

Image Retrieval in Forensics: Tattoo Image Database Application

Jung-Eun Lee, Rong Jin, and Anil K. Jain
Michigan State University

Wei Tong
Carnegie Mellon University

Two modifications to the Tattoo-ID system, based on automatic content-based image retrieval, help improve its retrieval accuracy, particularly for low-quality tattoo image queries.

Whether for passports, credit cards, laptops, or mobile phones, automated methods of identifying citizens through their anatomical features or behavioral traits have become a common feature of modern life. Biometric recognition, or simply *biometrics*, refers to the automatic recognition of individuals based on their anatomical and/or behavioral characteristics.¹ One of the most well-known biometric traits is fingerprints. The success of automatic fingerprint systems in law enforcement and forensics around the world has prompted the use of biometrics in various civil identification systems. For example, in 2007 alone, the US Department of Homeland Security Immigration and Border Management System (US-VISIT, www.dhs.gov/files/programs/usv.shtm) collected fingerprint and face images of more 46 million visitors to the US.

Although tremendous progress has been made in biometrics and forensics, many situations exist where the primary biometric traits—fingerprint, face, and iris—alone cannot identify an individual with sufficiently

high accuracy. This is especially true when image quality is poor (for example, because a surveillance camera obtained a blurred image or off-central pose) or only a partial fingerprint is available (as in the case of latent fingerprints lifted at crime scenes). In the case of face recognition, the matching performance severely degrades under pose, lighting, and expression variations and because of occlusion and aging. In such cases, it is critical to acquire supplementary information to assist in the identification procedure.

On the basis of this rationale, the US Federal Bureau of Investigation is developing the Next-Generation Identification (NGI) system for identifying criminals.² In addition to using additional biometric modalities, such as a palm print and iris, to augment fingerprint evidence, the NGI system will include soft biometric traits, including scars, marks, and tattoos, collectively referred to as SMT.

The use of soft biometrics in forensics has been recognized as a valuable tool for solving crimes. This article focuses on one such soft biometric, namely tattoo images, which are routinely collected by law enforcement agencies and used in apprehending criminals and identifying suspects. The current practice of tattoo matching and retrieval, based on ANSI/NIST classes, is prone to significant errors due to limited vocabulary and the subjective nature of labeling. To improve the performance and robustness of keyword-based tattoo matching, we introduced the Tattoo-ID content-based image retrieval (CBIR) system. This system automatically extracts features from a query image and retrieves near-duplicate tattoo images from a database. In this article, we present two Tattoo-ID modifications that further improve the retrieval accuracy, particularly for queries with low quality. The modifications involve a robust similarity measure and metadata utilization in the form of free-keyword annotation in conjunction with the large lexical database WordNet. Experimental results on a database of 100,000 images show that the enhanced system achieves a top-20 retrieval accuracy of 90.5 percent.

Soft Biometrics in Use

Soft biometric traits are characteristics that provide some identifying information about an individual, but they lack the distinctiveness and permanence to sufficiently differentiate

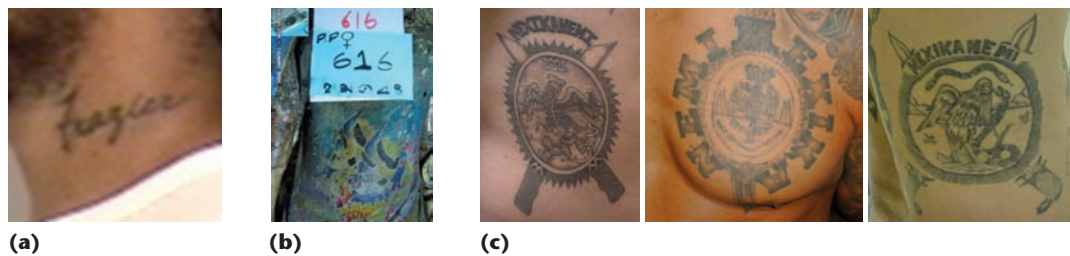


Figure 1. Tattoos for identification. Potential tattoo identification applications include (a) a crime suspect and (b) a victim of the 2004 Asian Tsunami. For gang membership tattoos of the well-known Mexikanemi Mafia gang in Texas, (c) through (e), we can see the large intraclass variability in the same gang's membership tattoos.



Figure 2. Tattoo images. These samples are from a database of 64,000 tattoo images from the Michigan State Police.

between two individuals.¹ Because soft biometric traits help narrow the identity of a suspect or a victim in forensics investigations, many law enforcement agencies collect and maintain such information in their databases. It is thus not surprising that the FBI collection standard includes prominent SMT present on a subject's body.

