Wavelet Based Texture Classification

Nicu Sebe Michael S. Lew
Leiden Institute of Advanced Computer Science,
Niels Bohrweg 1, 2333 CA, Leiden, The Netherlands
{nicu mlew}@liacs.nl

Abstract

Textures are one of the basic features in visual searching and computational vision. In the literature, most of the attention has been focussed on the texture features with minimal consideration of the noise models. In this paper we investigated the problem of texture classification from a maximum likelihood perspective. We took into account the texture model, the noise distribution, and the inter-dependence of the texture features. Our investigation showed that the real noise distribution is closer to an Exponential than a Gaussian distribution, and that the L_1 metric has a better retrieval rate than L_2. We also proposed the Cauchy metric as an alternative for both the L_1 and L_2 metrics. Furthermore, we provided a direct method for deriving an optimal distortion measure from the real noise distribution, which experimentally provides consistently improved results over the other metrics. We conclude with results and discussions on an international texture database.

1. Introduction

Texture analysis is important in many applications of computer image analysis for classification, detection or segmentation of images based on local spatial patterns of intensity or color. Textures are replications, symmetries and combinations of various basic patterns or local functions, usually with some random variation. Textures have the implicit strength that they are based on intuitive notions of visual similarity. This means that they are particularly useful for searching visual databases and other human computer interaction applications. However, since the notion of texture is tied to the human semantic meaning, computational descriptions have been broad, vague and sometimes conflicting.

The method of texture analysis chosen for feature extraction is critical to the success of the texture classification. However, the metric used in comparing the feature vectors is also clearly critical. Many methods have been proposed to extract texture features either directly from the image statistics, e.g. co-occurrence matrix, or from the spatial frequency domain [13]. Ohanian and Dubes [8] studied the performance of four types of features: Markov Random Fields parameters, Gabor multi-channel features, fractal-based features and co-occurrence features. Comparative studies to evaluate the performance of some texture measures were made in [10], [9]. Recently there has been a strong push to develop multiscale approaches to the texture problem. Smith and Chang [12] used the statistics (mean and variance) extracted from the wavelet subbands as the texture representation. To explore the middle-band characteristics, tree-structured wavelet transform was used by Chang and Kuo in [2]. Ma and Manjunath [7] evaluated the texture image annotation by various wavelet transform representations, including orthogonal and bi-orthogonal, tree-structured wavelet transform, and Gabor wavelet transform (GWT). They found out that Gabor transform was the best among the tested candidates, which matched the human vision study results [1].

Most of these previous studies have focussed on the features, but not on the metric, nor on modeling the noise distribution. In this paper, we study the effect of the noise and the metric and their interrelationship within the maximum likelihood paradigm, using Gabor and wavelet features.

1.1. Texture Features

Gabor filters produce spatial-frequency decompositions that achieve the theoretical lower bound of the uncertainty principle. They attain maximum joint resolution in space and spatial-frequency bounded by the relations $\Delta_x^2 \cdot \Delta_u^2 \geq \frac{1}{4\pi}$ and $\Delta_y^2 \cdot \Delta_v^2 \geq \frac{1}{4\pi}$, where $[\Delta_x^2, \Delta_y^2]$ gives resolution in space and $[\Delta_u^2, \Delta_v^2]$ gives resolution in spatial-frequency. In addition to good performances in texture discrimination and segmentation, the justification for Gabor filters is also supported through psychophysical experiments. Texture analyzers implemented using 2-D Gabor functions produce a strong correlation with actual human segmentation [11]. Furthermore, the receptive visual field profiles are adequately modeled by 2-D Gabor filters [3].

Gabor functions are Gaussians modulated by complex sinusoids. In two dimensions they take the form [3]:

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

The Gabor filter masks can be considered as orientation and scale tunable edge and line detectors. The statistics of these microfeatures in a given region can be used to characterize the underlying texture information. A class of such self similar functions referred to as Gabor wavelets is discussed in [6]. This self-similar filter dictionary can be obtained by appropriate dilations and rotations of $g(x,y)$ through the generating function,

$$g_{mn}(x,y) = a^{-m} g(x', y'), \quad m = 0, 1, \cdots, S-1 \quad (2)$$

$$x' = a^{-m}(x\cos\theta + y\sin\theta), \quad y' = a^{-m}(-x\sin\theta + y\cos\theta)$$

where $\theta = n\pi/K$, K the number of orientations, S the number of scales in the multiresolution decomposition, and $a = (U_h/U_l)^{-1/(S-1)}$ with U_l and U_h the lower and the upper center frequencies of interest.

