Automating Footwear Impressions Retrieval through Texture

Neil Farrugia (✉), Joseph Vella

*Department of Computer Information Systems, University of Malta,
https://www.um.edu.mt*

A B S T R A C T :

This study aims to provide an automatic footwear extraction and correlation system. The artefact proposed is able to apply pre-processing, extract key features and retrieve the relevant matches from a footwear impression repository.

In order to compare images, a comparison function was utilised. This function creates a MPEG-1 movie out of the images and employs the size of the movie in order to calculate the similarity. For pre-processing of the prints, apart from common techniques, an original concept of tessellations was applied. A publicly available dataset of footprints, a subset of which come from crime scenes, was used. The results obtained from the development of this project have shown that the accuracy of matching depends on the quality of the images that are being used. Comparisons are done in two batches: first, all crime scene prints (i.e. 170) are compared with all the reference prints (i.e. 1175), then the procedure is repeated with various pre-processing methods being applied to the input prints. Accuracy averaged at 55 % (without pre-processing) and at 65 % for a particular method of pre-processing (i.e. based on 43 prints).

Introduction

In most crime scenes, a large number of footwear impressions are collected. In fact, it is estimated that over 30 % of burglaries produce viable footprints that can

✉ Corresponding Author: E-mail: neilf95@gmail.com

be efficiently obtained from a scene of the crime.[1] In the majority of cases, footwear impressions are more commonly found at crime scenes when compared to other biometric data such as fingerprints.[2] This could be because fingerprints are usually prevented from being left behind with the use of gloves, but this covering is not prevalent with footwear.

From the footwear impressions investigators can gain insights regarding the scene of the crime. The variety of prints can determine the number of people at a crime scene, the sharpness of the print can tell if the person was walking, running, sneaking, and even show the path a perpetrator traversed in the scene of the crime. The prints can also link multiple crimes together.

New shoes of the same model will leave behind the same or very similar print. However, as time passes, natural erosion known as "wear" occurs on the shoe, which damages it in a unique way, hence forming a distinctive print. Where and to what extent the wear occurs depends on the owner of the shoe. A famous judicial case where footprints were involved during the investigation was the OJ Simpson Trial of 1995.[3] Here the investigators found many blood-soaked footprints at the scene of the crime and managed to match these to a pair of designer Bruno Magli shoes owned by Simpson.

Aims and Objectives

Since footprints are still mostly checked manually by investigators, there is a need for a system that can sift prints in accurately and efficiently.

The aim of this study is to create a system that can automatically compare footwear impressions and return relevant matches similar to the footwear impression of interest. The system must be automated in order to make the job of the investigator acceptable and consistent. Since prints from real crime scenes are not clear and contain high amounts of noise, the system must be robust too. It needs to cope with different types of images and with images of varying amounts of noise including a shoe's wear.

The similarity measure to compare prints is based on image (whole or part) texture.

Furthermore, in order to complete this task, importance will be given to possible approaches that can be taken in order to develop an efficient method of accurately comparing images. Techniques that help improve image quality and reduce noise are to be looked into. Furthermore, a known and available dataset of shoe prints is adopted for reporting results.

Background and Related Work

Footwear Impressions

A footwear impression, known variously as a shoeprint, footprint, imprint and footmark, is an outline or indentation left by a foot on a surface. These patterns are particular for every shoe which makes them very useful pieces of evidence for crime scene investigations. The difference in the patterns can depend on many factors such as the material of the shoe, the brand and also the shoe's size.

Each and every shoe, except newly purchased ones are said to be unique in some way. Even if the shoe is of the same brand and size, when worn by two different people they will have their own unique imprint given a substantial amount of use. This is because as a shoe is worn, it is naturally damaged and eroded as a result of particular use and aging. The area of most wear on the sole of a shoe will be unique for every person due to the wearer's gait, the environment they walk in and the material of the shoe. Due to this uniqueness of outsole patterns, investigators are interested in shoe prints as collaborating evidence. A study by W. Adair et al. proves that the characteristics of each footwear impression (even that of the same shoe brand and model) are distinctive to that specific print at that specific point in time.[4] As time passes the insole is further degraded and thus changing its print.

