JOURNAL OF INFORMATION, KNOWLEDGE AND RESEARCH IN COMPUTER ENGINEERING

CONTENT BASED FORENSICS TATTOO IMAGE RETRIEVAL AND IDENTIFICATION METHOD FOR DATABASE APPLICATIONS

MOUNESHWAR KANAMADI [1] , DR D.S.BHOSLE [2]

Department of Computer science and Engineering, AMGOI Vathar Kolhapur Maharastra
Department of Information Science and Engineering, AMGOI Vathar Kolhapur Maharastra

ABSTRACT: *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. Although tremendous progress has been made in biometrics and forensics, many situations exist where the primary biometric traits such as fingerprint, face, and iris alone cannot identify an individual with sufficiently high accuracy . Next Generation Identification (NGI) system for identifying criminals by 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 the current work 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. The proposed system uses the Tattoo-ID content-based image retrieval (CBIR) system which automatically extracts features from a query image and retrieves near-duplicate tattoo and also improves the retrieval accuracy particularly for queries with low quality images from a database with 90.5 percent.*

Key words: Tattoo image, Next Generation, Identification (NGI), Content-based image retrieval (CBIR) system.

INTRODUCTION

There has been an increase in the placement of body art involving puncturing of the skin in recent years. In the past more men than women sported tattoos, body piercings were mainly in the soft part of the ear lobes of women. During the 1940s military personnel often had tattoos with patriotic designs, together with hearts and the names of loved ones. 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. 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 or only a partial fingerprint is available. People have used tattoos in order to represent themselves and the size of the tattooed population is rising rapidly.

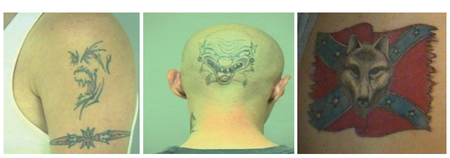


Figure 1: Tattoo images. These samples are from a database of tattoo images.

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 and Asian tsunami in 2004. Criminal identification using tattoos is another important application. Many individuals acquire tattoos to individuate themselves, display their personality, or exhibit a group membership (see Figures2), the analysis of tattoos often leads to a better understanding of an individual's background and membership in various organization.

(a) (b) (c)

Figure 2: Tattoos for identification. (a) a crime suspect and (b) a victim of the 2004 Asian Tsunami. (c) For gang membership tattoos.

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. 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. 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. we constructed a database tattoo images. 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. We used one of the duplicates as a query to retrieve other duplicates in the database.(see fig 3)

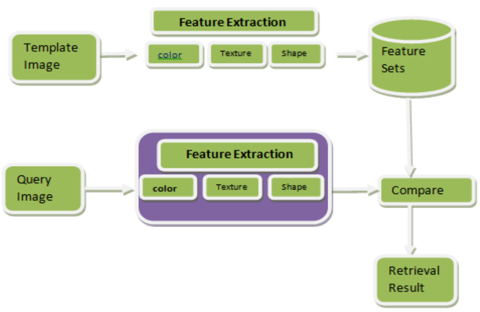


Fig 3: Image retrieval system.

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.

RELATED WORK

Tattoo-id: automatic tattoo image retrieval for suspect & victim [1], matching and retrieval of tattoo images: active contour CBIR and glocal image features [2], content-based image retrieval: an application to tattoo [3], scars, marks and tattoos: automatic matching & retrieval [4], Unsupervised Tattoo Segmentation Combining Bottom-Up and Top-Down Cues [5]

A. Tattoo-id: automatic tattoo image retrieval for suspect & victim:

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 to compute the similarity. Given two normalized histograms $H1$ and $H2$, the similarity is defined as:

$$S_{H1H2} = \sum_{i}^{B} \min(H^1(i), H^2(i))$$

Where B is the number of bins. 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. To evaluate the retrieval performance of the Tattoo-ID system, two sets of experiments are conducted with following scenarios:

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.4 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.

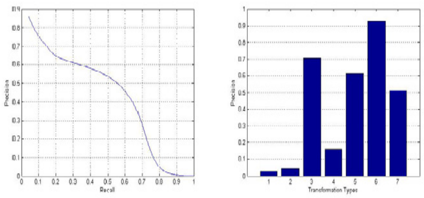
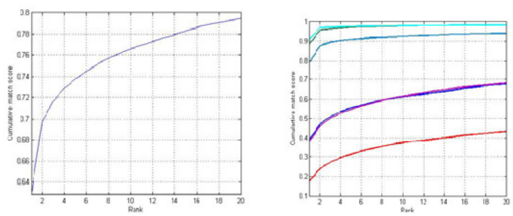


Fig. 4: Experiment I results: (a) Precision & recall curve, (b) precision of each transformation at rank 20

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. Since there is only one truly "similar" image in the database for every transformed query image, we adopted the cumulative matching curve (CMC) as the evaluation metric for this experiment. Fig. 5 shows the CMC curve



.

