To appear in PCM 2007, Dec 2007

Tattoo-ID: Automatic Tattoo Image Retrieval for Suspect & Victim Identification

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

Computer Science and Engineering, Michigan State University,
East Lansing, Michigan 48824, USA
{jain, leejun11, rongjin}@cse.msu.edu

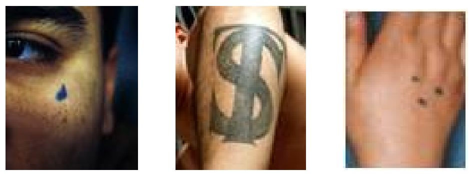
Abstract. Tattoos are used by law enforcement agencies for identification of a victim or a suspect using a false identity. Current method for matching tattoos is based on human-assigned class labels that is time consuming, subjective and has limited performance. It is desirable to build a content-based image retrieval (CBIR) system for automatic matching and retrieval of tattoos. We examine several key design issues related to building a prototype CBIR system for tattoo image database. Our system computes the similarity between the query and stored tattoos based on image content to retrieve the most similar tattoos. The performance of the system is evaluated on a database of 2,157 tattoos representing 20 different classes. Effects of segmentation errors, image transformations (e.g., blurring, illumination), influence of semantic labels and relevance feedback are also studied.

Keywords: Human Identification, Content-based Image Retrieval, Tattoos, Forensics

1 Introduction

People have used tattoos in order to represent themselves and to be identified as distinct from others for over 5,000 years [1]. Until recently, practice of tattooing was limited to particular groups, such as motor bikers, sailors and members of criminal gangs. But, now, tattoos are no longer associated with such unsavory reputations, and as a result, the size of the tattooed population is rising rapidly. The rising popularity of tattoos among the younger section of the population is even more surprising. A study published in the Journal of the American Academy of Dermatology in 2006 reported that about 36% of Americans in the age group 18 to 29 have at least one tattoo [2].

Tattoos are a useful tool for person identification in forensic applications. There has been an increased emphasis on the use of “soft biometric” traits [3] (e.g., tattoos) in identification tasks when primary biometric traits (e.g., fingerprints) are either no longer available, or corrupted. Tattoo pigments are embedded in the skin to such a depth that even severe skin burns often do not destroy a tattoo; tattoos were used to identify victims of 9/11 attacks [4] and Asian tsunami in 2004 [5]. Criminal identification using tattoos is another important application, because tattoos often contain useful information, such as gang membership, religious beliefs, previous convictions, years spent in jail, etc. [1] (Fig. 1). A study by Burma [6] suggested that delinquents are significantly more likely to have tattoos than non-delinquents.



(a) (b) (c)

Fig. 1. Examples of Criminal Tattoos: (a) Teardrop tattoo (person has killed someone or had a friend killed in prison), (b) Texas Syndicate (TS) gang member tattoo, (c) Three dots tattoo found on the prisoner's back of the hand [7-8].

(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 2. Sample tattoos from the eight major classes in ANSI/NIST-standard: (a) Human, (b) Animal, (c) Plant, (d) Flag, (e) Object, (f) Abstract, (g) Symbol, and (h) Other.

Law enforcement agencies routinely photograph and catalog tattoo patterns for the purpose of identifying victims and convicts (who often use aliases). The ANSI/NIST-ITL 1-2007 standard [9] defines 8 major class labels, e.g., human face, animal, and symbols for tattoos (Fig. 2). Manual searches are performed by matching the class labels of query and database tattoos. The tattoo matching process based on human-assigned class labels is subjective, has limited performance and time-consuming. Further, simple class description in textual query ("find a dragon tattoo") does not include all the semantic information in tattoos [10] as evident by the large intra-class variability (Fig. 3). Finally, the classes in ANSI/NIST standard are not adequate to describe increasing variety of new tattoo designs.