Shoe prints can be frequently identified at the scene of a crime. High quality impressions which are recovered during a forensic investigation (see Fig. 1) and accurately matched to another print, can be as crucial to a case as other strong biometric data such as fingerprints.[5]

Existing Systems

In academia there are prototypes of automated shoe print analysis, however most of these cannot be applied to real life cases. Presently the systems that do exist are not automated, they are semi-automatic, meaning they still rely on human interaction to classify the shoeprint. Also, many rely on unavailable datasets (e.g. shoe prints per supplier per model). Initial advances of semi-automatic methods that had been made [5,6] involve manual interaction of the user to create a description of the print using a "codebook" of basic shapes such as circles, lines, wavy patterns. The user must perform the same procedure for the query image.

Several problems arise with these systems adoption. The main drawback of this is that the process is tedious and very time-consuming. The query image is constructed by the user which can lead to inconsistency between users' query images (as different examiners might not construct the image in the same way i.e. using the same shapes). These problems push the need for automatic footprint analysis systems to be further researched and developed.

Many different studies have been conducted regarding automatic shoeprint recognition systems.[7, 8, 9, 10] Determining which implementation is a better one is
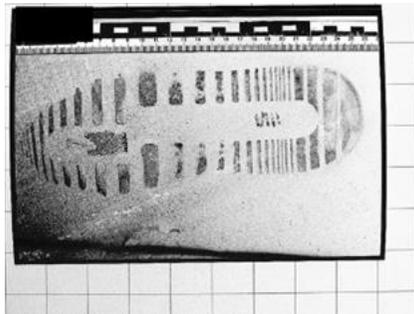

Figure 1: Footwear Impression obtained from a Crime Scene.[8]

very challenging because all the tests were conducted on different data sets, using different types of images, and adopting different testing environments.

Luostarinen and Lehmussola compare and contrast seven different automated footprint impression recognition prototypes in order to determining which is the most reliable.[7] The evaluated approaches are with a set-up closely following the respective literature source: Power Spectral Density (PSD), Hu's Moment Invariants, Gabor Transform, Fourier-Mellin Transform, Mahalonobis Distance Map, Local Interest Points with RANSAC, and Spectral Correspondence of Local Interest Points. All these methods have been tested under the same environment in order to get a fair estimate of how they perform when compared together. The authors also collected 499 pairs of foot prints and used these to generate and partition them into three datasets marked as 'good', 'bad' and 'ugly' over which the seven methods were applied. Each data set differed in quality from excellent quality in the good set, to noisy damaged images (actual crime scene prints) in the ugly set. The results showed significant differences in the system's accuracy when applied on the good and bad datasets however they all performed poorly when making use of actual prints taken from a crime scene. Results for the bad dataset range from 10 % to 85 % probability of good match when comparing to 5 % of the dataset.

Another system is by Kortylewski et al.,[8] who propose an unsupervised system to extract periodic patterns from the footprints and then attempt to match the prints based on these patterns. The authors make use of pre-processing techniques such as local and rotation normalization. The system is said to convert the periodic patterns which are extracted using a Fourier Transform and compare these transformations in order to check if two footwear impressions are similar. Results here range from 10 % to 55 % probability of good match when comparing to 5 % of the dataset for crime scene impressions.

Riba et al.[17] compiles a survey on a good number of systems (and a subset of which are also mentioned by Luostarinen and Lehmussola [7]). The authors consider forty works from the literature and divides these in modes of recognition based on the whole impression (i.e. holistic) or based on some feature (i.e. local). This surveys reports, citing the respective primary source, an accuracy ranging from 27 % to 100 % for 1 % request retrieval. The authors indicate that primary sources vary in methodology, datasets, and reporting of accuracy.

The area of automatic recognition of crime-scene shoeprints is still a developing area, which requires more attention and more thorough research. Some attention must be devoted to datasets and indicative requirements of functionalities expected.

Dataset

The dataset selected to support this study is maintained by Kortylewski et al.[8] The dataset is called the Crime Scene Footwear Impression Database (CSFID-170) and is a dataset the authors used in creating an Automated Footwear Impression Re-

trieval System (AFIRS), making it an appropriate selection for this study. The dataset, which is available to the public was created by the German State Police which confirms the validity of the prints being from real crime scenes.

