

A Texture Descriptor for Image Retrieval and Browsing

P. Wu, B. S. Manjunanth, S. D. Newsam, and H. D. Shin*

Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106-9560

*Samsung Electronics Co.

{peng, manj, snewsam, hdsin}@iplab.ece.ucsb.edu

Abstract

Image texture is an important visual primitive in image search and retrieval applications. To characterize texture information of images, a texture descriptor is proposed to capture both the perceptual information and the statistics of textures. The descriptor consists of two parts: a Perceptual Browsing Component (PBC) and a Similarity Retrieval Component (SRC). The emphasis of this paper is to introduce the extraction method for the PBC which is based on multidimensional decomposition of images using Gabor filters. By analyzing the one dimensional projections resulting from filtered images, PCB gives a quantitative measure about the structuredness of textures and also identifies the directionality and coarseness (scale). Our experimental results demonstrate that PBC can capture these texture properties quite well.

Keywords: texture descriptor, perceptual image browsing, structuredness, multiresolution decomposition, Gabor filtering

1. Introduction

In recent years, Image texture has emerged as an important visual primitive to search and browse through large collections of similar looking patterns. An image can be considered as a mosaic of textures and texture features associated with the regions can be used to index the image data. For instance, a user browsing an aerial image database may want to identify all parking lots in the image collection. A parking lot with cars parked at regular intervals is an excellent example of a textured pattern when viewed from a distance, such as in an airphoto. Similarly, agricultural areas and vegetation patches are other examples of textures commonly found in aerial and satellite imagery. Examples of queries that could be supported in this context could include “Retrieve all Landsat images of Santa Barbara which have less than 20% cloud cover” or “Find a vegetation patch that look

like this region”. To support image retrieval or browsing applications as mentioned above, an efficient representation of textures is needed.

One of the widely used representations of textures is the texture feature proposed in [1] and its improved version [2]. The texture feature used in [1] and [2], to some extent, are based on models of human texture perception. More recently, several random field based texture models and multi-scale filtering methods [6, 11] have been studied. Use of texture for content based retrieval has been explored by several researchers [3, 7, 8]. Among these, features computed from Gabor filtered images appear quite promising. A comprehensive evaluation on using Gabor features can be found in [7, 9]. More recent evaluation and comparison using other texture features also support the observation that the orientation and scale selective Gabor filtered images capture relevant texture properties for applications such as image retrieval very well [11].

In this paper, we propose a novel texture descriptor based on our prior work on using Gabor filters. The descriptor has two parts: The first part relates to a perceptual characterization of texture, similar to a human characterization, in terms of structuredness, directionality and coarseness (scale). This representation is useful for browsing type applications and coarse classification of textures. We call this part the Perceptual Browsing Component (PBC). The second part provides a quantitative description that can be used for accurate search and retrieval. This is referred to as the Similarity Retrieval Component (SRC). Both of them are derived from multiresolution Gabor filtering. This framework is illustrated in Fig. 1.

The SRC is the same texture feature as that was introduced in [7]. Since the extraction and application of SRC has been investigated and evaluated in a very detailed way by different researchers [7, 8, 9, 11], the main emphasis of this paper is to introduce an extraction method for the PBC and demonstrate its usefulness in image classification and browsing.

The paper is organized as follows: Section 2 introduces the formation of PBC feature vector and its

semantics; Section 3 will present the extraction method for PBC. In Section 4, some experimental results and an example of using PBC in image browsing is given. Section 5 concludes with discussions.

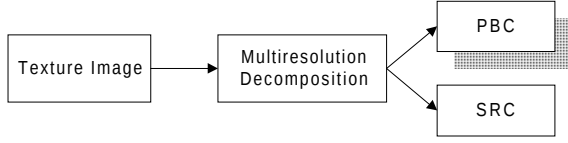


Figure 1: A framework for feature extraction

2. Perceptual Browsing Component (PBC)

The Perceptual Browsing Component has the following format:

$$PBC = [v_1 \ v_2 \ v_3 \ v_4 \ v_5] \quad (1)$$

where

- v_1 is an integer and $v_1 \in \{1, \dots, N_v\}$, the larger the value is, the more structured the corresponding texture is. Thus, v_1 provides a confidence measure on the structuredness of the texture.
- $v_2, v_3 \in \{1, \dots, K\}$ give two quantized directions which best capture the structuredness. For example, '1' may correspond to 0° , and K correspond to 150° when the direction are quantized to 30° intervals.
- $v_4, v_5 \in \{1, \dots, S\}$ give two quantized scales which best capture the structuredness. S correspond to the number of levels in a wavelet decomposition, see [7].

