Image Retrieval Based upon Directional Fields

Che-Yen Wen,1,* Ph.D.; Zheng-Xian Li,2 M.S.; Chiu-Chung Yu,3 M.S.

1 Department of Forensic Science, Central Police University, No.56, Shu Jen Road, Taoyuan 333, Taiwan, R.O.C.
2 Crime Prevention Affairs Section, No.5, Lane 553, Chung Hsiao East Road Section 4, XinYi area, Taipei, 11002, Taiwan, R.O.C.
3 Forensic Science Center, Kaohsiung City Government Police Bureau, No.260, Jhongjheng 4th Road, Cianjin District, Kaohsiung City 80147, Taiwan, R.O.C.

Received: October 15, 2008 / Accepted: November 10, 2008

Abstract

Image database has been used in forensic science recently, such as fingerprint, cartridge cases, drugs tablets, tool marks, suspect face photos, shoeprints, handwriting, etc. There is a growing interest in finding images in large collections or from remote databases. In order to retrieve an image by its contents, the image has to be described or represented by features. This kind of databases is called the Content Based Image Retrieval (CBIR) system. CBIR uses features to be as retrieval "keywords".

Shape is an important visual feature in CBIR. In this paper, we propose a robust representation and description method for images. This method can be applied to CBIR.

Keywords: CBIR, Forensic science, Shape.

Introduction

There is a growing interest in finding images in large collections or from remote databases. In forensic sciences, photographic images have been increasingly stored and transmitted in the digital format recently. The forensic images include crime scene pictures, fingerprint, shoeprints, tool marks, shoeprints, faces, handwriting, cartridge cases, drugs tablets, and so on. Those forensic images have been built in different databases. Forensic experts usually use these images to construct the relevancy of evidences and provide useful information for detectives. However, it is tedious to search target images by human visual in large database.

Recently, CBIR attracts forensic examiners’ attention. Many image retrieval techniques have been proposed to assist forensic experts in identification. The developments of various available image databases and ways of searching these databases on image contents have been provided. These visual features were evaluated and compared with existed forensic databases [1]. Visual features include color, texture, shape and spatial relationships. We have used these features to retrieval crime scene image database [2]. In visual features, many researches still attach great importance to shape feature. For examples, pistol images left in crime scene can be represented by shape descriptors and retrieved the target image from pistol images [3]. In order to find an image, the image has to be described or represented by certain features. Shape is an important visual feature and it has been designed in various forms, including shape signature, signature histogram, shape invariants, moments, curvature, shape context, spectral features, etc [4].

An edge orientation histogram method for shape similarity measure is proposed. This method applies an edge detector to an image and constructs the histogram of local tangent orientation of all pixels that lie on edges. Since this method uses the edges individually and ignores correlation between neighboring edges, its effectiveness is limited [5-7]. The direction histogram of similar edges was considered to establish the correlation between neighboring edges using a weighting function.

* Corresponding author, e-mail: cwen@mail.cpu.edu.tw
and adds this extra information to the edge direction histogram. The edge orientation autocorrelogram (EOAC) was proposed [8-10]. The correlation between neighboring edges in a window around the kernel edge is designed to get the correlogram of the global distribution of local edges. This feature classifies image edges based on two factors: edge orientation and correlation between neighboring edges. Hence, it includes information of continuous edges and lines of images and describes major shape properties of images. EOAC propose a gradient based method, called “average squared gradients”, to calculate the average orientation of window. Besides, it also gives a measurement called “coherence” that indicates how well the gradients are pointing in the same direction. Rotated human face images and cartridges’ breach face impression were used to examine the retrieval capability of EOAC [11-13]. Since the orientation is averaged in a window, it has a better ability of noise resistance than tangent orientation.

In this paper, we propose an edge orientation histogram based method to extract features of images, and we apply those features to CBIR. The edge orientation gives the shape information and the coherence gives both correlation and texture information around edges.

**Methods**

Shape representation and description techniques can be generally grouped into two main categories: contour-based methods and region-based methods [4, 14]. The classification is according to the shape features that extracted from the contour or the whole shape region. For each class, methods are further divided into structural approaches and global approaches. This subclass is based on whether the shape is represented as the whole object or represented by segments. The flow chart of the classification is shown in Fig.1.

