Advances in Social Science, Education and Humanities Research, volume 204
4th International Conference on Economics, Social Science, Arts, Education and Management Engineering (ESSAEME 2018)

Research on Image Retrieval Algorithm Based on Shape Analysis

Lei Ao[1], Cheng Zhang[2], Yuchen Wei[3]

Network Engineering Teaching and Research Section

City Institute, Dalian University of Technology

Dalian, China

E-mail: 154001958@qq.com

Keywords: Texture, Co-occurrence matrix, Roughness, Contrast, Directionality

Abstract: The texture in digital images is the spatial correlation of the gray or color of adjacent pixels, or the visual representation of the change of image gray level and color space position, which can be expressed as edges, shapes, stripes, color blocks and so on. Therefore, to extract the texture features by the method of mathematics or information theory, the image grayscale or color space overall or each side with digital feature representation, then using texture analysis method to detect the texture element texton and access to relevant information distribution arrangement.

1. Introduction

Shape, as one of the most important image features, is an important basis for people to understand the image. In the survey of users, people tend to use shape based image retrieval method in image retrieval system, because it is more suitable for people's visual perception. However, there are many problems to be solved in the real application of shape based image retrieval. For example, the QBIC retrieval system that is considered to be more successful is also based on color and texture algorithm, which is able to achieve satisfactory results in IBM retrieval. Shape matching is a central task in visual information system, computer vision, pattern recognition and robotics. Image retrieval based on shape is also a problem of how to effectively shape matching.

But the shape is a two-dimensional feature. Although the distance between two shapes can be calculated in multidimensional space, such as Hausdorff distance [12], the computation amount of matching in multidimensional space is very large, and it is very sensitive to noise. So the usual method is to reduce the dimension of two dimensional features into one dimensional feature, and then match one dimensional feature. Image retrieval must satisfy translation, rotation and scale invariance. Meanwhile, because image databases are usually very large, operational efficiency is very important.

2. Arithmetic Invariant Moment

Invariance moment is a statistical feature of image, which satisfies translation, rotation and scale invariance, and has been widely used in image recognition field. The invariant moments used to describe the shape have arithmetic invariant moments [13] and Zerrnike moments and Legendre moments.

For the integral region S, a two dimensional continuous function F (x, y) is given, and the (p+q) order moment is defined as:

$$mpq=\iint xpyqF(x,y)dxdy,(p,q=0,1,2\ldots\ldots)$$

If the digital image function F (x, y) is piecewise continuous and has a non 0 value in the finite part of XY plane, it can be proved that all the moments exist, and the moment sequence MPQ is uniquely determined by F (x, y), otherwise MPQ also uniquely identifies F (x, F). Its central moment can be expressed as follows:

$$Mpq=\iint(x-\bar{x})^p(y-\bar{y})^qF(x,y)dxdy$$

In the formula of the formula: $\bar{x}=m_x/m\ \infty$, $\bar{y}=m_y/m\ \infty$
For digital images, it can be used to replace integrals:

$$Mpq=\sum_x\sum_y x^p y^q f(x,y)$$

$$\mu_{pq}=\sum_x\sum_y(x-\bar{x})^p(y-\bar{y})^q f(x,y)$$

In 1962, Hu[15] proposed the invariant moment theory of image recognition, and established a statistical feature extraction method for image recognition, which has been widely used. Although arithmetic invariant moments are originally proposed to be used to calculate global image data, it can also be used to describe the shape features. Li[16] gives 52 calculation methods of invariant moments, and the seven invariant moments are discussed in this paper.

$$\mu_{pq}=\sum_x\sum_y(x-\bar{x})^p(y-\bar{y})^q f(x,y)$$

order normalized center moment:

$$\eta pq=\mu pq/\mu_{00}^{r} \qquad r=1+(p+q)/2 \qquad p+q=2,3\ldots$$

$\varphi_1=\mu_{20}+\mu_{02}$
$\varphi_2=(\mu_{20}-\mu_{02})^2+4\mu_{21}^{\ 2}$
$\varphi_3=(\mu_{30}-3\mu_{12})^2+(3\mu_{21}-\mu_{03})^2$
$\varphi_4=(\mu_{30}+3\mu_{12})^2+(\mu_{21}+\mu_{03})^2$
$\varphi_5=(\mu_{30}-3\mu_{12})(\mu_{30}+\mu_{12})[(\mu_{30}+\mu_{12})^2-3(\mu_{21}+\mu_{03})^2]+ (3\mu_{21}+\mu_{03})$
$(\mu_{21}+\mu_{03})[3(\mu_{30}+\mu_{12})^2-(\mu_{21}+\mu_{03})^2]$
$\varphi_6=(\mu_{20}-3\mu_{02})[(\mu_{30}+\mu_{12})^2-(\mu_{21}+\mu_{03})^2]+4\mu_{11}(\mu_{30}+\mu_{12})(\mu_{21}+\mu_{03})$
$\varphi_7=(3\mu_{21}-\mu_{03})(\mu_{30}+\mu_{12})[(\mu_{30}+\mu_{12})^2-3(\mu_{21}+\mu_{03})^2]+ (\mu_{21}+\mu_{03})$
$(\mu_{30}-3\mu_{12})[3(\mu_{30}+\mu_{12})^2-(\mu_{21}+\mu_{03})^2]$

Although these 7 moments can describe the shape features well, it is not enough to query the image by 7 scalars when the image database is large. Because the information they carry is very limited, so these invariant moments are usually used to pre filter the retrieved pictures, and then combine [19] with other methods.

3. Edge Comparison

The edge refers to the most significant part of the change in the local brightness of the image. The edge mainly exists in the target and target, target and background, target and region (including different colors). It is an important basis for image analysis, such as image segmentation, texture feature extraction and shape feature extraction. The first step in image analysis and understanding is the edge detection (edge detection). Because edge detection is very important, it has become one of the most active topics in the field of computer vision research.

3.1 Edge detection and tracking

In an image, the edge has two characteristics of direction and amplitude. Along the edge to the gray change smoothly and gray change perpendicular to edge acuteness, this change may be the step or ramp type, gray at the edge of the derivative greatly, while the two derivative at the edge of a value of zero, which respectively are positive and negative two peaks, but also that is to say, the edge points corresponding to the first derivative of the points of maximum amplitude, zero crossing point also corresponds to the two order differential. Therefore, it is a powerful means to extract boundary points by using the gradient maximum or the two derivative zero crossing point.

First order differential is the most basic method of image edge and line detection. The gradient of the image function f (x, y) at the point (x, y) (i. e., first order differential):

$$\nabla f(x,y)=[G_X G_Y]^r=[\frac{\partial f}{\partial x}\ \frac{\partial f}{\partial y}]^r$$

The amplitude and direction angle of this vector are as follows:

$$mag(\nabla f)=[G_x^2+G_y^2]^2$$

$$\varphi(x,y)=\arctan(G_y/G_x)$$

The partial derivative in the above three formula needs to be calculated for each pixel position. In practice, the cell domain template convolution is used to approximate the calculation. For G_x and G_Y, each one uses a template, so two templates are required to be combined to form a gradient operator. According to the size of the template and the difference in the value of the element, many different operators have been proposed.

3.2 Roberts edge operator

1	0
0	-1

0	1
-1	0

Fig.1 Roberts edge operators

Roberts edge detection is a kind of operator using local difference operator to find edge. Its two 2 * 2 templates are shown in Figure 1. The Roberts operator has the best response to a steep low noise image.

3.3 Sobel edge operator

The two convolution kernel shown in the following diagram forms the Sobel edge operator. Each point in the image is convolution with the two cores. One checkout is the most common vertical edge response and the other is most responsive to the horizontal edge.

-1	-2	-1
0	0	0
1	2	1

-1	-2	-1
0	0	0
1	2	1

Fig.2 Sobel edge operators

4. Two Order Differential - Laplacian Operator

LAPLACIAN OPERATOR is a two order differential operator, which defines a continuous function f (x, y), which defines the Laplasse value of the position (x, y) as follows:

$$\nabla^2 f=\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}$$

The Laplacian operator is a two order derivative operator, which is very sensitive to the noise in the image. In addition, it often produces two - pixel wide edges and can not provide information on the edge direction. Because of the above reasons, the Laplacian operator is rarely directly used for edge detection, mainly for the known edge pixels to determine the image pixel is in the dark area or area.

