Multi-model Feature Integration For Texture Classification

Yong Huang †, Kap Luk Chan
School of Electrical and Electronic Engineering
Nanyang Technological University
Nanyang Avenue, Singapore 639798
†email: *p146457665@ntu.edu.sg*

Abstract

The Wold texture model considers a textured pattern being composed of two main types of homogenous random fields, namely the deterministic and the indeterministic fields. The two fields can be represented by different type of models. It is known that, for textures, the multi-channel model based on Gabor function(Gabor model) is very effective for representing the deterministic fields, and a Gaussian Markov Random Field(GMRF) model is very effective in representing the indeterministic fields. In this paper, we propose to use both models for texture classification, in which features based on each model are integrated according to the consensus theory. A weighting parameter, the deterministic energy ratio determined from the spectrum distribution function, is used as the flexible weight in the consensus theory. In this way, a wider variety of textures can be better-represented and hence lead to better classification of the textures.

1. Introduction

The classification problem is basically the problem of identifying an observed textured sample as one of several possible texture classes by a reliable but computationally attractive texture classifier. This implies that the choice of the textural features should be as compact as possible, and yet as discriminating as possible. In other words, the extraction of texture features should efficiently embody information about the textural characteristics of the image pattern. Early methods on texture classification were based on single model feature which reflected the statistical or structural properties of texture image. The statistical feature characterizes the texture by statistics of image pixel gray scale values. These methods give relatively high recognition rate when the test pattern has random looking texture, but not for large scale and more structured patterns. The structural feature assumes that the texture is generated by the placement of the primitives according to certain placement rules. These methods usually give good results in classifying structural textures[12][3].

Natural texture usually contains both structural and statistical properties. Single feature set is insufficient to describe texture image completely. Combining texture features has been suggested by some authors[10][11][9]. Conceptually, the simplest method is the stacked-vector approach in which multi-model features are concatenated together into a single feature vector and input to a classifier. This method is very straightforward and work well if the features are similar. However, the method is not applicable when the features cannot be described by a common model, e.g. the multivariate Gaussian model.

In this paper, we proposed a texture classification method based on multi-model feature integration by consensus theory[1]. Two feature sets are extracted from the two components of the texture image decomposed from the image by the Wold-like decomposition[5]. Based on the consensus rule, two feature sets are combined with a flexible weight. 112 texture classes from Brodatz database[2] were used for analysis of classification performance.

2. Multi-model Texture Description

A texture field is generally a realization of a regular homogeneous random field. On the basis of a 2-D Wold-like decomposition theory for homogeneous random fields[4], the texture field can be decomposed into two mutually orthogonal components: a deterministic component and a purely indeterministic component. The deterministic component corresponds to the structural texture pattern while the purely indeterministic component corresponds to the random texture pattern.

2.1. Texture decomposition

Since the deterministic component gives rise to singularities(1-D and 2-D delta-functions) in the spectral domain, the estimation of the deterministic components can be carried out by detecting the spectral harmonic peaks and their support region.

Harmonic peaks can be found as local maxima in the power spectrum. The local maxima of the magnitude are found by searching a 5×5 neighborhood using a similar process as in[7]. The size of the neighborhood is chosen to match the resolution of the estimated spectra so that the resulting local maxima are separated from each other by at least two frequencies sample points. Starting from each harmonic peaks, a region is growing outwards continuously until the value of the magnitude is lower than a small portion of this peak value(10% in this work). This region is regarded as the peak support region of a harmonic peak.

The decomposition of a homogeneous random field is based on the decomposition of its spectrum. Denote the image's 2D DFT as $F(u,v)$, the corresponding frequency plane as $\Omega((\Omega_s, \Omega_g) \in \Omega$, where Ω_s is the set of frequencies corresponding to the harmonic peaks and evanescent lines). The spectrum of the random field can be then decomposed into the deterministic component

$$F_s(u,v) = \begin{cases} F(u,v) & , \quad (u,v) \in \Omega_s \\ 0 & , \quad otherwise \end{cases} \tag{1}$$

and the indeterministic component

$$F_g(u,v) = \begin{cases} F(u,v) & , \quad (u,v) \in \Omega_g \\ 0 & , \quad otherwise \end{cases} \tag{2}$$

The deterministic field $s(m,n)$ and the indeterministic field $g(m,n)$ are obtained by computing the inverse DFT of $F_s(u,v)$ and $F_g(u,v)$, respectively.