---

**Fig.1** Classification of shape representation and description techniques.
The edge orientation autocorrelogram (EOAC) belongs to the global approach. EOAC produces a feature vector derived from the integral boundary is used to describe the shape. The measure of shape similarity is usually a metric distance between the acquired feature vectors. The highlights of this approach includes (1) the correlation between neighboring edges in a window around the kernel edge; (2) the global distribution of local correlation of edges; (3) the shape aspects of an image is not sensitive to color and illumination variation; (4) the factors of translation, scaling, and small rotation are invariant; (5) image can be computed easy. Our experiments show that this new feature outperforms non-segmentation based methods in retrieval by shape similarity. The EOAC classifies edges based on their orientations and correlation between neighboring edges; hence it contains major shape properties of the image. In the below sections, the algorithm of generating EOAC will be proposed.

**Color to Gray-scale Transformation**

Our method is based upon the intensity information of an image. We transform color images to gray-scale images before applying our method. Here we use the HSB (hue, saturation, brightness) color model which also known as HSI (hue, saturation, intensity) or HSV (hue, saturation, value) to explain the transformation.

Hue is a color attribute that describes a pure color (pure yellow, orange, or red), whereas saturation gives a measure of the degree to which a pure color is diluted by white light. Brightness embodies the achromatic notion of intensity. The HSB color model decouples the intensity component (brightness) from the color-carrying information (hue and saturation) in a color image. Hence, brightness is used to minimize the affection of color variation extensively and the applicability of this new technique has been verified [15-16].

Since the RGB model is the mostly used model, we review the process of transforming the RGB model to the HSB model:

\[
H = \begin{cases} 
60 \times \frac{g-b}{\text{MAX} - \text{MIN}} + 0, & \text{if } \text{MAX} = r \text{ and } g \geq b \\
60 \times \frac{g-b}{\text{MAX} - \text{MIN}} + 360, & \text{if } \text{MAX} = r \text{ and } g < b \\
60 \times \frac{b-r}{\text{MAX} - \text{MIN}} + 120, & \text{if } \text{MAX} = g \\
60 \times \frac{r-g}{\text{MAX} - \text{MIN}} + 240, & \text{if } \text{MAX} = b 
\end{cases}
\]

\[
S = \text{MAX} - \text{MIN} \\
B = \text{MAX}
\]

where \( \text{MAX} \) is the maximum of the \((r,g,b)\) values, and \( \text{MIN} \) is the minimum of those values. Fig.2 shows each component of RGB and HSB color model of a color image respectively.
Fig. 2 All components of RGB and HSB color models for a color image. (a) A color image downloaded from internet (http://home.online.no/~rhagen2/privat/graphics/here/colorful.jpg). (b), (c), and (d) are the R, G, B components of (a) represented in gray-scale images respectively. (e), (f), and (g) are the H, S, B components of (a) represented in gray-scale images respectively.
**Edge Detection**

The “edge” is a “local” variation that is based on a measure of gray-level discontinuity at a point. We define a point in an image as being an edge point if its two-dimensional first-order derivative is greater than a specified threshold. The gradient of $f(x,y)$ is the first-order derivative at location $(x,y)$, it is defined as the vector [16]

$$
\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} .
$$

(2)

The magnitude of this vector is also defined as

$$
\text{mag}(\nabla f) = \left( G_x^2 + G_y^2 \right)^{1/2} .
$$

(3)

We use two times of the root mean square of gradient magnitudes of the whole image as the threshold to decide if the point belongs to the edge part. Fig.3(a) is the landscape image, the result of edge detection is displayed as Fig.3(b).

**Fig.3** (a) The landscape image; (b) The result of edge detection; (c) The coherence map of edge image.
The value of \( W \) is between 0 and 1. For a strongly oriented pattern and isotropic regions, its value is close to 1 and 0, respectively. The coherence map of an image is displayed as another image in Fig.3(c).

**Edge Orientation Quantization**

The orientation of edge computed by the average squared gradients method provides continuous values (real number). Although continuous values describe more details of the orientation of edges, it also requires more storage space and is difficult to compare as a feature. For this reason, we have to quantize the orientation into segments. In contrast, insufficient segments of orientation of edges will result in bad discrimination. In our method, we quantize the orientation of edge uniformly into 36 segments and each segment is equal to \( \frac{\pi}{36} \) radian or five degrees which we found is sufficient in representing orientation of edges.

