Automatic extraction and classification of footwear patterns

Maria Pavlou and Nigel M. Allinson

University of Sheffield, Sheffield, S1 3JD, UK
{m.pavlou, n.allinson}@sheffield.ac.uk

Abstract. Identification of the footwear traces from crime scenes is an important yet largely forgotten aspect of forensic intelligence and evidence. We present initial results from a developing automatic footwear classification system. The underlying methodology is based on large numbers of localized features located using MSER feature detectors. These features are transformed into robust SIFT or GLOH descriptors with the ranked correspondence between footwear patterns obtained through the use of constrained spectral correspondence methods. For a reference dataset of 368 different footwear patterns, we obtain a first rank performance of 85% for full impressions and 84% for partial impressions.

1 Introduction

Recent changes of UK police powers allows for collected footwear marks and evidence to be treated in the same way as fingerprint and DNA evidence. This generally untapped forensic source can be used to identify linked crime scenes, can link suspects in custody to other crime scenes and can sometimes provide strong courtroom evidence. Footwear evidence is quite common at crime scenes, frequently more so than finger prints [1], and of which approximately 30% is usable for forensic purposes [2]. In the UK the recovery rate of footwear evidence from crime scenes is expected to increase greatly from the current average of 15%. Changes in police procedures are expected to expand the current work load and there is need for practical systems to allow effective matching of footwear patterns to national databases. The provision of the underpinning technology is the focus of this study.

Automatic matching of footwear patterns has been little explored in the literature. Early works [2, 4–7] have employed semi-automatic methods of manually annotated footwear print descriptions using a codebook of shape and pattern primitives, for example, wavy patterns, geometric shapes and logos. Searching for an example print then requires its encoding in a similar manner as that used for the reference database. This process is laborious and can be the source of poor performance as similar patterns may be inconsistently encoded by different users. One automated approach proposed in [3] employs shapes automatically generated from footwear prints using various image morphology operators. The spatial positioning and frequencies of these shapes are used for classification with a neural network. The authors did not report any performance statistics for their system. Work in [8, 9] makes use of fractals to represent the footwear prints and a mean square noise error method is used for classification. They report a 88% success in classifying 145 full-print images with no spatial or rotational variations. More recently in [10], Fourier Transforms (**FT**) are used for the classification of full and partial prints of varying quality. The **FT** provides invariance to translation and rotation effects and encodes spatial frequency information. They report first rank classification results of 65% and 87% for rank 5 on full-prints. For partial prints, a best performance of 55% and 78% is achieved for first and fifth ranks respectively. Their approach is promising and shows the importance of encoding local information. Although the footwear prints are processed globally they are encoded in terms of the local information evident in the print. Finally in [11] pattern edge information is employed for classification. After image de-noising and smoothing operations, extracted edge directions are grouped into a quantized set of 72 bins at 5 degree intervals. This generates an edge direction histogram for each pattern which after applying a Discrete **FT** provides a description with scale, translational and rotational invariance. On a dataset of 512 full-print patterns which were randomly noised (20%), randomly rotated and scaled they achieve rank 20 classification of 85%, 87.5%, and 99.6% for each variation group. Their approach deals well with these variations and a larger dataset, however their query examples originate from the learning set and no performance statistics are provided for partial prints.

2 Approach

From discussion with police forces, two main aspects of footwear processing have been identified. The first regards the automatic acquisition, categorization/encoding and storage of footwear patterns at police custody suites. The second is the identification and verification of scene evidence with stored reference samples or with other scene evidence.

Our work has initially approached the task of footwear categorization and encoding. In this case footwear patterns of good quality can easily be obtained and digitized either from scanning prints from specialist paper or by directly scanning or imaging. The former is a good representation and is similar to ones obtained from a scene. The direct scan or image can however contain more information but is slow and suitable equipment is not always easily available.