The dataset consists of three different classes of footwear impressions in greyscale. These classes being Real Crime Scene Prints labelled as 'tracks_original.' Processed Crime Scene Prints labelled as 'tracks_cropped,' and the Reference Prints labelled as 'references.' The Crime Scene Prints dataset consists of 170 prints. These prints were extracted by forensic experts using gelatine filters or by photographing the impression. The Processed Crime Scene Prints are the same exact prints of the crime scene but with adjustments made to them. The prints have all been rotated into an upright position, and cropped to relatively the same size and are in JPEG file format. The Reference Prints set consists of 1175 prints and are in PNG file format. These prints have been collected from known shoe manufacturers including Nike, Reebok and Adidas.

Design

A Content Based Image Retrieval (CBIR) system required for the purpose of this study focuses on the area of footwear impressions. This area of focus was chosen as it is still a relatively new area of research in the field of digital forensics. The design and development of the system is influenced by several works. All the systems reviewed [5, 6, 7, 8] produced promising results. However, these results are specific to the particular dataset, and also the particular evaluation techniques that was used.

The proposed system for this study aims to provide an AFIRS that is automatic and robust with the goal of making the work of the forensic analyst faster, less laborious, less subjective, and more efficient.

System Architecture

The AFIRS is designed to be data driven, meaning it relies heavily on the dataset's content. The main goal of the system being designed is to have it correctly match the Crime Scene Prints to their equivalent Reference Print. The workflow is in Figure 2.

Implementation

Two main tools were required in order to complete this study. PostgreSQL was made use of to manage the Footprints crime scene database. PostgreSQL was primarily chosen because of its ability of loading images and its effective use of stored procedures for managing batch jobs on the server rather than on the client accessing the database.

MATLAB R2017b was made use of to implement the majority of the system and it was primarily chosen due its Image Processing Toolbox; this is an extensive module. It was also selected because of its ease of use and its active community.
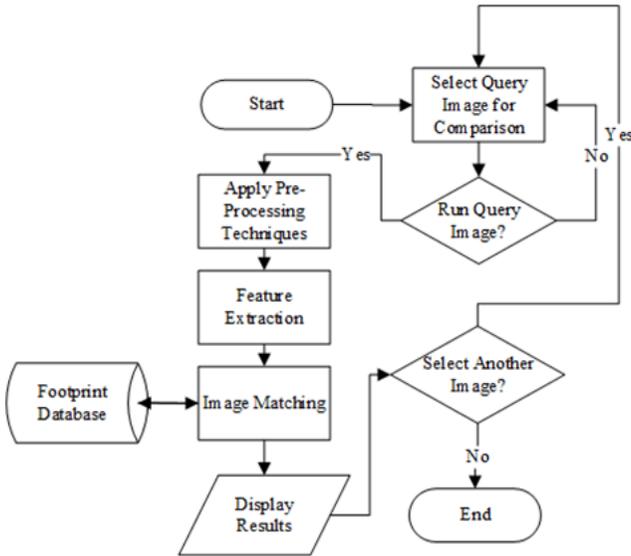
Figure 2: Footwear Retrieval System for Image Matching Workflow.

Implementation of Pre-Processing and Feature Extraction

Pre-Processing and Feature Extraction are two extremely important processes that must be utilised in any system that involves digital image processing especially where querying by images is incorporated, as is this case of CBIR.

Both these processes where grouped into one stage for convenience and more efficient coding practices. The function therefore takes in two images, applies techniques of pre-processing and feature extraction with the aim of improving the image quality. It consists of three main phases:

The first phase involves the application of the pre-processing techniques. The first technique that is applied to the images is Local Normalisation. Local Normalisation was applied because Patil et al.[11] showed that it is a vital step in the processing of footwear impressions as it helps in greatly improving the quality of the footprint processing. The other pre-processing technique that is used in this phase is filtering. Applying filters is beneficial in removing noise from the images which results in improved accuracy of the classification system. The filter chosen for the purpose of the footwear impressions was the Wiener Filter.[12] Since the Weiner Filter is well known to have promising results on grayscale images, it was therefore chosen as an ideal filter for this situation.

Following these two pre-processing techniques, feature extraction is then performed from the processed images. The form of Feature Extraction chosen to be implemented is edge detection. Edge detectors extracts patterns from footwear impressions as they enhance the lines of the print making detection less complicated and more accurate. The detector that was implemented was the Roberts Cross Edge Detector.[13] Like the Wiener filter, the Roberts Edge Detector was utilized because of its excellent results with grayscale images.