An alternative to gain in the trade-off between space and spatial-frequency resolution without using Gabor functions is with a wavelet filter bank. A two-band quadrature mirror filter (QMF) bank utilizes orthogonal analysis filters to decompose

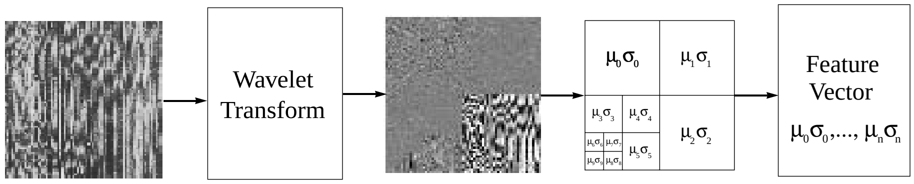


Figure 1. Texture classifier for Brodatz textures samples using QMF-wavelets based features

data into low-pass and high-pass frequency bands. Applying the filters recursively to the lower frequency bands produces wavelet decomposition as illustrated in Figure 1.

1.2. Texture Classification

The wavelet transformation involves filtering and subsampling. A compact representation needs to be derived in the transform domain for classification and retrieval. The mean and the variance of the energy distribution of the transform coefficients for each subband at each decomposition level are used to construct the feature vector (Figure 1). Let the image subband be $W_n(x,y)$, with n denoting the specific subband. Note that in the case of GWT the index n is regarded as mn with m indicating a certain scale and n a certain orientation. The resulting feature vector $f = \{\mu_n, \sigma_n\}$ with,

$$\mu_n = \int |W_n(x,y)|\,dx\,dy \tag{3}$$

$$\sigma_n = \sqrt{\left(\int |W_n(x,y)| - \mu_n\right)^2 dx\,dy} \tag{4}$$

Consider two image patterns i and j and let $f^{(i)}$ and $f^{(j)}$ represent the corresponding feature vectors. The distance between the two patterns in the features space is:

$$d(i,j) = \sum_n \left(\left| \frac{\mu_n^{(i)} - \mu_n^{(j)}}{\alpha(\mu_n)} \right|_{L_k} + \left| \frac{\sigma_n^{(i)} - \sigma_n^{(j)}}{\alpha(\sigma_n)} \right|_{L_k} \right) \tag{5}$$

where $\alpha(\mu_n)$ and $\alpha(\sigma_n)$ are the standard deviations of the respective features over the entire database and L_k is a notation for all possible metrics that can be used, e.g. L_1, L_2, L_c.

2. Maximum likelihood approach

Suppose we extract N pair sample blocks from the textures in the database. Consider x_i and y_i, $i = 1, \cdots, N$ the feature vectors of two samples extracted from the same texture M. Considering n_i as the "noise" vector obtained as the absolute difference between corresponding elements in x_i and y_i, namely $n_i = |x_i - y_i|$, the similarity probability can be defined:

$$P(X,Y) = \prod_{i=1}^{N} \{\exp[-\rho(n_i)]\} \tag{6}$$

where ρ is the negative logarithm of the probability density of the noise and X and Y are the sets of all feature vectors x_i and y_i extracted from the same texture.

According to (6) we have to find the probability density function of the noise that maximizes the similarity probability: *maximum likelihood* estimate for the noise distribution [4]. Taking the logarithm of (6) we have to minimize the expression:

$$\sum_{i=1}^{N} \rho(n_i) \tag{7}$$

Note that when the Exponential and Gaussian distributions are used in (7), we arrive at the L_1 and L_2 metrics, respectively. A distribution with more extensive tails is the Cauchy distribution, and the corresponding metric L_c is given by the expression:

$$L_c(X,Y) = \sum_{i=1}^{N} \log(\mathbf{a}^2 + (x_i - y_i)^2) \tag{8}$$

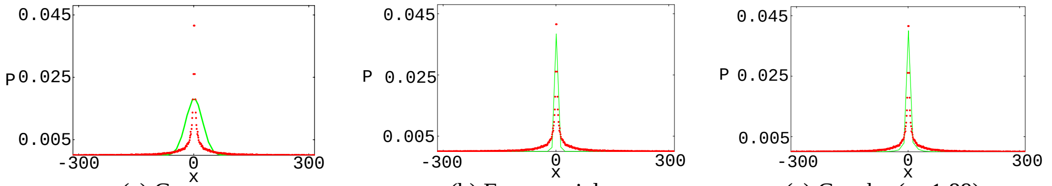
where **a** is a parameter which determines the height and the tails of the distribution. For a general noise distribution, considering ρ as the negative logarithm of the probability density of the noise, the corresponding metric is given by (7). In practice, the probability density of the noise can be estimated from the normalized histogram of the absolute differences.