2.1 Local Feature Detection and Description

Research on covariant region detectors and their descriptors is now well advanced and have been used extensively as building blocks in general recognition systems. Based on recent research [12,13] it is possible to select a number of affine invariant feature extractors suitable for footwear patterns. From these studies the Harris-Affine (**HA**) corner detector and Maximally Stable Extremal Region (**MSER**) detector are identified as being robust and having a high repeatability

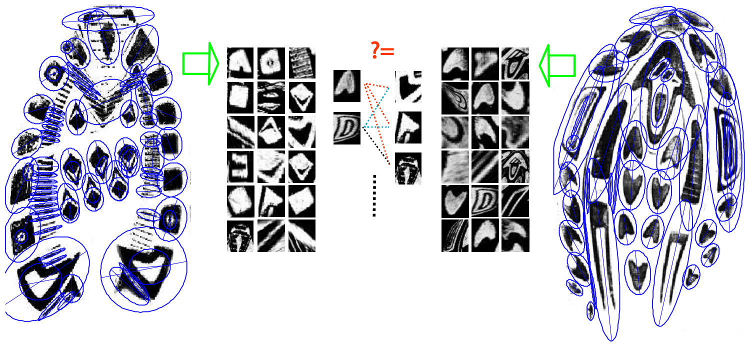

Fig. 1. Pair of footwear prints and some of their detected MSER features

under varied conditions such as affine transformations and image degradations (lighting, blurring, etc.).

The **MSER** detector (a watershed-based segmentation algorithm) performs well on images containing homogeneous regions with distinctive boundaries. Near-binary images of footwear patterns exhibit this set of characteristics. The **HA** detector provides a higher number of affine stable regions centered on corner features. These features sets are complementary as they have different properties and their overlap is usually small if not empty. Additionally their abundance is useful in matching images with occlusion and clutter. The **HA** detector can be used for footwear verification as its properties are well suited to corner-like features such as small cuts and grooves which are abundantly found in footwear patterns and can be the 'unique' features need to provide courtroom evidence. As a first step however, for footwear pattern matching and classification the **MSER** detector is employed as it is better suited for discriminating general patterns or shapes of footwear marks into classes.

Once a number of features have been found a suitable feature descriptor is needed to code the appearance or properties of the local features. In [14] the performance of a number of feature descriptors was evaluated using the above feature detectors and others. In most of the tests performed, the Gradient Location and Orientation Histogram (**GLOH**) descriptor provided the best results, closely followed by the Scale Invariant Feature Transform (**SIFT**) descriptor. The **SIFT** descriptor, computed for a normalized image patch, is a 3D gradient location and orientation histogram constructed using 8 quantized orientations and a 4×4 patch location grid. The resulting descriptor is of dimension 128. The **GLOH** is an extension of the **SIFT** descriptor designed to increase robustness and distinctiveness using a log-polar location grid with 3 bins in radial direction and 8 in angular direction. The gradient orientations are quantized into 16 bins. This gives a 272 bin histogram which is reduced in size using PCA to 128 dimensions.

2.2 Feature Matching

A combination of the above studies suggests that good matching performance is possible using **MSER** features encoded with **SIFT** or **GLOH** descriptors. Given images of footwear patterns, then verifying their similarity can be achieved by finding their matching features (Figure 1). The similarity of features can be determined using a suitable metric. In our case a Gaussian weighted similarity metric has been used as this allows a similarity threshold to be easily set. However, matching on descriptors alone is not sufficient as some features may be mismatched or a many-to-many mapping of features may occur. Furthermore, different footwear patterns may contain very similar features and so further steps are required to disambiguate matches. These steps depend on the application, but generally use methods of geometric filtering based on the local spatial arrangement of the regions.

2.3 Spectral Correspondence Matching with Constraint Kernels

The problem of finding feature correspondence between two or more images is well known and is of crucial importance for many image analysis tasks. A number of techniques can be used to tackle this problem and can be broadly categorized into three groups based on their application and approach. These are Point Pattern Matching, Graphical Models and Spectral Methods.

