Optik 126 (2015) 2144–2148

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Content-based image retrieval technology using multi-feature fusion

Min Huang [a,b], Huazhong Shu [a], Yaqiong Ma [b,*], Qiuping Gong [b]

[a] Laboratory of Image Science and Technology, School of Computer Science and Engineering, Southeast University, Nanjing 210096, China
[b] School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou 450002, China

ARTICLE INFO

Article history:
Received 11 April 2014
Accepted 21 May 2015

Keywords:
Image retrieval
Multi-feature
Similarity
Normalization
Relevance feedback

ABSTRACT

Due to the diversity of the image content, different images have different focuses, image retrieval system based on single feature has a lower performance, and it cannot apply to all images, so an image retrieval method using multi-feature fusion is proposed. In this method, the color moment in RGB color space in combination with the color histogram in HSV color space is used for color feature extraction, the improved Zernike moments are used for shape feature extraction, and the gray level co-occurrence matrix is used for texture feature extraction, then combining these three features. Finally, respectively using color features, shape features, texture features as well as the fused features for image retrieval, the experimental results show that the image retrieval method based on multi-feature fusion has better retrieval performance.

1. Introduction

With the rapid development and popularization of digital technology, computer and network technology, people increasingly come into contact with a lot of image information, images have become a common carrier to describe and store the information. Traditional text-based image retrieval technology has been unable to satisfy people's needs; therefore, content-based image retrieval technology is getting more and more attention of people, it has become a hot research topic [1,2]. It uses the image's color, texture, shape and other basic features for retrieval, the early content-based retrieval system widely used the retrieval method based on single feature, but because of the image itself contains a wealth of information, there exists no image feature which can effectively describe and distinguish all kinds of images. In order to overcome the problems brought by using single feature and improve the retrieval accuracy, an image retrieval method which fuses color, texture and shape these three basic features is proposed, users can finally get satisfied query results according to the relevant feedback results.

2. Main content-based retrieval techniques

2.1. Retrieval based on color features

Color feature is the most intuitive and obvious feature of the image, it has certain stability, and shows a very strong robustness to the change of noise, image size, direction and resolution [3]. The method used in this paper is color moment in RGB color space in combination with 72bin color histogram in HSV color space. Color moment of the image can be extracted in different space; through experimental comparison, it can be found that the one in RGB color space has better retrieval effect. Color moment has the advantages of simplicity; moreover, it can be used to represent the distribution of each color in the image, color information of the image is mainly concentrated in the lower order moments, only using the first moment, second moment and third moment can express the color distribution of the image. Color moment features in RGB color space require nine characteristic components (each pixel has three color components: each color component has three low order moments), the expressions of three low moments are as follows:

$$\begin{cases} u_i = \dfrac{1}{M \times N} \displaystyle\sum_{i=1}^{M} \sum_{j=1}^{N} p_{ij} \\ \sigma_i = \left[\dfrac{1}{M \times N} \displaystyle\sum_{i=1}^{M} \sum_{j=1}^{N} (p_{ij} - u_i)^2 \right]^{1/2} \\ s_i = \left[\dfrac{1}{M \times N} \displaystyle\sum_{i=1}^{M} \sum_{j=1}^{N} (p_{ij} - u_i)^3 \right]^{1/3} \end{cases}$$

Because the color moment information is too simple, its single retrieval effect is not very satisfactory, the method of combining color moment with color histogram is adopted, color histogram

* Corresponding author. Tel.: +86 13783493389.
E-mail address: myq1988726@163.com (Y. Ma).

http://dx.doi.org/10.1016/j.ijleo.2015.05.095
0030-4026/

using the non-uniform quantization scheme of 8:3:3, the specific quantitative rules are as follows:

$$H = \begin{cases} 0, & h \in [315, 20] \\ 1, & h \in [20, 40] \\ 2, & h \in [40, 75] \\ 3, & h \in [75, 155] \\ 4, & h \in [155, 190] \\ 5, & h \in [190, 271] \\ 6, & h \in [271, 295] \\ 7, & h \in [295, 315] \end{cases} \qquad S, V = \begin{cases} 0, & s, v \in [0, 0.2] \\ 1, & s, v \in [0.2, 0.7] \\ 2, & s, v \in [0.7, 1] \end{cases}$$

Making $L = 9H + 3S + V$, the three components will be combined into a color vector with 72 color values [4]. When using this method to quantify the color value in quantitative critical edge, it will produce certain quantitative error, in order to minimize this error, introducing a kind of non-truncated quantitative method, the calculation formula of L can be rewritten as: $L = round(9H + 3S + V)$, in which *round* is the rounding function, and then carrying on normalization processing.

