JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, January 2010, Vol. 55, No. 1
doi: 10.1111/j.1556-4029.2009.01209.x
Available online at: interscience.wiley.com

PAPER

CRIMINALISTICS

Nicholas D. K. Petraco,[1,2] *Ph.D.; Carol Gambino,*[1] *M.S.; Thomas A. Kubic,*[1,2,3] *Ph.D.; Dayhana Olivio,*[1]*; and Nicholas Petraco,*[1,3,4] *M.S.*

Statistical Discrimination of Footwear: A Method for the Comparison of Accidentals on Shoe Outsoles Inspired by Facial Recognition Techniques

ABSTRACT: In the field of forensic footwear examination, it is a widely held belief that patterns of accidental marks found on footwear and footwear impressions possess a high degree of "uniqueness." This belief, however, has not been thoroughly studied in a numerical way using controlled experiments. As a result, this form of valuable physical evidence has been the subject of admissibility challenges. In this study, we apply statistical techniques used in facial pattern recognition, to a minimal set of information gleaned from accidental patterns. That is, in order to maximize the amount of potential similarity between patterns, we only use the coordinate locations of accidental marks (on the top portion of a footwear impression) to characterize the entire pattern. This allows us to numerically gauge how similar two patterns are to one another in a worst-case scenario, i.e., in the absence of a tremendous amount of information normally available to the footwear examiner such as accidental mark size and shape. The patterns were recorded from the top portion of the shoe soles (i.e., not the heel) of five shoe pairs. All shoes were the same make and model and all were worn by the same person for a period of 30 days. We found that in 20–30 dimensional principal component (PC) space (99.5% variance retained), patterns from the same shoe, even at different points in time, tended to cluster closer to each other than patterns from different shoes. Correct shoe identification rates using maximum likelihood linear classification analysis and the hold-one-out procedure ranged from 81% to 100%. Although low in variance, three-dimensional PC plots were made and generally corroborated the findings in the much higher dimensional PC-space. This study is intended to be a starting point for future research to build statistical models on the formation and evolution of accidental patterns.

KEYWORDS: forensic science, footwear, shoes, multivariate, principal component analysis, linear discriminant analysis, pattern recognition, accidental marks, accidentals

Footwear impression evidence is present at many crime scenes and can be found visible or latent on a variety of surfaces such as glass, carpet, paper, wood, dirt, concrete, tile, and snow (1). Shoe impressions can be more of a challenge for a criminal to avoid leaving than fingerprints, and like fingerprints, they can link a person to a crime scene (1). Nevertheless, footwear impression evidence is much less utilized because it is more difficult to spot and collect and more prone to contamination. Also, the ability to make positive identifications between a suspect's shoes and crime scene impressions is not as well known to those in criminal law (1).

Shoe impressions can be identified based on class characteristics like manufacturer, brand, model, and shoe size (2). There are thousands of different shoe designs for men and women, as well as a variety of sizes. The rapid rate that shoe designs are replaced adds to the discriminating power of shoe print evidence. Aside from design, the possible imperfections, variations, and random characteristics introduced during the manufacturing process can significantly reduce the number of possible candidates in the identification of an unknown impression (1).

Accidental characteristics (accidentals) are nonreproducible cuts, tears, punctures, and the like that accumulate on the outsole as the shoe is worn (2). Much like minutiae on fingerprints, footwear accidentals are identified based on agreement in a feature's appearance and position. Fingerprint minutiae, however, only have a finite number of descriptors. The possible shapes of a shoe accidental mark are infinite (1,3,4). Hence, if the shape of an accidental has enough complexity, it is theorized that just one would be enough to make a positive identification (1).

While it is a strongly held belief by many footwear examiners that the patterns of accidental marks on shoes are unique, this is an inductive conclusion that has not been thoroughly studied using controlled experiments. This poses a problem in the wake of the Daubert decision in which the U.S. Supreme Court rejected the Frye "general acceptance rule" concerning the admissibility of certain evidence submitted as scientific (5,6). As a result, various forms of physical evidence have been the subject of Daubert and other admissibility challenges. While the use of footwear impression evidence in criminal trials has recently been upheld by the United States Court of Appeals in a 2006 Daubert challenge case, future challenges are inevitable (7).

Taken literally, the adjective "unique" applied to accidental patterns means that there is one and only one pattern like it in the

[1]Department of Science, John Jay College of Criminal Justice, City University of New York, 899 10th Avenue, New York, NY 10019.
[2]Faculty of Chemistry, Graduate Center, City University of New York, 365 5th Avenue, New York, NY 10016.
[3]Faculty of Forensic Science, Graduate Center, City University of New York, 365 5th Avenue, New York, NY 10016.
[4]New York City Police Department Crime Laboratory, Trace Evidence Unit, 150-14 Jamaica Avenue, Jamaica, NY 11432.

Received 6 Oct. 2007; and in revised form 12 Jan. 2009; accepted 31 Jan. 2009.

world. This is a conclusion impossible to prove unless the accidental pattern of every shoe in use in the world is known at every point in time. With such a seemingly impossible task at hand, i.e., to "prove the uniqueness" of an accidental pattern, does this mean all is lost? Of course not! There exists a vast array of statistically based methodologies which, when given complicated pattern data, can be used to gauge similarity in a statistical sense. Many of these methods are known to be robust even using small to medium sample sizes (8,9). Other methods, given enough data, should be able to yield random match probabilities (9–11). Unfortunately, few attempts have been made to apply actual mathematical formulas to the study of accidental patterns. A way to buffer against admissibility challenges (e.g., Daubert challenges) is to analyze the data with statistical methods. Once implemented, these forensic pattern comparison systems can be extensively tested and identification error rates can be established. In this way, one can generate quantitative comparisons based on sound scientific (statistical) principles and lend objectivity and reliability to a field seen largely as subjective.

