Neurocomputing 171 (2016) 673–684

Contents lists available at ScienceDirect

Neurocomputing

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

Image retrieval using spatiograms of colors quantized by Gaussian Mixture Models

Shan Zeng [a], Rui Huang [b,*], Haibing Wang [c], Zhen Kang [a]

[a] College of Mathematics and Computer Science, Wuhan Polytechnic University, Wuhan, Hubei 430023, China
[b] NEC Laboratories China, Beijing 100084, China
[c] College of Food Science and Engineering, Wuhan Polytechnic University, Wuhan, Hubei 430023, China

ARTICLE INFO

Article history:
Received 7 March 2015
Received in revised form
17 May 2015
Accepted 3 July 2015
Available online 15 July 2015

Keywords:
Image retrieval
Color quantization
Spatiogram
Gauss Mixture Model
Jensen–Shannon Divergence

ABSTRACT

In this paper we propose a novel image representation method that characterizes an image as a spatiogram–a generalized histogram–of colors quantized by Gaussian Mixture Models (GMMs). First, we quantize the color space using a GMM, which is learned by the Expectation-Maximization (EM) algorithm from the training images. The number of Gaussian components (i.e., the number of quantized color bins) is determined automatically according to the Bayesian Information Criterion (BIC). Second, we incorporate the spatiogram representation with the quantized Gaussian mixture color model. Intuitively, a spatiogram is a histogram in which the distribution of colors is spatially weighted by the locations of the pixels contributing to each color bin. We have modified the spatiogram representation to fit our framework, which employs Gaussian color components instead of discrete color bins. Finally, the comparison between two images is achieved by measuring the similarity between two spatiograms, for which purpose we propose a new measurement adopting the Jensen–Shannon Divergence (JSD). We applied the new image representation and comparison method to the image retrieval task. The experiments on several publicly available COREL image datasets demonstrate the effectiveness of our proposed image representation for image retrieval.

1. Introduction

Color histograms have been widely used to capture image properties for various computer vision problems. Because color histograms are relatively invariant to changes in image size and orientation, they have been used extensively for the image retrieval task. However, there are still some open problems with the color histogram based image representations. The first problem is the high dimensionality of color histograms in general. For instance, the histogram in the RGB color space with 256 bins in each channel has 2^{24} dimensions. Another obvious problem with color histograms is that they do not consider any spatial distribution of the colors in an image, possibly ignoring the shape and texture of the objects depicted in the image.

To address the high dimensionality problem, one can effectively reduce the dimensionality of the histogram by quantizing the color space into fewer distinct bins. Color quantization methods have been studied for decades. *Image-independent* methods that use a fixed palette are simple and fast but often give poor approximation because they do not adapt well to specific image contents. One popular image-independent color quantization method is to divide the range of values in each color channel into uniformly spaced levels [1]. Different color channels may have different numbers of bins, though. For example, in [2], the HSV color space is quantized into 3 Saturation and Value levels respectively and 8 Hue levels, because the Hue component is considered more important for some specific tasks. The performance of image-independent quantization applied to image retrieval has been reported in [3], with different quantization schemes such as RGB ($8 \times 8 \times 8$), Lab ($4 \times 8 \times 8$), Lu*v* ($4 \times 8 \times 8$). Because these image-independent methods do not take into account the actual colors appearing in the specific images one is dealing with, they usually give relatively poor image retrieval results.

Image-dependent methods, on the other hand, build a custom palette based on the color distribution of the given images, taking into account the image contents. Most such techniques treat color quantization as a problem of clustering three-dimensional points in the color space, where the points represent colors found in the images and the three dimensions represent the three color channels. Hence almost any clustering algorithm can be applied to color quantization. *Pre-clustering* algorithms such as median-cut [4] and octree [5] compute the palette only once while *post-clustering* algorithms such as k-means [6] and fuzzy c-means [7]

* Corresponding author.
E-mail address: hrzs2015@163.com (R. Huang).

http://dx.doi.org/10.1016/j.neucom.2015.07.008
0925-2312/

make an initial guess at the palette and then iteratively refines it. Post-clustering methods can obtain better color quantization results, although at the expense of increased computational cost, compared to pre-clustering methods. Therefore, most recent studies focus on achieving a better balance between computational efficiency and visual quality of the quantization results. We propose a color quantization method based on Gaussian Mixture Models (GMMs), learned by the Expectation-Maximization (EM) algorithm. The number of color bins, i.e., Gaussian components, is automatically determined according to the Bayesian Information Criterion (BIC). The GMM based color model can be seen as a probabilistic version of the above mentioned k-means algorithm for color quantization, and it allows the soft assignment of the pixels to multiple color bins.

