Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition Letters 28 (2007) 974–983
Pattern Recognition Letters
www.elsevier.com/locate/patrec

An Integrated Color and Intensity Co-occurrence Matrix

A. Vadivel [a], Shamik Sural [b,*], A.K. Majumdar [a]

[a] Department of Computer Science and Engineering, Indian Institute of Technology, Kharagpur 721302, India
[b] School of Information Technology, Indian Institute of Technology, Kharagpur 721302, India

Received 29 March 2005; received in revised form 21 September 2006
Available online 23 January 2007

Communicated by R. Manmatha

Abstract

The paper presents a novel approach for representing color and intensity of pixel neighborhoods in an image using a co-occurrence matrix. After analyzing the properties of the HSV color space, suitable weight functions have been suggested for estimating relative contribution of color and gray levels of an image pixel. The suggested weight values for a pixel and its neighbor are used to construct an Integrated Color and Intensity Co-occurrence Matrix (ICICM). We have shown that if the ICICM matrix is used as a feature in an image retrieval application, it is possible to have higher recall and precision compared to other existing methods.

Keywords: Co-occurrence matrix; HSV color space; ICICM; Image retrieval

1. Introduction

Color and texture are two low-level features widely used for image classification, indexing and retrieval. Color is usually represented as a histogram, which is a first order statistical measure that captures global distribution of color in an image (Swain and Ballard, 1991; Gevers and Stokman, 2004). One of the main drawbacks of the histogram-based approaches is that the spatial distribution and local variations in color are ignored. Local spatial variation of pixel intensity is commonly used to capture texture information in an image. Grayscale Co-occurrence Matrix (GCM) is a well-known method for texture extraction in the spatial domain (Haralick et al., 1973). A GCM stores the number of pixel neighborhoods in an image that have a particular grayscale combination. Let I be an image and let p and N_p respectively denote any arbitrary pixel and its neighbor in a given direction. If GL denotes the total number of quantized gray levels and gl denotes the individual gray levels, where, $gl \in \{0, \ldots, GL-1\}$, then each component of GCM can be written as follows:

$$\mathrm{gcm}(i,j) = \Pr((\mathrm{gl}_p, \mathrm{gl}_{N_p}) = (i,j)) \tag{1}$$

gcm(i,j) is the number of times the gray level of a pixel p denoted by gl_p equals i, and the gray level of its neighbor N_p denoted by gl_{N_p} equals j, as a fraction of the total number of pixels in the image. Thus, it estimates the probability that the gray level of an arbitrary pixel in an image is i, and that of its neighbor is j. One GCM matrix is generated for each possible neighborhood direction, namely, 0°, 45°, 90° and 135°. Average and range of 14 features like Angular Second Moment, Contrast, Correlation, etc., are generated by combining all the four matrices to get a total of 28 features (Haralick et al., 1973). In the GCM approach for texture extraction, color information is completely lost since only pixel gray levels are considered.

To incorporate spatial information along with the color of image pixels, a feature called color correlogram has recently been proposed. It is a three dimensional matrix that represents the probability of finding pixels of any two given colors at a distance '*d*' apart (Huang et al.,

* Corresponding author. Tel.: +91 3222 282330; fax: +91 3222 282206.
E-mail addresses: vadi@cc.iitkgp.ernet.in (A. Vadivel), shamik@cse.iitkgp.ernet.in (S. Sural), akmj@cse.iitkgp.ernet.in (A.K. Majumdar).

0167-8655/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2007.01.004

1997). Auto correlogram is a variation of correlogram, which represents the probability of finding two pixels with the same color at a distance '*d*' apart. This approach can effectively represent color distribution in an image. However, correlogram features do not capture intensity variation. Many image databases often contain both color as well as gray scale images. The color correlogram method does not constitute a good descriptor in such databases.

Another method called Color Co-occurrence Matrix (CCM) has been proposed to capture color variation in an image (Shim and Choi, 2003). CCM is represented as a three-dimensional matrix, where color pair of the pixels p and N_p are captured in the first two dimensions of the matrix and the spatial distance '*d*' between these two pixels is captured in the third dimension. This approach is a generalization of the color correlogram and reduces to the pure color correlogram for $d = 1$. CCM is generated using only the Hue plane of the HSV (Hue, Saturation and Intensity Value) color space. The Hue axis is quantized into HL number of levels. If individual hue values are denoted by hl, where $\mathrm{hl} \in \{0, \ldots, \mathrm{HL} - 1\}$, then each component of CCM can be written as follows:

$$\mathrm{ccm}(i,j) = \Pr((\mathrm{hl}_p, \mathrm{hl}_{N_p}) = (i,j)) \quad (2)$$

