Applied Soft Computing 17 (2014) 127–139

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Rotation-invariant texture image retrieval using particle swarm optimization and support vector regression

Hung-Hsu Tsai [a], Bae-Muu Chang [b,*], Shin-Hung Liou [a]

[a] Department of Information Management, National Formosa University, Hu-Wei, Yun-Lin 632, Taiwan, ROC
[b] Department of Information Management, Chienkuo Technology University, Chang-Hua, Chang-Hua 500, Taiwan, ROC

ARTICLE INFO

Article history:
Received 18 November 2011
Received in revised form 9 November 2013
Accepted 8 December 2013
Available online 20 January 2014

Keywords:
Content-based image retrieval
Log-polar mapping
Fast Fourier transform
Zernike moment
Particle swarm optimization
Support vector regression

ABSTRACT

This paper presents a novel rotation-invariant texture image retrieval using particle swarm optimization (PSO) and support vector regression (SVR), which is called the RTIRPS method. It respectively employs log-polar mapping (LPM) combined with fast Fourier transformation (FFT), Gabor filter, and Zernike moment to extract three kinds of rotation-invariant features from gray-level images. Subsequently, the PSO algorithm is utilized to optimize the RTIRPS method. Experimental results demonstrate that the RTIRPS method can achieve satisfying results and outperform the existing well-known rotation-invariant image retrieval methods under considerations here. Also, in order to reduce calculation complexity for image feature matching, the RTIRPS method employs the SVR to construct an efficient scheme for the image retrieval.

1. Introduction

With the advancement in information technology, people can readily take pictures using digital cameras, camcorders, web camera videos and smart mobile phones anytime and anywhere. This phenomenon makes image databases get an explosive growth. Thus, users need to employ efficient and effective tools to search for the images they want in the huge image database [7]. Current popular websites, such as Google and Yahoo, provide the function for image searching based on the text annotations of images. As a result, users often spend a lot of time and effort to edit these text annotations for each image. However, in many cases, text annotations cannot clearly describe the image contents. Moreover, a query request is difficult to be precisely described by text annotations. Therefore, the content-based image retrieval (CBIR) technique has been proposed to overcome these limitations mentioned above. The index in every image in the CBIR is popularly formed by the composition of image's own visual contents [16]. Subsequently, researchers improve the CBIR with rotation-invariant. At present, rotation-invariant texture image retrieval becomes important issue in image retrieval.

The texture descriptor of images plays an important role in computer vision and image pattern recognition, especially in representing the contents of an image [4,18]. Several image retrieval systems, which extract the rotation-invariant texture features of images, have been developed recently [5,7,8,11,19]. The feature extractions of several proposed methods have been devised in frequency domain for extracting rotation-invariant features [5,7,11,16]. In Ref. [7], Kokare combines a dual-tree rotated complex wavelet filter (DT-RCWF) and a dual-tree complex wavelet transform (DT-CWT) to obtain the texture features for rotation-invariant from 12 different angles. However, the similarity measurement formula is not optimized. Thus, the method cannot get better discrimination between two different images [7]. In [16], Tzagkarakis presents the kullback-leibler-divergence (KLD) method which employs the Gaussianized steerable pyramids to extract the texture features of images. Nevertheless, it is insufficient to search for the optimal number of outputted images. In Ref. [5], the rotation-invariant Gabor (RIG) method is proposed which combines the Gabor filters with same scales and different angles to extract the rotation-invariant texture features of images [5]. In Ref. [11], Rallabandi presents wavelet-based hidden Markov trees (WBHMT) which combines the wavelet transformation and the hidden Markov tree to extract the rotation-invariant texture features of images. Unfortunately, the feature extraction algorithm of the method requires high computational complexity. In Ref. [12], Sim presents the modified Zernike moments (MZM) which combine the discrete Fourier transformation and the Zernike

* Corresponding author. Tel.: +886 937 749299; fax: +886 4 711 1142.
E-mail addresses: thh@nfu.edu.tw (H.-H. Tsai), bmchang@cc.ctu.edu.tw (B.-M. Chang), afflatus_zz@hotmail.com (S.-H. Liou).

1568-4946/$ – see front matter
http://dx.doi.org/10.1016/j.asoc.2013.12.003

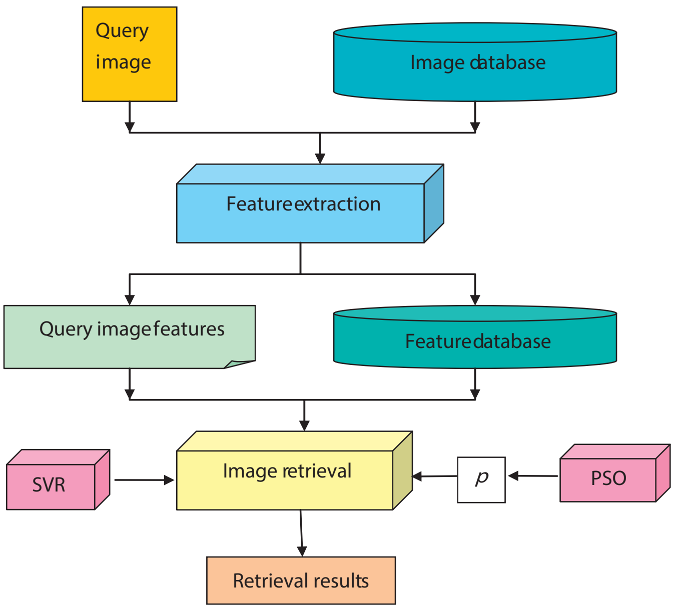

Fig. 1. The conceptual structure of the RTIRPS method.

moments to construct the rotation-invariant texture features of images. Nevertheless, the similarity measurement formula is not optimized. Hence, the method cannot get nearly optimal solutions in the number of outputted images, various feature weights, etc.