**Feature Computation**

The coherences of the same quantized orientation of edges are summed through the entire image. The formula is as follows [11]

\[
\theta(x, y) = \frac{1}{2} \tan^{-1} \left( \frac{2G_xG_y}{\sum_{i} (G_x^2 - G_y^2)} + \frac{\pi}{2} \right)
\]

\( \theta(x, y) \) is the angle of orientation ranging from \(-\pi/2\) to \(\pi/2\), and \( \tau \) is the neighboring area of position \((x, y)\). We set the neighboring area as a square box of odd sizes (such as 3x3). When width or height of an image is less than the predefined threshold, \( t_P \) (depending on the image size in the database for retrieval), we must normalize the box size \( BS \) by

\[
BS = \begin{cases} 
3 & \text{if } sL \leq tP \\
3 + 2 \times \text{floor} \left( \frac{sL \times 3}{tP \times 2} \right) & \text{otherwise} 
\end{cases}
\]

where \( sL \) is the smaller side of the image size, \( \text{floor}(\cdot) \) as a function that rounds the element to the nearest integers less than or equal to the element. Since the neighboring area varies with the image size, this area is naturally invariant to image scaling. In averaging squared gradients method, there is still another measurement called coherence, \( W(x, y) \). It indicates how well the gradients are pointing in the same direction. The formula is

\[
W(x, y) = \frac{\left( \sum_{i} (G_x^2 - G_y^2) \right)^2 + 4 \left( \sum_{i} G_xG_y \right)^2}{\left( \sum_{i} (G_x^2 + G_y^2) \right)^2}.
\]

The value of \( W \) is between 0 and 1. For a strongly oriented pattern and isotropic regions, its value is close to 1 and 0, respectively. The coherence map of an image is displayed as another image in Fig.3(c).

**Computation of Orientation and Coherence**

We use the method called “averaging squared gradients” to calculate the average orientations of edges within their neighboring areas [10]. Gradients cannot directly be averaged in some local neighborhood since opposite gradient vectors will then cancel each other, although they indicate the same orientation. A solution to this problem is to double the angles of the gradient vectors before averaging them. After doubling the angles, opposite gradient vectors will be in the same direction, and their magnitude will be reinforced. Meanwhile, the magnitude of perpendicular gradients will be cancelled. After averaging, the gradient vectors have to be converted back to their single-angle representation. This method not only doubles the angle of the gradients but also squares the length (magnitude) of the gradient vectors. This has the effect that strong orientations have a higher vote in the average orientation than weaker orientations. The formula is as follows

\[
\theta(x, y) = \frac{1}{2} \tan^{-1} \left( \frac{2G_xG_y}{\sum_{i} (G_x^2 - G_y^2)} + \frac{\pi}{2} \right)
\]

\( \theta(x, y) \) is the angle of orientation ranging from \(-\pi/2\) to \(\pi/2\), and \( \tau \) is the neighboring area of position \((x, y)\). We set the neighboring area as a square box of odd sizes (such as 3x3). When width or height of an image is less than the predefined threshold, \( t_P \) (depending on the image size in the database for retrieval), we must normalize the box size \( BS \) by

\[
BS = \begin{cases} 
3 & \text{if } sL \leq tP \\
3 + 2 \times \text{floor} \left( \frac{sL \times 3}{tP \times 2} \right) & \text{otherwise} 
\end{cases}
\]

where \( sL \) is the smaller side of the image size, \( \text{floor}(\cdot) \) as a function that rounds the element to the nearest integers less than or equal to the element. Since the neighboring area varies with the image size, this area is naturally invariant to image scaling. In averaging squared gradients method, there is still another measurement called coherence, \( W(x, y) \). It indicates how well the gradients are pointing in the same direction. The formula is
space provided as the characteristic property of target image and the excellent retrieval index is established consequently.

Fig. 4 Some examples are shown to construct the feature space.
Normalization

An effective image retrieval similarity-matching algorithm is independent of translation, rotation, scaling, color, and illumination variations.