$$s(m,n) = \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F_s(u,v) e^{j\frac{2\pi}{N}(mu+nv)} \tag{3}$$

$$g(m,n) = \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F_g(u,v) e^{j\frac{2\pi}{N}(mu+nv)} \tag{4}$$

After the above procedure, a texture image is decomposed into two component images: deterministic component $s(m,n)$ which exhibits regular(or structural) texture pattern and indeterministic component $g(m,n)$ which exhibits random(or stochastic) texture pattern. The deterministic energy ratio can be defined as:

$$\alpha = \frac{F_s(u,v)}{F_s(u,v) + F_g(u,v)} \tag{5}$$

Figure 1 shows the decomposition of Brodatz texture D017.

2.2. Feature Extraction

It has been demonstrated that for textures, the multi-channel model based on Gabor function(Gabor model) is

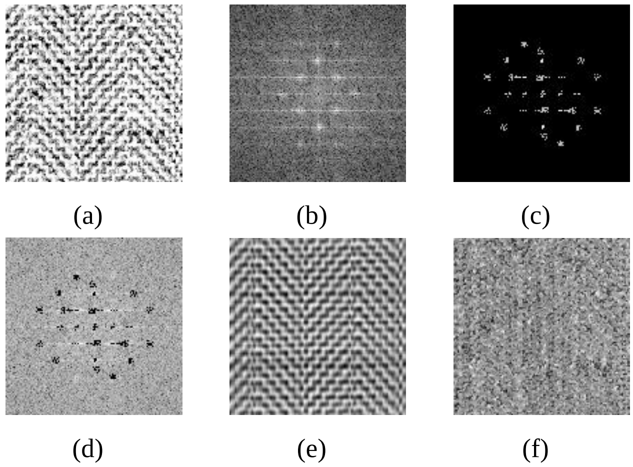

Figure 1. Decomposition of Brodatz texture D017: (a) Original image; (b) Fourier magnitude; (c) spectral harmonic peaks; (d) spectral indeterministic component; (e) reconstruction of deterministic component; (f) reconstruction of indeterministic component;

very effective for representing the deterministic fields, and a Gaussian Markov Random Field(GMRF) model is very effective in representing the indeterministic fields.

In this paper, a Gabor filter bank with four scales ($S = 4$) : $0.05, 0.1, 0.2$ and 0.4 cycles/image-width, and six orientations ($K = 6$) at $0°, 30°, 60°, 90°, 120°$ and $150°$ is used to extract the features of the deterministic component. Two Gabor features were extracted from each filtered images, which are the mean, μ_{mn}, and standard deviation, σ_{mn}[8]. So, $2SK$ features can be extracted, forming a $2SK$-dimensional feature vector

$$\mathbf{F_g} = [\mu_{00}\sigma_{00} \cdots \mu_{(S-1)(K-1)}\sigma_{(S-1)(K-1)}] \tag{6}$$

To extract the features from the indeterministic component, a second order GMRF model was considered. The parameters of GMRF model $\Theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ were estimated from [6] using the least square method. The feature vector is denoted by

$$\mathbf{F_m} = (\theta_1, \theta_2, \theta_3, \theta_4, \bar{\mu}, \bar{\sigma}^2) \tag{7}$$

where $\bar{\mu}$ and $\bar{\sigma}^2$ are the mean and variance of the gray level texture.

Both $\mathbf{F_g}$ and $\mathbf{F_m}$ are used to describe the characteristics of a texture. The problem now is how can these two feature vectors be integrated.

3. Feature Integration by The Consensus Theory For Classification

3.1. Consensus Theory

The consensus theory[1] defines general procedures for combining probability distributions to summarize estimates from multiple experts with the assumption that the experts make decisions based on Bayesian probabilities. The combination formula obtained is called a consensus rule.

One of the consensus rule is logarithmic opinion pool(LOGP). It is described by:

$$L_j(Z) = \prod_{i=1}^{n} p(\omega_j | z_i)^{\lambda_i} \tag{8}$$

or

$$\log(L_j(Z)) = \sum_{i=1}^{n} \lambda_i \log(p(\omega_j | z_i)) \tag{9}$$

where $Z = [z_1, \cdots, z_n]$ is a compound vector consisting of observations from all the data sources, $p(\omega_j|z_i)$ is a source-specific posterior probability and λ_i's$(i = 1, \cdots, n)$ are source-specific weights which control the relative influence of the data sources. The weights are associated with the sources in the global membership function ($L_j(Z)$) to express quantitatively the goodness of each source.