Everett, Lambert, and Buckleton have suggested a Bayesian approach to interpreting footwear marks (12). They advocate applying their method to footwear identification when the acquired features alone are not overwhelming enough to warrant a sound positive identification. Another study was performed by Geradts et al. who used algorithms to construct a footwear database called REBEZO in cooperation with the Dutch police. The data consisted of shoeprints found at crime scenes, shoes obtained from suspects, and store-bought shoes (13,14). The algorithm segments shoe sole profiles and attempts to identify and classify distinguishable shapes for comparison against a database of known shoe sole profiles. The authors note that currently the system has difficulty comparing complex shapes and that more research is needed (15).

Computational models of facial recognition have proved extremely successful in criminal investigation and security systems. These numerical pattern comparison techniques are fast, reliable, and relatively easy to understand. While there are some differences from one model to the next, they all attempt to represent a facial image as a data set which is compared to other data sets stored in a database (16–18). Such data sets will obviously be very complex, and distinguishing between them requires use of a computer that can sort through vast amounts of data and perform complicated pattern recognition tasks. Using computers in pattern recognition has the added benefit of lessening human bias introduced in gauging how "similar" two patterns of data are to one another (10,11,15,18,19).

In facial pattern recognition, a particular scheme stemming from information theory decomposes the data set representation of an image (a facial image, an accidental pattern on a shoe sole, etc.) into a smaller set of characteristic "features" known as principal components (PCs) (20,21). Principal component analysis (PCA) essentially eliminates information which varies little within all the accidental patterns included in the analysis, and captures the variation within the data independently of any human judgment. The method accomplishes this task using new statistically uncorrelated and orthogonal variables constructed from the old variables. PCA serves to reduce the dimension of the data, which can be enormous, to a manageable level.

The purpose of this study is to use facial pattern recognition techniques to demonstrate that accidental patterns found on footwear outsoles can be compared against each other within a statistical pattern recognition framework. Using such a framework we then show how identification error rates of the system can be estimated. In this study, we only used the accidental pattern from the top of the sole, and only the positions of the accidentals were recorded. Size and shape were not used in comparisons due to the fact that characteristics having amorphous properties are very difficult to treat computationally (22,23). While our programs are evolving to take these features into account, as of yet, they cannot. Second, our minimal treatment of the accidental patterns examined in this study allow us to show how robust statistical discriminations can be made even with a minimal amount of information, i.e., using only the distribution of the accidentals on the top sole impression.

In this study, five pairs of shoes having the same manufacturer and model were worn by the same person for a period of 30 days each. Although there have been studies monitoring the appearance of accidental characteristics over time (2), there have been none on the same model shoe worn by the same person. Under these circumstances, one would expect the greatest positional agreement of accidentals as many of the usual variables will be constant such as manufacturer, material, foot morphology, weight distribution, walking pattern, and routine. Aside from observing how similar the footwear patterns will be under these conditions, this study will provide a starting point for future research upon which to build statistical information on the formation and evolution of accidental patterns.

Methodology

Five pairs of ladies Lands' End, size 7 med (model D86 M30400 565) shoes were worn for a period of 30 days each. An initial shoe print was recorded before wear. Shoe prints were subsequently recorded on days 1 through 7, 14, 16, 18, 20, 24, 28, and 30. A total of 15 patterns were recorded for each shoe. Four replicates were made of each print for evaluating repeatability and because the first print made was always too dark to see fine detail. The naming convention used for each shoe distinguished order of pair worn and left or right. For example, the first shoe pair worn is called P1. The left shoe of P1 is called P1L and the right shoe is called P1R.

Unfortunately most of the accidental patterns for the left shoe of pair 5 (P5L) were not readable and thus all the patterns for P5L were dropped from this study. Hence, there are nine shoes in this study: P1L (shoe 1), P1R (shoe 2), P2L (shoe 3), P2R (shoe 4), P3L (shoe 5), P3R (shoe 6), P4L (shoe 7), P4R (shoe 8), and P5R (shoe 9).

Generation of Outsole Prints

The magna brush method was used to record the prints on to 8″ × 11″ white copy paper as it is found to be superior to dusting fingerprint powders when dealing with nonsmooth and porous surfaces (24). Black magnetic flake powder (all particles magnetic) was used over magna powder (iron particles mixed with fine powder) as it enhances wet prints better, and produces much lower background and smudge levels (24).

Recording Accidental Marks

Only the position and quantity of accidentals were considered in this study. Neither their size nor morphology was evaluated. Accidentals were recorded using a charting method adapted from the Abbott grid locater and the method adapted from a paper involving statistical analysis of barefoot impressions performed by Kennedy et al. (25–27). Figure 1 shows the grid used in this study to record the accidental patterns.