3. Extraction Method for the PBC

3.1 Multiresolution Decomposition

Using Gabor filters designed in [7], an image is decomposed at K orientations and S scales. For a given image $I(x, y)$, its decomposed image at scale m ($m = 1, \dots, S$) and direction n ($n = 1, \dots, K$) is defined to be

$$W_{mn}(x, y) = \int I(x, y) g_{mn}^*(x - x_1, y - y_1) dx_1 dy_1 \quad (2)$$

where $*$ indicates the complex conjugate, $W_{mn}(x, y)$ is the filtered image at a specific scale m and a specific direction n . In this paper, we set $S = 4$ and $K = 6$.

3.2 Overview of the Extraction Method

From the multiresolution decomposition, we get a set of filtered images. Each of them represents the image information at a certain scale and at a certain orientation. The objective of structuredness analysis is to get a quantitative measure of structuredness from these filtered images. The analysis is based on the following observations.

- Structured textures usually consist of dominant periodic patterns. The more regular the periodicity is, the stronger the structuredness is.
- A periodic or repetitive pattern, if it exists, could be captured by filtered images. This behavior is usually captured in more than one filtered output.
- The dominant scale and orientation information can also be captured by investigating the distribution of those filtered images which constitute the perception of structuredness.

The analysis of structuredness is a two step procedure. The first step is the analysis on each filtered image. The objective of this step is to determine the existence of a repetitive pattern. The second step is performed on all the filtered images that are identified as having some kind of periodicity. From the second step, a quantitative measure is derived to characterize the structuredness. The information about the dominant scale and the dominant direction, in terms of contributing to the formation of structuredness, are also derived.

3.3 Analysis of Each Filtered Image

The analysis procedure for each filtered image is shown in Fig. 2.

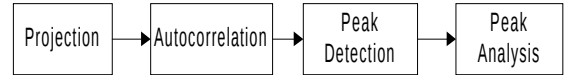


Figure 2: The analysis of each filtered image

• Projection

For each filtered image, the projections along horizontal and vertical directions are computed. These are denoted as P_H and P_V , respectively. (if a texture has a dominant orientation other than horizontal and vertical, it is possible to compensate for this by taking the Radon projections along the dominant direction. The dominant direction can be estimated using the power spectrum analysis [6], for example. For notational simplicity, we consider only the horizontal and vertical projections here.)

• Autocorrelation

For each filtered image, its horizontal or vertical projection is a one dimensional sequence, which can be denoted as $p(l)$, where $l = 1, \dots, N$. Consider now the horizontal projections, on which the following analysis is performed (same analysis can then be performed on the vertical projections as well.) The normalized autocorrelation function (NAC) is defined to be

$$NAC(k) = \frac{\sum_{m=k}^{N-1} p(m-k)p(m)}{\sqrt{\sum_{m=k}^{N-1} p^2(m-k) \sum_{m=k}^{N-1} p^2(m)}} \quad (3)$$

Fig. 3 shows the horizontal projections of image T001.01 from the Brodatz album (the texture image is shown in Figure 5).

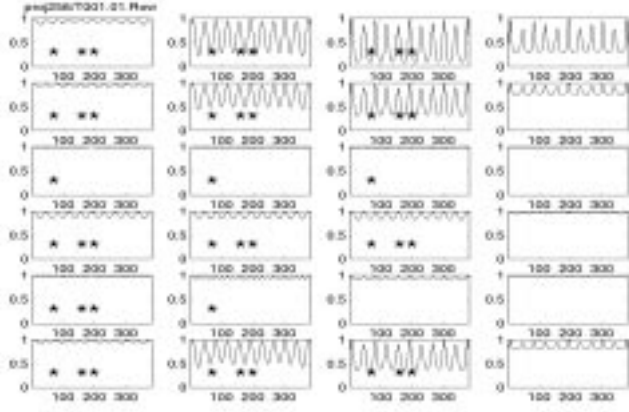


Figure 3: *NAC* of horizontal projections of all the 4×6 filtered images from image T001.01. The projections labeled with '*' only are the detected potential candidates and those also labeled with '**' are the final candidates after clustering