The consensus theory has shown the advantages of using multiple data sources in remote sensing application[1].

3.2. Classification using Multi-model Feature Integration

Suppose we have n sets of feature vectors due to n models, $\mathbf{F} = \{F_1, \cdots, F_n\}$ for each texture image F, where $F_i = (f_1, \cdots, f_{s_i}), i = 1, \cdots, n$, s_i is the dimension of the feature vector. Let there be k classes denoted $\omega_j, j = 1, \cdots, k$; The goal of the classification is to assign an input image to class ω_j for which the probability $P(\omega_j|F_1, F_2, \cdots, F_n)$ is maximum, i.e.,

$$\hat{c} = \arg\max_j (P(\omega_j | F_1, F_2, \cdots, F_n)) \tag{10}$$

where $P(\omega_j|F_1, F_2, \cdots, F_n)$ is the conditional probability that ω_j is the correct class given the observed data $F_1, F_2, \cdots, F_n$.

In particular, for the minimum-error-rate classification, the classification rule can be expressed as the discriminate function $G_j(F) = \log(P(\omega_j|F_1, F_2, \cdots, F_n))$. Applying to the consensus theory, we have:

$$G_j(F) = \sum_i \lambda_i \log(P(\omega_j | F_i)) \tag{11}$$

According to the Bayesian theory:

$$P(\omega_j | F_i) = \frac{p(F_i|\omega_j)P(\omega_j)}{\sum_j p(F_i|\omega_j)P(\omega_j)} \tag{12}$$

If we discard the prior probabilities $P(\omega_j)$ by treating them equal, the classification rule become to maximize

$$\hat{c} = \arg\max_j (\sum_i \lambda_i p(F_i | \omega_j)) \tag{13}$$

where λ_i are the weights that can be derived from the energy ratio, α, of deterministic component of a texture class; and, $p(F_i|\omega_j)$ is the conditional density of individual model feature vector in class ω_j.

3.3. Classification Scheme

Texture classification based on multi-model feature integration is illustrated in Figure 2. The inputs to the system are images from one of the M texture classes. The images are separated into test and training sets. In the training stage, two prototypes are trained. They correspond to the deterministic component(C_{1j}) and indeterministic component (C_{2j}). The training sets are used to estimate the sample means, covariance matrix of the two prototypes and also the deterministic energy ratio α. In this case, a class ω_j can be expressed by its two prototypes.

In this paper, an assumption was made that all the texture images in the database have homogeneous patterns. Under this assumption, a Gaussian model is assumed for a particular family of texture features. In this case, the conditional densities $P(F_i|C_{ij}), i = 1, 2; j = 1, \cdots, k$ is distributed according to a multivariate normal distribution:

$$P(F_i|C_{ij}) = \frac{1}{(2\pi)^{d/2}|\Sigma_{ij}|^{1/2}} \exp\left[-\frac{1}{2}(F_i - \mu_{ij})^T \Sigma_{ij}{}^{-1}(F_i - \mu_{ij})\right] \tag{14}$$

where μ_{ij} is the mean vector and Σ_{ij} the covariance matrix associated to the class ω_j:

$$\begin{aligned} \mu_{\mathbf{ij}} &= \frac{1}{M}\sum_{i=1}^{M} F_{ij} \\ \Sigma_{ij} &= \frac{1}{M}\sum_{i=1}^{M} (F_{ij} - \mu_{ij})(F_{ij} - \mu_{ij})^T \end{aligned} \tag{15}$$

In this case, the decision rule can be shown to be equivalent to

$$\begin{aligned} \hat{c} &= \arg\min_j G_j(F_1, F_2) \\ &= \arg\min_j \{\lambda_{1j} D_{1j}(F_1) + \lambda_{2j} D_{2j}(F_2)\} \end{aligned} \tag{16}$$