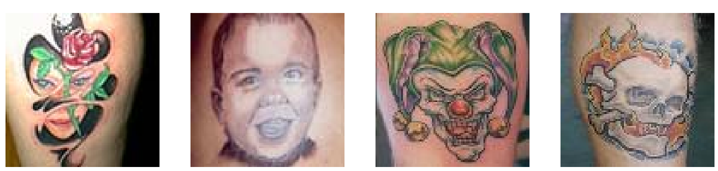


Fig. 3. Different images belonging to HUMAN category.

We describe the preliminary design and implementation of a CBIR system for tattoo images, called Tattoo-ID. While, successful CBIR systems for a variety of application domains such as satellite, trademark, vacation, and medical images have been reported [12], [13], to our knowledge, this is the first attempt to build a system for tattoo images. Compared to the general CBIR systems, which retrieve images similar to the given query, the goal of our Tattoo-ID system is driven by the specific application requirements. Since tattoo images are used for human identification, the system is expected to retrieve all the images in the database that are some minor variations of the specific tattoo in the given query image. In this sense, each tattoo is unique. Any variation in its multiple images is caused by the imaging environment and the condition of the tattoo on the skin (e.g., fading over time). Therefore, the challenges of this application domain are not only the large intra-class variability, but also a large number of potential image variations. We currently limit ourselves to variations introduced by image transformation, such as blurring and illumination (Fig. 4).

(a) (b) (c) (d)

Fig. 4. Examples of difficult tattoos: (a) and (b) same tattoo from two different viewpoints, (c) and (d) multi-object tattoos.

2 Tattoo Image Database

We have downloaded 2,157 tattoo images from the Web [14] belonging to eight main classes and 20 subclasses based on ANSI/NIST standard [9]. Tattoo images are often captured under non-ideal conditions (e.g., by a surveillance video camera). We classify the resulting image transformations as: blurring, additive noise, changes in illumination, color, aspect ratio and rotation (Fig. 5). For each image transformation, we generate 20 different variations of a tattoo as follows: two different intensities of blurring, additive noise and illumination; six different aspect ratios, four different rotations, and four different color changes. This results in a total of 43,140 transformed images of 2,157 original tattoo images.

(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 5. Examples of tattoo image transformations: (a) original, variations due to (b) blurring, (c) illumination, (d) aspect ratio, (e) and (f) color, (g) additive noise, and (h) rotation.

3 Tattoo-ID System

In our design of a tattoo CBIR system (Fig. 6), we examine and evaluate the following issues: (i) types of transformations that mimic the reality in capturing tattoo images and their effects on image retrieval, (ii) choice of image features and similarity measure, and (iii) effect of user relevance feedback.

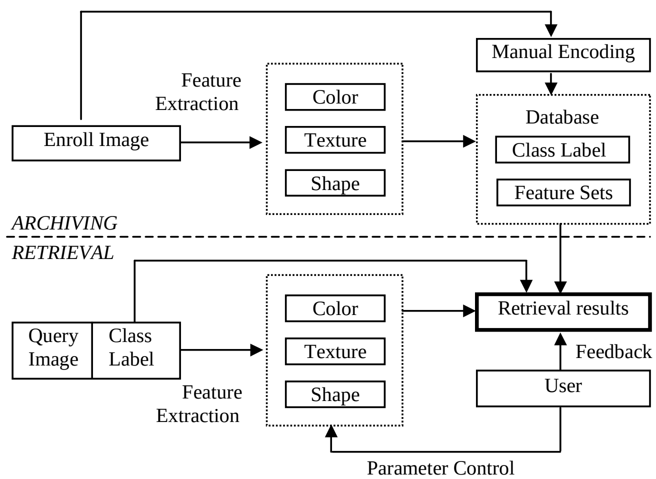


Fig. 6. Tattoo-ID system.

3.1 Preprocessing

Since tattoos appear on uniform regions of skin, edge based operations generally perform well for foreground segmentation (region of interest). We used 3×3 Sobel operator to obtain the magnitude and direction of gradient at each pixel. By thresholding the gradient, followed by a morphological closing and opening operations, foreground is obtained (Fig. 7).

