Scars, marks, and tattoos: a soft biometric for identifying suspects and victims

Anil Jain and Jung-Eun Lee

An automatic system for matching and retrieving tattoo images will help advance the investigative capabilities of law enforcement and forensics agencies.

A wide variety of biometric systems have been developed for automatic recognition of individuals based on their anatomical (e.g., fingerprint, face, and iris) and behavioral (e.g., signature and gait) characteristics. Despite tremendous technological progress, there are many situations where primary biometric traits are either not available or are difficult to capture, or where the quality of the sensed images is poor. In such cases, ‘soft’ biometric traits such as height, sex, eye color, ethnicity, scars, marks, and tattoos can assist in identifying a person. Although these alternative traits cannot differentiate uniquely, they do contain discriminatory information that helps to narrow down the possibilities. Accordingly, law enforcement agencies collect and maintain this kind of demographic information in their databases. Among the many soft biometric traits, scars, marks, and tattoos (SMTs) have been particularly useful in law enforcement and forensics. Criminal identification is an important application because tattoos often contain subtle clues to a suspect’s background and history, such as gang membership, religious beliefs, previous convictions, and years spent in jail (see Figure 1). Tattoos are also useful in establishing the identity of a nonskeletalized body of a victim. This is due not only to the increasing prevalence of tattoos but also their impact on other methods of human identification based on visualization, pathology, and trauma.

Law enforcement agencies routinely photograph and catalog tattoo patterns when booking suspects (who often use aliases). Based on the ANSI/NIST-ITL (Information Technology Laboratory) 1-2007 standard, each image is manually labeled into one of 70 categories and then stored with a suspect’s criminal history record. A database search of tattoos involves matching the class label of a query tattoo with those in the database. This matching process is subjective, has limited performance, and is very time-consuming. Furthermore, simple class description in a textual query (e.g., ‘dragon’) does not include all the semantic information present in an image. Finally, the classes in the ANSI/NIST standard are not adequate to describe the increasing variety of new tattoo designs.

To efficiently and accurately match tattoo images, we have created an automatic system called Tattoo-ID. Our approach is one of content-based image retrieval using features (e.g., color,
shape, and texture), instead of labels or keywords, to compute the similarity between two images. Given a query, Tattoo-ID retrieves the top-N (say, N=20) images in the database that visually resemble it, and presents them to the user in order of similarity. ‘User feedback’ or ‘preference’ based on retrieved images could be used to improve both feature extraction and the similarity measure of the matching module. To keep our system compatible with current practice in law enforcement, Tattoo-ID also employs class and subclass labels. In other words, a user can specify both the tattoo image and its ANSI/NIST category information as part of the query. The current version of our system uses scale-invariant feature transformation\(^6\) to extract characteristic ‘keypoints’ from a tattoo image and to represent the image using descriptors (vectors) associated with the points. Matching is then performed by comparing the descriptors in two images. The Tattoo-ID system currently contains 64,000 tattoo images (courtesy of the Michigan State Police). Keypoints extracted from these images are stored into the database. In the matching and retrieval stages, users provide a tattoo image with optional ancillary information such as labels and location of the tattoo on the body. Initial results based on 1000 queries against the full database show 83.5% rank-1 and 91.2% rank-20 retrieval accuracies (see Figure 2).\(^7\) In other words, 835 out of 1000 queries found a true match in the first image returned out of a database of 64,000. Similarly, 912 queries were correctly answered from among the first 20 images retrieved.

With the rapidly growing use of tattoos for victim and suspect identification in forensics and law enforcement, the Tattoo-ID system could be of great value in apprehending suspects and identifying victims. Although the system already has excellent retrieval performance, matching of severely distorted and noisy images (e.g., those that show nonuniform illumination and blurring) is still a major challenge. Our long-term goal is to improve the system with respect to such queries. Accordingly, we are investigating additional salient features and more robust image matching. Algorithms to speed up searching are also being explored.

We thank Tattoo-ID team members Rong Jin, Fengjie Li, Unsang Park, and Nick Gregg for their support.

Author Information

Anil Jain and Jung-Eun Lee
Department of Computer Science and Engineering
Michigan State University
East Lansing, MI
http://biometrics.cse.msu.edu/

References