Among the various soft biometric traits, tattoos have been considered one of the most important pieces of evidence. Tattoos provide more discriminative information for identifying a person than the traditional demographic indicators such as age, height, race, and gender.³ In addition, because many individuals acquire tattoos to individuate themselves, display their personality, or exhibit a group membership (see Figures 1c through 1e), the analysis of tattoos often leads to a better understanding of an individual's background and membership in various organizations.

Despite the value of soft biometrics in forensics, putting them to practical use has been difficult. Unlike primary biometric traits, much variability exists in pattern types in many of the soft biometric traits. Whereas a primary biometric trait has its own unique physical representation (including ridge patterns and minutiae in fingerprints; eyes, nose, and lips on faces; and texture in irises), tattoo images often consist of objects with varying shapes,

colors, and textures (see Figure 1), making it challenging to effectively represent them. This is the main reason why researchers have made relatively little effort in automatic matching and retrieval of tattoo images.

Tattoo Image Retrieval

Tattoos engraved on the human body have been successfully used to assist human identification in forensics. This is not only because of the increasing prevalence of tattoos, but also due to their impact on other methods of human identification such as visual, pathological, or trauma-based identification. (A study published in the *Journal of the American Academy of Dermatology* in 2006 reported that about 36 percent of Americans 18 to 29 years old have at least one tattoo.⁴) Tattoo pigments are embedded in the skin to such a depth that even severe skin burns often do not destroy a tattoo. For this reason, tattoos helped identify victims of the 9/11 terrorist attacks and the 2004 Asian tsunami³ (Figure 1b). Criminal identification is another important application because tattoos often contain hidden meanings related to a suspect's criminal history, such as gang membership, previous convictions, years spent in jail, and so forth (see Figures 1 and 2).

Law enforcement agencies routinely photograph and catalog tattoo patterns for the

purpose of identifying victims and suspects (who often use aliases). The ANSI/NIST-ITL1-2011 standard defines eight major classes (human, animal, plant, flag, object, abstract, symbol, and other) and a total of 70 subclasses (including male face, cat, narcotics, American flag, fire, figure, national symbols, and wording) for categorizing tattoos.⁵ A search of a typical tattoo image database currently involves matching a query tattoo's class label with the labels for the tattoos in the database. The current practice of matching tattoos according to the manually assigned ANSI/NIST class labels has the following limitations:

- Class label does not capture the semantic information in tattoo images.
- Law enforcement agencies maintain millions of tattoo images.
- Tattoos often contain multiple objects and cannot be classified appropriately into the ANSI/NIST classes.
- Tattoo images have large intraclass variability.
- The ANSI/NIST classes are incomplete for describing new tattoo designs.

To overcome the limitations of the current practice of keyword-based tattoo matching, we have developed the Tattoo-ID automatic tattoo matching and retrieval system.^{6–8} This system has been licensed to MorphoTrak, which plans to release a commercial version.⁹ To the best of our knowledge, Tattoo-ID is the first prototype of an operational system for tattoo-image matching and retrieval. Although Scott Acton and Adam Rossi also proposed a tattoo matching and retrieval system based on global features (such as color and shape),¹⁰ their system was evaluated on high-quality Web-downloaded images, where query images were synthetically generated from the gallery images. In a previous work,⁶ we showed that such global features are inadequate for matching tattoo images in operational databases.

The Tattoo-ID System

Tattoo-ID is based CBIR,¹¹ in which the goal is to find the images from a database that are nearly duplicates of the query image.

Although general-purpose CBIR systems have only limited retrieval performance because of the well-known semantic-gap problem,¹¹ CBIR systems have been shown to be effective for near-duplicate image retrieval,¹¹ which fits well with tattoo-image retrieval. Tattoo-ID extracts keypoints from images using scale-invariant feature transform (SIFT)¹² and uses a matching algorithm^{7,8} to measure the visual similarity between two images. It then retrieves the database images with the largest similarities to the query. We choose SIFT because it yields the best performance for tattoo matching and retrieval compared to both the global image features (color, shape, and texture) and the other local descriptors (such as speeded up robust features [SURF], gradient location and orientation histogram [GLOH], and Harris Laplace¹³).

To objectively evaluate Tattoo-ID's performance, we constructed a database of 64,000 tattoo images provided by the Michigan State Police (see Figure 2). We cropped the tattoo images to extract the foreground and suppress the background. To construct the query set, we manually identified 1,000 images in the database that had near duplicates. These duplicates are introduced in the database as a result of multiple arrests of the same person at different times or multiple photographs of the same tattoo taken during a booking (see Figure 3). We used one of the duplicates as a query to retrieve other duplicates in the database.