Four matrices representing neighbors at angles 0°, 90°, 180° and 270° are considered. This approach was further extended by separating the diagonal and the non-diagonal components of CCM to generate a Modified Color Co-occurrence Matrix (MCCM). MCCM, thus, may be written as follows:

$$\mathrm{MCCM} = (\mathrm{CCM_D}, \mathrm{CCM_{ND}}) \quad (3)$$

Here, $\mathrm{CCM_D}$ and $\mathrm{CCM_{ND}}$ correspond to the diagonal and off-diagonal components of CCM. The main drawback of this approach is that, like correlogram, it also captures only color information and intensity information is completely ignored.

An alternative approach is to capture intensity variation as a texture feature from an image and combine it with color features like histograms using suitable weights (Manjunath et al., 2001). One of the challenges of this approach is to determine suitable weights since these are highly application-dependent. In certain applications like Content-based Image Retrieval (CBIR), weights are often estimated from relevance feedback given by users (Aksoy and Haralick, 2000; Wu and Zhang, 2002). While relevance feedback is sometimes effective, it makes the process of image retrieval user-dependent and iterative. There is also no guarantee on the convergence of the weight-learning algorithms. In order to overcome these problems, researchers have tried to combine color and texture features together during extraction.

Palm (2004) proposed two approaches for capturing color and intensity variations from an image using the LUV color space. In the Single-channel Co-occurrence Matrix (SCM), variations for each color channel, namely, L, U and V are considered independently. In the Multi-channel Co-occurrence Matrix (MCM), variations are captured taking two channels at a time – UV, LU and LV. Since the LUV color space separates out chrominance (L and U) from luminance (V), SCM in effect, generates one GCM and two CCMs from each image independently. As a result, correlation between the color channels is lost. However, in MCM, the count of pairwise occurrences of the values of different channels of the color space is captured. Thus, each component of MCM can be written as follows:

$$\mathrm{mcm_{UV}}(i,j) = \Pr((u_p, v_{N_p}) = (i,j)) \quad (4a)$$

$$\mathrm{mcm_{LU}}(i,j) = \Pr((l_p, u_{N_p}) = (i,j)) \quad (4b)$$

$$\mathrm{mcm_{LV}}(i,j) = \Pr((l_p, v_{N_p}) = (i,j)) \quad (4c)$$

Here, $\mathrm{mcm_{UV}}(i,j)$ is the number of times the U chromaticity value of a pixel p denoted by u_p equals *i*, and the V chromaticity value of its neighbor N_p denoted by v_{N_p} equals *j*, as a fraction of the total number of pixels in the image. Similarly, $\mathrm{mcm_{LU}}(i,j)$ and $\mathrm{mcm_{LV}}(i,j)$ are defined. One MCM matrix is generated for each of the four neighborhood directions, namely, 0°, 45°, 90° and 135°.

Deng and Manjunath (2001) proposed a two-stage method called JSEG, which combines color and texture after image segmentation. In the first stage, colors are quantized to the required levels for differentiating between various regions of an image. Pixel values of the regions are then replaced by their quantized color levels to form a color map. Spatial variation of color levels between different regions in the map is viewed as a type of texture composition of the image. Yu et al. (2002) suggested the use of color texture moments to represent both color and texture of an image. This approach is based on the calculation of Local Fourier Transformation (LFT) coefficients. Eight templates equivalent to LFT are operated over an image to generate a characteristic map of the image. Each template is a 3×3 filter that considers eight neighbors of the current pixel for LFT calculation. First and second order moments of the characteristic map are then used to generate a set of features.

In this paper, we propose an integrated approach for capturing spatial variation of both color and intensity levels in the neighborhood of each pixel using the HSV color space. In contrast to the other methods, for each pixel and its neighbor, the amount of color and intensity variation between them is estimated using a weight function. Suitable constraints are satisfied while choosing the weight function for effectively relating visual perception of color and the HSV color space properties. The color and intensity variations are represented in a single composite feature known as Integrated Color and Intensity Co-occurrence Matrix (ICICM). While the existing schemes generally treat color and intensity separately, the proposed method provides a composite view to both color and intensity variations in the same feature. The main advantage of using ICICM is that it avoids the use of weights to combine individual color and texture features. We use ICICM feature in an