Shoe Print Recognition

David Wark, 0457866

Introduction

Shoe print recognition is an often overlooked piece of forensic evidence, but since the Serious Organised Crime Police Act (2005) *[1]* gave it the same legal status as DNA and finger-print evidence, more attention is being turned to this surprisingly prolific type of evidence.

The impression of a piece of footwear can potentially lead to 3 identifications, of increasing accuracy; a shoe of a certain size, a shoe of a certain make or a specific shoe. The quality of the print will usually dictate which of these is possible. A vague outline is unlikely to provide conclusive evidence, but could offer some information about the size, and in some cases the brand of a shoe. A high quality print that reveals defects and shape-differences specific to an individual's shoe is very rare, and the fact that there will only be one sample of any specific shoe means the chances of a conclusive classification here is low. The most common type of recognition will be that of the shoe's make and model, due to the availability of this data (from shoe manufacturers), the number of samples and the ability to classify according to identifiable features.

In the past, shoe prints captured at a crime scene using photography or a cast (for imprints) would be processed manually against a database. The user would use guides to identify features on the sole of the shoe, and would then be presented with candidates which matches this description until a match could be visually verified. This is both more time consuming and less reliable than automated methods. The description of features may be difficult to standardize, and without appropriate guidance it could be difficult for the user to know how to describe what they see. As new shoes are produced with different features, these modes of description will have to be updated explicitly. It is also hard to guarantee that a user will be able to spot subtle similarities, especially if the image quality is low, coupled with the fatigue that will no doubt come after cycling through many images in one sitting.

An automated approach, therefore, is desirable. There has been some research in recent years to this effect, and drawing on several of these studies I have composed the following outline for a basic automated system, followed by some examples and discussion of some of the key areas.

Basic Approach

- *Capturing data*

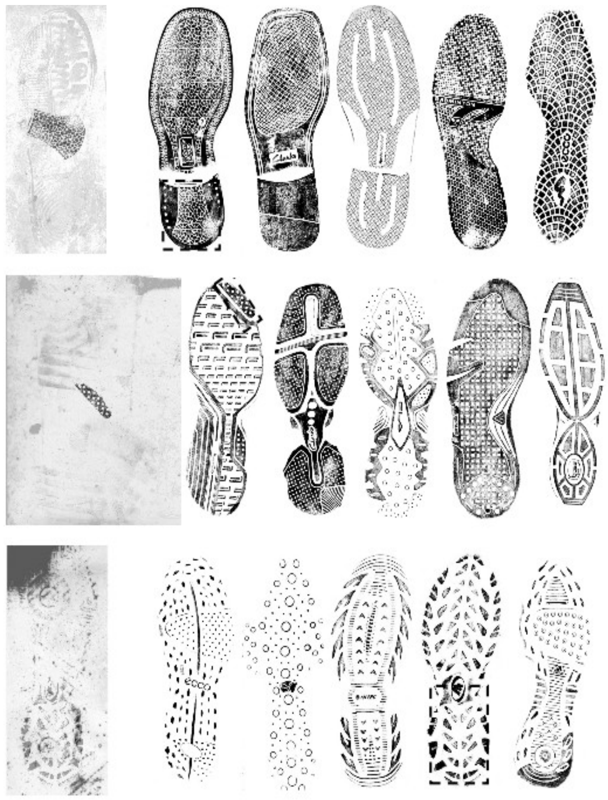
  The test sample will be captured by forensic scientists at the crime scene. There are two kinds of database that will be useful in the classification procedure. A database of shoe prints found at other crime scenes will allow cases and suspects to be linked, and possibly identified. A database of shoe prints from footwear manufacturers will allow the identification of the brand and model of shoe that was being worn, and could be used as evidence if a suspect is found to posses this shoe.

- *Cleaning data*

  Shoe print images are almost universally noisy. Prints left behind are often faint or partial, and the wear and tear on the sole of a shoe through daily use can make the resulting features even harder to identify. Normally, humans can identify the required features despite a certain degree of noise, but techniques used by automated image classification can fall apart if the level of noise is too great. Simple background/foreground thresholding will not be sufficient to preserve the integrity of shoe print images, because the patterns are complicated (making the loss of data after thresholding more likely) and the samples often incomplete (making the loss of data even more serious). An example approach to noise removal that preserves the integrity of image features is discussed below.

