Image retrieval using 2D Gabor functions

RYSZARD S. CHORAS
Faculty of Telecommunications and Electrical Engineering
University of Technology and Agriculture
S. Kaliskiego 7, 85-791 Bydgoszcz
POLAND

Abstract: - The retrieval of images from a large database of images is an important and emerging area of research. High-resolution images in the physical database are decomposed into sets of image features which are stored in the logical database. Currently a few image retrieval systems use texture as features to search images. In this paper 2D image is processed using a set of Gabor filter to derive a feature vector representing the image. This method is useful for processing large collections of image data.

Key-Words: - Image retrieval, Texture, Gabor functions, Feature vector.

1 Introduction

In wide domains the majority of data is archival in the form of images. For the management of archived image data, an image database system which supports the analysis, storage and retrieval of images is needed (Fig.1). Much attention is done to the problems of how to retrieve images efficiently. Image matching and retrieval is based on some characteristic features. The input images are analyzed to extract the features and these features are stored in the database, along with the original images. Whenever an image is submitted for search its features are extracted. These extracted features are matched against those in the database.

Image data retrieval can be done using following features:

- low level features e.g. color, histogram, texture information,
- intermediate level features including shape primitives,
- highest level features including domain specific information.

We consider a texture information as the basic feature which is then used to retrieve images.

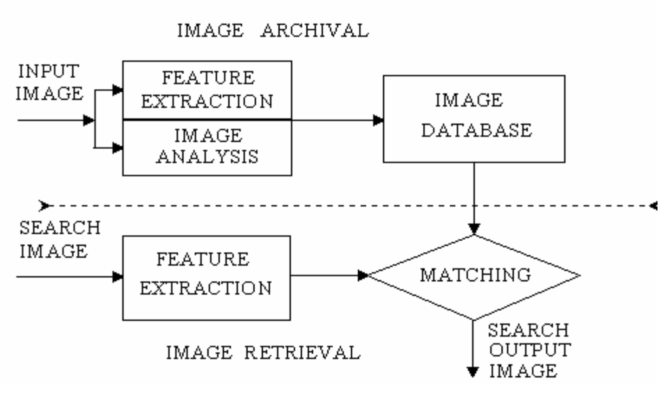

Fig.1. An image retrieval model.

Because, this paper describes a texture feature descriptor that is proposed to image similarity-based retrieved, an image can be considered as a mosaic of textures.

Each image is represented as a function of two variables $I(x, y)$ and stored as a two-dimensional array. If $M = \{1,2,\ldots,x,\ldots m\}$ and $N = \{1,2,\ldots,y,\ldots n\}$ are the spatial domains, then $D = M \times N$ is the set of resolutions cells and the digital image I is a function which assigns gray level value to each and every resolutions cells, i.e. $I : M \times N \to G$.

Pixels do not participate directly in the search process, and are only used for extraction of the features. For each image a feature vector is extracted and stored. Searching is done by sequentially going through each feature vector stored in the database and computing its distance to the input images feature vector.

We propose the use of oriented multi-resolution Gabor filters to extract features from the database images.

2 Gabor filter in spatial and frequency domain

Gabor filters produce spatial-frequency decompositions that achieve the theoretical lower bound of the uncertainty principle. The receptive visual field profiles are adequately modeled by 2D Gabor filters and Gabor filters give good performances in texture discrimination and segmentation.

We consider a Gabor function as texture description function and use Gabor function to derive a set of filter banks for use in texture description.

A two-dimensional Gabor function and its Fourier transform can be written as

$$g(x,y)=\left(\frac{1}{2\pi\sigma_x\sigma_y}\right)\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi jWx\right] \quad (1)$$

$$G(u,v)=\exp\left\{-\frac{1}{2}\left[\frac{(u-W)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right]\right\} \quad (2)$$

where $\sigma_u = \frac{1}{2\pi\sigma_x}$ and $\sigma_v = \frac{1}{2\pi\sigma_y}$.