(a) (b) (c) (d)

Fig. 7. Examples of tattoo segmentation: (a) and (c) are input images, (b) and (d) are corresponding segmented images.

3.2 Image Features

Our choice of features for capturing low-level image attributes (color, shape and texture) is based on the extensive literature on content-based image retrieval [11], [12], [13], [16], [19]. Specific features chosen depend on our application domain of tattoo images.

Color. We used the RGB space to extract two color descriptors, namely color histogram and color correlogram [11], [15]. A color correlogram is a table indexed by a pair of colors, where the k-th entry for (i, j) specifies the probability of finding a pixel of color j at a distance k from a pixel of color i in the image. In our experiments, the color histogram and correlogram are calculated by dividing each color component into 20 and 63 bins, respectively, resulting in total of 60 and 189 total bins for the color histogram and correlogram, respectively. For computational efficiency, we compute color autocorrelogram between identical colors in a local neighborhood, i.e., $i = j$ and $k = 1, 3, 5$.

Shape. Classical 2D shape representation uses a set of moment invariants. Based on 2nd and 3rd order moments, a set of seven features that are invariant to translation, rotation, and scale are obtained [16]. Two feature sets are extracted from the segmented grayscale and the gradient tattoo images, respectively [11].

Texture. Edge Direction Coherence Vector that stores the ratio of coherent to non-coherent edge pixels with the same quantized direction (within an interval of 10°) is used [12]. A threshold (0.1% of image size) on the edge-connected components in a given direction is used to decide the region coherency. This feature discriminates structured edges from randomly distributed edges.

3.3 Matching

To decide a match between two tattoo images, we first compute a similarity score for each attribute (color, shape and texture) separately. Since each of the features is in the form of a vector, we regarded the vectors as histograms and apply histogram intersection method [17] to compute the similarity. Given two normalized histograms H^1 and H^2, the similarity is defined as:

$$S_{H^1,H^2} = \sum_{i}^{B} \min(H^1(i), H^2(i)) \qquad \textbf{(1)}$$

where B is the number of bins. When the two histograms are completely disjoint (overlapping), the similarity score equals 0 (1). For color histogram, the similarity scores are calculated by averaging the similarities in individual color components (R, G, B). We currently assign the same weight to all the features, so that the overall matching score between two images is calculated as the sum of similarity scores from individual attributes.

3.4 Relevance Feedback

A simple relevance feedback approach is implemented to improve the retrieval accuracy. A straightforward approach for relevance feedback is to modify the query by the centroid of the images that are marked as relevant [18]. However, the problem with this approach is that it assumes a single mode for the set of images that the user considers relevant to the query. Given the large image variation, it is more likely for the relevant images to exhibit multi-modal distribution than uni-modal distribution. To address this problem, we first compute the similarity of every image in the database to each of the relevant images and then retrieve images based on this new similarity. The user can continue refining the retrieved images by running the feedback procedure iteratively until all the matched tattoo images are found. A graphical user interface is designed to allow a user to provide relevance judgments for the retrieved images.

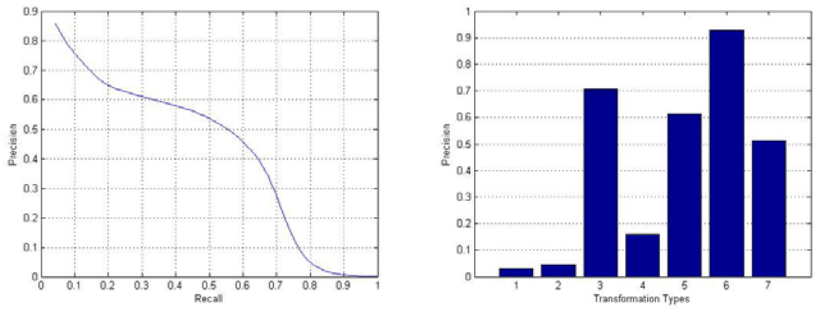
4 Experimental Results

To evaluate the retrieval performance of the Tattoo-ID system, three sets of experiments are conducted with following scenarios: (I) high quality query images that are taken by police officers when booking suspects in prison, and (II) query images of victims/suspects that are captured at crime scenes. We simulate the second category of query images by applying the image transformations (see Fig. 5) to the high quality query images. In the third experiment, (III) we repeated experiments I and II by including semantic (class) labels associated with query and database to determine if retrieval performance can be improved. Table 1 summarizes the database and the query size for the experiments I and II.