• Peak Detection

Given the autocorrelation function $NAC(k)$, the peaks and valleys are detected and their positions and magnitudes are recorded. Let M be the number of peaks and N be the number of valleys, then $p_posi(i)$, $p_magn(i)$ are the positions and magnitudes of those peak points, where $i = 1, \dots, M$, and $v_posi(j)$, $v_magn(j)$ are the positions and magnitudes of the valley points, where $j = 1, \dots, N$. The contrast of the projection is defined to be

$$contrast = \frac{1}{M} \sum_{i=1}^M p_magn(i) - \frac{1}{N} \sum_{j=1}^N v_magn(j) \quad (4)$$

• Peak Analysis

Given a peak sequence $p_posi(i)$ with length M , two measures are extracted, the average of the distances among the peaks, dis , and the square root of the standard deviation of distances, std . Then

$$\gamma = \frac{std}{dis} \quad (5)$$

If γ is smaller than a pre-selected threshold T_p , the corresponding projection is considered to represent periodic information. The projection would be chosen for further analysis and represented by a vector $p_{mn}(dis, std)$, where m and n denote the scale and direction information associated with the filtered output. These are the potential candidates for further analysis.

3.4 Analysis of Potential Candidates

Among the P_h or P_v coming from all the filtered images, some of them would be selected as potential candidates. They form a set of vectors $\{p_{mn}(dis, std)\}$. A consistency check is now performed to identify the dominant candidates from the list of potential candidates. A *modified agglomerative clustering* algorithm [5] is used to select the final candidate projections containing the dominant structuredness information, see Fig 3.

Let us denote the final set of selected candidates as $\{C_{mn}(dis, std)\}(H)$ and $\{C_{mn}(dis, std)\}(V)$, for the horizontal and vertical projections, respectively, which are then used to compute the PBC.

Once the final candidates are selected, four elements of PBC, v_2 , v_3 , v_4 and v_5 could be determined.

Denoting $C(m^*(H), n^*(H))(H)$ as the candidate projection from P_H that has the maximum contrast and $C(m^*(V), n^*(V))(V)$ as the candidate projection from P_V that has the maximum contrast. We have

$$PBC[v_4] = m^*(H) \text{ and } PBC[v_2] = n^*(H)$$

$$PBC[v_5] = m^*(V) \text{ and } PBC[v_3] = n^*(V)$$

3.5 Computing $PCB[v_1]$

The method of measuring the degree of the structuredness is based on the following observations on the distribution of candidate vectors.

- For strong structured textures, their periodicity could be captured by multiple projections --- the candidates chosen from the above procedure. These candidates are usually close to each other.
- If the texture is not structured or only weakly structured, the distribution of the candidates, if they exist, usually is sparse and the neighboring relationship can rarely be detected.

If such a consistency in the neighboring projections is detected from the projections in the candidate set, this would result in a larger credit, indicating a stronger structure. Based on these observations, the candidate projections are further classified as follows:

Candidate Classification

C_1 : For a specific candidate, we can find at least one another candidate at its neighboring scale or orientation. The credit value associated with this class is $V_1 = 1.0$.

C_2 : For a specific candidate, we can find at least one another candidate distributed at the same scale or orientation but no candidate is located at its neighboring scale or orientation. The credit value associated with this class is $V_2 = 0.5$.

C_3 : The candidate is the only one distributed at its scale and orientation. The credit value associated with this class is $V_3 = 0.2$.

Quantitative Measurement

At this stage, each of the candidate projections has an associated value computed based on the above classification. Let

$$M = \sum_{i=1}^3 N_i * V_i \quad (6)$$

where N_i is the number of candidate projections classified as C_i . M is calculated independently for the horizontal (M_H) and vertical (M_V) projections. Let

$$M_{img} = M_H + M_V \quad (7)$$

M_{img} is classified into N_v bins by using *option decision tree classifier* [12]. The larger the value of M_{img} is, the more structured the corresponding texture is. In our current implementation, $N_v = 4$. Consequently, each image is associated with a number B_{img} ($B_{img} \in \{1, \dots, N_v\}$) to indicate which bin an image belongs to.

$$PBC[v_1] = B_{img}$$

4. Experimental Results

To demonstrate the performance of PBC, each of 512×512 image texture from the Brodatz database [4] is divided to 4 256×256 subimages.