To examine our system's robustness, we further augmented the 64,000 tattoo images with 36,000 randomly selected images from the ESP game database (www.gwap.com/gwap/gamesPreview/espgame). We evaluated Tattoo-ID's retrieval performance using cumulative matching characteristics (CMC). That is, for a given rank position N , its CMC score is computed as the percentage of queries with matched images that are found in the top- N retrieved images. Our previous work⁸ showed that Tattoo-ID can correctly retrieve the duplicate tattoos in the top 20 images ($N = 20$) for 85.6 percent of queries, and the average retrieval time per query is approximately 191 seconds on an Intel Core 2, 2.66-GHz, 3-Gbyte RAM processor (see Figure 3). In addition, we proposed an unsupervised ensemble ranking approach to manage the scalability problem.⁸ The approach achieves similar retrieval accuracy (85.9 percent for rank-20 accuracy) at a



Figure 3. Tattoo-ID retrieval examples. Each row shows a query tattoo (with the number of keypoints), the top-7 retrieved images, and the associated matching score (number of matching keypoints). Three duplicates were retrieved from the database for the first query, and two duplicates were retrieved for the second query.

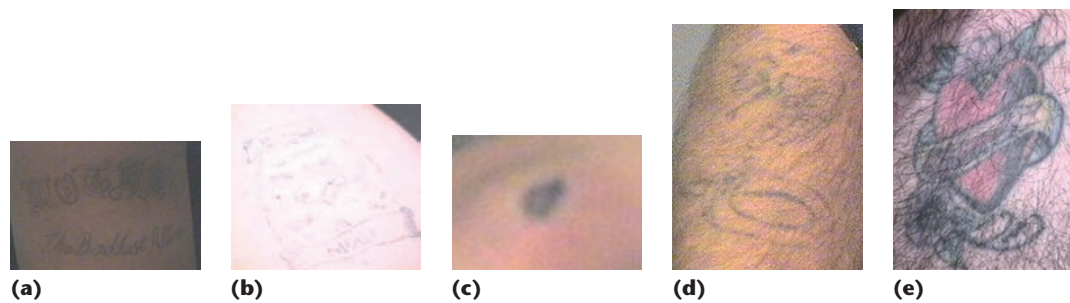


Figure 4. Examples of ugly tattoo images: (a) tattoo with low contrast (0), (b) tattoo with uneven illumination (11), (c) small tattoo size (2), (d) tattoos faded and covered with hair (15), and (e) tattoo covered by substantial body hair (381). The numbers in parentheses indicate the number of extracted keypoints.

significantly reduced retrieval time (14.7 seconds per query).

Although Tattoo-ID's overall retrieval accuracy is good, the performance drops off significantly if query images are of low quality (see Figure 4). For example, when images have low contrast, uneven illumination, or small tattoo size, only a few keypoints are extracted from the images, making it difficult to perform the matching. If tattoo images are covered by heavy body hair, the majority of keypoints are extracted from the body hair rather than the tattoos. These noisy keypoints lead to several false matches and, consequentially, low retrieval accuracy. We refer to the images with limited retrieval performance as *ugly tattoo images*, following the nomenclature introduced for poor-quality latent fingerprint images in the NIST-SD27 database.

To systematically evaluate the performance of Tattoo-ID for ugly tattoos, we extracted a subset of 252 ugly tattoo images from the 1,000 query images. The extracted images were query tattoos for which either

- the correct duplicate could not be retrieved in the top-20 ranks or
- the matching score of the first retrieved image was small (less than 10) and the top-10 retrieved images had similar matching scores (the standard deviation of the top-10 matching scores was less than 0.1).

Figure 5 compares Tattoo-ID's retrieval performances against 748 typical quality and 252 ugly-quality queries. Compared to the typical

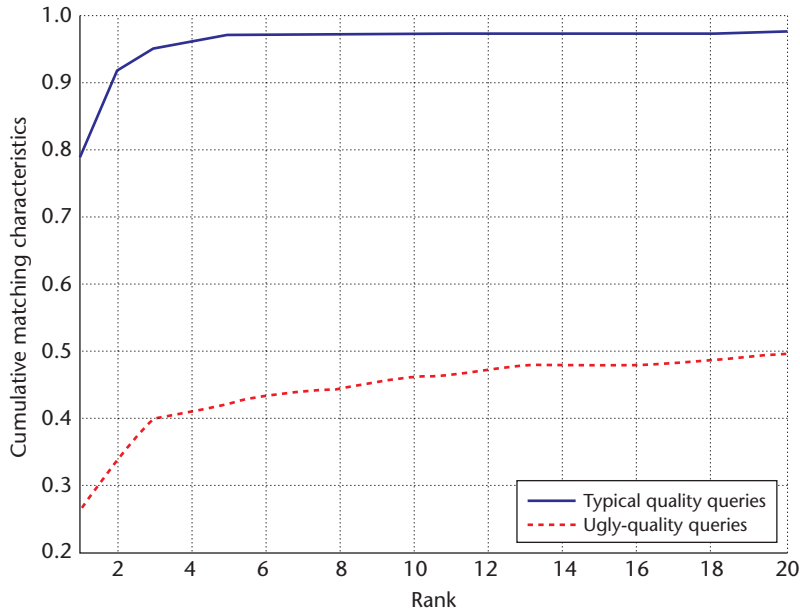


Figure 5. Retrieval performance for typical and ugly-quality queries. The 252 ugly-quality queries show significantly lower retrieval performance.

quality queries (97.7 percent for rank-20 accuracy), the 252 ugly-quality queries show significantly lower retrieval performance (49.6 percent for rank-20 accuracy). In this article, we aim to improve the system's robustness, especially for low-quality images, and consequently, overall retrieval performance.

Tattoo-ID Enhancements

We have improved the system performance by developing more robust similarity measures and using the metadata associated with tattoo images.

Robust Similarity Measures

Because of the low image contrast and vagueness of faded tattoos, the numerous spurious keypoints extracted often lead to many false matches. To address this challenge, we developed two strategies to improve the robustness of the similarity measure: symmetric matching and weighted keypoint matching.

To measure the similarity between a query image I_q and a database image I , denoted by $S(I_q, I)$, we compute the number of keypoints from I_q that match the keypoints from I .¹² A keypoint K_q^i from I_q is considered to match a keypoint from I if the ratio of the shortest and the second-shortest distance from K_q^i to the keypoints from I is smaller than a predefined threshold γ ($\gamma = 0.49$). This similarity measure is asymmetric; that is, $S(I_q, I) \neq S(I, I_q)$.

One shortcoming of the asymmetric similarity measure is that it might produce many false

matches, particularly if there is a keypoint in the database image I with a descriptor that is very similar to that of several keypoints in I_q . We address this limitation by developing a symmetric similarity measure for a pair of images I_q and I as follows:

1. Compute the asymmetric match scores between I_q and I and between I and I_q , resulting in two sets of matched keypoint pairs, denoted by $M(I_q | I)$ and $M(I | I_q)$.
2. Compute the symmetric similarity measure, denoted by $S_S(I_q, I)$, as the number of matched keypoint pairs that appear in both sets; that is, $S_S(I_q, I) = |M(I_q | I) \cap M(I | I_q)|$.

Note that $S_S(I_q, I) = S_S(I, I_q)$. This symmetric similarity measure lets us remove some of the false matches.

The weighted keypoint matching approach tries to reduce the effect of false matches by introducing two sets of weights to the keypoints in a query image. This approach is based on two intuitions. First, if a keypoint K_i in a gallery image I is matched to multiple keypoints from a query image, we consider these multiple keypoints in the query image to be indistinctive and assign them low weights in the similarity measure. We refer to this weight as *local distinctiveness*. Second, if a keypoint K_q^i finds its matches from many different gallery images, we consider it to be indistinctive and assign it a low weight. We refer to this weight as *global distinctiveness*. More specifically, suppose a query image I_q has l keypoints,

$$K_q = \{K_q^1, K_q^2, \dots, K_q^l\}$$

and there are N_G images in the gallery G . Let $m^i(I)$ be the number of keypoints in K_q that are mapped to the same keypoint in a gallery image I as K_q^i , and let n^i be the number of images in the gallery G where K_q^i finds its matched keypoints. Given $m^i(I)$ and n^i , the similarity between a query image I_q and a database image I , denoted by $S_W(I_q, I)$, is computed as follows:

$$S_W(I_q, I) = \sum_i^l x_i \left(\frac{1}{m^i(I)} \log \frac{N_G}{n^i} \right) \quad \text{where}$$

$$x_i = \begin{cases} 1, & \text{if } K_q^i \text{ is matched} \\ 0, & \text{otherwise} \end{cases}$$