Using the prints from the magna brush technique, the best replicate from each interval was selected. Two lines were drawn tangent

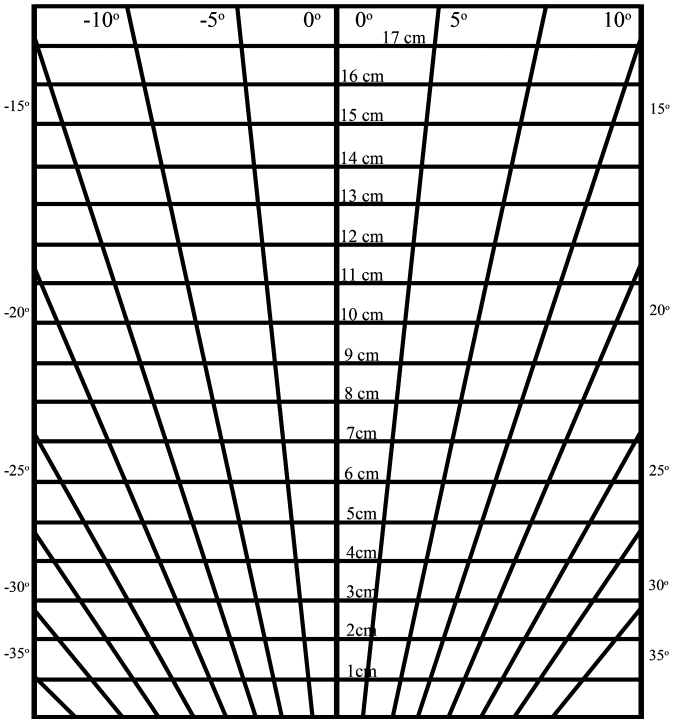

FIG. 1—*Grid used to record the accidental patterns.*

to the widest part of the print on the ball portion of the shoe's sole, and another two lines tangent to the widest part of the print on the heel of the shoe. The midpoint between each of these was marked. Another line was made perpendicular to the tangent lines, directly under where the shoe pattern of the sole ended. The two midpoints of the ball and heel were then connected.

The prepared grid (cf. Fig. 1) was printed on transparent film so that it could be laid directly over the shoe print. The *X* axis of the grid was placed over the perpendicular line drawn below the end of the sole's pattern. The *Y* axis was placed over the line that connects the two midpoints. In this way, all accidentals in the ball part of the shoe could be characterized. In this study, only the top portion of the shoe sole was examined, not the heel.

The number of accidentals found in each quadrant was recorded into charts for each shoe and day of wear. In order for an accidental to be counted it must have appeared clearly in more than one replicate. If an accidental extended into two different quadrants, the quadrant was selected in which the majority of the accidental was found. Because of wear and variations in the manufacturing process and in making each print (the amount of oil used, how one stepped on the paper, etc.), the shoeprint proportions were not exactly the same for each print. Thus, using the above method of preparing the print for the transparent grid may yield different results (the area on the sole in which the boxes lay may be different). In order to deal with this problem a shoeprint for both the right and left foot was selected to be used as a master template whose proportions were used for all other prints regardless of their own proportions. In this way, the boxes of the grid were positioned over the same place for each print.

Statistical Methods

In facial recognition, a two-dimensional image is numerically represented as a vector (a one-dimensional list) of pixels that make it up. For example, a 256×256 pixel image is rearranged into a 65,536 unit long vector. Each pixel is a box of varying color and intensity. An accidental pattern on the outsole of a shoe may also be "pixilated" or rather divided up into boxes and rearranged into a vector. Each box that makes up the accidental pattern contains a varying number of accidental marks. The accidental patterns are compared by first decomposing them into their PCs and then using a metric function to measure their distance apart in PC-space. We will use the method of maximum likelihood Gaussian linear classification analysis (sometimes also called linear discriminant analysis, LDA) to numerically gauge the similarity between patterns based on their proximity in PC-space (10,11,28).

The grid for recording the positions of accidental marks was 18×18 boxes as shown in Fig. 1. The boxes, with the number of accidentals they contained, were translated into feature vectors $\mathbf{X}_i$ (9,19). In this study, the components of the feature vector are the number of accidental marks appearing in a particular grid box of a given shoe on a given day. Data were stored in Excel and algorithms for the accidental pattern comparisons were written using the Mathematica computer algebra system (29). Initially, a total of fifteen accidental patterns were to be recorded for each of ten shoes (five pairs) for a total of 150 accidental patterns, but because accidental marks for the left shoe of the fifth pair came out unreadable, only 135 accidental patterns were recorded. Furthermore, only those accidental patterns with at least one accidental mark were included in the PCA and maximum likelihood Gaussian—linear classification analysis (MLG-LCA), leaving a total of 116 accidental patterns.

The translation of the accidental patterns into feature vectors was performed by stacking each 18-unit-long column of the grid beneath the one after it starting from the leftmost column. This procedure yielded feature vectors 324 units long (18×18 boxes = 324 boxes) which was then assembled into an $n \times P$ data matrix $\mathbf{X}$, where n is the number of accidental patterns (feature vectors, here 116) to be used in a given analysis, and $P = 324$.

$$\mathbf{X} = \begin{bmatrix} X_{1,1} & \cdots & X_{1,j} & \cdots & X_{1,324} \\ \vdots & & \vdots & & \vdots \\ X_{i,1} & \cdots & X_{i,j} & \cdots & X_{i,324} \\ \vdots & & \vdots & & \vdots \\ X_{116,1} & \cdots & X_{116,j} & \cdots & X_{116,324} \end{bmatrix}$$

Every box in the feature vector represents a random variable and every row in the data matrix is a vector of number of accidental marks observed. The symbol $\mathbf{X}_i$, designates a (row) vector of data representing accidental pattern *i*. A data matrix with *n* rows contains *n* accidental patterns. The average of all row vectors in $\mathbf{X}$ is the average vector $\overline{\mathbf{Z}}_i$. The multivariate analyses of data set ($\mathbf{X}$) undertaken in this study were PCA and MLG-LCA. For details on these methods see reference (30) and references therein. The Mathematica notebooks developed for this study are available upon request from the authors.