Table 1. Database and query size for Experiments I and II.

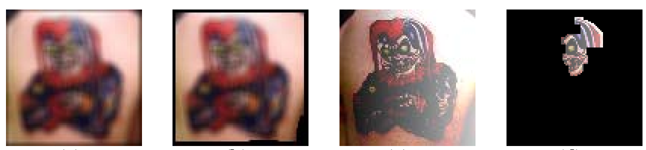
Experiment	Number of queries	Database size
I	2,157 high quality tattoos	43,140 transformed images
II	43,140 transformed images	2,157 high quality tattoos

Experiment I aims to evaluate the Tattoo-ID system with high quality query images. A retrieved image is deemed to be relevant when it is a transformed version of the query image. Both precision and recall are used as the evaluation metrics. Fig. 8(a) shows the average precision and recall curve of this experiment; the precision at ranks 1, 10 and 20 are 84.6%, 60.4%, and 51.2%, respectively. Among all the images in the database, the transformed images that are blurred and undergo changes in illumination and color component modification are the most difficult to retrieve (Fig. 8(b)). This is because both blurring and uneven illumination makes it difficult to correctly automatically segment a tattoo image from its background (Fig. 9). This issue can be partially addressed by the user relevance feedback. Overall, the average precision at rank 20 is improved from 51.2% to 64.7%, 68.7%, and 69.2% after the first three iterations of user relevance feedback.



(a) (b)

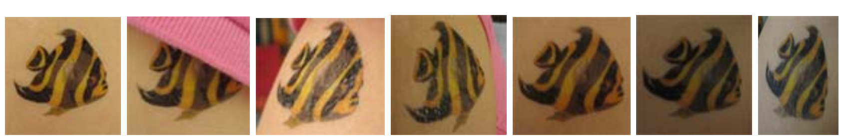
Fig. 8. Experiment I results: (a) Precision & recall curve, (b) precision of each transformation at rank 20 (transformations are labeled as follows. 1: blurring, 2: illumination, 3: additive noise, 4: color, 5: rotation, 6: aspect ratio, and 7: overall).



(a) (b) (c) (d)

Fig. 9. Examples of segmentation errors: (a) blurred image, (b) segmentation, (c) uneven illuminated image and (d) its segmentation. For the blurred image in (a), the entire image is found as foreground.

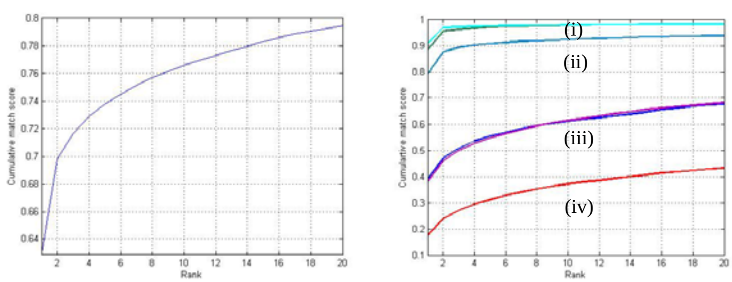
In addition to the synthetic transformation of the tattoo images, we also evaluated our approach on multiple images of a tattoo captured under different imaging environments. One such example is shown below (Fig. 10). Seven different images of the fish tattoo were captured and added to our tattoo database of 2,157 images; each image was then used as a query to find the other six images among the 2,163 images. Our system successfully found one similar image at rank 1 for five out of the seven queries; the average number of correctly retrieved images at rank 20 is three. Queries that posed problems are those that have severe distortions (Figs. 10(c) and (d)).



