

CONTENT-BASED IMAGE RETRIEVAL: AN APPLICATION TO TATTOO IMAGES

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

Department of Computer Science & Engineering
Michigan Sate University, East Lansing, Michigan 48824

ABSTRACT

Tattoo images on human body have been routinely collected and used in law enforcement to assist in suspect and victim identification. However, the current practice of matching tattoos is based on keywords. Assigning keywords to individual tattoo images is both tedious and subjective. We have developed a content-based image retrieval system for a tattoo image database. The system automatically extracts image features based on the Scale Invariant Feature Transform (SIFT). Side information, i.e., body location of tattoos and tattoo classes, is utilized to improve the retrieval time and retrieval accuracy. Geometrical constraints are also introduced in SIFT keypoint matching to reduce false retrievals. Experimental results on 1,000 queries against an operational database of 63,593 tattoo images show a rank-20 accuracy of 94.2%; the average matching time per query is 2.9 sec. on Intel Core 2, 2.66 GHz, 3 GB RAM processor.

Index Terms — Image retrieval, tattoo images, SIFT, point matching, geometric constraints

1. INTRODUCTION

The popularity of digital cameras and advances in storage capabilities have led to tremendous growth in image and video databases. Flickr, a well-known photo sharing website hosts more than three billion images with over 2.5 million new images uploaded everyday [1]. These archives require tools for efficient image search and retrieval. Most of the available image search tools, such as Google Images and Yahoo! Image search, are based on textual annotation of images and their performance is not satisfactory. In these tools, images are manually annotated with keywords and then retrieved using text-based search methods. However, this manual image annotation is subjective, cumbersome, inconsistent and time consuming, especially when the image database is large and diverse. These shortcomings have led to the development of content-based image retrieval (CBIR) techniques [2] to simplify and improve the image retrieval performance.

The goal of CBIR is to extract visual content of an image in terms of simple low-level attributes (e.g., color, texture, and shape). These attributes or features are then used for representing and matching images without the use of any

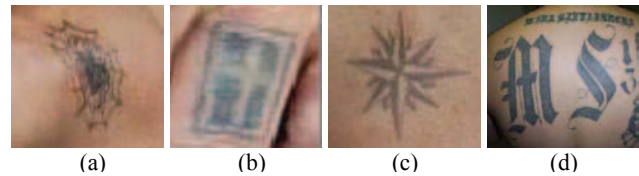


Fig. 1. Criminal tattoos of specific gangs or groups. (a) Spider Web: a drug addiction or a thief, (b) St. Petersburg Cross: the wearer had been in a prison, (c) Star: each point indicates a year served in jail, (d) tattoo of a gang, called MS-13 [3].



Fig. 2. Tattoos were found to be useful for (a) and (b) victim and (c) suspect identification [4, 5].

textual keywords. Despite the large amount of work on CBIR, the performance of current systems is not satisfactory, and needs substantial improvement [6]. The main limitation of the CBIR systems is the well-known *semantic gap* problem. Pavlidis [6] notes that “*Because there is little connection between pixel statistics and the human interpretation of an image (the “semantic gap”) the use of large number of generic features makes highly likely that results will not be scalable, i.e. they will not hold on collections of images other than the ones used during the development of the method*”. There have been many attempts to compensate for the semantic gap based on relevance feedback, learning, salient feature matching, composite querying and probabilistic approaches [2]. Another difficulty is caused by the size of the image database. Suppose matching two images takes only 0.001 sec, an optimistic figure. Then matching a query against a database of, say, one million images will take approximately 16 minutes. Therefore, an indexing method is necessary for an efficient search of large image databases.

This paper builds on our earlier work [7] on CBIR in the domain of tattoo image database. Tattoo image matching is receiving significant attention from law enforcement agencies due to their value in determining criminal history, such as gang membership, previous convictions, years spent in jail, etc. (Fig 1). Tattoo pigments are embedded in skin to such a depth that even severe skin burns or humidity



Fig. 3. Large intra-class variability in tattoo images. All the images shown here belong to the “Symbol” category.