Since the feature vector of an accidental pattern is simply a point (in a high dimensional space) the similarity between patterns can be gauged by an appropriate distance metric and decision algorithm. The degree of "sameness" between two arbitrary accidental patterns was determined numerically by using MLG-LCA (28). The PC-derived data matrix $\mathbf{Z}$ was used in place of the original data matrix of accidental patterns $\mathbf{X}$, due to its significantly smaller size (116×32 at most for $\mathbf{Z}$ vs.116×324 for $\mathbf{X}$ in one case) and due to the fact that direct application of MLG-LCA to $\mathbf{X}$ was impossible due to problems encountered with singular pooled covariance matrices required by the algorithm.

The MLG-LCA decision model uses a distance function to find the mean feature vector $\overline{\mathbf{Z}}_i$ that is closest to the "test" feature vector $\mathbf{Z}_j$.

The actual discriminant function constructed for the patterns from shoe i is given as

$$L_i(\mathbf{Z}_j) = \overline{\mathbf{Z}}_i^{\mathrm{T}} \mathbf{S}_{\mathrm{pl}}^{-1} \mathbf{Z}_j - \frac{1}{2} \overline{\mathbf{Z}}_i^{\mathrm{T}} \mathbf{S}_{\mathrm{pl}}^{-1} \overline{\mathbf{Z}}_i$$

where $\mathbf{S}_{\mathrm{pl}}^{-1}$ is the inverse of the pooled covariance matrix. See reference (30) for a detailed description. A total of $k = 9$ discriminant functions were constructed, one for each shoe. The algorithm decides that the accidental pattern j is most similar to the predefined set of accidental patterns from shoe i according to the decision rule

$$\arg \max_i L_i(\mathbf{Z}_j).$$

The predefined sets of accidental patterns were chosen to be all those patterns recorded for a particular shoe over the course of the 30 days. Thus, there are nine sets of accidental patterns, one set for each shoe. Each set consists of 15 patterns, one for each day that data was recorded. In words, the decision rule above means: "the (PCA reduced) accidental pattern $\mathbf{Z}_j$ is most similar to the set of (PCA reduced) accidental patterns from shoe i whose discriminant function yields the largest value" (28).

The ability of discriminant functions to accurately predict the sample identity of a pattern which they have not been trained with, is called classification error analysis (31). This is a very important topic whenever statistical pattern recognition techniques are applied to forensic evidence. This is because discriminant functions, while trained on a finite (probably small) set of data, will be expected to classify or identify new pieces of evidence which they have not been trained with. Thus, rigorously derived accurate estimates for error rates of computed sets of discriminant functions are critical in forensic science applications. For this study, we estimate the error rates of the k discriminant functions in three different ways. We actually compute estimates of the "correct classification rate" which is one minus the error rate and is reported as a percentage.

The first estimate used is the "apparent" correct classification rate computed by determining the number of accidental patterns assigned to their correct sample (by the discriminant functions) divided by the total number of accidental patterns. This performance estimate is known to be biased and tends to yield an overly optimistic correct classification rate (28).

The second estimate is the overall "hold-one-out" correct classification rate (10,32). This is computed by first recalculating the linear discriminant functions omitting a single accidental pattern from the data set. Thus, the recalculated discriminant functions are not trained to identify the held out accidental pattern. This omitted accidental pattern is then classified with the recalculated linear discriminant functions and the process is repeated sequentially for each accidental pattern in the data set. The number of correctly classified "held-out" accidental patterns is divided by the total number of accidental patterns in the entire data set (116 in this study) to yield the overall hold-one-out correct classification rate.

Finally, the "average hold-one-out" correct classification rate is computed. This process involves replicating a data set composed of n observation vectors, n-times. Each replicate data set, however, contains all but one of the original data vectors (33). The n data sets are then used to recalculate a statistic on that data set in the absence of the deleted data vector, producing a set of estimates of the statistic. The set can then be used to produce an average and standard deviation for the statistic (33). The average hold-one-out correct classification rate is mathematically the least biased estimation of the discrimination functions' classification performance (13).

Here, we compute the average hold-one-out correct classification rate by first computing all the samples' hold-one-out correct classification rates and recording them in a "cross-validation table." Next, the average and standard deviation of the samples' hold-one-out correct classification rates is found yielding the average hold-one-out correct classification rate (10,32,33).

Results and Discussion

PCA of All Accidental Patterns with At Least One Accidental Mark

Out of 135 accidental patterns recorded for nine shoes (cf. Methodology section, paragraph two), 116 contained at least one accidental mark. These 116 accidental patterns were processed with PCA. It was found that for the shoes in this study (all worn by the same person and all the same make and model), 32 PCs described 99.5% of the total variance in the data set. One can think of variance as the overall structure of the data. Thus, the 292 dimensions excluded collectively only accounted for 0.5% of the data's structure. This 32D data set was then subject to MLG-LCA, also called linear discriminant analysis (LDA, cf. [9]), in order to quantitatively probe the differences between the accidental patterns generated by each shoe. Table 1 shows the hold-one-out cross validation results for the correct classification of each accidental pattern using MLG-LCA. The overall hold-one-out correct classification rate was 92% (97% apparent correct classification rate). The average hold-one-out correct classification rate was 92 ± 9%. We were surprised at these high correct classification rates, especially considering the fact that evaluation of the data did not include details of the accidentals (such as size and shape) or something like the outsole topography.

