Expert Systems with Applications 38 (2011) 11412–11420

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Fast color-spatial feature based image retrieval methods

Chuen-Horng Lin [a], Der-Chen Huang [c], Yung-Kuan Chan [b,*], Kai-Hung Chen [c], Yen-Jen Chang [c]

[a] Department of Information Science, National Taichung Institute of Technology, Taiwan, ROC
[b] Department of Management Information Systems, National Chung Hsing University, Taiwan, ROC
[c] Department of Computer Science, National Chung Hsing University, Taiwan, ROC

ARTICLE INFO

Keywords:
Color-based image retrieval
Color-histogram
K-means
CBIR

ABSTRACT

In this paper, three types of image features are proposed to describe the color and spatial distributions of an image. In these features, the *K*-means algorithm is adopted to classify all of the pixels in an image into several clusters according to their colors. By measuring the spatial distance among the pixels in a same cluster, the three types of color spatial distribution (CSD) features of the image is obtained. Based on the three types of CSD features, three image retrieval methods are also provided. To accelerate the image retrieval methods, a fast filter is also presented to eliminate most undesired images in advance. A genetic algorithm is also given to decide the most suitable parameters which are used in the proposed image retrieval methods. The proposed image retrieval methods are simple. Moreover, the experiments show that the proposed methods can provide impressive results as well.

1. Introduction

Content based image retrieval (CBIR) has received much attention due to the great demand for multimedia information systems (Flickner, Sawhney, Niblack, & Ashley, 1995; Pentland, Picard, & Scarloff, 1994; Smith et al., 1996). Efficiently indexing images in database is an essential issue in image applications. Traditional image retrieval systems (Rui & Huang, 1999; Stricker et al., 1995) use textual features, such as captions, filenames, keywords to annotate and index images. When applied to a huge image database, however, textual information fails to capture sufficient image content. Low-level features, such as textures (Huang & Dai, 2006; Jhanwar, Chaudhuri, Seetharaman, & Zavidovique, 2004), colors (Qiu, 2004), and shapes (Ko & Byun, 2005) are extensively used to index image features in the CBIR system.

Color attribute analysis has been proven to be very successful in retrieving images, such as color histograms (Swain & Ballard, 1991). However, since it ignores the spatial distribution of pixel colors on an image, the retrieved results often make little sense. Consider the two quite different images in Fig. 1, which produce the same color histogram. A retrieval method that relies completely on color histograms will consider them to be the same image. Motivated by this shortcoming, in this paper, three types of color spatial distribution (CSD) features are presented to characterize the color and spatial distributions of all the pixels in an image. Based on these three types of CSD features, this paper still proposes three image retrieval methods; we call them the CSD-based image retrieval methods. To speed up these image retrieval methods, this paper also provides a filter fast to eliminate most undesired images in advance during retrieving. It provides a genetic algorithm to decide the most suitable the parameters taken by the CSD-based image retrieval methods.

2. Related works

This section will briefly introduce some techniques that will be adopted by the CSD-based image retrieval methods or will be compared with the CSD-based image retrieval method in performances.

2.1. Genetic algorithm

A genetic algorithm (GA) (Man, Tang, & Kwong, 1999) is a heuristic optimization method that operates through a determined and randomized search. The set of possible solutions for an optimization problem is considered as a population of individuals, and the degree of adaptation of an individual to its environment is specified by its fitness. A chromosome, essentially a set of character strings, represents the coordinate of an individual in the search space. A gene is a subsection of a chromosome that encodes the value of a single parameter being optimized. Typical encoding for a gene could be binary or integer.

Derived from the biological model of evolution, a genetic algorithm operates on the Darwinian principle of natural selection,

* Corresponding author. Address: Department of Management Information Systems, National Chung Hsing University, No. 250, Kuokuang Rd., Taichung, Taiwan, ROC. Tel.: +886 4 2284 0864x755; fax: +886 4 22857173.

E-mail addresses: linch@ntit.edu.tw (C.-H. Lin), huangdc@cs.nchu.edu.tw (D.-C. Huang), ykchan@nchu.edu.tw (Y.-K. Chan), fzworld@gmail.com (K.-H. Chen), ychang@cs.nchu.edu.tw (Y.-J. Chang).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.03.014

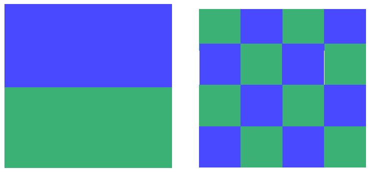

Fig. 1. Two quite different images with the same color histograms.

which holds that, given a certain population, only the individuals that adapt well to their environment will be more likely to survive and transmit their characteristics to their descendants. A genetic algorithm consists of three major operations: selection, crossover, and mutation. The selection operation evaluates all individuals, and only those best fitted to their environment survive. The crossover operation recombines the genetic material of two individuals to produce new combinations with the potential for better performance. The mutation operation induces changes in a small number of chromosomal units with the goal of maintaining sufficient population diversity during the optimization process.