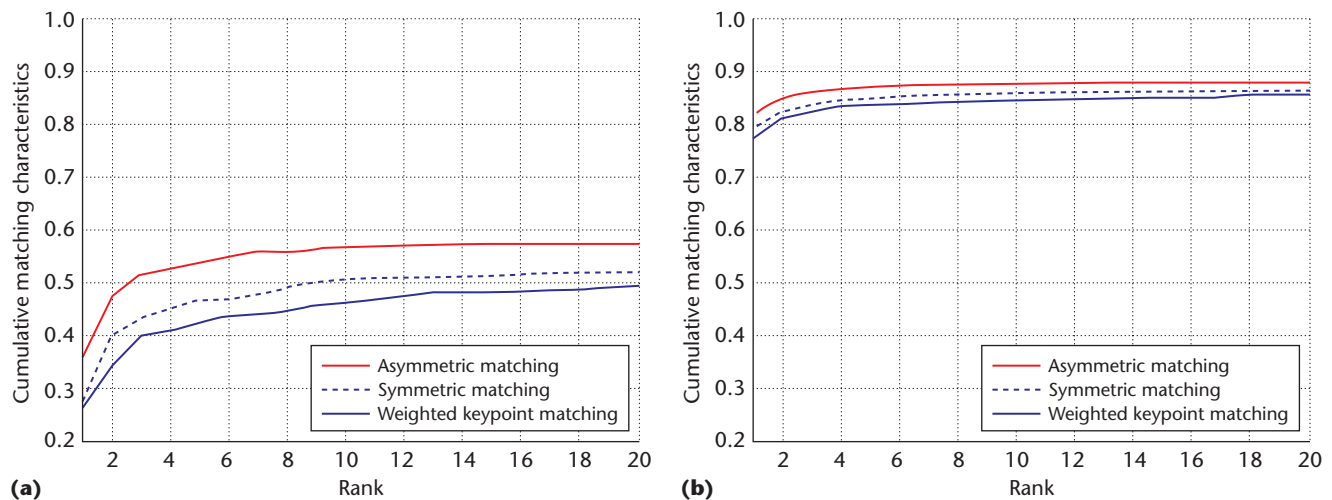


Figure 6 compares the retrieval performance of the asymmetric similarity, the symmetric similarity, and the weighted keypoint matching on the database of 100,000 images with the 1,000 query images that we described earlier. Both the symmetric matching and weighted keypoint matching improve the retrieval performance. The average rank-20 accuracy improved from 85.6 to 86.3 percent by the symmetric matching and to 88 percent by the weighted keypoint matching (Figure 6b). More noticeable improvements are clear for the ugly query images (Figure 6a), where the average rank-20 accuracy improved from 49.6 to 51.8 percent by the symmetric matching and to 57 percent by the weighted keypoint matching. Finally, compared to the symmetric matching, the weighted keypoint matching is more effective. According to the student t test (at the level of 5 percent), all the improvements are statistically significant. Overall, our results indicate that a soft-weighting approach is more robust to false matches than a hard-threshold approach such as in symmetric matching.

Metadata Utilization

To further improve retrieval performance, we evaluated the utility of metadata for tattoo-image retrieval. We created a collection of tattoo images with manually assigned metadata. Because of the substantial manual labor needed to label the images, we randomly selected 21,000 tattoo images from the 64,000 tattoo images in our database, including the 1,000 queries and their near-duplicate images, for manual annotation. The labeling

was done by 12 subjects who were Michigan State University students. On average, each subject was asked to annotate approximately 3,500 images in two ways: using up to four ANSI/NIST major classes and using his or her own keywords. The average number of classes assigned per tattoo image was two, and the average number of free keywords was approximately 3.5. Each image was annotated by two subjects, and we formed the final annotation by merging the annotations from the two subjects. After performing spell check and word stemming, the final number of unique free keywords for this database was 2,019. (Recall that the number of ANSI/NIST major classes is only eight.) We used this collection of manually annotated tattoo images to examine the effect of metadata on the retrieval performance.

To utilize the ANSI/NIST-based metadata (eight major classes), we implemented a two-stage matching scheme:

1. Select a subset of database tattoos that shared at least one class label with the query tattoo.
2. Perform keypoint-based image matching only for the selected subset.

Figure 7 shows the retrieval results for 252 ugly-quality tattoo queries and all 1,000 tattoo queries. In both cases, the introduction of ANSI/NIST class labels leads to a significant drop in retrieval performance because each ANSI/NIST class covers a broad range of tattoo types. Consequently, “similar” tattoo images

Figure 6. Retrieval performances for (a) 252 ugly-quality queries and (b) all 1,000 tattoo queries with the robust similarity measures. Both symmetric matching and weighted keypoint matching improve retrieval performance.