(a) (b) (c) (d) (e) (f) (g)

Fig. 10. Seven different images of the fish tattoo taken under different imaging conditions.

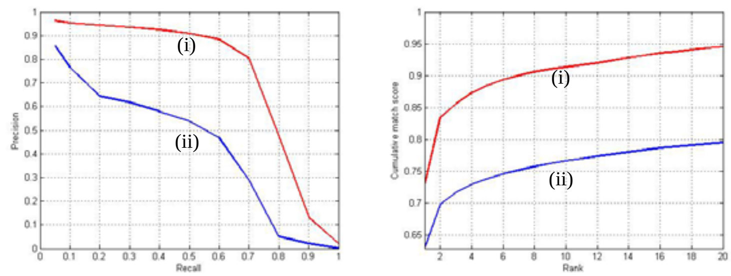
Experiment II aims to evaluate the retrieval accuracy of the Tattoo-ID system when the query images are noisy. Now a retrieved image is deemed to be relevant when the query image is generated from the retrieved one by one of the image transformations shown in Fig. 5. Since there is only one truly "similar" image in the database for every transformed query image, we adopted the cumulative matching curve (CMC) [20] as the evaluation metric for this experiment. Fig. 11(a) shows the CMC curve. Overall, the chance of retrieving the correct image in the database is ~60% when we only look at the first retrieved image, and is increased to ~80% when we examine the top 20 retrieved images. Similar to the Experiment I, queries that are blurred, have uneven illumination and whose color components have been changed are the most difficult (Fig. 11(b)) for the system and we attribute the problem to the errors in automatic image segmentation. We confirm this hypothesis by manually segmenting the two images in Fig. 9 (that could not be segmented correctly by our algorithm) and running the retrieval experiment using the manually segmented images as queries. The rank of the correct database image is now increased from 83 to 21 for 9(a) and from 1,157 to 32 for 9(c) when using the manually segmented images.



(a) (b)

Fig. 11. Experiment II results: (a) Cumulative Matching Curve, (b) CMC for each transformation types: (i) noise and aspect ratio, (ii) rotation, (iii) color and blurring, (iv) illumination.

Experiment III. Tattoo-ID system can optionally accept the semantic class label(s) of a query image from users. By using the class label(s), if available, as a part of the query, we are able to narrow down the candidate tattoo images in the database and therefore enhance the retrieval performance. Fig. 12 compares the retrieval performances of our CBIR system with and without using class labels as a part of the query: Figures 12 (a) and (b) show the precision and recall curves of experiment I and CMC of experiment II with and without class labels. By including class labels in the query, we are able to improve the retrieval performance significantly for both experiment I and II. In particular, for experiment I, the average precision at rank 20 is improved from 51.2% to 90.8% by using class labels, and the cumulative match score at rank 20 in experiment II is improved from 79.4% to 94.5%. We thus conclude that the incorporation of class labels in the query can significantly improve the retrieval performance. However, it is important to note that query class label alone (without image attributes) is not enough to retrieve the correct match in the database since there are ~100 images/class in our database. Hence, a combination of image based features and class label is needed for good retrieval performance.



(a) (b)

Fig. 12. Comparison between retrieval performance with and without class label: (a) Average Precision and Recall curve ((i) with class label and (ii) without class label), and (b) CMC ((i) with class label and (ii) without class label).