In the process of color retrieval, calculate the color moments and the 72bin color histogram feature values of the image to be retrieved, respectively find the similarity distance between them and the corresponding feature values of each picture in the picture library, and then calculate the total similarity value according to their weights.

2.2. Retrieval based on shape features

Shape is one of the basic characteristics of depicting the objects; using shape features for image retrieval can improve the efficiency and accuracy. Generally speaking, there are two kinds of representation methods for shape features: one is based on the contour, the other is based on the region; the typical representatives of these two methods are respectively the Fourier descriptors and invariant moments. In this paper, the Zernike moments are selected for shape feature extraction; they have good rotation invariance and simple calculation, at the same time they are widely used as a kind of shape descriptor.

2.2.1. Zernike moments

Zernike moments are a special kind of complex moments, they are orthogonal functions based on Zernike polynomials, Zernike polynomials are orthogonal in the unit circle, and their orthogonalities make Zernike moments independent, they have large superiority in characteristic expression ability [5]. The definition of Zernike orthogonal polynomials are as follows:

$$V_{nm}(x, y) = V_{nm}(\rho, \theta) = R_{nm}(\rho)e^{jm\theta}$$

where n and m are the orders of the orthogonal Zernike polynomials, n is a positive integer or zero, m is a positive or negative integer, they are subject to the conditions $n - |m| = even$ and $n \geq |m|$; ρ is the vector length between circle dot and the pixel (x, y), θ is the angle between vector ρ and the x-axis of counterclockwise direction [6]; $R_{nm}(\rho)$ is an orthogonal radial polynomial of real value, it is given by the following formula:

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s[(n-s)!]\rho^{n-2s}}{s!\left(\frac{n+|m|}{2} - s\right)!\left(\frac{n-|m|}{2} - s\right)!}$$

Zernike moments of the image refer to the projection of image function $f(x, y)$ on the orthogonal polynomial $\{V_{nm}(x, y)\}$, n order Zernike moment with the repetition of m is defined as:

$$Z_{nm} = \frac{n+1}{\pi} \iint_{x^2+y^2 \leq 1} f(x, y)V^*_{nm}(x, y)dxdy$$

Zernike moments in polar coordinates can be defined as:

$$Z_{nm} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 f(\rho, \theta)\rho d\rho d\theta$$

2.2.2. Rotation invariance of Zernike moments

Assuming that an image is rotated through angle α, by the definition of Zernike moments, the relationship between the Zernike moments of the rotated image $\hat{Z_{nm}}$ and the original one Z_{nm} can be deduced by the formula:

$$\hat{Z_{nm}} = Z_{nm}e(-jm\alpha)$$

It can be seen from the formula that the Zernike moments before and after the image rotation only have phase changes, their magnitudes remain unchanged, so the magnitude $|Z_{nm}|$ can be taken as the rotation invariant feature of the target [7].

2.2.3. Improved Zernike moments

In the discrete case, the rotation and scale transformation will cause resampling and requantization of digital images, it makes their invariance cannot remain strictly invariant. So the Zernike moments are improved in order to get better invariance, considering first to carry on shape normalization to the target area of the image, and then normalize the Zernike moments [8]. The concrete steps are as follows:

(1) Take the barycenter of the target as the center of the polar coordinates and the distance between the center and the outermost pixel in the target area as the radius, thus the pixels in the target area are resampled into the unit circle.
(2) Find the zeroth order geometric moment of the target, namely:

$$m_{00} = \sum\sum f(\rho, \theta)$$

(3) Calculate each order of the Zernike moments in the unit circles:

$$Z_{mn} = \frac{n+1}{\pi} \sum\sum f(\rho, \theta)V^*_{nm}(\rho, \theta)$$

(4) Normalize the Zernike moments by using m_{00}:

$$Z^*_{nm} = \frac{Z_{nm}}{m_{00}}$$

(5) Evaluate the magnitudes $|Z^*_{nm}|$ of Z^*_{nm}.

Two second moment modulus values $|Z^*_{20}|$ and $|Z^*_{22}|$ of the improved Zernike moments are selected as the feature values.

2.3. Retrieval based on texture features

Texture features are a kind of internal visual features which do not based on color or brightness, they reflect the homogeneity and contain the surface information as well as surrounding environment of the image, and spatial information of the image can be described quantitatively. Haralick et al. defined fourteen feature parameters of gray level co-occurrence matrixes (GLCM) for texture analysis [9], the study found that only four features are not