- *Feature Detection*

  Due to the design of modern shoes (particularly trainers), for grip and for aesthetics, there are many identifiable types of feature found on the soles of common shoes. To guarantee usefulness in an identification task, important attributes of any features extracted are "locality, repeatability, distinctiveness, and robustness to different degradations" *[3]*. This means any features should be small enough that they might reoccur, seperable, reliably identifiable and invariant to transformations and a degree of noise. There are some well-defined mathematical approachs that yield features which satisfy these conditions. A couple of these are discussed below. Once features have been detected they need to be described in a suitable way. Two common feature descriptors are SIFT (Scale Invariant Feature Transform) and GLOH (Gradient Location and Orientation Histogram). These ensure the features are described in a robust way, that can be invariant to all kinds of transformation and variations in different image properties, such as contrast, viewpoint etc.



Sample shoe prints and features (Image from [3])

- *Classification*

  Once the images have been processed and the features detected, standard classification techniques can be used to identify potential matches against the database mentioned above. Similarity between features will be computed and all the samples with a similarity above a certain threshold will be considered candidates. Currently, the automation of shoe print recognition goes this far and no further. The set of potential matches is then analysed by a forensic scientist . There have been some attempts to improve this process by sorting the result matches according to some similarity measure, and in *[5]* a technique for this is outlined which uses the Fourier transform to sort candidate images.

Noise Removal

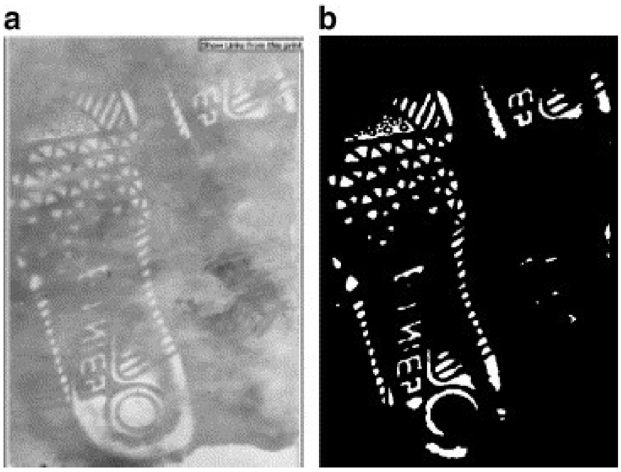
There are several methods that attempt to reduce noise in a captured image. The classic approach to noise removal is thresholding, which aims to classify pixels in an image as either part of the object or part of the background. If the object is sufficiently different in colour or texture from the noise in the image, finding the correct threshold value can allow the object to be discriminated fully. However, if object are more complex, with different parts varying in texture and colour, this global thresholding approach may not be sufficient, since the threshold necessary to discriminate one part of the object from the background may not be usable for other parts. Some information about the key features in the sample could be thresholded out, and this will reduce the chances of finding a match.

In *[2]*, the results of global thresholding are augmented and adapted using results from a technique called non-local mean filtering. For each pixel in the image, its non-local neighbourhood is identified. This consists of every other pixel in the image with a similar local neighbourhood window. It is proposed that the result of applying an operation to a pixel "can be estimated by performing the same operation on all of its reference pixels in the same image", where reference pixels are all those in the non-local neighbourhood.

The expected result of applying operation F(V) is given by applying the operation to all reference pixels j in the non-local neighbourhood

$$\widehat{F}(V) = \sum_{j} r_j^n \cdot F(V),$$

For thresholding, the operation is represented by F(V) = V (i.e. the identity operation), and if we set r as the Gaussian weighted correlation between the local neighbourhoods of the two pixels it can be used as a probability that the two pixels will have the same category in thresholding (object or background). By summing these for each pixel in the non-local neighbourhood, a probability for the test pixel belonging to either category is achieved and a thresholding decision can be made. Some form of decision rule is needed to combine this result with the result of the global thresholding to give the final image. Experiments in *[2]* revealed that this was more effective at noise removal than other trusted thresholding methods.