5 Conclusions and Future Work

We have presented the design and development of a prototype CBIR system for tattoo images. With the growing use of tattoos for person identification by forensics and law enforcement agencies, such a system will be of great societal value. While tattoo matching alone may not be sufficient to uniquely identify a person, it can provide useful information about the person's identity such as the group (religious, criminal, etc.) to which he belongs. Our preliminary experimental results based on a set of relatively simple image features are promising. However, the complex nature of the tattoo images requires that we develop robust image segmentation and feature extraction algorithms to further improve the retrieval performance. We have also shown that including the tattoo class information in the query improves the retrieval performance. Our ongoing work addresses (i) expanding the database (to 10,000 tattoos), (ii) capturing multiple images of a tattoo under different imaging conditions, (iii) finding additional salient features and designing a robust matcher, (iv) building a complete user interface, and (v) developing a more sophisticated user relevance feedback mechanism.

Acknowledgements

The authors would like to thank Yi Chen, Unsang Park and Chris Samiadji-Benthin for their assistance in preparing this manuscript. This research was supported by NSF IUC on Identification Technology Research (CITeR).

References

1. A Brief History of Tattoos, http://www.designboom.com /history/tattoo_history.html
2. Tattoo Facts and Statistics, Oct. 2006, http://www.vanishingtattoo.com/tattoo_facts.htm.
3. Jain A. K., and Dass, S. C. and Nandakumar, K.: Can Soft Biometric Traits Assist User Recognition? Proc. SPIE, vol. 5404, pp. 561-572, Orlando, Apr., 2004
4. Lipton, E. and Glanz, J.: Limits of DNA Research Pushed to Identify the Dead of Sept. 11, NY Times, Apr. 22, 2002
5. Decay Challenges Forensic Skills, The Standard-Times, Jan. 8 2005
6. Burma, J. H.: Self-Tattooing among Delinquents: A Research Note, Sociology and Social Research, vol. 43, 341-345, 1959
7. Texas Prison Tattoo, http://www.foto8.com/issue09/reportage/AndrewLichtenstein/prisontattoos01.html
8. Prison Tattoos and Their Meaning, http://www.tattoo-designs.dk/prison-tattoos.html
9. ANSI/NIST-ITL 1-2007 standard: Data Format for the Interchange of Fingerprint, Facial, & Other Biometric Information, 2007
10. GangNet: A 21st Century Solution to the Gang Problem, Dec. 2006 http://psd.orionsci.com/Products/Gangnet.asp
11. Long, F., Zhang, H. J., and Feng, D. D.: Fundamentals of Content-Based Image Retrieval, in Multimedia Information Retrieval and Management- Technological Fundamentals and Applications, Springer, 2003
12. Jain, A. K. and Vailaya A.: Shape-Based Retrieval: A Case Study with Trademark Image Databases, Pattern Recognition, vol. 31, No.9, pp. 1369-1390, 1998
13. Shih, P. and Liu, C.: Comparative Assessment of Content-Based Face Image Retrieval in Different Color Spaces, International Journal of Pattern Recognition and Artificial Intelligence, vol. 19, no. 7, pp. 873-893, 2005
14. Online Tattoo Designs, http://www.tattoodesing.com/gallery/
15. Huang, J., Kumar, S. R, Mitra, M, Zhu, W. and Zabih, R.: Image Indexing using Color Correlogram, Proc. IEEE Computer Society Conf. on CVPR, pp. 762-768, 1997
16. Hu, M. K.: Visual Pattern Recognition by Moment Invariants, IEEE Trans. Information Theory, Vol. 8, pp. 179-187, 1962
17. Swain, M. J. and Ballard, D. H.: Indexing via Color Histograms, Proc. 3rd International Conference on Computer Vision, pp. 309-393, 1990
18. Rui, Y., Huang, T. S., Ortega, M., Mehrotra, S.: Relevance Feedback: A Power Tool in Interactive Content-based Image Retrieval, IEEE Trans. on Circuits and Systems for Video Technology 8(5), pp. 644-655, Sep. 1998
19. Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovicand D., D. Steele, D., and Yanker. P.: Query by Image and Video Content: The QBIC system. IEEE Computer, volume 38, pp. 23-31, 1995
20. Moon. H, and Phillips. P. J., Computational and Performance Aspects of PCA-based Face Recognition Algorithms, Perception, Vol. 30, pp. 303-321, 2001