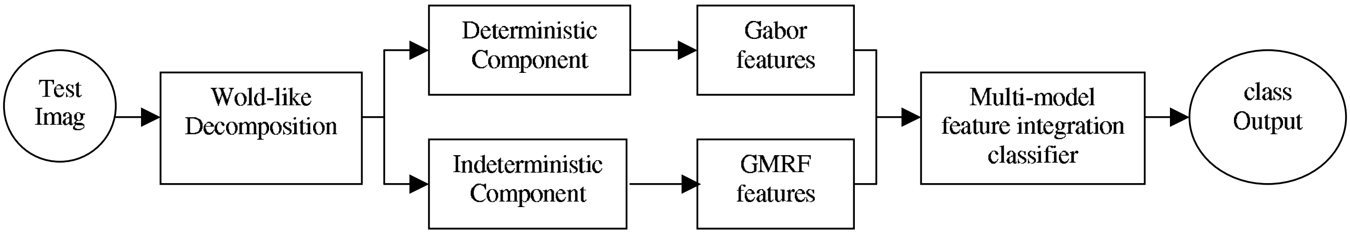

Figure 2. Block diagram of the classification scheme

where $D_{ij}(F_i), i = 1, 2$ is given by

$$D_{ij}(F_i) = (F_i - \mu_{ij})\Sigma_{ij}^{-1}(F_i - \mu_{ij})^T + \log\{|\Sigma_{ij}|\} \quad (17)$$

Since the deterministic energy ratio α represents the share of energy in the spectrum by the deterministic components, $(1-\alpha)$ thus represent the share by the indeterministic component. Hence, in equation (16), λ_{1j} can be replaced by α_j, and λ_{2j} can be replaced by $(1 - \alpha_j)$. Then, the final decision rule can be rewritten as:

$$\hat{c} = \arg\min_j \{\alpha_j D_{1j}(F_1) + (1 - \alpha_j) D_{2j}(F_2)\} \quad (18)$$

If a texture is regular, the deterministic component will be emphasized. Otherwise, if a texture is random, the indeterministic component will be emphasized.

4. Experiments

This section presents experiments that test the recognition capability of the method and compared its performace with the single model feature methods.

The texture images used in this experiment are from the Brodatz texture album [2], in which each of the 112 Brodatz texture is considered to form one texture class. Each classes have 200 sample images of size 128×128 which are sampled from the original 512×512 image. All these 200 sample images are divided into two equal subsets, for training and testing. Each training set consists of 100 images which are sampled from the top half of the original 512×512 texture image, another 100 images sampled from the bottom half of the texture image for testing.

The classification accuracy is calculated from the confusion matrix which contains information about the correct classification and misclassification of all classes. Confusion matrix is a $M \times M$ matrix, where M is the number of classes. For 100% classification this matrix should be diagonal. Results for the classification were calculated based on individual feature vector and the combination of them by consensus theory.

Two sets of experiments were performed. In the first experiment, 16 images were selected from Brodatz album.

Table 1. the deterministic energy ratio α of each texture class

class	α	class	α
D021	0.78056	D080	0.21912
D053	0.76357	D019	0.21766
D020	0.70527	D078	0.21152
D006	0.65546	D057	0.12348
D055	0.45248	D092	0.08768
D083	0.42371	D024	0.05528
D085	0.37742	D029	0.04076
D084	0.29789	D009	0.01876

They exhibit more homogenous and have various appearance with different value of deterministic energy ratio (Table 1). Figure 3 shows the 16 images sorted by the deterministic energy ratio. It can be concluded that the larger deterministic energy ratio generally corresponds to structural pattern while the small one corresponds to stochastic pattern. It substantiates that the use of the deterministic energy ratio as the flexible weight in the multi-model feature combination is reasonable. The classification results are shown in Table 2.

Table 2. Comparison of the average correct classification rate of different method for 16 Brodatz textures (%)

method	Gabor	GMRF	Combination
Rate	97.36	94.88	99.63

The second experiment used the whole Brodatz database (112 classes). In order to evaluate the effectiveness of the multi-model feature integration, the average correct classification rates for different types of texture images in the database are calculated. 112 classes of textures are partitioned into three groups by the deterministic energy ratio α.