The final phase of this module involves the image matching. Here the distance function is called to calculate the similarity between the two input images. Before applying the distance function, the two images being compared are divided into three equal sections. This is called tessellation and was a concept created for the purpose of this study. The separation of each image into three parts allows for the distance function to analyse the image in more detail since it will be focusing on a smaller area of the image rather than the whole. The distance function is run on each segment individually.

Implementation of Image Matching

The distance function that was chosen for this study is one that is provided by Campana and Keogh.[14] The function, named the CK1 Distance Function, was selected because of three reasons. Firstly, the main motivation for making use of this function was because of the wide variety of datasets that the function was applied to. These varied from sets of insects such as moths [15] to furniture using a dataset of wood.[16] The second motivation is that it had not been used with footwear impressions offering an opportunity to analyse and test the robustness of this function for this study. The third reason is that the function is parameter-free. The benefit of having no parameters is that it makes the function open to more general use rather than specific to one scenario and would not rely on parameters that need to be indicated and if not properly set can easily cause accuracy issues.

This CK1 Distance function is used to calculate the similarity of two footwear impressions. It takes two images as input, i.e. the images that will be compared, and returns a value for their distance. Another feature of this function is that it is a symmetric function meaning that the order in which the images are passed to it does not affect the distance computed. The CK1 Distance Function works by creating a movie out of the two images passed in, which is rendered using an MPEG-1 encoder. Once the movie is created, a calculation is performed based on the size of the movie. The concept behind using the size is that, if two images are similar then the movie created by these two frames would be small in size. This is because in the generation of movies, frames are compressed as much as possible between one and another, which means that the more similar the frames are to each other, the smaller the size of the movie is. Therefore, it is indicated that the smaller the distance returned by the function the more similar are the images. The literature asserts that the function is robust and it can detect subtle differences too.[14]

Result Analysis

In order to make the analysis of the results more efficient a method was created to automatically analyse the results and calculate the accuracy.

The approach taken to calculate the accuracy of the image matching on the footwear impression, is to use a table of results provided in the dataset as a benchmark to compare our results to. This results table contains the Crime Scene Prints and their matching Reference Print. It is important to note that the results are assumed to be correct by Kortylewski et al.,[8] which might not be the case for each combination of prints.

Every individual Crime Scene Print is compared to all the Reference Print instances. The results of each prints are written to a CSV file, one file for each Crime Scene Print. The results are in ascending order by distance, from lowest distance being the most similar, to the highest distance being the least similar.

Thus, to calculate the accuracy by using the results table and the files of results, a formula was used. The assumption that is being made is that, if the actual matching Reference Print is in the first position of the list, then the accuracy for that Crime Scene Print is said to be 100 %. The formula created is:

$$accuracy = \left(\frac{1175 - (\text{index} - 1)}{1175}\right) 100$$

Where 1175 is the total number of Reference Prints and index is the position of the actual correct match. To test the validity of the example mentioned before, if the position of the actual correct match is in the first position this means index = 1 then the accuracy should be 100 %.

Proof:
$$accuracy = ((1175 - (\text{index} - 1))/1175)\ 100$$
$$= ((1175 - (1 - 1))/1175)\ 100$$
$$= (1)\ 100 = 100$$

Results and Evaluation

The following section explain methods and their merits with regards to the performance of the developed system. Several tests have been run, which have placed the system under different working conditions in order to analyse how these affect its behaviour. The results of these tests are evaluated and discussed. Initial tests were required to tune and apply solutions in order to produce the final version of the shoeprint recognition system.

Evaluation of Distance Function

Prior to testing for image matching, preliminary tests were conducted on the distance function alone in order to see where its limitations lie and to investigate the different conditions that affect it. Several conditions where applied to the distance function adopted. These being: using identical images, using almost identical images (i.e. with slight alterations) and images with major alterations; using it after applying noise and rotation; applying images of different sizes and using tessellations. Results are found in Table 1.

Evaluation of the system without Pre-Processing

During this test, every single Crime Scene Print (i.e. one of the 170) was compared to all the Reference Print (i.e. 1175) without the application of pre-processing and feature extraction. The accuracy was then calculated for each individual Crime Scene Print.