**Fig. 5: 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.**

B.Matching And Retrieval Of Tattoo Images: Active Contour Cbir And Glocal Image Features

Active Contour Segmentation

Before image segmentation via active contours is attempted, the image is simplified using an area open-close filter. The filter removes small scale objects that might impede the extraction of the dominant tattoo shape. We use an active contour to extract the tattoo. Essentially, an active contour is a curve that moves based on two forces. First, an internal force makes the contour smooth. Second, an external force stops the contour at peaks in edge magnitude, so that the contour comes to rest at image boundaries. Unfortunately, active contours must be very close to such boundaries before the contour can sense the boundary and rest on the edge. In other words, active contours are very sensitive to initialization and can be sensitive to image noise and clutter. In order to capture tattoo boundaries in a robust manner, we apply vector field convolution (VFC) as the external force. VFC provides robustness to noise and initialization, flexibility and reduced computational cost.VFC operates on the premise of convolving a vector field kernel with the edge information. The convolution extends the field of influence of a given edge, "guiding" the active contour to strong edges. In VFC, k is the vector field kernel and is defined as

$$\mathbf{k}(x, y)=m(x, y) - \mathbf{n}(x, y) ,$$

Where $m(x, y)$ is the magnitude of the vector at (x, y) and $\mathbf{n}(x, y)$ is the unit vector pointing to the origin. The VFC external force $\mathbf{v}(x, y) = [u(x, y), v(x, y)]$ is given by calculating the convolution of the vector field kernel $\mathbf{k}(x, y)$ and the edge map $f(x, y)$ (which could be normalized gradient magnitude)

$$\mathbf{v}(x, y)=f(x, y)*\mathbf{k}(x, y) ,$$

Where * denotes convolution. Since the edge map value $f(x, y)$ is larger near the image edges, edges contribute more to the VFC than homogeneous regions. An example of VFC contour evolution is shown in Fig.6

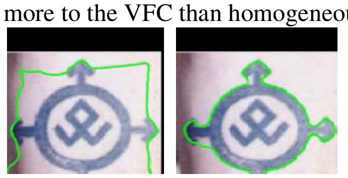


Fig. 6: VFC active contour evolution progressing from left to right.

Skin Detection

We also attempted to improve performance via skin detection. Here, the idea is to detect skin regions, and then to increase the speed of the VFC active contour proportional to the probability of skin presence at a given position. We computed such probabilities using the classical approach in HSV space. A typical result is shown in Fig.7

Fig. 7: Skin detection: a) original image; b) skin shown in white and potential tattoo area in black.

C. Content-Based Image Retrieval: An Application to Tattoo Image Matching

Given an images *Ii*, a set of SIFT key points *Ki={ki1, ki2, …, kin}* is detected. In matching a pair of images *Ii* and *Ij*, all the key points *Ki* from *Ii* and *Kj* from *Ij* are compared to measure how many key points are successfully matched. The Euclidean distance from *kia* to all the key points in *Kj* are calculated to obtain the closest distance *d1* and the second closest distance *d2*. When the ratio *d1/d2* is sufficiently small (less than a threshold *th_ratio* which is 0.2 in our system), *kin* is considered to have a matching key point in *Kj*. By using the ratio of *d1* and *d2*, both similarity and uniqueness of point pair are considered. This algorithm is summarized below.

SIFT-Matching (*Ki= {ki1, ki2… kin}*, *Kj*,*th_ratio*)
1. **for** *a* ← 1 **to** *n*
2 .**do** *d1* ← closest-key point-distance (*kia*, *Kj*)
3 .**do** *d2* ← second-closest-key point-distance (*kia*, *Kj*)
4 .**if** *d1/d2* < *th_ratio*
5. **do** add-matching-key point (*kia*)

Geometric Constraints on SIFT Matcher

We use local geometric constraints to reduce false matching points. Let *Mij* represent the matching key points between image *Ii* and *Ij*. Then, *Mij* can be expressed with two different sets as *Mij* = *Mij,T* ∪ *Mij,F*, where *Mij,T* represents true matching points and *Mij,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. There is a large possibility of false matchings in the presence of viewpoint variations or blurring in the image. When a key point belongs to *Mij,F*, it is likely to match to many other key points. On the other hand, a key points in *Mij,T* is likely to match to 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 *Lm*, *m=1, 2, 3, …*, 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 *Lm*, and regard *Lm* as belonging to *Mij,F* if the size is larger than a threshold *t* (value is set at 0.2). All the matching key points not in *Mij,F* are regarded as true matching points. Finally, the number of key points that belong to *Mij, T* is used to retrieve the top-N candidate images.

D. Scars, Marks And Tattoos :Automatic Matching & Retrieval

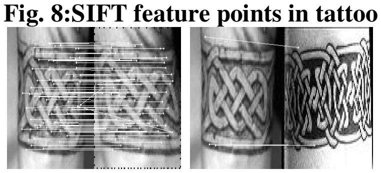
Sift Features

Scale Invariant Feature Transform (SIFT) extracts 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 viewpoint, addition of noise and change in illumination.