In this experiment, 15 different image textures are selected which are

class1={T001,T006,T014,T020,T095}
class2={T002,T007,T009,T012,T097}
class3={T018,T055,T067,T075,T107}

These images are selected subjectively so that class1 includes images that are considered clearly as strong structured, class2 contains images that are strong non structured and images in class3 are neither strongly structured nor completely random. Figure (4) shows the subimages and their corresponding PBC (Note, for each image texture, only one subimage (Txx.01) is shown as sample of the image texture).

In the experiment, we set $N_v = 4$, $S = 4$ and $K = 6$. So each element of PBC vector has the meaning as following:

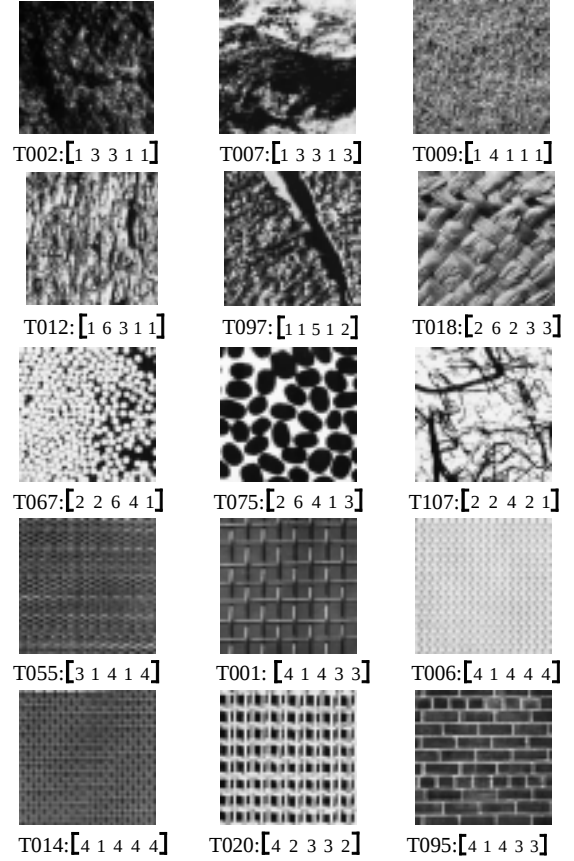


Figure 4 Example images and their associated PBC

$v_1 \in \{1, \dots, 4\}$, the larger the v_1 is, the more structured the texture is. For example, a image with $v_1 = 4$ would be one of the more structured images, like T001.01; $v_2, v_3 \in \{1, \dots, 6\}$ give two quantized orientations which best capture the structuredness. The orientation from 0° to 180° is quantized to 6 orientations with a 30° interval. So ‘1’ corresponds to 0° , ‘2’ corresponds to 30° and so on. v_2 and v_3 could be same; $v_4, v_5 \in \{1, \dots, 4\}$ give two quantized scales which best capture the structuredness. v_4 and v_5 could be same.

For example, for image T001.01 which has

$$PBC_{T001.01} = [4 \ 1 \ 4 \ 3 \ 3]$$

we can tell that it is a strong structured image, its structuredness can be best captured from orientation 1 and 4, which are 0° and 90° respectively, and in scale 3.

From the experiment, we observe the following

- $PBC[v_1]$ gives a good description of the structuredness in good agreement with human characterization.

- $PCB[v_1]$ also provides a confidence measure of v_i , $i = 2,3,4,5$. The larger the v_1 is, the more trustful the v_i , $i = 2,3,4,5$, are.
- For images with large $PCB[v_1]$, v_i ($i = 2,3,4,5$) provide good estimation about at which orientations and which scales the structuredness can be best captured.

We have developed a browsing system that uses PBC to browse through images using texture features. Fig 5(a) shows an interface. Fig 5(b) provides an example browsing scenario.