Let $g(x,y)$ be the mother wavelet. Then a self-similar filter dictionary can be obtained by appropriate dilations and rotations of $g(x,y)$ through the generation function [2,1]:

$$g_{mn}(x,y)=a^{-m}g(x',y'),\quad a>1,\quad m=0,1,\ldots,S-1$$
$$x'=a^{-m}(x\cos\theta+y\sin\theta),$$
$$y'=a^{-m}(-x\sin\theta+y\cos\theta) \quad (3)$$

where $\theta=\frac{n\pi}{K}$ and K is the total number of orientations, S is the number of scales in the multiresolution decomposition, and $a=\left(\frac{f_h}{f_l}\right)^{\frac{-1}{S-1}}$, $W=f_h$ and f_l, f_h are the lower and upper center frequencies of interest.

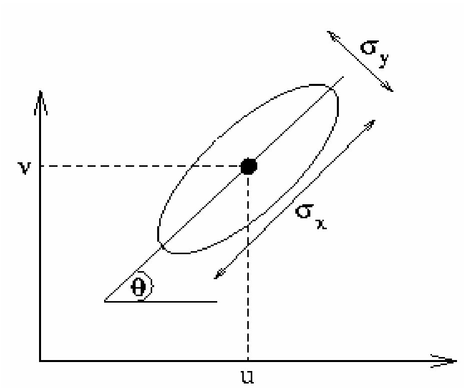

Fig. 2. Parameters defining filter in the frequency domain.

The spectral extents of filters ($K=6$ and $S=4$) are described in Fig.3. See Fig.4 for examples of the filters at various orientations and scales.

From the multiresolution decomposition, a given image is decomposed into a set of filtered images. Each of these images represents the image information at a certain scale and at a certain orientation.

3 Texture features

Texture is one such property of an image which describes the spatial relationships between gray levels of pixels in a region. We have applied two-dimensional Gabor filters to compute texture features. Gabor filters provide a powerful representation space to describe image texture. Gabor filters exhibit optimal localization properties in the spatial domain as well as in the frequency domain. In essence, Gabor filters are ideally suited for texture discrimination problems. In particular, we consider two main steps: (i) Filter the input image through a bank of SK Gabor filters to obtain SK filtered images. (ii) Compute the feature vectors consisting of the local energy estimates.

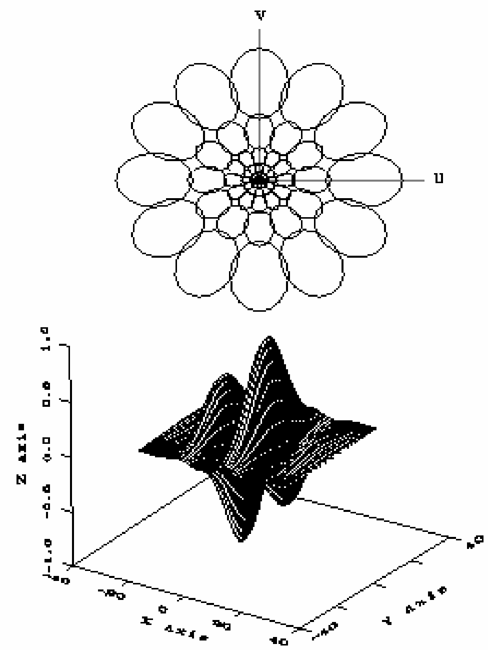

Fig.3. The set of 2D variable window Gabor filters.

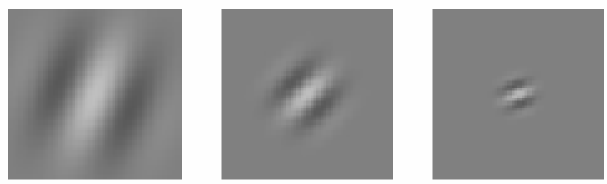

Fig.4. Real components of Gabor wavelet filters at different orientations and scales.