Table 1. Showing the Different conditions and characteristics of the CK1 Distance Function.

| Condition	Results
Identical Images	Distance function returns 0 when identical images are used. This indicates that the distance is directly proportional to the dissimilarity of two images.
Identical Images: Slight Alterations	Slight alterations to the image do not greatly affect the distance function output as it still returns a small distance.
Identical Images: Major Alterations	Major alterations affect the distance function more than the previous test. It is expected that as more alterations are made the less similar it is becoming to the original one.
Noise	Noise greatly affects the Distance function. Four different types of noise where applied to the images: which were Gaussian, Poisson, Salt & Pepper and Speckle. Poisson was the least disruptive on the function as it barely affected the distance whereas Speckle was the most disruptive with a substantial increase.
Rotation	When rotating the same image four times (i.e. at the angles (in degrees) of 30, 90, 180 and 270) the results showed that the distance function is highly sensitive to rotation, especially at differing angles.
Size	Size does not greatly affect the distance function. However, the results showed that the larger the image is, the more accurate the match is. Thus, it can be indicated that the size of the image is directly proportional to the similarity of the images.
Tessellation	The application of tessellations onto the images resulted in a slightly higher distance than without tessellation. This does not mean that tessellation is not suitable, in fact it is better, because the segments are focused on individually and thus a more accurate result is returned.

As it can be seen in Figure 3, the results of this test where quite promising since no pre-processing was being applied. Several prints achieved high accuracy, with the best being 99.9 %. There were also low accuracies that were achieved, however, the overall accuracy achieved is recorded at 55.2 %. This test indicates that the accuracy is highly dependent on the quality of the image that is being used.

Evaluation of the system with Pre-Processing

This test builds upon the previous one. The same test is conducted but pre-processing is applied. The Crime Scene Prints used for this test are the Re-Processed Crime Scene Prints. The results of pre-processing where more spread out than those without pre-processing as it can be seen in Figure 4. The best match that was recorded using pre-processing was at 98.2 % and the overall accuracy being at 51.4 %.
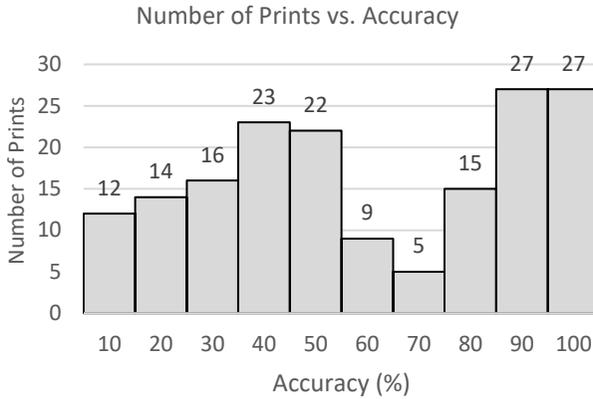
Figure 3: Graph of the Results of the system without Pre-Processing.
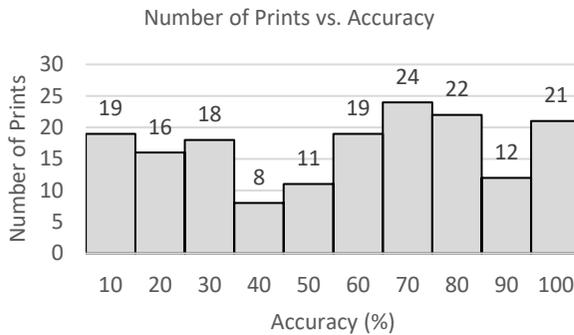
Figure 4: Graph of the Results of the system with Pre-Processing.

Comparison of Results

In order to compare the results of two tests, the accuracies of the prints were inspected individually.

Table 2 shows the top 3 least accurate and most accurate prints obtained during the test without pre-processing. The original distance is shown as well as the distance of the same prints, after pre-processing. Two observations can be made from this table. When pre-processing is applied to prints that achieved high accuracy without it, the distance is increased and when it is applied to prints which have achieved a high accuracy the distance is decreased. Thus, it can be said that pre-processing should be applied to prints that achieved a lower accuracy during the test without pre-processing.

Table 2. Comparing Evaluations of No Pre-Processing to Pre -Processing.