The 32D structure of this data is obviously too high in dimensionality to plot. It is none-the-less very instructive to have a physical picture of the data. For this reason projection of the 116 accidental patterns into three-dimensional (3D) PC-space is plotted in Fig. 2 which accounts for 59.7% of the data's variance. While these first three PCs only account for a small portion of the overall

TABLE 1—*Hold-one-out cross-validation table for maximum likelihood Gaussian linear classification of the accidental patterns examined in this study.*

Shoe ID	No. Accidental Patterns Recorded*	No. Misidentified Patterns for Shoe	Incorrectly Predicted Shoe	Individual Shoe "Hold-One-Out" Correct Identification Rates (%)
P1L	12	0		100
P1R	11	0		100
P2L	12	1	P1R	92
P2R	12	0		100
P3L	11	2	P1L, P2L	82
P3R	15	2	P4R, P5R	87
P4L	15	0		100
P4R	13	3	P1L, P5R × 2	77
P5L	0	Omit	Omit	Omit
P5R	15	1	P3R	93

The 324-dimensional accidental patterns were reduced to 32 dimensions (99.5% of total variance) using PCA. This table shows that the average correct identification rate for accidental patterns found on these shoes is estimated to be 92%.

*The accidental patterns used in the classification analysis were those that had at least one accidental mark.

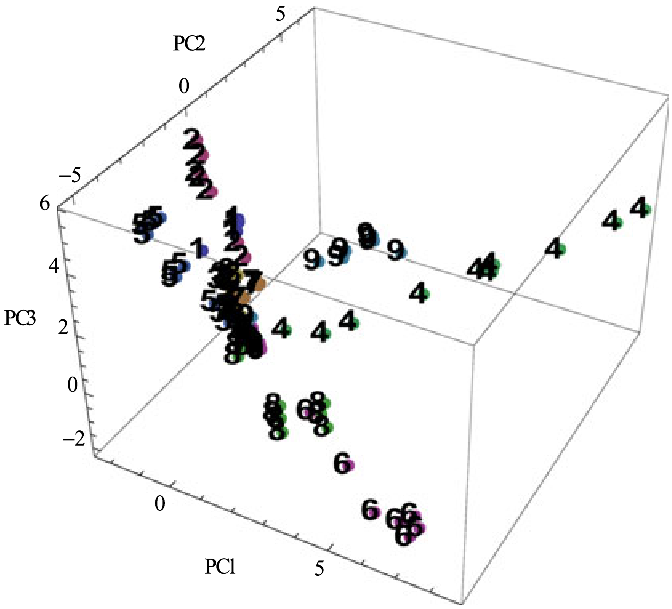

FIG. 2—*All accidental patterns with at least one accidental mark, projected into the space of the first three PCs (59.7% of total variance). The numbers adjacent to each data point label the shoe.*

variance in the data, some clustering of the accidental patterns for particular shoes is evident. This is consistent with the findings in 32D PC-space, i.e., most of the patterns generated by the same shoe are in close proximity. Within the classification theory we employ in this study, MLG-LCA, the more proximate data points are in space, the more likely their identity is the same, i.e., the more likely they are drawn from the same distribution (28,32).

In general, the accidental patterns appear to evolve by starting as a clump of points (i.e., patterns with only one or two accidental marks) and then spreading out linearly in the PC1-PC2 plane (cf. Fig. 2). Interestingly, the accidental patterns of shoes 2 (P1R), 4 (P2R), 5 (P3L), 6 (P3R), 8 (P4R), and 9 (P5R) trace out fairly linear paths in 3D PC-space (Alternative viewpoints of Fig. 2 are available from the authors upon request.) Patterns from shoe 4 (P2R), while somewhat spread out are nonetheless strikingly distinct from the other accidental patterns. These patterns seem to follow a very linear path through 3D PC-space as they change over time. Shoe 9 (P5R) shows accidental patterns that are much closer to each other than those for shoe 4 (P2R) and are also distinct from most of the other patterns. Unfortunately, we cannot infer much about the linear paths traced out by some of the patterns as they evolve over time since these 3D plots account for a relatively low amount of the data's overall variance. The intention of these figures is only to examine if the accidental patterns for the same shoe are relatively close together in 3D PC-space and form distinct clusters.

PCA of Accidental Patterns from Days 14 to 30

The accidental patterns from the first 7 days contained few or no accidental marks. Thus, these patterns are necessarily similar and all tightly clustered (cf. lower left of Fig. 2). When these patterns are removed and the remaining 63 accidental patterns from days 14 to 30 are dimensionally reduced with PCA, the first 28 PCs account for 99.5% of the data's variance. Table 2 shows the hold-one-out cross validation results for correct classification of these accidental patterns using MLG-LCA. The overall hold-one-out

TABLE 2—*Hold-one-out cross-validation table for maximum likelihood Gaussian linear classification of the accidental patterns for days 14–30.*

Shoe ID	No. Accidental Patterns Recorded	No. Misidentified Patterns for Shoe	Incorrectly Predicted Shoe	Individual Shoe "Hold-One-Out" Correct Identification Rates (%)
P1L	7	0		100
P1R	7	2	P2R × 2	71
P2L	7	0		100
P2R	7	0		100
P3L	7	2	P1L × 2	71
P3R	7	1	P4R	86
P4L	7	0		100
P4R	7	0		100
P5L	0	Omit	Omit	Omit
P5R	7	0		100

The 324-dimensional accidental patterns were reduced to 28 dimensions (99.5% of total variance) using PCA. This table shows that the average correct identification rate for accidental patterns in this time period, is estimated to be 92%. This is consistent with the correct identification rate derived from Table 1.

correct classification rate was 92% (100% apparent correct classification rate). The average hold-one-out correct classification rate was 92 ± 13%. These rates are in general quite good although the individual correct classification rates for shoes 2 (P2R) and 5 (P3L) are at 71%. Considering that there are only seven patterns for each shoe, just one misidentification will strongly impact the shoe's average correct classification rate. Given the good overall correct classification rates obtained for this data we would expect the rates for shoes 2 (P2R) and 5 (P3L) would increase if more accidental patterns had been recorded between days 14 and 30.