To overcome color histograms' lack of spatial information, some researchers have proposed to work on the higher level of image representation without adapting color histogram itself. The simplest way is to divide an image into small regions and concatenate the color histograms computed in all the regions into one feature vector. The spatial pyramid matching method [8] further partitions an image into increasingly finer spatial sub-regions and computes histograms from each of the sub-region, and the resulting "spatial pyramid" has shown excellent performance on the image classification and object detection tasks. Other researchers have tried to incorporate spatial information into color histograms more directly. Gauss Mixture Vector Quantization (GMVQ) [9], as opposed to all the previously mentioned *scalar quantization* methods, provides a better way of exploiting the spatial characteristics of the images. This technique matches Gaussian models to observed vectors by matching the model covariance and mean to the observed sample covariance and mean, thereby incorporating the second-order moment structure into the matching and better capturing spatial relationships among pixel intensities. Birchfield et al. [10] proposed the concept of spatial histograms, or *spatiograms*, as a generalization to histograms to include potentially higher-order spatial moment information, which describes not only the occurrence of pixel values but also the mean and covariance of the pixel coordinates in each histogram bin. The potential image comparison capabilities of the spatiogram have been validated in the field of object tracking [10] and video retrieval [11], and have attracted increasing research attentions in recent years.

To this end, we propose a new image representation method to address both of the above mentioned limitations of color histograms by combining the GMM based color quantization with spatiograms. We first quantize the color space using a GMM, which is learned by the Expectation-Maximization (EM) algorithm from the training images. Unlike the other Gaussian model based algorithms [9,10], in which the number of Gaussian components is pre-defined, we use the Bayesian Information Criterion (BIC) [12] based EM algorithm, namely EM-BIC, to automatically determine the number of quantized color bins. Second, based on the quantized color bins, we compute the spatiogram of the image. Intuitively, the spatiogram is a histogram in which the distribution of colors is spatially weighted by the locations of the pixels contributing to each color bin. A typical second-order spatiogram of an image includes not only the numbers of pixels belonging to individual color bins (i.e., the frequency component), but also the mean vectors and covariance matrices of the coordinates of the pixels associated with different bins (i.e., the spatial component). Since our GMM based color quantization allows each pixel to belong to multiple color bins with different probabilities (summing up to 1), our modified spatiogram includes the same statistics as the original spatiogram does, but the computation of such statistics is also weighted by the probabilities of each pixel belonging to different color bins. This *soft assignment* strategy leads to more accurate characterization of the image.

Besides the computation of spatiograms, the comparison between two spatiograms is also an important component of our image representation, especially for the image retrieval task. In [10] the similarity between two spatiograms is computed as the similarity between two traditional histograms (i.e., the frequency component), but each bin is weighted by the similarity between the spatial distributions of the pixels in that bin from the two spatiograms (i.e., the spatial component). Conaire et al. [13] proposed an improvement over the original spatiogram similarity measure by converting the 2nd order spatiogram back to a histogram with an extra dimension of space and deriving the Bhattacharyya coefficient for an infinite number of spatial feature bins. However, it has been shown that the Bhattacharyya coefficient based measure has fairly poor discriminative power [14]. On the other hand, the Kullback–Leibler (KL) Divergence, well-motivated and widely used in information theory, has shown promising performance in probabilistic retrieval methods [15]. Ulges et al. [11] treat the spatiograms as joint distributions of attributes (i.e., the frequency component) and locations (i.e., the spatial component), and propose a distance measure that computes the divergence between two joint distributions based on the Jensen–Shannon (JS) Divergence, a symmetrized and smoothed version of the KL divergence. We follow similarity measure scheme in [10] for its simplicity, but also adopt the JS divergence to more accurately estimate the similarity between the spatial distributions.

In summary, the main contributions of this paper are as follows: First, we improve the standard GMM based color quantization method by automatically determining the number of color bins using the EM-BIC algorithm. Our model allows the soft assignment of the pixels to multiple color bins. Second, we compute spatiograms of the images in the reduced color space using a revised formulation to accommodate to the soft assignment of the pixels. Finally, we proposed a new distance measure adopting the JS divergence to compare two spatiograms. This way we avoid both the dimensionality problem and the lack of spatial information problem of the traditional color histograms, and the soft assignment of the pixels leads to a more accurate description of the image. Section 2 explains the entire framework in detail. In Section 3 we apply the proposed method to the image retrieval task, and the effectiveness is demonstrated on several publicly available image datasets, as shown in Section 4. We conclude the paper in Section 5 with some discussions of the future work.

2. Proposed method

In this section, we introduce our proposed image representation method, including the following three parts: (1) the GMM based color quantization method, (2) the spatiograms defined in the reduced color space with the soft assignment of the image pixels and (3) the modified JS divergence method to compare two above mentioned spatiograms.

2.1. Quantizing the color space using GMMs

The basic idea of GMM based color quantization is that all the colors in a certain color space can be approximated by a mixture of M Gaussian distributions. In image-dependent color quantization, we need to find a GMM to fit the characteristics of the colors in certain images. Given a set of colors $\mathbf{c} = (\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_N)$, assuming the color space is 3-dimensional, e.g., $\mathbf{c}_l = (h_l, s_l, v_l)^T$, the probability of $\mathbf{c}_l$ originating from the k^{th} $(k = 1, 2, \cdots, M)$ Gaussian