Print ID	No Pre- Processing (Distance)	Pre- Processing (Distance)
118	1.70	89.36
12	3.40	36.34
23	4.253	19.23
122	99.31	89.83
148	99.65	55.49
19	99.91	77.27

Evaluation of new Pre-Processing method

A new pre-processing method was created with the goal of improving the overall accuracy of the system. This method makes use of different pre-processing and feature extraction techniques, which are the median filter for pre-processing and the canny edge detector for feature extraction.

Unlike the initial tests, this was run on a subset of the original dataset. A sample of 25 % was used, which is 43 prints. The highest achieved accuracy using this method was of 98.9 %. The overall achieved accuracy using this new method resulted at 64.6 % which is an increase of 13.2 % when compared to the other pre-processing method and an increase of 9.4 % when compared to using no pre-processing.
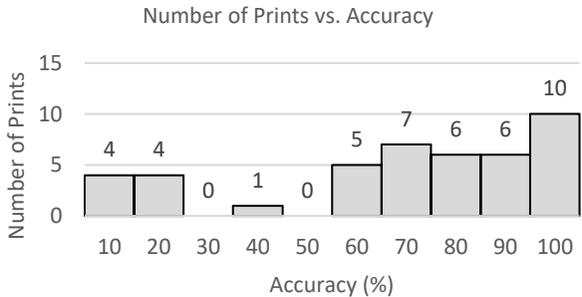
Figure 5: Results of the system with second pre-processing method but applied to a smaller set of prints (i.e. 43).

Summary of Results

From the various tests that have been conducted it can be concluded that the CK1 Distance function is sensitive to noise, size, rotation and alterations present within the image.

The system has performed well with an accuracy of 55 % (170 prints) when apply no pre-processing to the prints. Through testing it has been found that pre-

processing is useful in certain scenarios but not for all. If the prints have achieved a good accuracy score, then there is no need for pre-processing as it might interfere with the print identification. However, when the prints did not achieve acceptable accuracy, pre-processing techniques can improve the accuracy of the print identification; i.e. the investigator can run the same print with either pre-processing enabled. From the two different pre-processing techniques implemented, the second proved to be more accurate overall with an accuracy of 64.6 % (43 prints).

Conclusions and future work

Future work

As repeatedly indicated in the literature, e.g. by Rida et al.,[17] the current problem impeding the fruition of research in footprint and shoe-print identification, is the lack of publicly available data and well-defined evaluation techniques. Most experiments carried out are evaluated on synthetically generated data, adding artificial distortion to them, such as noise and blur, to try and imitate a real crime scene print. Unfortunately, this is not realistic and hints at untenable effectiveness. In other cases where real data is used then this data is not available to the public.

With that being said, the fact that to some extent, an AFIRS has been developed is step in the right direction. Further improvements and functionalities would be developed.

The following list is of future additions that can be added to this study:

• Making the system more robust in order to apply it to other areas of Digital Forensics such as motor vehicle wheel tracks;

• Making use of other distance functions;

• Running the system on the recently available footprint dataset FID- 300 [8];

• Improve the computational time performance during pre-processing and matching.

Conclusions

The main scope of this study was to develop a CBIRS in the area of Digital Forensics. The selected area was of Footwear Impressions (e.g. shoeprints) as it is one of the lesser researched when compared to other forensic data that can be found at a crime scene. Thus, the system that has been designed and implemented is an AFIRS. Furthermore, a publicly accessible and realistic dataset was adopted.

The resulting artefact that has been developed manages to match real scene of the crime footprints at the best accuracy of 55 % (no pre-processing) and 65 % (with a particular pre-processing method). With the accuracy of the match is dependent on the quality of the image and the amount of noise present.

The Campana Keogh Distance Function (CK1),[14] based on footprint image texture, was appropriate for the purpose of this study. It has proved to be have a tenable basis and is reflected in the results that were achieved.

MATLAB proved to be a very effective tool. It provides a robust development environment and the Image Processing Toolbox was also extensively used. The third-party libraries which are made available also helped in making the implementation more efficient. PostgresSQL was an effective data handling tool. Querying and retrieving results was concise.

References

1   Girod Alexandre, "Computerized Classification of the Shoeprints of Burglar's Soles," *Forensic Science International* 82, no. 1 (1996): 59-65.

2   E. Murray, "Footprints Can Play an Important Role in Solving Crimes," *Footprint Lab*, December 1, 2011, Available at http://science.marshall.edu/murraye/Footprint%20 Lab.html.