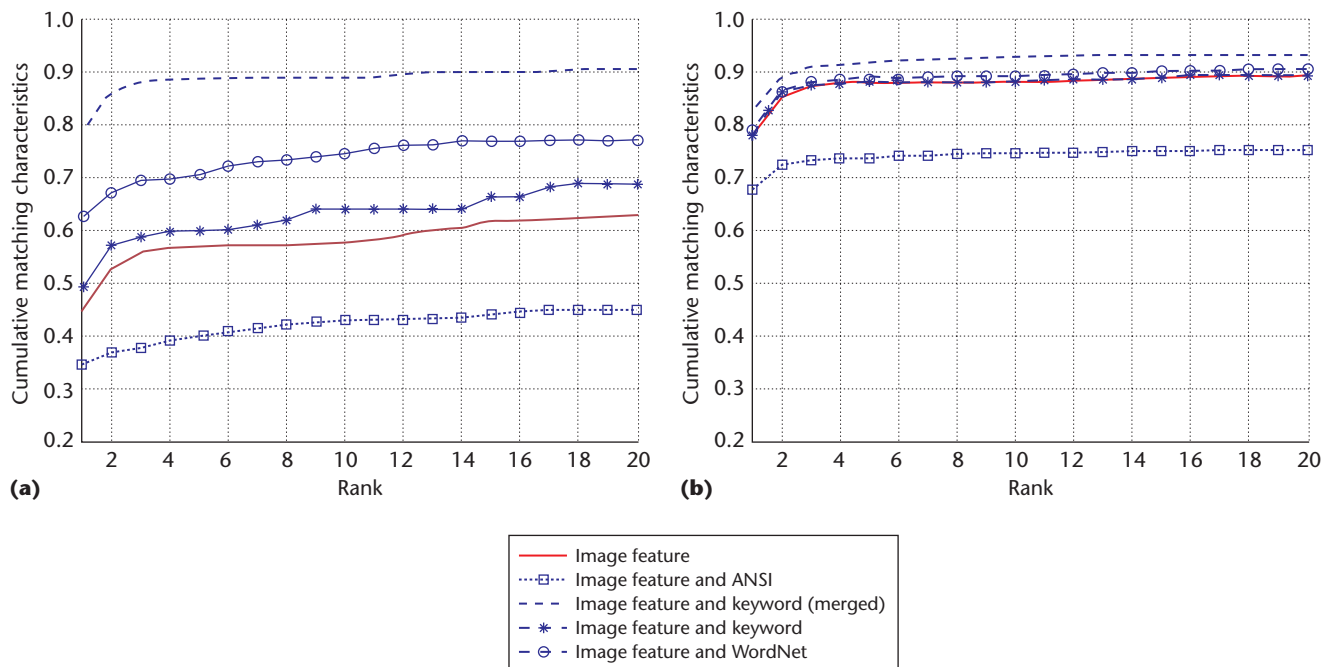


Figure 7. Retrieval performance for (a) 252 ugly quality queries and (b) all 1,000 tattoo queries with and without metadata information against the database of 21,000 images. The use of ANSI/NIST class labels leads to a significant drop in retrieval performance.

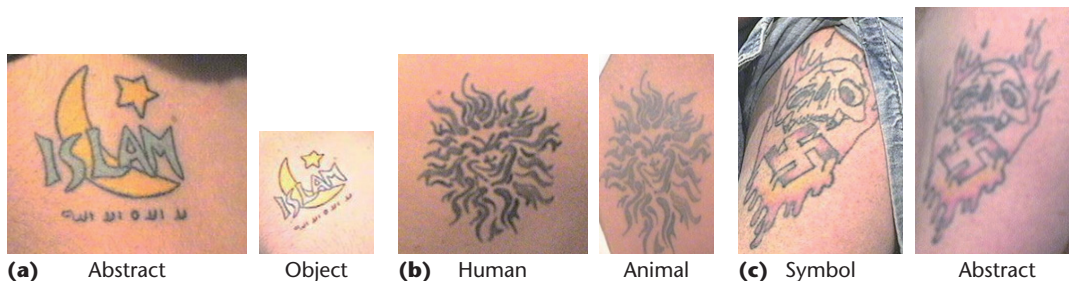


Figure 8. Examples of inconsistent assignment of ANSI/NIST classes to near-duplicate tattoo pairs. Although each set shows near-duplicate images of the same tattoo, the subjects in our experiment annotated them differently on the basis of the ANSI/NIST classes (shown under each image).

could be assigned to different classes, making it difficult to match tattoo images according to their ANSI/NIST class assignments (see Figure 8). This limitation of the ANSI/NIST major classes led us to explore the free-keyword annotation for improving tattoo-image retrieval performance.

Metadata Generated by Free-Keyword Annotations

We treat the keyword annotations as free text and apply the standard text-retrieval methods to compute the similarity score for metadata. More specifically, we use the *tf-idf* weighting

scheme for text retrieval and the Lemur text search engine (www.lemurproject.org) to efficiently compute the matching scores between free-keyword annotations. Given the similarity $S_W(I_q, I)$ based on the weighted keypoint matching, and the similarity $S_T(I_q, I)$ based on keyword matching, we compute the combined similarity score as $S(I_q, I) = S_W(I_q, I) + w \times S_T(I_q, I)$, where the weight parameter w is empirically tuned to optimize the retrieval performance.