The first three PCs accounted for 63.6% of the variance for the accidental patterns from days 14 to 30 and their projection into 3D PC-space is shown in Fig. 3. Even at this relatively low variance, accidental patterns for shoes 2 (P1R), 3 (P2L), 5 (P3L), and 9

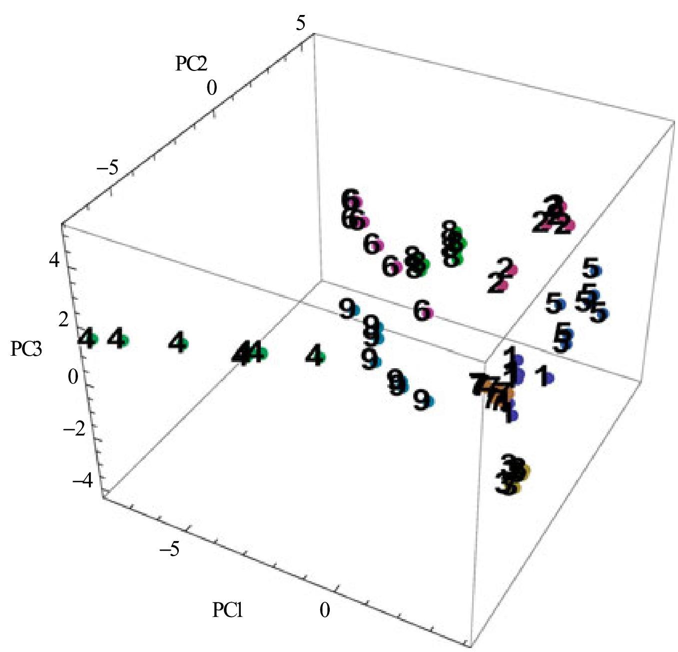

FIG. 3—*Accidental patterns for days 14–30, projected into the space of the first three PCs (63.6% of total variance). The numbers adjacent to each data point label the shoe.*

TABLE 3—*Hold-one-out cross-validation table for maximum likelihood Gaussian linear classification of the accidental patterns for days 20–30.*

Shoe ID	No. Accidental Patterns Recorded	No. Misidentified Patterns for Shoe	Incorrectly Predicted Shoe	Individual Shoe "Hold-One-Out" Correct Identification Rates (%)
P1L	4	0		100
P1R	4	2	P2R × 2	50
P2L	4	0		100
P2R	4	0		100
P3L	4	2	P1L × 2	50
P3R	4	2	P1L, P4R	50
P4L	4	1	P2L	75
P4R	4	0		100
P5L	0	Omit	Omit	Omit
P5R	4	0		100

The 324-dimensional accidental patterns were reduced to 28 dimensions (99.6% of total variance) using PCA. This table shows that the average correct identification rate for accidental patterns in this time period is estimated to be 81%. See text for discussion of this much lower average correct identification rate.

(P5R) are clearly distinct from each other and easy to pick out by eye. Also, when viewing Fig. 3 straight up the PC1 axis (not shown) shoes 6 (P3R) and 8 (P4R) clearly form distinct clusters. The accidental patterns for shoe 7 (P4L) are intermingled, however, with those for shoe 1 (P1L) from any point of view.

PCA of Accidental Patterns from Days 14 to 20

Next, we examine the 36 accidental patterns from days 14 to 20. The first 21 PCs account for 99.6% of the variance in this data set. Table 3 shows the hold-one-out cross validation results for classification of these accidental patterns using MLG-LCA. The overall hold-one-out correct classification rate was 81% (100% apparent correct classification rate). The average hold-one-out correct classification rate was 81 ± 24%. Note that while the overall and average correct classification rates for this data set are low there are only four patterns for each shoe. Each misidentification by MLG-LCA would thus be expected to impact these correct classification rates by a wide margin.

Figure 4 shows a plot of the first three PCs for accidental patterns from days 14 to 20. The plot accounts for 65.7% of this data set's variance. Good clustering of patterns (points) stemming from the same shoe can be seen during this third week of wear. The patterns of shoe 3 (P2L) appear intermingled with those for shoe 1 (P1L). However, if Fig. 3 is viewed straight down the PC1-axis one would see that this is not the case. Similarly, viewing the data straight down the PC3-axis reveals that the patterns for shoes 5 (P3L), 6 (P3R) and 8 (P4R) are all distinct and well separated in space (alternative views of Fig. 3 are available from the authors upon request). Only shoes 1 (P1L) and 7 (P4L) are too close to visually differentiate in 3D PC-space.

