Forensic Science International 275 (2017) 102–109

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

Classification of footwear outsole patterns using Fourier transform and local interest points

Nicole Richetelli [a], Mackenzie C. Lee [b], Carleen A. Lasky [a], Madison E. Gump [a], Jacqueline A. Speir [a,*]

[a] West Virginia University, 208 Oglebay Hall, PO Box 6121, Morgantown, WV 26506, United States
[b] Cedar Crest College, 100 College Drive, Allentown, PA 18104, United States

ARTICLE INFO

Article history:
Received 17 November 2016
Received in revised form 9 February 2017
Accepted 23 February 2017
Available online 4 March 2017

Keywords:
Footwear
Phase only correlation
SIFT
RANSAC
Fourier–Mellin transform

ABSTRACT

Successful classification of questioned footwear has tremendous evidentiary value; the result can minimize the potential suspect pool and link a suspect to a victim, a crime scene, or even multiple crime scenes to each other. With this in mind, several different automated and semi-automated classification models have been applied to the forensic footwear recognition problem, with superior performance commonly associated with two different approaches: correlation of image power (magnitude) or phase, and the use of local interest points transformed using the Scale Invariant Feature Transform (SIFT) and compared using Random Sample Consensus (RANSAC). Despite the distinction associated with each of these methods, all three have not been cross-compared using a single dataset, of limited quality (*i.e.*, characteristic of crime scene-like imagery), and created using a wide combination of image inputs. To address this question, the research presented here examines the classification performance of the Fourier–Mellin transform (FMT), phase-only correlation (POC), and local interest points (transformed using SIFT and compared using RANSAC), as a function of inputs that include mixed media (blood and dust), transfer mechanisms (gel lifters), enhancement techniques (digital and chemical) and variations in print substrate (ceramic tiles, vinyl tiles and paper). Results indicate that POC outperforms both FMT and SIFT + RANSAC, regardless of image input (type, quality and totality), and that the difference in stochastic dominance detected for POC is significant across all image comparison scenarios evaluated in this study.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

With the increased popularity of crime solving dramas on television, the public is much more aware of what crime scene investigators are looking for while processing a scene. This knowledge, whether accurate or not, has altered the jury's expectations regarding the analysis and exhibition of forensic evidence presented during a criminal trial [1]. If this knowledge has affected the jury, it is equally likely to have altered how criminals attempt to conceal their crimes, putting greater importance on evidence types currently outside the limelight of the media. These alternative forms of evidence, such as the classification and source individualization of shoeprints, provide information that is critical in linking suspects to victims, crime scenes, and even multiple crime scenes to each other. In fact, if the crime scene is void of all other forms of evidence, footwear impressions may very well represent the only remaining item of probative value available at the scene.

When properly preserved and documented, footwear class features can be extremely valuable [2], affording the analyst the ability to focus a criminal investigation, link high volume crimes together, or otherwise provide information vital to the successful resolution of a case. Additionally, if the potential pool of source/exemplar footwear is limited, then the human observer's exceptional pattern recognition skills can be used to easily link a questioned impression to a database exemplar with a known brand and manufacturing history. Moreover, the authors assert that this is true even in the presence of overwhelmingly low and/or mismatched signal-to-noise ratios (SNR), temporal resolutions, spatial resolutions, spectral resolutions, distortions, perspectives, scale variations, rotations, translations, substrates, mixed media and dimensionality. In other words, the human observer is innately able to create meaningful linkages when presented with numerous

* Corresponding author.
E-mail address: Jacqueline.Speir@mail.wvu.edu (J.A. Speir).

http://dx.doi.org/10.1016/j.forsciint.2017.02.030
0379-0738/© 2017 Elsevier B.V. All rights reserved.

variations in media and content, provided the total number of exemplars for comparison is limited in size [3–5].

However, the rapid rate of production and turnover of manufactured footwear in the United States challenges the human observer's pattern recognition efficiency [6]; for all but the most commonly encountered shoes and questioned impressions with the lowest SNR, manual classification methods are likely to become increasing inefficient in today's forensic discipline. With this in mind, research concerning the *automated* or *unsupervised* comparison and retrieval of known match imagery has gained momentum within the field of forensic footwear analysis over the last few decades. Unfortunately, *automatically detecting* a linkage between questioned impressions from crime scenes and reference shoes from a database with known brand and manufacturing features remains a challenge, and the reason for this difficulty can be ascribed to two primary factors. First, the signal-to-noise ratio (SNR) in typical crime scene impressions is extremely low. This is often the result of a host of additional variables, including but not limited to partial or incomplete patterns, the influence of additional wear, variations in image modality such as media (blood, dust, soil) and substrate (ceramic tiles, vinyl tiles, paper), as well as artifacts from transfer mechanisms (gel or electrostatic lifters, casts) and enhancement techniques (physical, chemical, optical). In addition to low SNR, the second primary factor that inhibits the successful linkage of a questioned impression to a known source impression from a database is image registration, or the presence of affine variations in scale, rotation and translation [7], (as well as more extreme and non-linear deformations such as perspective, distortion and skew) that cause confusion when attempting to automatically identify correspondences between two images. As a result of both noise and mis-registration, automated systems aimed at the *unsupervised* retrieval of possible matches require intentional and robust mathematical solutions in order to avoid recognition loss when faced with both real, and sometimes with what seem like trivial, image differences.