3   Sargur N. Srihari, "Analysis of Footwear Impression Evidence," US DoJ Report, 2011.

4   T. W. Adair, J. Lemay, A. Mcdonald, R. Shaw, and R. Tewes, "The Mount Bierstadt Study: An Experiment in Unique Damage Formation in Footwear," *Journal of Forensic Identification* 57, no. 2 (March/April 2007): 199-205.

5   Wayne Ashley, "What Shoe Was That? The use of Computerised Image Database to Assist in Identification," *Forensic Science International* 82, no. 1 (1996): 7-20.

6   N. Sawyer and C. Monckton, "'Shoe-fit' - A Computerised Shoe Print Database," European Convention on Security and Detection, Brighton, UK (1995): 86-89, https://doi.org/10.1049/cp:19950475.

7   Tapio Luostarinen and Antti Lehmussola, "Measuring the Accuracy of Automatic Shoeprint Recognition Methods," *Journal of Forensic Sciences* 59, no. 6 (May 2014): 1627-1634.

8   Adam Kortylewski, Thomas Albrecht, and Thomas Vetter, "Unsupervised Footwear Impression Analysis and Retrieval from Crime Scene Data," in C. Jawahar and S. Shan, eds., Computer Vision - ACCV 2014 Workshops. ACCV 2014. *Lecture Notes in Computer Science*, vol. 9008 (Cham: Springer, 2015), 644-658.

9   Yi Tang, Sargur N. Srihari, Harish Kasiviswanathan, and Jason J. Corso, "Footwear Print Retrieval System for Real Crime Scene Marks," in F. Sako, K.Y. Franke, and S. Saitoh, eds., Computational Forensics, IWCF 2010. *Lecture Notes in Computer Science*, vol. 6540 (Berlin, Heidelberg: Springer, 2011), 88–100.

10   Omar Nibouche, Ahmed Bouridane, M. Gueham, and Moussadek Laadjel, "Rotation Invariant Matching of Partial Shoeprints," *13th International Machine Vision and Image Processing Conference*, 2009, pp. 94–98.

11   Pradeep M Patil and Jayant V Kulkarni, "Rotation and Intensity Invariant Shoeprint Matching Using Gabor Transform with Application to Forensic Science," *Pattern Recognition* 42, no. 7 (2009): 1308–1317.

12   Mukesh Motwani, Mukesh C Gadiya, Rakhi Motwani, and Frederick Harris, "Survey of Image Denoising Techniques," *Proceedings of GSPX* (2004), 27-30.

13   Mark Nixon and Alberto S. Aguado, *Feature Extraction & Image Processing*, Second Edition (Academic Press, 2008).

14   Bilson J. L. Campana and Eamonn J. Keogh, "A Compression Based Distance Measure for Texture," *Statistical Analysis and Data Mining* 3, no. 6 (2010): 381-398.

15   Michael Mayo and Anna T. Watson, "Automatic Species Identification of Live Moths," *Knowledge-Based Systems* 20, no. 2 (March 2007): 195-202.

16   Olli Silvén, Matti Niskanen, and Hannu Kauppinen, "Wood Inspection with Non-supervised Clustering," *Machine Vision and Applications* 13, no. 5–6 (March 2003): 275–285.

17   Imad Rida, Sambit Bakshi, Hugo Proença, Lunke Fei, Amine Nait-Ali, and Abdenour Hadid, "Forensic Shoe-print Identification: A Brief Survey," *arXiv preprint arXiv:1901.01431* (2019).

About the Authors

Neil **Farrugia** is a Software Engineer with interests in the area of image processing and digital forensics. He graduated with a Bachelor's in Software Development at the University of Malta and continued his studies with a Master of Enterprise in Knowledge Based Entrepreneurship from the University of Malta. Neil has also participated in a local startup competition held by the Malta Information Technology Agency, where he and his team were among the winners with an application that applied image processing techniques to satellite data to generate environmental assessments.

Dr. Joseph G. **Vella** lectures and researches in the areas of database technology at the University of Malta. His first degree was a BSc in Mathematics and Computing (UM) and a doctoral degree from the University of Sheffield Engineering Faculty. Dr. Vella has participated in numerous projects at national and EU levels mainly dealing with data integration and consolidation for data warehousing, data science, and cloud services.