5. Conclusions and Future Work

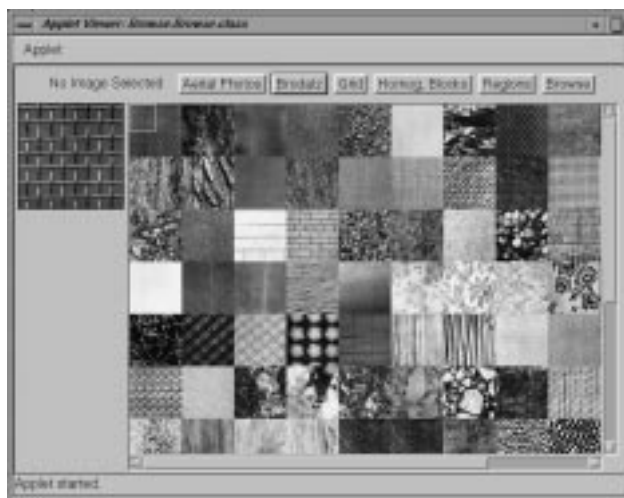
In this paper, a texture descriptor is proposed which can characterize both the perceptual and statistic features of textures. The extraction method of the Perceptual Browsing Component is introduced. Our experiment results demonstrate that PBC can capture human's perception on structuredness quite well.

We are currently investigating the use of PBC and SRC in improving retrieval performance.

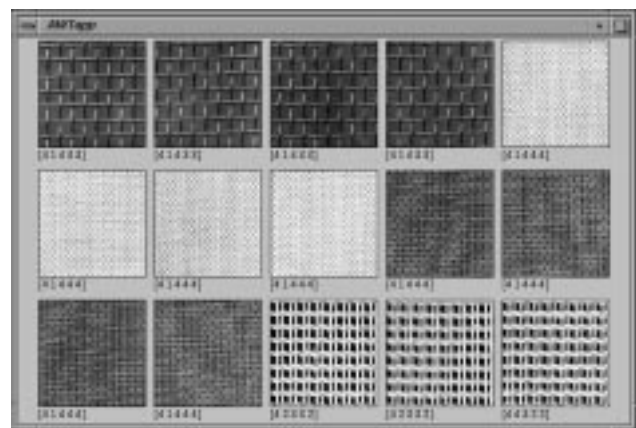
Acknowledgements: This research is supported in part by Samsung Electronics Co..

References

- [1] Hideyuki Tamura, Shunji Mori, and Takashi Yamawaki, "Texture features corresponding to visual perception", IEEE Trans. On Sys. Man, and Cyb, SMC8(6), 1978.
- [2] Will Equitz and Wayne Niblack, "Retrieving images from a database using texture-algorithms from the QBIC system", Technical Report RJ 9805, Computer science, IBM Research Report, May 1994.
- [3] John R. Smith and Shih-Fu Chang, "Automated binary texture feature sets for image retrieval", Proc. ICASSP-96, Atlanta, GA, 1996.
- [4] P. Brodatz, "Textures: A photographic album for artists & designers," Dover, NY, 1966.
- [5] R.O. Duda and P.E. Hart, Pattern classification and scene analysis, John Wiley & Sons, New York, 1970.
- [6] F. Liu, R. W. Picard, "Periodicity, directionality, and randomness: Wold features for image modeling and retrieval," MIT Media Lab Technical Report No. 320, March 1995.
- [7] B.S. Manjunath, W.Y. Ma, "Texture features for browsing and retrieval of image data", IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(8), 837-842, 1996.
- [8] W.-Y. Ma and B.S. Manjunath, "A texture thesaurus for browsing large areal photographs", Journal of The American Society for Information Science, 49(7):633-648, 1998.
- [9] W.Y. Ma, HongJiang Zhang, "Benchmarking of Image Features for Content-based Retrieval", Proc. of 32nd Asilomar Conference on Signal, System & Computers, 1998.
- [10] J. Mao, A. Jain, "Texture classification and segmentation using multiresolution simultaneous autoregressive models," Pattern Recognition Journal, 25(2), pp. 173-188, 1992.
- [11] C.S. Li, J.R. Smith, V. Castelli and L. Bergman, "Comparing Texture Feature Set for Retrieving Core Images in Petroleum Applications", Proceedings of SPIE Storage and Retrieval for Image and Video Databases VII, San Jose, CA, Jan. 26-29, 1999, SPIE vol. 3656, pp.2-11.
- [12] W. Buntine, "Learning classification trees", Statistics and Computing 2(2), pp. 63-73, 1992.



(a) The interface of a perceptual browsing system.



(b) Images browsed according to the structuredness of the top left hand corner image in (a)

Figure 5: A perceptual image browsing system