Point Pattern Matching attempts to decide whether a pattern or spatial arrangement of points appears in an image. This involves the matching of isometries where a mapping is sought which transforms the query pattern onto a gallery pattern. Graphical Models also find mappings of graph structures and are based in representations of factored joint probability distributions. Both approaches and early Spectral Methods have been successful in graph matching problems. However they lack the ability to incorporate additional properties of the points being matched.

A simple and direct approach of associating features of two arbitrary patterns was proposed by Scott and Longuet-Higgins [16]. Applying singular value decomposition (SVD) to a suitable proximity matrix of feature locations it is possible to find good correspondences. This result stems from the properties of the SVD to satisfy *exclusion* (one-to-one mappings) and *proximity* principles [17]. One of the limitations of such spectral methods is their particular susceptibility to the effect of size differences between point samples and structural errors. To improve performance Pilu [15] included a feature similarity constraint based on the local gray patches around any feature point.

Similarly a number of feature similarity constraints are used in this work. As a first step a basic feature similarity constraint is enforced. Assume that two pattern images $\mathcal{I}_A$, $\mathcal{I}_B$ are given along with their set of feature descriptors $\mathcal{F}_A$ and $\mathcal{F}_B$. By employing a Gaussian function,

$$K(i,j) = e^{\frac{-\|\mathcal{F}_i - \mathcal{F}_j\|^2}{2\sigma^2}} \tag{1}$$

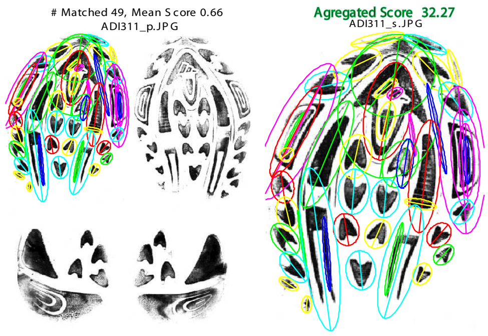

Fig. 2. An example of a matched partial to its reference image. Corresponding coloured ellipses indicate matched features with corresponding neighbourhoods

between every pairing of the features in each image a Gaussian feature similarity matrix $G^{\mathcal{F}}_{ij}$ can be formed. Multiplying $G^{\mathcal{F}}$ with the Gaussian proximity matrix $G^{\mathcal{D}}$ (based on the coordinate positions of features) results in a similar formulation to that used in [15]. Applying the SVD based algorithm of [16] at this point gives a high proportion matching of features. However, as the number of features can be high, strong correspondences cannot be found.

To enhance the performance of the algorithm, locality and neighbourhood constraint kernels are also applied. This allows the SVD algorithm to consider features which are strongly matched in terms of their neighbours. The neighbourhood constraint kernel $G^{\mathcal{N}}$ enforces features matches whose neighbouring features are also similar. Constructing $G^{\mathcal{N}}$ is straight forward since $G^{\mathcal{F}}$ has already been obtained. All that is required is to construct an index list h^i of suitable neighbours for every feature i. Using feature coordinates the nearest N surrounding features are selected as neighbours. $G^{\mathcal{N}}$ can now be constructed as follows:

$$G^{\mathcal{N}}_{ij} = \frac{1}{N} \sum_{p,q=1}^{N} G^{\mathcal{F}}_{h^i_p, h^j_q}. \tag{2}$$

The locality constraint kernel enforces feature pairings whose neighbouring features are similarly positioned around the central feature. This positioning is defined in terms of the angle between neighbours relative to the nearest neighbour. For each feature i the angle to its first nearest neighbour is found and the remaining neighbours angles,θ, are recorded relative to the first. The locality constraint kernel is then constructed as follows:

$$G^{\mathcal{L}}_{ij} = \frac{1}{N-2} \sum_{p=1}^{N-2} e^{\frac{\|\theta^i_p - \theta^j_p\|^2}{2\sigma^2}}. \tag{3}$$

With a suitable selection of σ's for $G^{\mathcal{D}}$, $G^{\mathcal{F}}$ and $G^{\mathcal{L}}$ based on the maximum feature distance, similarity and neighbour angle deviations a final constrain ma-