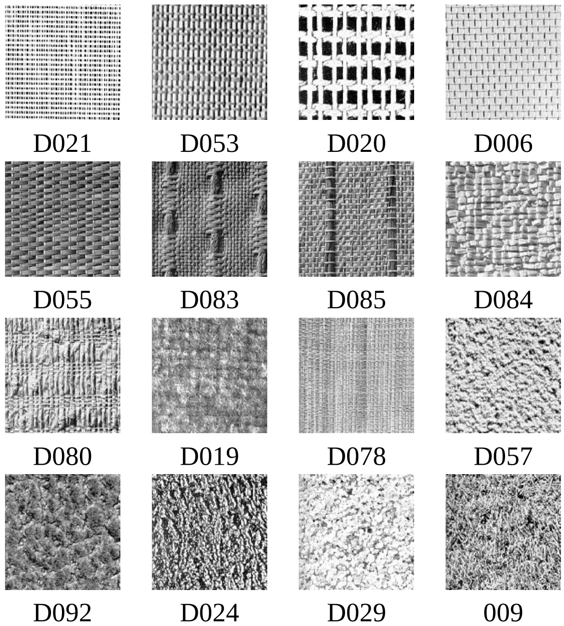

Figure 3. Texture images from Brodatz album used in this experiment

Group 1 contains 41 texture classes with $\alpha > 0.4$. The texture images in this group exhibit structural property. Group 3 contains 19 texture classes with $\alpha < 0.18$. The texture images in this group exhibit random property. The rests are grouped in the Group 2. The classification results are shown in Table 3 and Table 4. The group limits are determined manually by observing the α values for all 112 textures.

Table 3 gives an overview of the average correct classification rate of all the 112 texture classes. As expected that the multi-model feature integration improved the correct classification rate significantly.

Table 3. Comparison of the average correct classification rate of different method for 112 Brodatz textures (%)

method	Gabor	GMRF	Combination
Rate	51.84	47.21	61.32

Two conclusions can be drawn from the results shown in Table 4. One is that the Gabor model is more efficient than GMRF model in describing the structural textures (results of Group 1). In contrast, the GMRF model is more efficient than Gabor model in describing the stochastic textures (results of Group 3). Yet, the multi-model feature integration method can improve the correct classification rate significantly in both kinds of textures.

Table 4. Comparison of the correct classification rate of each groups (%)

Method	Gabor	GMRF	Combination
Group1	82.20	63.80	87.02
Group2	58.23	51.00	68.72
Group3	69.23	74.11	80.21

5. Conclusions

Since the information provided by a single model feature could be incomplete or imprecise, it is of interest to integrate multi-model features to obtain a better description of a texture image. A texture classification method using multi-model feature integration by consensus theory is proposed in this paper. The texture field is assumed to be a realization of a regular homogeneous random field, which is characterized in general by a mixed spectral distribution. Different from existing methods, the feature sets are extracted from the different components of texture which are modelled as different texture models and a weighting parameter. The deterministic energy ratio determined from the spectrum distribution function is used as the flexible weight based on the consensus theory. Experimental results indicate that this method improved the classification accuracy significantly for a wider variety of texture types.

References

[1] J. A. Benediktsson and P. H. Swain. Consensus theoretic classification methods. *IEEE Transactions on Systems Man and Cybernetics*, 22(4), July/August 1992.

[2] P. Brodatz. *Textures: A Photographic Album for Designers and Artists*. Dover, 1966.

[3] F.M.Vilnrotter, R. Nevatia, and K. E. Price. Structural analysis of natural textures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8, 1986.

[4] J. M. Francos, A. Z. Meiri, and B. Porat. On a wold-like decomposition of 2-d discrete random fields. In *Proc. of the IEEE Conf. on Acoustics, Speech and Signal Processing*, volume 5, pages 2695–2698, 1990.

[5] J. M. Francos, A. Z. Meiri, and B. Porat. A unified texture model based on a 2-d Wold-like decomposition. *IEEE Transactions on Signal Processing*, 41(8):2665–2678, 1993.

[6] R. L. Kashyap and R. Chellappa. Estimation and choice of neighbors in spatial interaction models of images. *IEEE Transactions on Information Theory*, IT-29, Jan. 1983.

[7] F. Liu and R. W. Picard. Periodicity, directionality, and randomness: Wold features for image modeling and retrieval.

IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(7):722–733, 1996.

[8] B. S. Manjunath and W. Y. Ma. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(8), August 1996.

[9] L. S. Ng, M. S. Nixon, and J. N. Carter. Texture classification using combined feature sets. In *Image Analysis and Interpretation, IEEE Southwest Symposium on*, volume 2, pages 103–108, 1998.

[10] A. H. S. Solberg. Texture fusion and classification based on flexible discriminant analysis. In *Proceedings of 13th IEEE International Conference on Pattern Recognition*, volume 2, pages 596–600, 1996.

[11] A. H. S. Solberg and A. K. Jain. Texture fusion and feature selection applied to sar imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 35(2), March 1997.

[12] T.Matsuyama, S. Miura, and M. Nagao. Structural analysis of natural textures by fourier transformation. *Computer Vision, Graphics, Image processing*, 21, 1983.