0	-1	0
-1	4	-1
0	-1	0

-1	-1	-1
-1	8	-1
-1	-1	-1

Fig.3 Form board of Laplacian operator

5. Segmentation Threshold Algorithm

In edge detection, the edge can be detected according to a predetermined threshold. This can simplify the algorithm, but because the edge features of the image are different, the pre set threshold can not be applied to all the images. This paper uses "Otsu" algorithm to calculate the threshold value of each image. The basic principle of [22] algorithm is the Otsu ":

If the threshold is selected to increase the variance of the gray level of the object part and the background part, the more accurate the image object and the background can be separated. An image of the gray level of M, B (I, J) is a pixel (I, J) gray, P (k) is the image gray value of the frequency of K, P (k) = set {(I, J) |b (I, J) the number of elements =k} t, such as segmentation of background and object to threshold gray value, namely background {(I, J) | B (I, J) >t}, objects (I, J) {|b (I, J) = t), so:

The number of pixels in a part of an object: $\overline{w}_0(t)=\sum_{i=0}^{t} p(i)$

The number of pixels in the background part: $_1(t)$ $\overline{w}=\sum_{i=t-1}^{m-1} p(i)$

The mean value of the grayscale of an object: $\mu_0(t)=\mu_0(t)=\sum_{i=0}^{t}\frac{i\times p(i)}{\overline{w}_0(t)}$

The gray mean of the background: $\mu_1(t)=\sum_{i=i+1}^{m-1}\frac{i\times p(i)}{\overline{w}_1(t)}$

The average gray mean of the whole image:$\mu=w_0(t)\times\mu_0(t)\times w_0(t)\times\mu_0(t)$

By "Otsu algorithm calculation formula for the best threshold value of image g:

$$G=\mathrm{Argmax}[\overline{w}_0(t)\times(\mu_0(t)-\mu)^2\times\overline{w}_0(t)\times(\mu_0(t)-\mu)^2]^2$$

To sum up, there are two kinds of model [23], namely, four face-model and eight face-model when using edge tracking. The four face-model is to judge the adjacent pixels in the four directions of the current pixel, such as non - background points, then the point is non - boundary, and the other is the boundary. The eight face-model is to judge the adjacent pixels in eight directions. It seems that the face-model is more accurate, but in fact it has a greater error in the calculation of the length of the contour arc.

As shown in Figure 4, for the four Pro model, 1, 3, 5, 6, and 8 points for the boundary; for the eight Pro model, 1-8 is the boundary, in the calculation of arc length, four face-model on the 1-3 point of arc length as the square root of 2, eight face-model on the 1-3 point of the arc length is 2. It is clear that the former is closer to the real value.

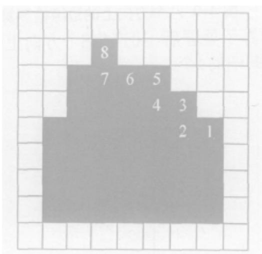

Fig.4 Four face-model of Border tracking

Figure 5 shows the result of using this method to trace the edge of the original graph:

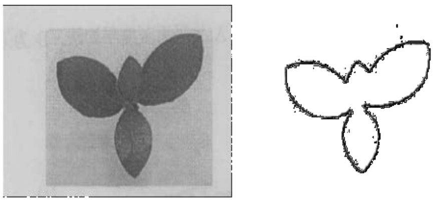

Fig.5 The picture of Border tracking

References

[1] Aalick R,Shanmugam K,Dinstein.Texture features for image classification, IEEE Transactions [J],Systems Man and Cybernetics,1973,3(6):17-20

[2] Li Xiancheng. Research on image retrieval technology based on main colors [J], traffic and computer, 2006, 24 (4): 8-10

[3] Manjunath B S.Texture Features for Browsing and Retrieval of Image Data[J], IEEE trans on image processing,1993,2(4):29-41

[4] Sun Junding, Xiao Wu. Image retrieval based on color distribution features [J], optoelectronics. Lasers, 2006,17 (8): 17-19

[5] Huang Yuanyuan, He Yunfeng. An image retrieval method based on color features [J] Chinese Journal of image and graphics, 2006,11 (12): 33-37