Gabor filtered output of the image is obtained by the convolution of the image with Gabor function for each of the orientation/spatial frequency (scale) orientation. Given an image $I(x,y)$, we filter this image with $g_{mn}(x,y)$

$$T_{mn}=\sum_k\sum_l I(x,y)g_{mn}^{*}(x-k,y-l) \quad (4)$$

where * indicates the complex conjugate.

This process results in a vector of SK complex values at each point of an image.

The transform vector (TV) for the given image $I(x, y)$ is the vector of matrices defined as

$$TV = [T_{11}, \ldots, T_{SK}] \tag{5}$$

The power spectrum, the mean and the variance at each transform coefficients position for each T_{mn} are used to construct the feature vector (Fig. 5).

Let m indicate a certain scale and n a certain orientation. The mean μ_{mn}, the power spectrum p_{mn} and the variance var^2_{mn} of the magnitude of the transform coefficient, can be defined:

$$\mu_{mn} = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N} T_{mn}(x, y) \tag{6}$$

$$p_{mn} = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N} T^2_{mn}(x, y) \tag{7}$$

and

$$\mathrm{var}^2_{mn} = \sum_{x=1}^{M}\sum_{y=1}^{N} [T_{mn}(x, y) - \mu_{mn}]^2 = \\ = \sum_{x=1}^{M}\sum_{y=1}^{N} T^2_{mn}(x, y) - MN\mu^2_{mn} \tag{8}$$

The feature vector (FV) is represented as follows

$$FV = [p_{11}, \mathrm{var}^2_{11}, \ldots, p_{SK}, \mathrm{var}^2_{SK}] \tag{9}$$

Two images i and j have two corresponding feature vectors $FV^{(i)}$ and $FV^{(j)}$. A normalized distance between the two images in the feature space is

$$d^{(i)(j)}(p, \mathrm{var}^2) = \sum_m \sum_n d^{(i)(j)}_{mn} \tag{10}$$

where

$$d^{ij}_{mn} = \left| \frac{p^{(i)}_{mn} - p^{(j)}_{mn}}{\Xi(p_{mn})} \right| + \left| \frac{\mathrm{var}^{2(i)}_{mn} - \mathrm{var}^{2(j)}_{mn}}{\Xi(\mathrm{var}^2_{mn})} \right| \tag{11}$$

where $\Xi(p_{mn})$ and $\Xi(\mathrm{var}^2_{mn})$ are respectively the standard deviations of power spectrum and the variance of the transform coefficients over the database.

4 Image retrieval process

The image data is represented as a $I_{input} = \langle I, FV \rangle$ where I is the image matrix, FV the feature vector. The retrieval process is defined as follows:

1. Image $I(x, y)$ in the database is represented by image matrix I and feature vector,
2. Search image is represented by feature vector,
3. Similarity between search image and the database image is then evaluated by combining individual feature vector representation similarity,
4. The database images are ordered by their similarity to search image,
5. The image with higher similarity is marked.

5 Experimental results

To compute the effectiveness of the similarity retrieval a test was designed as follows: 10 query images were used and 500 images in the database were found. The above described feature vector of the query image was then compared to the corresponding feature vector of the images in the database to obtain a list of similar images. Then the efficiency of retrieval $\eta = \frac{k}{L}$ where k is the number of similar images retrieved and L is the total number of similar images in the database was calculated. Average retrieval efficiency computed over 10 queries for list of similar images with 10 elements is 87,6%.

5 Concluding remarks

We have presented a content based image retrieval system which is based on texture features. Textured images have space-varying local properties. The feature extraction is based on the mechanism of multichannel representation of the retinal images in the biological visual system. The local properties of the texture can be obtained using a set of Gabor filters. The proposed texture descriptor provides a robust representation of many images.