Noisy sample and result of noise removal using pixel context technique. (Image and equation above from [2]).

Feature Detection

One method for detecting such features is Harris Corner Detection, and since it is "well suited to corner-like features such as small cuts and grooves which are abundantly found in footwear patterns"*[4]*, it is an obvious candidate for recognizing repeating patterns in shoeprints. It works by computing a score for a section of an image by finding the sum of square differences between this section and its neighbouring regions (shifting the area of interest by (x,y)).

$$S(x,y) = \sum_u \sum_v w(u,v) \; \left(I(u,v) - I(u-x, v-y)\right)^2$$

By approximating this value using a second order function, we can find a second moment matrix, describing the gradients in the image,

$$S(x,y) \approx \frac{1}{2}(x \quad y) A \begin{pmatrix} x \\ y \end{pmatrix}$$

where,

$$A = \sum_u \sum_v w(u,v) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

The eigenvectors and eigenvalues of this matrix are calculated and used to assess the variation in each dimension. If both are large, a corner has been detected, and if only one is large an edge has been detected. These detections can be made with a reasonable indifference to rotation, but not scale. A Gaussian scale-space representation of an image allows invariance to scale by describing variations at multiple different scales, at which the feature may be encountered in subsequent images. The combination of these two techniques is called a Harris-Laplace detector and provides a scale and rotation invariant method for identifying corners, edges or other points of interest.

Another robust method for feature detection is MSER, which is "better suited for discriminating general patterns or shapes of footwear marks into classes" *[4]*. MSER stands for Maximally Stable Extremal Regions. An extremal region in an image (also referred to as a blob) is a region all of whose surrounding pixels have a higher or lower intensity. Maximally stable extremal regions have the desirable property of being affine invariant. In *[3]* a Harris-Laplace detector is modified to incorporate a similar idea to that of MSER. Instead of using scale information purely from the identified features themselves, the "blob" surrounding the feature is identified and it's scale is computed. Corners are only considered candidates if there is some predefined relationship between these scales.

References

[1] **Serious Organised Crime Police Act 2005** - Part 3 section 118 of this document, "Impression of Footwear", deals with shoe-prints, http://www.opsi.gov.uk/acts/acts2005/ukpga_20050015_en_1, accessed 31/01/08

[2] *H.Su, D.Crookes, A.Bouridane,* **Thresholding of Noisy Shoeprint Images based on pixel context**; *Pattern Recognition Letters 28(2): 301-307 (Jan. 2007)*

http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6V15-4KXDW8K-1&_user=809099&_rdoc=1&_fmt=&_orig=search&_sort=d&view=c&_acct=C000043939&_version=1&_urlVersion=0&_userid=809099&md5=fc28c9132984f379e9af56cd889287d6, accessed 31/01/08

[3] *H.Su, D.Crookes, A.Bouridane, M.Gueham,* **Local image Features for Shoeprint Image Retrieval**; *Information Assurance and Security, 2007. IAS 2007. Third International Symposium on , 29-31 Aug. 2007 Page(s):387 – 392*

http://ieeexplore.ieee.org/iel5/4299731/4299732/04299804.pdf, accessed 31/01/08

[4] *M.Pavlou, N.M.Allinson,* **Automatic extraction and classification of footwear patterns;** *In Lecture Notes in Computer Science, Proc. Intelligent Data Engineering and Automated Learning (IDEAL '06), Burgos, Spain, 2006, p721-728.*

http://www.shef.ac.uk/eee/vie/IDEAL06_MPavNAllin.pdf, accessed 31/01/08

[5] *P. de Chazal, J.Flynn, R.B.Reilly,* **Automated Processing of Shoeprint Images Based on the Fourier Transform for Use in Forensic Science**; *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on, Volume 3, Issue , 14-17 Sept. 2003 Page(s): III - 569-72 vol.2*

http://ieeexplore.ieee.org/iel5/8824/27938/01247308.pdf, accessed 31/01/08

Equations from “Feature Detection” section taken from *http://en.wikipedia.org/wiki/Corner_detection*