(a) Wording, Insect (b) Plant, Rose, Red Flower

Fig. 4. Examples of images in the Michigan Police tattoo database and their corresponding cropped images. Tattoo classes assigned by us are also shown.

do not completely destroy a tattoo, making it possible to help identify a victim (Fig. 2). As a result, matching a tattoo present on a suspect or victim against a known database of tattoos associated with criminal gangs, or other organized groups could help identify the person¹. The current tattoo retrieval procedure in law enforcement agencies is based on keyword matching. A set of 70 keywords have been defined in the ANSI/NIST ITL 1-2000 standard [8] and each tattoo is assigned one of these labels. This approach is time-consuming and subjective with limited performance due to the complexity and large intra-class variability of tattoo images (Fig. 3).

In this paper, we present a CBIR system for a database of tattoo images that is based on the well-known local image descriptors (SIFT keypoint [9]) as opposed to global image features (i.e., shape, color and texture). The goal is to find tattoo images in the database that are *visually similar* to the query image. The requirement of *visually similar* (as opposed to *semantically similar*) tattoo retrieval narrows the semantic gap dramatically. Indeed, this is a well-known and active research topic in document and information retrieval community known as *known-item retrieval* [10]. To overcome the computational cost of retrieval, we utilize side information, if available in the form of (i) the location of tattoo on the body, and (ii) tattoo class labels. We also introduce geometric constraints between the SIFT keypoints in the two images to reduce the number of false matches. Experimental results based on over 1,000 image queries on a database of 63,592 images show very encouraging results.

2. TATTOO IMAGE RETRIEVAL

Given a query image, the goal of the tattoo retrieval system is to retrieve all visually similar tattoos in the database. Typically, the top-N (say, N=20) most similar tattoo images

are retrieved and presented to the user. If the user wishes to see additional tattoos, the system retrieves the next N-most similar images from the database, and so on.

2.1. Tattoo Image Database

We have access to 63,593 operational tattoo images (640×480 color images). The tattoo images were cropped to extract the foreground and suppress the background (see Fig. 4). About 3% of the images in the database are duplicates of the same tattoo. To evaluate the retrieval performance of our CBIR system, one of the duplicates is used as a query to retrieve the other duplicate(s) in the database (see Fig. 5). The retrieval experiments were done in a leave-one-out fashion in which 1,000 queries were searched against a gallery of 63,592 images.

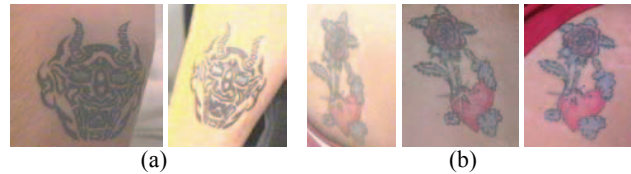


Fig. 5. Multiple copies of the same tattoo in the database. (a) and (b) show copies of the same tattoo. These tattoo copy images were captured at different times.

2.1.1 Tattoo Image Indexing

Our indexing approach first makes use of the location of the tattoo on body, if available. Tattoo location is a reliable piece of information because it can be tagged precisely and objectively. Hence, searching similar images at the same body location can significantly reduce the matching time without any loss of matching accuracy. The National Crime Information Center (NCIC) [11] has defined 31 major categories (i.e., arm, calf, and finger), and 71 sub categories (i.e., left upper arm, right calf, and left hand finger) for body location. The images in our database cover 42 of these categories in the NCIC standard; the most frequent body locations for tattoos in our database are upper right arm (19.6%) and upper left arm (15%).

In addition to tattoo location, tattoo class labels, if available, based on the ANSI/NIST standard are also utilized to index tattoo images. Each tattoo was manually annotated with keywords defined in the ANSI/NIST standard (eight major class labels and 70 subclass labels). In contrast to current practice in law enforcement, where only one class label is assigned per tattoo, we assigned each tattoo with as many applicable class labels as possible. The average number of class labels in our database is about two per image.

2.2. Image Features: SIFT

Scale Invariant Feature Transform (SIFT) [9] is a well-known and robust local feature based approach used for object recognition. In [7], we showed that SIFT features surpassed low-level attributes, (e.g., color, texture, and shape) in tattoo image matching and retrieval. SIFT extracts

¹ A tattoo does not uniquely identify a person, but it can indicate membership in a group which, in turn, facilitates establishing the person’s identity.

repeatable characteristic feature points from an image and generates descriptors representing the texture around the feature points. These feature points are invariant to image scale and rotation, and are shown to provide robust matching across a substantial range of affine distortion, change in 3D viewpoints, additive noise and change in illumination. A more detailed description of this transform can be found in [9].