References:

[1] P. Wu, B.S. Manjunath, S. Newsam, H.D. Shin, A texture descriptor for browsing and similarity retrieval, *Signal Processing: Image Communications*, Vol.16, 2000, pp. 33-43.

[2] B.S. Manjunath, W.Y. Ma, Texture features for browsing and retrieval of image data, *IEEE Transaction on Pattern Analysis and Machine Intelligence,* Vol.18, No. 8, 1996, pp. 837-842.

[3] T.P. Weldon, W.E. Higgins, D.F. Dunn, Efficient Gabor filter design for texture segmentation, *Pattern Recognition,* Vol. 29, No. 12, 1996, pp.2005-2015

[4] A. Jain, F. Farrokhma, Unsupervised texture segmentation using Gabor filters, *Journal of Pattern Recognition,* Vol.24, 1991, pp.1167-1186

.

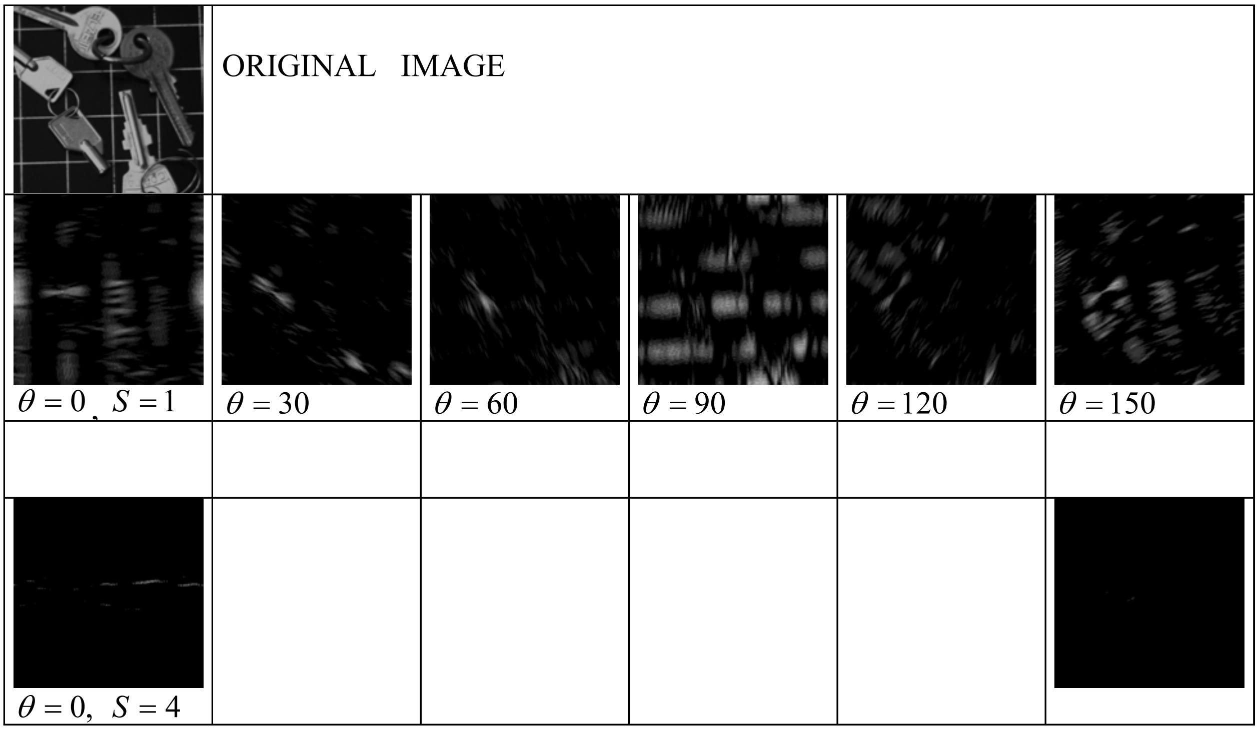

Fig.5. Filtered images T_{mn}