2.2. ANMRR

This paper will use the MPEG-7 retrieval metric NMRR (normalized modified retrieval rank) (Stehling, Falcao, & Nascimento, 2001) to describe the effectiveness of an image retrieval method. NMRR indicates not only how many of the correct items are retrieved, but also how highly they are ranked among the retrieved items. NMRR is defined by

$$NMRR(q) = \frac{\left(\sum_{k=1}^{NG(q)} \frac{Rank(k)}{NG(q)}\right) - 0.5 - \frac{NG(q)}{2}}{M(q) + 0.5 - 0.5 \times NG(q)},$$

where $NG(q)$ is the size of the ground truth set for a query image q, $Rank(k)$ is the ranking of the ground truth images by the retrieval algorithm and $M(q)$ specifies the "relevance rank" for each query. As the size of the ground truth set is normally unequal, a suitable $M(q)$ is determined by

$$M(q) = \min(4 \times NG(q), 2 \times GTM),$$

where GTM is the maximum of $NG(q)$ for all queries. If $Rank(k) > M(q)$, then $Rank(k)$ is changed into $M(q) + 1$. The $NMRR$ is in the range of $[0,1]$, and the smaller $NMRR$ is, the better the retrieval performance will be. $ANMRR$ is defined as the average $NMRR$ over a range of queries, and is given by

$$ANMRR = \frac{1}{NQ} \sum_{q=1}^{NQ} NMRR(q),$$

where NQ is the number of query images.

2.3. Texture similarity based image retrieval method

Huang and Dai (2003) proposed a texture based image retrieval method. The method associates a coarse and a fine feature descriptors with each image. Both descriptors are derived from the coefficients of wavelet transform of the original image. The coarse feature descriptor is used at the first stage to quickly screen out non-promising images; the fine feature descriptor is subsequently employed to find the truly matched images.

Discrete wavelet transform (DWT) can be implemented by a low-pass filter (L) and a high-pass filter (H) to decompose an image into four different frequency bands; the lowest frequency band is then split in the same way at half the rate of the previous frequency. First, Huang and Dai divides an original image into four bands (LL, HL, LH and HH) by one-level DWT. Second, they construct the gradient vectors of the four frequency bands as the feature vectors of the image. These bands are named SVG1, SVG2, SVG3 and SVG4, respectively. Then, Huang and Dai apply a filter called energy distribution pattern string (EDP-string) to quickly prune off most of non-promising images. The EDR-string is computed based on the pyramidal wavelet decomposition. A l-level wavelet decomposition for a image texture can be represented as $\{C_j, D_{j,k}\}_{j=1,2,\ldots,l,k=1,2,3}$. Generally, after l-level wavelet decomposion, sub-image C_j becomes an indication of luminance for most textured images. Thus, sub-images $\{D_{j,k}\}_{j=1,2,\ldots,l,k=1,2,3}$ are considered only for the purpose of quick filtering.

2.4. Motif co-occurrence matrix based image retrieval method

Jhanwar et al. (2004) proposed an image retrieval method using a motif co-occurrence matrix (MCM). The MCM is derived from the motif transformed image. The whole image is divided into 2×2 pixel grids. Each grid is replaced by a scan motif that minimizes the local gradient while traversing the 2×2 grid to form a motif transformed image. The MCM is then defined as a 3D matrix MCMeach of whose entries $MCM(i,j,k)$ denotes the probability of finding a motif i at a distance k from the motif j in the transformed image. Conceptually, the MCM is quite similar to the color co-occurrence matrix (CCM) (Jhanwar et al., 2004); however, the performance of MCM is better than that of CCM since it captures the third-order image statistics in the local neighbouhood.

2.5. SamMatch-based image retrieval method

Vu, Hua, and Tavanapong (2003) proposed a SamMatch-based image retrieval method to process the query using a sampling-based matching approach called SamMatch (Hua, Vu, & Oh, 1999). This method resizes each database image into 256×256 pixels, quantifies it into 256 colors, and considers the average color of the pixels in each 16×16 block to be the color of the block. Consider two arbitrarily-shaped subimages Q and S, each represented by $n$sampled blocks. Their matching score is defined as follows:

$$Sim(Q, S) = \sum_{i=1}^{n} \frac{w_i}{1 + D(c_i^S, c_i^Q)},$$

where $D(c_i^S, c_i^Q)$ is the distance of c_i^S and c_i^Q. c_i^S is the color of block i of S and c_i^Q is that of Q. The parameter w_i is a weight factor. Since SamMatch compares the corresponding sampled blocks of subimages, it implicitly involves the shape, size, and texture features of the image objects. As a result, SamMatch has the benefits of region-based techniques without reliance on the highly variable accuracy of segmentation methods. However, this occurs by having the user point out the object areas at the time of the query.

3. CSD features

In this paper, three types of color spatial distribution (CSD) features, respectively denoted by CSD1, CSD2, and CSD3, are provided. The CSD features can depict the spatial distribution of the pixels with similar colors in an image. This section will introduce these three types of CSD features in detail.

In a full color image, a pixel color is generally described by a 24-bit memory space, so there are a total of 2^{24} possible pixel colors. Before extracting the CSD features of an image, the pixels of all database images are first categorized into K clusters by the K-means algorithm (Su & Chou, 2001) according to their colors; then the average color (R_i, G_i, B_i) of all the pixels in cluster i are computed by the following formula: