A database of two-dimensional images; footwear outsole impressions

Soyoung Park and Alicia Carriquiry

March 24, 2020

Abstract

Footwear outsole images were obtained from 150 pairs of used shoes. The motivation for constructing the database was to enable a statistical analysis of two-dimensional (2D) images of shoe outsoles, to understand within and between image variability, and to develop methods for the evaluation of forensic pattern evidence of shoeprints. Since we scanned the outsole of the used shoes, the images capture not only the outsole pattern design but also the marks that arise from wear and tear and that may help identify the shoe that made the impression. Each shoe in a pair was scanned five times, so that replicate images can be used to estimate within-shoe variability. In total, there are 1,500 2D images in the database. The EverOS footwear scanner was used to capture the outsole of each shoe. The scanner detects the weight distribution of the person wearing the shoe when he or she steps on the scanning surface. The database is a useful resource for forensic scientists or for anybody else with an interest in image comparison. The database we describe, was constructed by researchers in the Center for Statistics and Applications in Forensic Evidence (CSAFE) at Iowa State University.

Access and Funding

These data are available under license CC-BY 4.02. Users may share and adapt the dataset. Credit must be attributed to the authors for the original creation of the data. The full license can be found at `creativecommons.org`.

This work was partially funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) through Cooperative Agreement 70NANB15H176 between NIST and Iowa State University, which includes activities carried out at Carnegie Mellon University, University of California Irvine, and University of Virginia.

Data Collection Methods

To build the collection of shoes, we asked colleagues at Iowa State University to bring in their used shoes. Therefore, the sample of shoes is a sample of convenience, and is not meant to be representative of any specific population of shoes. By accident, the database includes several pairs of shoes of the same model, brand and size owned by the same person. It also includes shoes of the same model, brand and size but owned by different individuals. Further, several pairs of shoes of the same model and brand but different sizes are also included. Mostly, the images represent a variety of shoe models with different outsole designs. Therefore, we recorded the ID of the owner, the brand and model of the shoes, size and gender. The gender information is recorded because the sizes are a mix of men's and women's.

The outsole scanner used in this study is an EverOS laboratory footwear scanner (`https://www.shopevident.com/`). This scanner acquires the footwear outsole impression by detecting the weight of the wearer. To scan the shoes, we first clean the outsole to remove some dirt. Then, the person who is wearing the shoe steps on the surface of the scanner. As she steps on the scanner, the wearer shifts her weight from the heel to the toe so that the scanner captures the outsole pattern in detail. Once the scanner detects the outsole patterns, the image with the scanned outsole can be visualized on the screen in the special software provided by EverOS.

File Formats and Naming

As mentioned above, the 300 shoes in the collection were scanned five times each. Thus, for a pair of shoes, there are 10 images, five for each shoe in the pair. In all, the database consists of 1,500 images of shoe outsole impressions from 150 pairs of shoes.

The image files are labeled using the format AAA_BB_L/R_CC, where:

- AAA is a three-digit number between 001 and 028 that identifies the owner of the pair of shoes,
- BB is a two-digit number between 01 and 10 that identifies each pair of shoes belonging to the same individual,
- L/R denotes Left and Right and indicates whether the image corresponds to the left or the right shoe in a pair, and
- CC is a two-digit number between 01 and 05 that identifies the replicate image number for a shoe.

For example, the file labelled 015_01_L_01 is the image of the first replicate of the left shoe in the first pair owned by individual 15.