Available online at www.sciencedirect.com
ScienceDirect
J. Vis. Commun. Image R. 17 (2006) 947–957
JOURNAL OF VISUAL Communication & IMAGE Representation
www.elsevier.com/locate/jvci

Texture image retrieval and image segmentation using composite sub-band gradient vectors ☆

P.W. Huang [a,*], S.K. Dai [a], P.L. Lin [b]

[a] *Department of Computer Science, National Chung-Hsing University , Taichung 40227, Taiwan*
[b] *Department of Computer Science and Information Management, Providence University, Shalu, Taiwan*

Received 15 March 2002; accepted 11 August 2005
Available online 19 June 2006

Abstract

A new texture descriptor, called CSG vector, is proposed for image retrieval and image segmentation in this paper. The descriptor can be generated by composing the gradient vectors obtained from the sub-images through a wavelet decomposition of a texture image. By exercising a database containing 2400 images which were cropped from a set of 150 types of textures selected from the Brodatz Album, we demonstrated that 93% efficacy can be achieved in image retrieval. Moreover, using CSG vectors as the texture descriptor for image segmentation can generate very successful results for both synthesized and natural scene images.

Keywords: Texture descriptor; CSG vector; Image retrieval; Image segmentation; Wavelet decomposition

1. Introduction

Texture is one of the most important attributes used in image analysis and pattern recognition. It provides surface characteristics for the analysis of many types of images including natural scenes, remotely sensed data, and biomedical modalities and plays an important role in the human visual system for recognition and interpretation. Although there is no formal definition for texture, intuitively this descriptor provides measures of properties such as smoothness, coarseness, and regularity.

A common weakness of most of texture analysis schemes is that they analyze the image at one single scale. Beck et al. [1] found that visual cortex can be modeled as a set of independent channels, each with a particular orientation and frequency tuning. The multichannel processing is a strong motivation for multi scale texture analysis methods. Several multichannel texture analysis systems have been proposed [2–4]. In the last decade, wavelet theory has emerged and received the attention of the image processing society. Daubechies [5,6] provided the discretization method of wavelet transform. Mallat [7] established the

☆ This research work was sponsored by National Science Council of ROC under contract No. NSC 90-2213-E-005-015.
* Corresponding author. Fax: +886 4 287 3028.
E-mail addresses: powhei.huang@msa.hinet.net, huang@ amath.nchu.edu.tw (P.W. Huang), lan@pu.edu.tw (P.L. Lin).

1047-3203/$ - see front matter
doi:10.1016/j.jvcir.2005.08.005

connection between wavelet transform and multi resolution theory. Since then, wavelet theory has become a mathematical framework which provides a formal and unified approach to multi scale (multi resolution) image analysis.

The main reason of using multi scale methods for texture image analysis is based on the theory that "big" patterns can be better captured by lower resolution processes while "small" patterns can be better captured by higher resolution processes. However, big patterns are usually treated as objects rather than texture elements. Thus, to avoid such an ambiguity, we carefully define a texture image to be dealt with in this paper as the one having "homogeneous" texture elements which are small enough to be viewed as an integral texture surface rather than individual objects of the same pattern in regular placement.

Based on the above assumption, we propose a new texture descriptor called *Composite Sub-band Gradient Vector* (or CSG vector) for the texture images containing homogeneous small "texels." A CSG vector is formed by composing the gradient vectors generated from the sub-images of a wavelet transform of an image. We applied this new texture descriptor to image retrieval in image database systems and image segmentation to demonstrate its effectiveness. In the experiment of image segmentation, the CSG vector was used as a discriminatory feature to segment both synthesized and natural scene images. The results from segmenting both types of images were very successful. In image retrieval, the CSG vector was used as a texture descriptor to extract images that are similar in texture to the query image from the database. We established an image database containing 2400 texture images of size 128×128 pixels which were cropped without overlapping from a set of 150 texture images of size 512×512 pixels selected from the Brodatz Album [8]. Two images are regarded as similar only if they are cropped from the same original image. We adopt this rigorous criterion to prevent any possible influence due to subjective factors. Our experimental results showed that 93% efficacy can be achieved in image retrieval.

The paper is organized as follows. In Section 2, we briefly review the theory of wavelet decomposition from which a CSG vector is derived. A new texture descriptor, the CSG vector, is introduced in Section 3. The results of image retrieval and image segmentation by using CSG vectors as the texture descriptor are presented in Sections 4 and 5, respectively. In Section 6 we summarize the results of the study and draw conclusions.

2. Wavelet decomposition

The pyramidal wavelet transform uses a family of wavelet functions and the associated scaling functions to decompose the original signal into different sub-bands. The decomposition process is recursively applied to the low-frequency sub-band to generate the next level of the hierarchy. The wavelet and scaling filters are applied in both the horizontal and vertical directions, followed by a 2–1 sub-sampling of each output image. This generates three orientation selective detail images $D_{j,k}$ and a coarse or approximate image C_j, where $k = 1,2,3$ and j represents the level of decomposition. The next level of resolution in the hierarchy is produced by repeating the same process. Thus, the hierarchical wavelet decomposition of an image can be described by the following equations:

$$\begin{aligned} C_j &= [H_x * [H_y * C_{j-1}]_{\downarrow 2,1}]_{\downarrow 1,2}, \\ D_{j,1} &= [H_x * [G_y * C_{j-1}]_{\downarrow 2,1}]_{\downarrow 1,2}, \\ D_{j,2} &= [G_x * [H_y * C_{j-1}]_{\downarrow 2,1}]_{\downarrow 1,2}, \\ D_{j,3} &= [G_x * [G_y * C_{j-1}]_{\downarrow 2,1}]_{\downarrow 1,2}, \end{aligned}$$

where * denotes the convolution operator, $\downarrow$ 2,1 denotes down-sampling at every other pixel along the x-direction, $\downarrow$ 1,2 denotes down-sampling at every other pixel along the y-direction, and $C_0 = I$ is the original image. H_x and G_x are a low and high-pass filter, respectively, along the x-direction. H_y and G_y are a low and high-pass filter, respectively, along the y-direction. The original image is thus represented by a set of sub-images at several scales $\{C_J, D_{j,k}\}_{j=1,\ldots,J;\ k=1,2,3}$ which is a multi scale representation of depth J of the image I. In our prototype system, we choose DAUB4 as the wavelet basis because it has the best average performance [9].