PCA of Accidental Patterns from Days 20 to 30

The best classification results of accidental patterns were obtained for the last recorded week of wear, days 20–30 (36 patterns). From a footwear examiners point of view, this is not unexpected as the shoes have accumulated the most wear and therefore have developed the most elaborate patterns of random accidental marks. The first 22 PCs accounted for 99.5% of the data's variance. All hold-one-out cross validation results for correct classification of

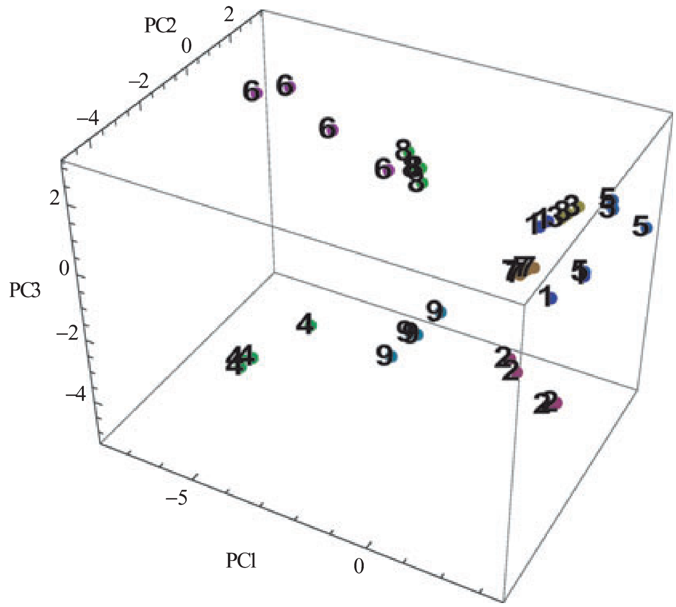

FIG. 4—*Accidental patterns for days 14–20, projected into the space of the first three PCs (65.7%). The numbers adjacent to each data point label the shoe.*

these accidental patterns using MLG-LCA in 22D PC-space were 100%.

For an approximate, although visual representation of how different the accidental patterns stemming from each shoe are at this point, the data from days 20 to 30 projected into the space of the first three PCs in Fig. 5 (accounts for 66.7% of the data's variance). Even at this relatively low variance value Fig. 5 conveys that the patterns for each shoe form absolutely distinct and well-separated clusters in 3D PC-space. Overall, as time passes it is exquisitely clear that the accidental patterns developed on the outsoles of these

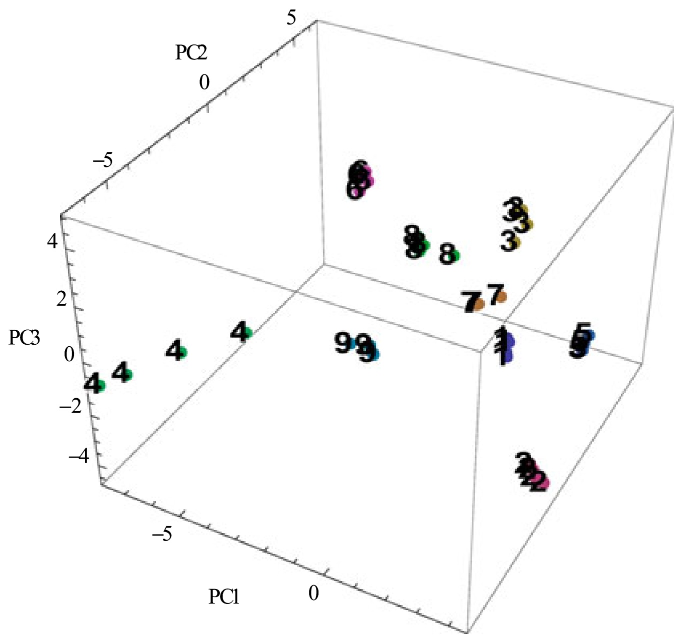

FIG. 5—*Accidental patterns for days 20–30, projected into the space of the first three PCs (66.7% of the total variance). The numbers adjacent to each data point label the shoe.*

shoes became more and more distinguishable, i.e., the inter-cluster spread in data space increases over time.

Conclusion

The footwear accidental pattern comparison technique presented in this study utilized only a tiny amount of the information typically available to the footwear examiner and yet it was usually able to correctly identify which shoe generated a particular pattern. The way the method works is to mathematically compare the distribution of accidental marks (accidental patterns) on the sole of an unknown shoe to multiple accidental patterns generated by shoes of known identity. The statistical comparison method used in this study was maximum likelihood Gaussian linear classification. The identity of the unknown pattern is assigned to the known shoe with statistically the most similar accidental patterns.

The high correct classification rates from our minimally detailed data lend a great deal of credence to the proposition postulated by imprint examiners of the "uniqueness" of accidental patterns. If data are also recorded for the physical characteristics of each accidental, the above results indicate that this method would be even more successful in identifying a shoe from one or more related accidental patterns.

Patterns from the same shoe although at different points in time tended to cluster closer to each other than patterns from different shoes. This was demonstrated (numerically) in high dimensional PC-space using MLG-LCA, and (graphically) in 3D PC-space. The 3D PC-space plots graphically show that generally there is no relationship between the patterns of the left and right shoes from the same pair, as might be expected. If there were such a relationship then one would expect to see tighter clustering between the patterns (i.e., points in the 3D plots) from the same pair of shoes.

Correct classification rates using MLG-LCA and the hold-one-out procedure ranged from 81% to 100% when 99.5% of the data's variance was retained. Two factors affected the correct classification rates, length of time the shoe was worn and the number of accidental patterns included in the analysis. Most notably, the longer the shoe was worn, the more different the patterns became. Although some of this information is well known to the trained footwear examiner, in a court of law if one can use sophisticated yet understandable statistical methods to draw conclusions about evidence and discuss statistical certainties, one can make a profound impression on the courts and support expert footwear examiner testimony that has been arrived at qualitatively.

By using the same manufacturer and model of shoe as well as having the same person as the wearer, many variables that contribute to the "unique" characteristics formed on shoe soles were eliminated or muted. Hence, the ability to still easily distinguish between such shoes with minimally detailed data strongly supports the claims of the great discrimination power of footwear impressions. Logic then dictates that the inclusion of accidental mark details, such as size and shape, will further add to this method's discriminating power.