Because translation causes no effect on the orientation and the coherence around edges, this method is naturally invariant to translation. Besides, image color and illumination variations change relative intensity difference between neighboring regions thus these variations alter only the gradient magnitudes of the image. Since we use twice the root mean square of gradient magnitudes of whole image as the threshold while detecting edges, our proposed method is also invariant to color and illumination variations. Since image scaling affects the total number of image pixels, it affects the number of edges in consequence. Theoretically speaking, the orientation and coherence around edges remain the same. But we found that is only true for ideal cases. Because scaling inherently adds or loses some information of the image, the orientation and coherence are also slightly affected. Despite the disturbance during image scaling process, we normalize the feature vectors by scaling with the edge numbers. Fig. 5 shows some examples. Rotation of an image only shifts the histogram of edge orientations and the coherence near an edge remains the same. Fig. 6 shows an example. In order to prevent the rotation variation, we shift the edge orientation histogram and use the minimal difference while retrieving an image.

Fig. 5 Scale invariant by normalization.

(a) 

(b) 

(c) 

(d)
The feature vectors are pre-computed and stored for entire images. Then the feature vector of query image is computed. We use $L1$ distance as comparison measures with the following formula

$$L1(x,y) = \min \sum_{i=1}^{n} |x_i - y_i|$$

where $(x,y)$ is the query image and the image from database respectively, $x$ and $y$ are the feature vectors of $(x,y)$, $i$ is the rank number of the feature vector, and $n$ is the number of elements of feature vector. Next chapter will show some retrieved examples ranking by similarity calculation.

**Experiments**

In this section, the shape feature of image by calculating the correlation of edge points is proposed to illustrate the main concept of our method. Therefore, these image databases used to scrutinize our methods include (1) forensic image databases include the element of tire mark and shoe print from SICAR (Tyre mark and shoe print evidence management system) and the CD-ROM from www.cartridge-corner.com; (2) shape image databases created by Latecki, Lakamper [17] and Sebastian, Klein [18] respectively; (3) texture image database that is downloaded from internet (http://www.pbase.com/pieval/structure).

**Shape image databases**

About the recognition of 2D and 3D objects, shape features played the leading role in image retrieval. That is
the fast methods to accept or reject what is our attention in images. Figs. 7 and Fig. 8 show some samples and the retrieval results respectively. From the retrieval results, all similar shapes of the image database are ranked by the degree of similarity. According to this method, the motion detection and recognition of objects will be implemented vigorously by automatic image process techniques.

![Fig.7 The retrieve result of human shape images.](image-url)
Fig. 8 The retrieve result of animal shape images.
**Texture image databases**

EOAC could also be used to classify texture images. As Fig.9, the texture images that have more horizontal, vertical and directional edges arranged will be classified as the same category. The individual texture of textile is constructed by different weaves, hence this method is suitable to classify and retrieve the individual texture of textile.

![Fig.9 The retrieve result of texture images.](image-url)
Forensic image database

Shoeprints are often found at crime scene. The elements of shoeprint have variety of patterns, such as lines, waves, zigzags, circles, diamonds and blocks etc. Regardless of complete or imperfect shoeprints left, we still discover some elements of shoeprint remain on the surface of an object at the crime scene. We can find suspect shoe patterns by recognizing the shoe element patterns, as Fig.10.

Fig.10 The retrieve result of the element of tire marks or shoe prints.
Another head stamp image database is used to experiment with our methods. The markings imprinted on the base of the cartridge case had constructed the special pattern, that usually contains information on the caliber and manufacturer or his logo. As Fig.11 shows the retrieval result, the target image listed in the top rank is the same as the query image.

Fig.11 The retrieve result of head stamp images.
Conclusion

In this paper, we propose an edge orientation histogram based method to extract features of images, and we apply those features to CBIR. The edge orientation gives the shape information and the coherence gives both correlation and texture information around edges. The “shape” and “texture” features are thus combined together to analyze images. Our method is translation, rotation, scaling, color, and illumination invariant. In our experiments, we also apply our method to some forensic image databases. From the experimental results, we can retrieve target images exactly from databases. In the future work, we will apply our method to wide forensic databases, such as shoeprints, centerfire headstamp.

Reference