Aforementioned phenomena motivate us to develop a novel rotation-invariant texture image retrieval method, called the RTIRPS method which can overcome the drawbacks of the above methods [5,7,11,12,16]. The RTIRPS method employs the PSO algorithm which searches for a set of nearly optimal thresholds and enlarging constants in order to improve the retrieval performance. Fig. 1 depicts the conceptual design for the RTIRPS method. Note that the RTIRPS method has following properties: (1) simultaneously employing three rotation-invariant features to design dis-similarity measurement algorithm, (2) optimizing the dis-similarity measurement algorithm by using the PSO algorithm, (3) searching for an optimal number of outputted images, (4) reducing the computational complexity of feature matching by employing the SVR.

The remainder of this paper is arranged as follows. Section 2 introduces the background knowledge including the PSO algorithm, the SVR, the LPM, the Gabor filter, and the Zernike moment. Subsequently, the RTIRPS method is described in Section 3. Section 4 presents the experimental results. Finally, conclusions are given in Section 5.

2. Background knowledge

2.1. The PSO algorithm

The PSO algorithm was proposed by Kennedy in 1995. It is employed to solve the optimization problems in many applications, such as, image processing and vehicle routing problem [3,10]. In the PSO algorithm, birds decide the next direction and the distance of movement by referencing the past directions of movement and the current location when they search for foods. A particle in the algorithm is simulated as a biological individual in a hyper-dimensional search space to find out nearly optimal solution of more complex problems [2,6]. More specifically, the training procedure for the PSO algorithm is briefly described as follows:

Step 1. Initialize the beginning parameters: swarm populations, weights, ranges of coordinates of locations for particles, and the number of training iterations. Also, the particles are randomly located and the movement vector is randomly assigned.
Step 2. Store Gbest and all Pbest locations at the current iteration by using an evaluation process employing the fitness function for all particles.
Step 3. If the number of training iterations is terminated or the accuracy is satisfied, then output Gbest and Pbest locations, and the algorithm terminates. Otherwise, go to Step 4.
Step 4. Calculate the movement vectors in Eq. (1) for all particles.
Step 5. Modify the locations of all particles utilizing Eq. (2) and then go to Step 2.

The movement vector (location) is specified as follows:

$$\mathbf{V}_i(t+1) = w\mathbf{V}_i(t) + c_1 \times r_1 \times (\mathrm{Pbest}_i - \mathbf{X}_i(t)) + c_2 \times r_2 \times (\mathrm{Gbest} - \mathbf{X}_i(t)), \quad (1)$$

where $\mathbf{V}_i = (V_{i1}, V_{i2}, \cdots, V_{im}) \in \Re^m$, $\mathbf{V}_i(t+1)$ represents the movement vector of particle i at the $(t+1)$th iteration, w indicates the inertia weight, c_1 and c_2 denote the acceleration coefficients which are random numbers in [0,1], r_1 and r_2 are also two randomly generated values in [0,1]. Moreover, in Eq. (1), the first term w$\mathbf{V}_i(t)$ denotes the particle's inertia, the second term $c_1 \times r_1 \times (\mathrm{Pbest}_i - \mathbf{X}_i(t))$ indicates the particle's cognition-only model, and the third term $c_2 \times r_2 \times (\mathrm{Gbest} - \mathbf{X}_i(t))$ stands for the particle's social-only model. The location of particle i is modified by Eq. (2).

$$\mathbf{X}_i(t+1) = \mathbf{X}_i(t) + \mathbf{V}_i(t+1) \quad (2)$$

where $\mathbf{X}_i(t+1)$ represents the location of particle i at the $(t+1)$th iteration, $\mathbf{V}_i(t+1)$ denotes the movement vector of particle i at the $(t+1)$th iteration. Hence, the new location of particle i is to add its current location vector to its movement vector.

2.2. A review of SVR

The SVR is a learning algorithm which is based on the statistical learning theory and the structural risk minimization principle [14]. Because the SVR possesses the feature of the global optimal solution and considers the structured risk, it has famously been utilized in the data mining and the pattern recognition [14].

Suppose that a set of training samples be represented as $\{(\mathbf{x}_k, y_k)\}_{k=1}^N$, $\mathbf{x}_k = (x_{k_1}, x_{k_2}, \ldots, x_{k_n})^t \in \Re^n$, $y_k \in \Re$, where $\mathbf{x}_k$ and y_k respectively indicate the input vector and its corresponding desired output of the kth training sample and t stands for the transpose operation, $\Re$ expresses the set of real numbers. The linear regression function, f, is described in Eq. (3).

$$f(\mathbf{x}) = (\mathbf{w}^t\mathbf{x} + b) \quad (3)$$

where $\mathbf{w} = (w_1, w_2, \ldots, w_n)^t \in \Re^n$ and denotes the weight vectors, and $b \in \Re$ indicates the bias quantity, respectively [14]. Here, the linear SVR can be optimized and the optimized problem can be rewritten in Eq. (4).

$$\text{Minimize } \frac{1}{2}\mathbf{w}^t\mathbf{w}$$
$$\text{Subject to } \begin{cases} y_k - \mathbf{w}^t\mathbf{x}_k - b \le \varepsilon \\ \mathbf{w}^t\mathbf{x}_k + b - y_k \le \varepsilon \end{cases}, \quad k = 1, 2, \ldots, N \quad (4)$$

where ε is nonnegative which denotes the maximum error between the tolerant prediction output and the desired output. Because errors and noises may exist in the sample data in the most applications, the extra items have to be included in the application and