We have already begun to expand upon this study by increasing the number of participants, frequency of data collection, and lengthening the total time over which accidental patterns are collected. We believe this study will further strengthen and elaborate on our findings here. In the future, we would like to implement automated methods of data collection for the shoes, in particular using high resolution laser scanning to map the surface topography of the outsole. Computer aided design software could then be used to make unbiased measurements and projections of the data for comparison to other shoe soles and shoe sole impressions.

Acknowledgments

We thank Ms. Helen Chan, Mr. Manny Chapparo, Mr. Peter Diaczuk, and Ms. Marta Ekstrom for many valuable discussions and for reading our manuscript.

References

1. Bodziak W. Footwear impression evidence. 2nd edn. Boca Raton, FL: CRC Press, 1995.
2. Cassidy MJ. Footwear identification. 1st edn. Salem, OR: Lightning, 1980.
3. Deskiewicz K, editor. Schallamach pattern on shoe outsole acknowledged by court in footwear identification. Proceedings of the International Symposium on Setting Quality Standards for the Forensic Community; 1999 May 3-7; San Antonio, TX. Washington, DC: Forensic Science Communications (FBI), 1999.
4. Davis RJ, Keeley A. Feathering of footwear. Sci Justice 2000; 40(4):273–6.
5. *Daubert v. Merrell Dow Pharmaceuticals,* 509 U.S. 579. (1993).
6. Bohan TL, Heels EJ. The new scientific evidence "standard" and the standards of several states. J Forensic Sci 1995;40(6):1030–44.
7. *United States v. Mahone*, 453 F.3d 68 (1st Cir. 2006).
8. Vapnik VN. Statistical learning theory. 1st edn. New York, NY: Wiley, 1998.
9. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning. New York, NY: Springer, 2001.
10. Duda RO, Hart PE, Stork DG. Pattern classification. 2nd edn. New York, NY: Wiley, 2001.
11. Theodoridis S, Koutroumbas K. Pattern recognition. 3rd edn. San Diego, CA: Academic Press, 2006.
12. Everett IW, Lambert JA, Buckleton JS. A Bayesian approach to interpreting footwear marks in forensic casework. Sci Justice 1998;38(4): 241–7.
13. Geradts Z, Keijzer J. The image-database REBEZO for shoeprints with developments on automatic classification of shoe outsole designs. Forensic Sci Int 1996;82:21–31.
14. Geradts Z, Keijzer J, Keereweer I, editors. Automatic comparison of striation marks and automatic classification of shoe marks. Proceedings of SPIE; 1995 July 9; San Diego, CA. Washington, DC: SPIE, 1995.
15. Geradts Z. Content-based information retrieval from forensic image databases [dissertation]. Utrecht: University of Utrecht, 2002.
16. Kirby M. Geometric data analysis: an empirical approach to dimensionality reduction and the study of patterns. 1st edn. New York, NY: Wiley, 2000.
17. Trucco E, Verri A. Introductory techniques for 3D facial recognition. 1st edn. New York, NY: Prentice-Hall, 1998.
18. Grenander U, Miller M. Pattern theory: from representation to inference. 1st edn. New York, NY: Oxford, 2007.
19. Fukunaga K. Statistical pattern recognition. 2nd edn. San Diego, CA: Academic Press, 1990.
20. Turk M, Pentland A. Eigenfaces for recognition. J Cognitive Neurosci 1991;3(1):71–86.
21. Jolliffe IT. Principal component analysis. 2nd edn. New York, NY: Springer, 2004.
22. Burger W, Burge MJ. Digital image processing: an algorithmic introduction using Java. 1st edn. New York, NY: Springer, 2008.
23. Gonzalez RC, Woods RE. Digital image processing. 3rd edn. Saddle River, NJ: Pearson, 2008.
24. Wilshire B, Hurley N. Development of two-dimensional footwear impressions using magnetic flake powders. J Forensic Sci 1996;41(4): 678–80.
25. Abbott JR. Footwear evidence. Springfield, IL: Charles C. Thomas, 1964.
26. Kennedy RB, Pressmann IS, Chen S, Petersen PH, Pressman AE. Statistical analysis of barefoot impressions. J Forensic Sci 2003;48(1):55–63.
27. Kennedy RB, Chen S, Pressmann IS, Yamashita AB, Pressman AE. A large-scale statistical analysis of barefoot impressions. J Forensic Sci 2005;50(5):1071–9.
28. Rencher AC. Methods of multivariate analysis. 2nd edn. Hoboken, NJ: Wiley, 2002.
29. Wolfram Research, Inc. Mathematica [computer program]. Version 5.1. Champaign, IL: Wolfram Research, Inc. 2005.

30. Petraco NDK, Gil M, Pizzola PA, Kubic TA. Statistical discrimination of liquid gasoline samples from casework. J Forensic Sci 2008;53(5): 1092–101.
31. Shawe-Taylor J, Cristianini N. Kernel methods for pattern analysis. 1st edn. London, UK: Cambridge University Press, 2004.
32. Mardia KV, Kent JT, Bibby JM. Multivariate analysis. 1st edn. Amsterdam, Holland: Academic Press, 1980.
33. Lanyon SM. Jackknifing and bootstrapping: important ''new'' statistical techniques for ornithologists. Auk 1987;104:144–6.

Additional information and reprint requests:
Nicholas D. K. Petraco, Ph.D.
Department of Science
John Jay College of Criminal Justice
899 10th Avenue
New York
NY 10019
E-mail: npetraco@jjay.cuny.edu