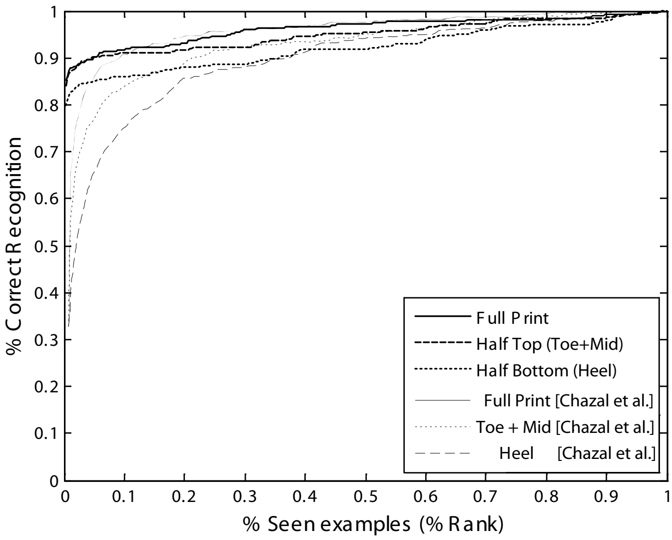

Fig. 3. Rank Recognition Performance of matching with full and partial prints.

trix is obtained as follows:

$$G_{ij} = \sqrt[4]{G_{ij}^{\mathcal{D}} \times G_{ij}^{\mathcal{F}} \times G_{ij}^{\mathcal{N}} \times G_{ij}^{\mathcal{L}}}, \tag{4}$$

for $i, j = 1 \ldots |\mathcal{F}_B|, |\mathcal{F}_B|$. After using G in the algorithm proposed in [16] and obtain pairings $\{i, j\}$ the matched pairs are further thresholded by keeping only those who's $G_{i,j} > e^{-1}$. This is a convenient threshold provided reasonable values of σ have been set for the constraint kernels. An example of a partial match using the above approach is shown in Figure 2.

3 Experiments

A subset of 368 different footwear patterns from the Forensic Science Service database is used [18]. Each pattern class consists of two images, a reference set image containing a whole left and right print and a test set image of either a complete left or right print. The test set print is a different print of varying quality of the same class as that in its corresponding reference image. In order to test on partial prints two additional test sets were produced from the approximate division of the test images into sole and heel sections.

Testing proceeds by applying the above approach for every pairing of reference and test image. The output of each matching attempt returns a list of paired features along with their individual match score G_{ij}, where i and j are feature indexes of the test and reference images. A total match score is taken as the sum of the feature match scores. A best match is that having the highest aggregate score. The value of $2\sigma^2$ is set at 0.3 for the feature similarity constraint and 5 for the locality constraint and the number of neighbours N is set to 5.

Table 1. Comparison of proposed algorithm with proposed approach of Chazal et al. [10]

Approach	Number of unique patterns	% Seen images (Rank)	Full-print	Partial-prints Toe-Mid	Partial-prints Heel
Chazal et al. [10]	140	0.7 **(1)**	65	55	41
	140	5 **(7)**	87	78	66
	140	20 **(28)**	95	89	86
Proposed Method Section 2	368	0.3 **(1)**	**85**	**84**	**80**
	368	1.6 **(6)**	**88**	**87**	**83**
	368	5.4 **(20)**	**90**	**90**	**85**

4 Results

Performance was measured on the observed reference images, in terms of the highest aggregate score, before the correct match was found. Figure 3 shows the correct recognition rate (CRR) for full and partial prints. It can be seen that a good matching performance is achieved, starting at 85% for first rank on full-prints and rising to 91% for the best 6 matches. The performance of our approach is strong even for partial prints. For example, when matching 'Half Top' partial prints our system returns a rank '1' CRR of 84% rising to 90% at rank '6'.