2.3 Image Matching

Given an images I_i , a set of SIFT key points $K_i=\{k_{i1}, k_{i2}, \dots, k_{in}\}$ is detected. In matching a pair of images I_i and I_j , all the keypoints K_i from I_i and K_j from I_j are compared to measure how many keypoints are successfully matched. The Euclidean distance from k_{ia} to all the key points in K_j are calculated to obtain the closest distance d_1 and the second closest distance d_2 . When the ratio d_1/d_2 is sufficiently small (less than a threshold th_ratio which is 0.49 in our system), k_{in} is considered to have a matching key point in K_j . By using the ratio of d_1 and d_2 , both the similarity and the uniqueness of point pairs are considered. This algorithm is summarized below.

```
SIFT-Matching ( $K_i=\{k_{i1}, k_{i2}, \dots, k_{in}\}, K_j, th\_ratio$ )
1  for  $a \leftarrow 1$  to  $n$ 
2      do  $d_1 \leftarrow$  closest-keypoint-distance ( $k_{ia}, K_j$ )
3      do  $d_2 \leftarrow$  second-closest-keypoint-distance ( $k_{ia}, K_j$ )
4      if  $d_1/d_2 < th\_ratio$ 
5          do add-matching-keypoint ( $k_{ia}$ )
```

2.3.1 Geometric Constraints in SIFT Matcher

We use local geometric constraints to reduce the number of false matching points. Let M_{ij} represent the matching key points between image I_i and I_j . Then, M_{ij} can be expressed in terms of two different subsets $M_{ij} = M_{ij,T} \cup M_{ij,F}$, where $M_{ij,T}$ represents the true matching points and $M_{ij,F}$ represents false matching points. It is expected that removing the false matching points will increase the retrieval accuracy. The SIFT descriptor, the basic matching attribute, is constructed as a fixed length (=128) histogram based on edge orientation. The number of false matchings in the presence of viewpoint variations or blurring in the image is likely to be large. When a key point belongs to $M_{ij,F}$, it is likely to match to many other key points. On the other hand, a key points in $M_{ij,T}$ is likely to match to either one or very small number of other key points. Given a query image I , it is matched with all the images in the gallery database D and the number of matching points is obtained for each gallery image. Let $L_m, m=1, 2, 3, \dots$, represent a set of key points in the query image that are matched into the same key point in D . We calculate the size of the area covered by L_m , and regard L_m as belonging to $M_{ij,F}$ if the size is larger than a threshold t (value is set at 0.2). All the matching key points not in $M_{ij,F}$ are regarded as true matching points. Finally, the

number of key points that belong to $M_{ij,T}$ is used to retrieve the top-N candidate images.

3. EXPERIMENTAL RESULTS

Fig. 7 compares the performance of retrieval experiments with and without the proposed geometric constraints. Without the constraints, the rank-1 accuracy is 36.4% and the rank-20 retrieval accuracy is 76.2%. The corresponding accuracies with the constraints are 85.9% and 89.5%, respectively. This shows that the proposed geometrically constrained SIFT matching is remarkably effective to overcome the limitations of the conventional SIFT operator. We collected 636 query-gallery image pairs that were incorrectly matched at rank-1 when constraints were not used, and observed the number of matching keypoints. While the average number of matching keypoint without the constraints is ~ 70 , it is dramatically reduced to ~ 8 when the constraints were applied. This shows that the geometric constraints help in removing falsely matched keypoints.

Fig. 8 shows the performances of retrieval experiments with and without the class label information. The highest accuracy is obtained when the tattoo location and tattoo class information are both utilized. The best rank-1 accuracy we obtained is 90.6%, and the best rank-20 accuracy is 94.2%. The tattoo location and class label information not only improved the retrieval accuracy, it also reduced the matching time significantly. The average time for matching a query to 63,592 images in our gallery is reduced from 150.32 sec to 2.9 sec, (on Intel Core 2, 2.66 GHz, 3 GB RAM processor), by utilizing the tattoo location and class label information. The longest matching time is 11.43 minutes and the shortest time is 0.02 sec.