Keypoint Matching

Matching is performed by comparing key points in two images based on the associated descriptors using the Euclidean distance metric. We use the number of matching key points as the matching score between two images (seeFig.8)

Fig. 8:SIFT feature points in tattoo images



(a) 87 matching points (b) 3 matching points

Fig.11:Two matching examples with the number of matching key points between a pair of (a) similar and (b) different tattoo images

E. Unsupervised Tattoo Segmentation Combining Bottom-Up And Top-Down Cues

Regarding the complexity and intra-class variance of tattoo, our main idea is to transfer the tattoo segmentation into skin detection followed by a figure-ground segmentation. The outline of our algorithm is depicted in Fig.9.

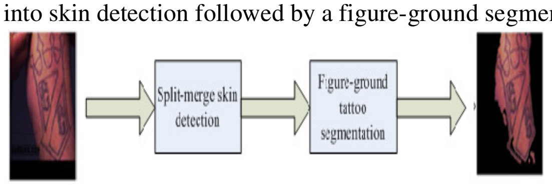


Figure 9: The outline of tattoo segmentation.

Initially a clustering technique is used on the range domain (color space) to separate the tattoo image into over segmented regions in an bottom-up manner. This is a very important step in that regions containing both skin and tattoo are much more non-homogeneous than these over segmented regions. Based on top-down cues learned from the image itself, a region merging step is introduced to group skin regions together. Through this split-merge process, skin and tattoo are distinguished from the background. Finally, K-Means algorithm is applied for figure-ground segmentation, where now the tattoo is the foreground and the skin is the background.

Split-merge skin detection:-To make our system clear, we simply begin by applying an initial clustering process on the gray-scale distribution f(I) of the image I so as to obtain an initial segmentation. Here f(I) evaluates the density estimate covering the range of intensity in the gray-scale form of I. Since pixels from the same cluster are more likely to have the similar intensity, we segment the tattoo images according to the local minima of distribution f(I). After applying the histogram-based clustering, $d-1$ local minima in f(I) indicates d clusters. Obviously, such Initial segmentation may suffer from under-segmentation due to background with similar intensity and over segmentation due to illuminant variation on the skin.

Figure-ground tattoo segmentation

To this point, we transferred a problem of tattoo segmentation with unknown number of clusters to a skin-tattoo binary segmentation. In this section, skin pixels should be distinguished from pixels belonging to tattoo. In that sense, skin pixel in this section indicates merely the skin pixel that is not covered by tattoo. Since we already know the number of potential clusters now, a k-means algorithm ($k = 2$) can be applied on the RGB color space of the foreground (skin). Now, the issue is which cluster should be tattoo. If distinction between tattoo and skin is needed, the pixels on the contour of the skin region are more likely to be skin pixels rather than pixels in tattoo. Because, otherwise, skin are fully covered by tattoo and distinguishing tattoo from skin is thus unnecessary. The cluster with more pixels on the contour of the foreground is labeled as the skin and the other the tattoo. If the structure of tattoo is preferred rather than the whole region containing tattoo in application, an alternative way of marking the tattoo is to apply a ridge on the skin region. This is reasonable since tattoo is a kind of man-made painting with clear boundaries while skin regions are more texture less in contrast.

CONCLUSION

This survey concludes that, CBIR system for tattoo image 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, geometrical constraints are applied to SIFT matching which dramatically reduces the number of false matchings. The SIFT key points described in this paper are particularly useful due to their distinctiveness, which enables the correct match for a keypoint to be selected from a large database of other key points.. The SIFT key points are particularly useful due to their distinctiveness, which enables the correct match for a keypoint to be selected from a large database of other key points.

REFERENCES

1. Tattoo-Id: Automatic Tattoo Image Retrieval For Suspect & Victim Identification Anil K. Jain, Jung-Eun Lee, and Rong Jin

2. Matching And Retrieval Of Tattoo Images: Active Contour Cbir And Glocal Image Features *Scott T. Acton2, Adam Rossi1* 1 Platinum Solutions, Reston, VA, USA

3 .Content-Based Image Retrieval: An Application to Tattoo Images *Anil K. Jain, Jung-Eun Lee, Rong Jin and Nicholas Gregg*

4. J.-E. Lee, A.K. Jain, And R. Jin, ‘‘Scars, Marks And Tattoos: Soft Biometric For Suspect And Victim Identification’’ Proc. Biometric Symp. (BSYM 08), IEEE Press, 2008

5. 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.R. Datta et al.

6. A.K.Jain, S.C.Dass, K.Nandakumar, ‘‘Can Soft Biometric Traits Assist User Recognition?’’ Proc. SPIE Conf. Biometric Technology for Human Identification, vol. 5404, 2004, pp. 561-572.

7. 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.