In Figure 7, the line plot labeled “Image feature and keyword (merged)” shows the retrieval results of combining the free-keyword-based

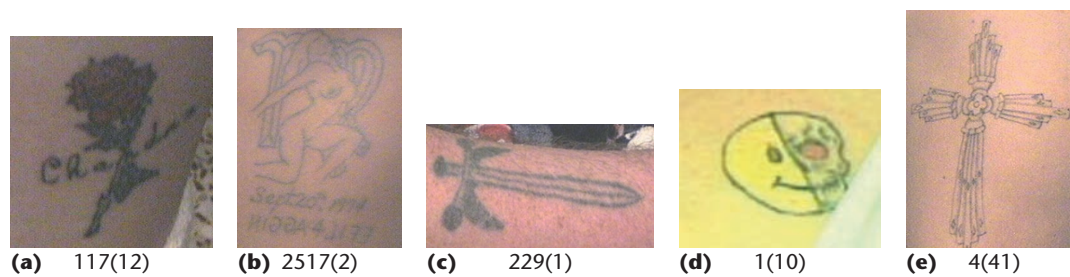


Figure 9. Comparison of retrieval results with and without free-keyword annotation. For both the successful cases, (a) through (c), and the failure cases, (d) and (e), the first number under each image is the ranking position for the correct retrieval based on the image feature alone, and the second number (in parentheses) is the ranking position for the correct retrieval based on both the image features and the merged free keywords.

matching with image matching. There is a significant improvement in retrieval performance for both ugly-quality queries (approximately 27 percent) and all the tattoo queries (approximately 10 percent). This indicates that the free-keyword annotation is more effective than the ANSI/NIST classes for retrieving near-duplicate tattoo images. This is because, unlike the small number of major classes in ANSI/NIST standard, which are often ambiguous in terms of labeling tattoos, most human subjects appear to be consistent in choosing free keywords for describing similar visual content.

One potential problem with this experiment is that the free-keyword annotations for query images were created by the same subjects who created the annotations for the gallery images. In an operational system, we would expect different subjects to perform keyword annotation for query images than for gallery images, which could degrade the retrieval performance. In fact, for the 21,000 annotated tattoo images, on average, less than 50 percent of the keywords are shared by two different subjects.

To accommodate this scenario, we changed the design of the metadata experiment by using the free-keyword annotations for query images by one subject, and the annotations for gallery images by a different subject. The line plot labeled “Image feature and keyword” in Figure 7 shows retrieval results for ugly-quality queries and all 1,000 queries. As expected, there is a significant drop in retrieval accuracy compared to the case when both query images and gallery images were annotated by the same subjects. On the other hand, compared to using image features alone, we still observe a significant improvement (approximately 7 percent) for ugly-quality queries and a marginal, but consistent improvement (approximately 1 percent) for all 1,000 tattoo queries.

Figure 9 shows examples of retrieval results based on a combination of free-keyword annotations and image features, where the images in

Figures 9a through 9c are successful retrievals and the images in Figures 9d and 9e are failure cases.

An analysis of the failure cases shows that subjects in our experiments assigned different free keywords to describe similar tattoos. For example, different subjects annotated the image in Figure 9d as “face” and “skull.” To address this problem, we expanded the annotation keywords using WordNet.¹⁴ WordNet is a large lexical database where nouns, verbs, adjectives, and adverbs are grouped into sets of cognitive synonyms, called *synsets*. Synsets interlink different conceptual-semantic and lexical relations. The underlying assumption is that different keywords used to describe similar tattoo images are likely to share the same semantic concept, and as a result, the concept expansion from WordNet might be able to bridge this gap.

In our study, we use the hypernym hierarchy in WordNet for keyword expansion. A hyponym shares a type-of relationship with its hypernym. For example, the hypernym of “dog” is “canine.” We chose the hypernym relation because two words sharing the same concept are likely to share common hypernyms in WordNet. Among the 2,019 different free keywords our subjects used in annotating the 21,000 tattoos, 1,737 keywords were found in WordNet and expanded with the corresponding hypernym hierarchy.

The line plot labeled “Image feature and WordNet” in Figure 7 shows the retrieval results using WordNet expansion for both ugly-quality queries and all 1,000 queries. For both cases, we observe up to an 8 percent improvement by using the WordNet expansion. The WordNet expansion clearly helps bridge the gap due to differences in free-keyword annotations. For example, for the query tattoo in Figure 9a, the correct retrieved image is found at rank-12 by fusion of the weighted keypoint matching and free-keyword matching scores. By expanding

the free keywords with WordNet, the correct retrieved image is found at rank-8 and the matching score is improved from 5 to 8.6. The WordNet expansion fails (see Figures 9d and 9e) when the gap between free keyword annotations by different subjects is too large. For example, the keyword annotation for the tattoo in Figure 9e is "symbol," whereas the keyword annotation for its true mate image in the database is "cross."

Conclusion

In this article, we took an unsupervised approach in designing appropriate similarity measures to explicitly address the challenge arising from low-quality tattoo image matching. In the future, we plan to improve the matching algorithm by exploring both supervised and semisupervised learning algorithms. Besides tattoos, Although Tattoo-ID focuses on tattoo image matching and retrieval, the underlying techniques developed in Tattoo-ID system can be adopted to other forensic image databases.¹⁵ Other types of soft forensic image evidence might include shoeprints and gang graffiti images. In the future, we plan to extend the Tattoo-ID system to different application domains.