1.1. Overview of automated methods

In general, there are two steps associated with automated shoeprint recognition. The first step is the transformation of the outsole geometry into some type of feature description, followed by a secondary step wherein the feature description is compared and ranked in terms of similarity with respect to a series of database reference descriptions. A brief literature review indicates a multitude of pattern recognition approaches to this problem; for instance, the *feature-description step* has been accomplished using Fourier methods [8–13], fractal decomposition [14], Zernike [10] and Hu's [15] moments, texture descriptives such as Gabor feature [16] and Mahalanobis distance maps [13], and local image features (such as Harris-Laplace or Maximally Stable Extremal Regions (MSER)) modified using the Scale Invariant Feature Transform (SIFT) [17–20]. In addition, the *similarity assessment step* has been accomplished using neural networks [8], minimization of the mean square noise error between reference and transformed images [14], correlation coefficients [9,10,13], peak height for composite, advanced and phase-only correlation [11,21], distance metrics such as Euclidean [10,15,16,22], point-matching using nearest neighbors [18] and Random Sample Consensus (RANSAC) [17,20], and finally, spectral correspondence matching [19]. Many investigators have also attempted to cross-compare the utility of various approaches, pitting Fourier methods against texture metrics [23], and point-matching *versus* Fourier and moment invariants [20], to name but a few illustrations.

Of the possible feature-based retrieval methods considered thus far [8–24], two algorithms seem to out-perform other competitors in terms of their applicability to footwear classification. This includes Fourier methods [20,23,24] and local interest points transformed using SIFT and evaluated using RANSAC inliers [20]. The Fourier methods of highest performance can be further subdivided into those that exploit magnitude/amplitude as a feature descriptor [20], and those that rely on image phase [23]. Despite this apparent clustering in success, it is extremely difficult to report on a *single* superior methodological approach for unsupervised footwear classification. Instead, it is *highly unlikely* that a *single* classification algorithm will out-perform all others in every scenario, and it is much more likely that each multi-phased metric will be susceptible to individual failure under *certain circumstances*, owing to one or more inherent weaknesses. To illustrate, consider a recent evaluation by Luostarinen and Lehmussola [20], comparing the accuracy of seven different automated algorithms, including Fourier, Hu's moments, Mahalanobis distance, Gabor transform and local interest points transformed using SIFT and mated using RANSAC. Based on the high quality imagery used in this comparison, the authors were able to conclude that Fourier–Mellin and local interest points with SIFT and RANSAC outperformed all other methods. However, when confronted with "non-ideal" input, such as crime scene imagery with low SNR, the performance of both methods decreased [20], wherein local interest points are believed to be heavily influenced by structured noise [24].

Interestingly, although not evaluated and observed by Luostarinen and Lehmussola [20], a decline in performance for local interest points (with SIFT and RANSAC) is also likely to exist for *high quality images* with *high SNR* if the outsole in question is comprised mostly of *repetitive patterns*. The reason for this is RANSAC's known susceptibility to creating false matches when presented with repetitive patterned imagery (such as a checkerboard) [25]. Ergo, an algorithm's quoted accuracy is also highly database (test-set) dependent.

In contrast to RANSAC's propensity for false matches when presented with repetitive patterns, one might argue that the strength of many Fourier methods lies in their ability to detect and characterize periodic variation. According to Kortylewski et al. [24], approximately 60% of outsoles show geometric periodicity in tread design (as assessed using a reference database of 1175 images), making Fourier methods attractive in terms of image characterization within the field of forensic footwear comparison. Moreover, it is also purported that this periodicity can be used to help extract relevant information from partial prints in noisy crime scene impressions (assuming that the background noise is not likewise periodic) [24]. However, if regular patterns are only present on 60% of the population of outsoles, then the success of Fourier methods on the remaining 40% of anticipated comparisons is less clear. Again, to repeat a previously voiced assertion, it is much more likely that each retrieval system (applied to this problem thus far) is susceptible to individual failure under certain circumstances, owing to one or more inherent weaknesses.

With this in mind, the aim of this study was to cross-compare the ability of three simple retrieval methods (Fourier–Mellin transform (FMT), phase-only correlation (POC) and local interest points + SIFT/RANSAC) using a common database that included 100 unique high quality Handiprint exemplars and 172 crime scene-like impressions created from 36 outsoles with variations in media type (blood and dust), substrate (ceramic tiles, vinyl tiles, acetate sheets and paper) and chemical/optical enhancement procedures (including contrast adjustment and the use of leucocrystal violet (LCV)). The results reported here represent an extension of the findings presented by Cervelli et al. [23], and Luostarinen and Lehmussola [20], by collectively pitting FMT, SIFT/RANSAC [20] and phase-only correlation [23] against each other using a single test-set.

Download English Version:

https://daneshyari.com/en/article/6462338

Download Persian Version:

https://daneshyari.com/article/6462338

Daneshyari.com