4. SUMMARY

We have presented a CBIR system for tattoo images. Such a system has great value in apprehending suspects and indentifying victims in forensics and law enforcement applications. Tattoo labels, in the form of its location on the body and its ANSI/NIST class label(s), are utilized to improve the matching time as well as retrieval accuracy. In addition, geometric constraints are applied to SIFT matching which dramatically reduce the number of false matchings. Our experimental results show the rank-20 accuracy of 94.2% with the use of tattoo labels and geometric constraints.

ACKNOWLEDGEMENT

This research was supported by the FBI Biometric Center of Excellence and CITeR at West Virginia University. We would like to thank Scott Swann and John Manzo of FBI and Lt. Gregory Michaud of Michigan Forensic Laboratory for their encouragement and assistance.

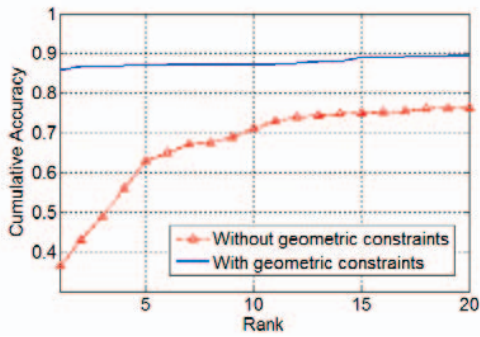


Fig. 7. CMC curves comparing the retrieval performances with and without the use of geometric constraints.

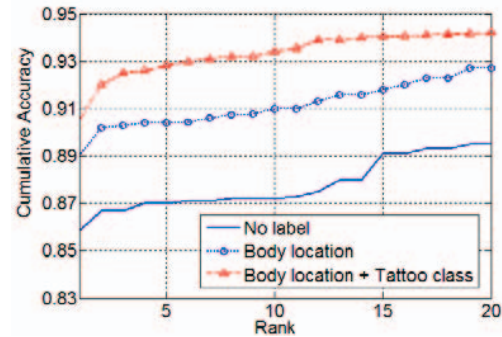


Fig. 8. CMC curves comparing the retrieval performances with and without the use of label information.



Fig. 9. Retrieval examples. Each row shows a query and its top-7 retrieved images and the associated match score

5. REFERENCES

- [1] Flickr, 3 Billion! <http://blog.flickr.net/en/2008/11/03/3-billion/>
- [2] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of the New Age", ACM Computing Surveys, Vol. 40, No. 2, pp. 1-60, 2008
- [3] The meaning of the tattoos, <http://community.livejournal.com/easternpromises/47809.html>
- [4] Cold case homicide team, police department, the city of San Diego, <http://www.sandiego.gov/police/about/coldcase95.shtml>
- [5] Violent Gangs, Federal Bureau of Investigation, <http://www.fbi.gov/hq/cid/ngic/gallery.htm>
- [6] T. Pavlidis, "Limitations of Content-based image retrieval", <http://www.theopavlidis.com/technology/CBIR/PaperB/vers3.htm>
- [7] J-E. Lee, A. Jain and R. Jin, "Scars, Marks, and Tattoos (SMT): Soft Biometric for Suspect and Victim Identification", Proc. Biometric Symposium, 2008. pp. 1-8.
- [8] ANSI/NIST-ITL 1-2000, Data Format for the Interchange of Fingerprint, Facial, & Scar Mark & Tattoo (SMT). <http://www.itl.nist.gov/ANSIASD/sp500-245-a16.pdf>, pp 34-36
- [9] D. Lowe, "Distinctive image features from scale-invariant keypoints", IJCV, Vol. 60, Issue 2, pp. 91-110, 2004.
- [10] P. Ogilvie, and J. Callan, "Combining document representations for known-item search", Proc. Int. Conf. on Research and Development in Information Retrieval, pp. 143-150, 2003.
- [11] Personal descriptors, National Crime Information Center (NCIC) code Manual, FBI, <http://www.itl.nist.gov/ANSIASD/1/sp500-245-a16.pdf>, pp 59-63