MM

Acknowledgments

This research was supported by the FBI Biometric Center of Excellence. Anil Jain's research was also supported in part by the World Class University (WCU) program funded by the Ministry of Education, Science, and Technology through the National Research Foundation of Korea (R31-10008).

References

1. A.K. Jain, S.C. Dass, and K. Nandakumar, "Can Soft Biometric Traits Assist User Recognition?" *Proc. SPIE Conf. Biometric Technology for Human Identification*, vol. 5404, 2004, pp. 561-572.
2. "Next Generation Identification," US Fed. Bureau of Investigation, www.fbi.gov/about-us/cjis/fingerprints_biometrics/ngi.
3. J.-P. Beauthier, P. Lefevre, and E.D. Valck, "Autopsy and Identification Techniques," *The Tsunami Threat-Research and Technology*, N.-A. Mörner, ed., InTech, 2011.
4. A. Laumann and A. Derick, "Tattoos and Body Piercings in the United States: A National Data Set," *J. Am. Academy of Dermatology*, vol. 55, no. 3, 2006, pp. 413-421.
5. ANSI/NIST-ITL 1-2011, *Data Format for the Interchange of Fingerprint, Facial, & Scar Mark & Tattoo (SMT)*, Nat'l Inst. Standards and Technology, 2011.
6. J.-E. Lee, A.K. Jain, and R. Jin, "Scars, Marks and Tattoos (SMT): Soft Biometric for Suspect and Victim Identification," *Proc. Biometric Symp. (BSYM 08)*, IEEE Press, 2008, pp. 1-8.
7. A.K. Jain et al., "Content-Based Image Retrieval: An Application to Tattoo Images," *Proc. 16th IEEE Int'l Conf. Image Processing (ICIP 09)*, IEEE Press, 2009, pp. 2745-2748.
8. J.-E. Lee, R. Jin, and A.K. Jain, "Unsupervised Ensemble Ranking: Application to Large-Scale Image Retrieval," *Proc. 20th Int'l Conf. Pattern Recognition (ICPR 10)*, IEEE CS Press, 2010, pp. 3902-3096.
9. "MorphoTrak Acquires Innovative Tattoo Matching Technology from Michigan State University," *Business Wire*, 15 Nov. 2011; http://findarticles.com/p/articles/mi_m0EIN/is_20100119/ai_n48674730.
10. S.T. Acton and A. Rossi, "Matching and Retrieval of Tattoo Images: Active Contour CBIR and Global Image Features," *Proc. IEEE Southwest Symp. Image Analysis and Interpretation*, IEEE CS Press, 2008, pp. 21-24.
11. R. Datta et al., "Image Retrieval: Ideas, Influences, and Trends of the New Age," *ACM Computing Surveys*, vol. 40, no. 2, 2008, pp. 1-60.
12. D. Lowe, "Distinctive Image Features from Scale Invariant Keypoints," *Int'l J. Computer Vision*, vol. 60, no. 2, 1999, pp. 91-110.
13. K. Mikolajczyk and C. Schmid, "A Performance Evaluation of Local Descriptors," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, 2005, pp. 1615-1630.
14. G.A. Miller, "WordNet: A Lexical Database for English," *Comm. ACM*, vol. 38, no. 11, 1995, pp. 39-41.
15. W. Tong et al., "Gang and Moniker Identification by Graffiti Matching," *Proc. Int'l ACM Workshop Multimedia in Forensics and Intelligence (MiFor)*, ACM Press, 2011.

Jung-Eun Lee is a PhD candidate in the Department of Computer Science and Engineering at Michigan State University. Her research interests include image retrieval and biometrics. Lee has a MS in computer science and engineering from Michigan State University. Contact her at leejun11@cse.msu.edu.

Wei Tong is a postdoctoral researcher in the Language Technologies Institute at Carnegie Mellon University. His research interests include image retrieval, image annotation, machine learning, and computer vision.

Tong has a PhD in computer science from Michigan State University. Contact him at tongwei@cs.cmu.edu.

Rong Jin is an associate professor in the Department of Computer Science and Engineering at Michigan State University. His research interests include statistical machine learning and its application to large-scale information management. Jin has a PhD in computer science from Carnegie Mellon University. Contact him at rongjin@cse.msu.edu.

Anil K. Jain is a university distinguished professor in the Department of Computer Science and

Engineering at Michigan State University. He is also a distinguished professor in the Department of Brain and Cognitive Engineering at Korea University. His research interests include pattern recognition, computer vision, and biometric authentication. Jain has a PhD in electrical engineering from Ohio State University. He is a fellow of IEEE. Contact him at jain@cse.msu.edu (corresponding author).



Selected CS articles and columns are also available for free at <http://ComputingNow.computer.org>.