5 Discussion and Conclusions

Our programme of work to develop robust footwear recognition systems will be expanded to much larger reference databases: the current national database contains over 13,000 footwear patterns. Nevertheless, we are greatly encouraged by these initial results, as good performance is possible using only a direct pattern matching approach with no explicit model of outsole appearance. It is difficult to compare different published approaches as different datasets and testing procedures are used. However, we attempt a comparison with the recent work reported in [10]. Table 1 shows this comparison based on similar rank numbers, while the best results of [10] are plotted in Figure 3. Bearing in mind that the number of different patterns used in our study is over twice that used in [10], a marked increase in performance was achieved. This is especially true in the early ranking figures where we report a 90% CRR from viewing only 5% of the database.

The foundations of our proposed automated footwear classification system are based on local shape and pattern structure. The selected feature and pattern descriptors are affine invariant and so can cope with relative translations and rotations. The abundance and localized nature of these features permit good recognition performance for partial impressions. Our on-going work will also explore the ability to match footwear marks retrieved from crime scenes to specified shoes recovered from suspects.

Acknowledgments

The work is financially supported by the UK EPSRC (Grant reference EP/D03633X) in collaboration with the UK Home Office Police Standard Unit and ACPO.

References

1. W. J. Bodziak, Footwear impression evidence detection, recovery and examination, Second ed. CRC Press, 2000
2. A. Girod, Computer classification of the shoeprint of burglars' shoes, Forensic Science Int., 82, 1996, p59-65.
3. Z. Geradts and J. Keijzer, The image-database REBEZO for shoeprints with developments on automatic classification of shoe outsole designs, Forensic Science Int., 82, 1996, p21-31.
4. N. Sawyer, SHOE-FIT A computerised shoe print database, Proc. European Convention on Security and Detection, 1995, p86-89.
5. W. Ashley, What shoe was that? The use of computerised image database to assist in identification, Forensic Science Int., 82(1), 1996, p7-20.
6. S. Mikkonen, V. Suominen and P. Heinonen, Use of footwear impressions in crime scene investigations assisted by computerised footwear collection system, Forensic Science Int., 82, 1996, p67-79.
7. S. Mikkonen and T. Astikainen, Databased classification system for shoe sole patterns - identification of partial footwear impression found at a scene of crime. Journal of Forensic Science, 39(5), 1994, p1227-1236.
8. A. Bouridane, A. Alexander, M. Nibouche and D. Crookes, Application of fractals to the detection and classification of shoeprints, Proc. 2000 Int. Conf. Image Processing, 1, 2000, p474-477.
9. A. Alexander, A. Bouridane and D. Crookes, Automatic classification and recognition of shoeprints, Proc. Seventh Int. Conf. Image Processing and Its Applications, 2, 1999, p638-641.
10. P. de 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 Analysis & Machine Intelligence, 27(3), 2005, p341-350.
11. L. Zhang and N. Allinson, Automatic shoeprint retrieval system for use in forensic investigations, UK Workshop On Computational Intelligence (UKCI05), 2005.
12. K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir and L. Van Gool, A comparison of affine region detectors, Int. Journal of Computer Vision, 65(1/2), 2005, p43-72.
13. K. Mikolajczyk and C. Schmid, Scale and affine invariant interest point detectors, Int. Journal of Computer Vision 60(1), 2004, p6386.
14. K. Mikolajczyk and C. Schmid, A performance evaluation of local descriptors, IEEE Trans. Pattern Analysis & Machine Intelligence, 27(10), 2004, p1615-1630.
15. M. Pilu, A direct method for stereo correspondence based on singular value decomposition, IEEE Conf. Computer Vision & Pattern Recognition (CVPR'97), 1997, p261.
16. G. Scott and H. Longuet-Higgins. An algorithm for associating the features of two patterns. Proc. Royal Society London, B244, 1991, p21-26.
17. S. Ullman. The interpretation of Visual Motion. MIT Press, Cambridge, MA, 1979.
18. Data taken from the UK National Shoewear Database, Forensic Science Service, Birmingham, B37 7YN, UK.