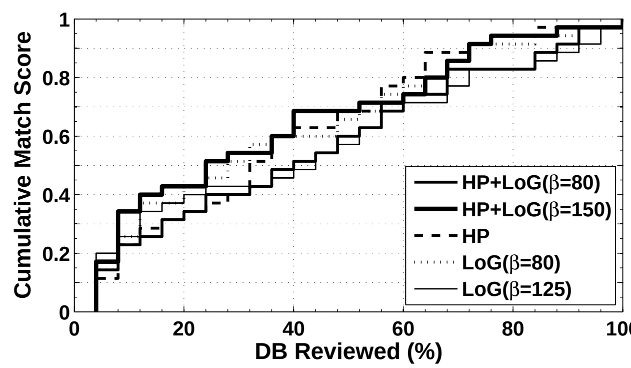

Figure 4: Results for real shoe marks as a function of the combinations of the high-pass emphasize filter (HP) and the weight function (LoG(β)).

Once chosen the best combination, we evaluated the employed system in comparison with the other MPOC, PSD and Mahalanobis methods, Fig. 5.

As can be seen, the translational and rotational invariance has a high price to be paid: analyzing the first 5% of the query results the probability to find the correct shoeprint is 50% for both MPOC and Mahalanobis methods, while it lowers to approximately 17% using the Fourier based method employed in this work.

6. CONCLUSIONS

In this paper we studied a system for the automatic retrieval of footwear that left the shoe marks found at the crime scene. The system exploits the translation and rotation properties of the Fourier transform to realize a translation and rotation invariant description suitable for image matching.

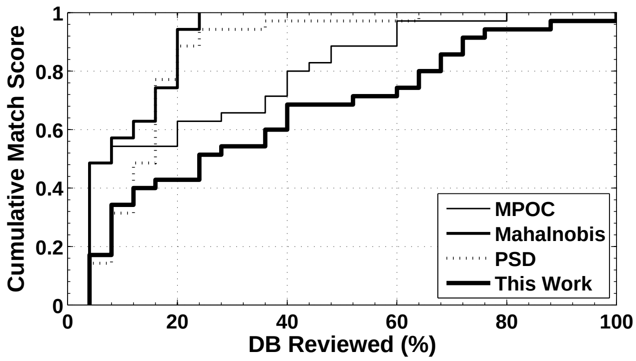

Figure 5: Results on real shoe marks for different systems.

Several different database where created: a reference DB of cropped shoeprints, a reference DB of full shoeprints, a synthetic shoe marks DB created adding gaussian noise or applying motion blur to the reference DB items, and a DB with real shoe marks coming from real crime scenes.

The synthetic DB has been tested both on a reference DB of cropped shoeprints, to allow a comparison to the performance of other published systems, and on the full shoeprints, to take advantage of the translation and rotation invariance.

We fine tuned the system with a combination of weighting functions, and then made a comparison of the method employed in this work and other methods available in literature, namely the MPOC, PSD and Mahalanobis methods.

Results show that the system performance degrades with the noisy real shoe marks, and its translational and rotational invariance has a price to be paid: analyzing the first 5% of the query results the probability to find the correct shoeprint is 50% for both MPOC and Mahalanobis methods, while it lowers to approximately 17% using the Fourier based method employed in this work.

Despite the results, the translation and rotation invariance of the system would make it more suitable in real cases, where the uncertainty of the aforementioned parameters would make both the MPOC and the Mahalanobis based systems less effective, if not useless. We are currently setting up a new test where both rotation and translation will be addressed on an improved database with more than 300 shoe sole images.

Future work will then be devoted to the study of a suitable pre-processing of the shoe mark images in order to increase the system performance, and other translation and rotation invariant descriptions will be analyzed and included for testing purposes as well.

REFERENCES

[1] G. Algarni and M. Amiane. A novel technique for automatic shoeprint image retrieval. *Forensic Sci. Int.*, 181:10–14, 2008.

[2] W. Ashley. What shoe was that? The use of computerized image database to assist in identification. *Forensic Sci. Int.*, 82:7–20, 1996.

[3] A. Bouridane, A. Alexander, M. Nibouche, and D. Crookes. Application of fractals to the detection and classification of shoeprints. In *Proc. Int. Conf. Image Processing*, volume 1, pages 474–477, 2000.

[4] P. Brodatz. *Textures: a photographic album for artists designers*. Dover, New York, 1966.

[5] D. Casasent and D. Psaltis. Position, rotation, and scale invariant optical correlation. *Appl. Optics*, 15:1795–1799, 1976.

[6] F. Cervelli, F. Dardi, and S. Carrato. A texture based shoe retrieval system for shoe marks of real crime scenes. In *Proc. Int. Conf. Image Analysis and Processing*, pages 384–393, 2009.

[7] F. Dardi, F. Cervelli, and S. Carrato. An automatic footwear retrieval system for shoe marks from real crime scenes. In *Proc. Int. Symp. Image and Signal Processing and Analysis*, pages 668–672, 2009.

[8] P. De Chazal, J. Flynn, and R. Reilly. Automated processing of shoeprint image based on the Fourier transform for use in forensic science. *IEEE Trans. Pattern Analysis Machine Intelligence*, 27:341–350, 2005.

[9] ENFSI Working Group on Marks. Website: www.intermin.fi/intermin/hankkeet/wgm/home.nsf/.

[10] European Network of Forensic Science Institutes. Website: www.enfsi.eu.

[11] Z. Geradts and J. Keijzer. The image-database REBEZO for shoeprint with developments for automatic classification of shoe outsole designs. *Forensic Sci. Int.*, 82:21–31, 1996.

[12] A. Girod. Computerized classification of the shoeprints of burglars' shoes. *Forensic Sci. Int.*, 82:59–65, 1996.

[13] A. Girod. Shoeprints: coherent exploitation and management. In *European Meeting for Shoeprint Toolmark Examiners*, The Netherlands, 1997.

[14] M. Gueham, A. Bouridane, and D. Crookes. Automatic recognition of partial shoeprints based on phase-only correlation. In *Proc. Int. Conf. Image Processing*, volume 4, pages 441–444, 2007.

[15] D. Maltoni, D. Maio, A. Jain, and S. Prabhakar. *Handbook of fingerprint recognition*. Springer, second edition, 2009.

[16] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. V. Gool. A comparison of affine region detectors. *Int. J. Computer Vision*, 65:43–72, 2005.

[17] M. Pavlou and N. Allinson. Automated encoding of footwear patterns for fast indexing. *Image Vision Computing*, 27:402–409, 2009.

[18] P. Phillips, H. Moon, S. Rizvi, and P. Rauss. The FERET evaluation methodology for face-recognition algorithms. *IEEE Trans. Patt. Anal. Mac. Intell.*, 22:1090–1104, 2000.

[19] N. Sawyer. SHOE-FIT a computerized shoe print database. In *Proc. Eur. Convention Secur. Detect.*, pages 86–89, 1995.

[20] B. Srivansa Reddy and B. N. Chatterji. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Trans. Image Proc.*, 5:1266–1271, 1996.

[21] The Mathworks. Website: www.mathworks.com.