3. Experiments

The textures used in the experiments are the 112 Brodatz textures. The database was formed by randomly subsampling 20 samples of 128×128 pixels in size from the 112 original textures, resulting in a classification in 112 different classes of 2240 random samples.

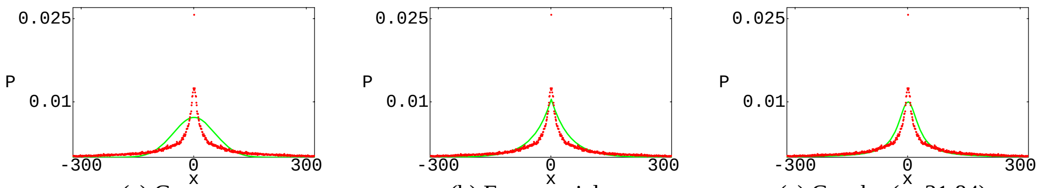
The setup of our experiments was the following. First the ground truth was known since the samples were extracted from the texture classes. The ground truth was split into two non-overlapping sets: the training set and the test set. In our experiments the training set consisted of 1000 samples from the ground truth. Second, for each sample in the training set a feature vector was extracted using the scheme in Figure 1. Note that in this experiments, the feature vector was composed from two independent features: the mean and the variance. For each of them the real noise distribution was estimated as the normalized histogram of the absolute difference of corresponding elements from the feature vectors in the training set. The Gaussian, Exponential and Cauchy distributions were fitted to each real noise distributions using the Chi-square test. We selected the model distribution which had the best fit and its corresponding metric (L_k) (see (5)) was used in ranking. The ranking was done using only the test set. It is important to note that for real applications, the parameter in the Cauchy distribution was found when fitting this distribution to the real distribution from the training set. This parameter setting was used for the test set and any further comparisons in the application.

As noted in Section 2 it is also possible to create a metric based on real noise distribution using maximum likelihood theory. Consequently, we denoted the maximum likelihood (ML) metric as (7) where ρ is the negative logarithm of the probability density function of the noise, approximated as the normalized histogram of the absolute differences from the training set. Note that there were two ML metrics calculated, one from



(a) Gauss (b) Exponential (c) Cauchy (**a**=1.89)

Figure 2. Noise distribution for mean feature in QMF-wavelets compared with the best fit Gaussian (a) (approximation error is 0.279), best fit Exponential (b) (approximation error is 0.207) and best fit Cauchy (c) (approximation error is 0.174)



(a) Gauss (b) Exponential (c) Cauchy (**a**=31.84)

Figure 3. Noise distribution for variance feature in QMF-wavelets compared with the best fit Gaussian (a) (approximation error is 0.036), best fit Exponential (b) (approximation error is 0.0255) and best fit Cauchy (c) (approximation error is 0.023)

the mean distribution and the other one from the variance distribution. It is also interesting to note that metric values were already normalized through the histogram so the normalization factors in (5) (the standard deviations) were not necessary.

We applied the theoretical results described in Section 2 in two experiments. First we determined the influence of the similarity noise model for texture classification using QMF-wavelets. Second, we investigated the Gabor wavelet transform applied to texture classification.

Recall that our database was composed by randomly extracting 20 subsamples from the 112 original textures. When doing classification, in the ideal case all the top 19 retrievals were from the same large image. The performance was measured in term of the average retrieval rate defined as the percentage of retrieving the 19 correct patterns when top n matches were considered.

3.1. Similarity Noise for QMF-wavelet transform

A QMF wavelet filter bank was used for texture classification by Kundu [5]. The authors identified several properties of the QMF filter bank as being relevant to texture analysis: orthogonality and completeness of basic functions, filter outputs that are spatially localized and the reduction of complexity afforded by decimation of filter outputs. In our implementation we used five levels of decomposition of the wavelet transform. We extracted the mean and the variance of each subband in a 32 (16 subbands $\times$ 2) dimensional feature vector.

As noted before, we had to compute two similarity noise distributions corresponding to mean and variance features. The similarity noise distributions were displayed in Figure 2 and 3. The similarity noise distribution was obtained as the normalized histogram of differences between the corresponding feature elements from the training set.

For both features, the Exponential had a better fit to the noise distribution than the Gaussian. Consequently, this implies that L_1 should have a better retrieval rate than L_2. The Cauchy distribution was the best fit overall and the results obtained with L_c reflect this. Figure 4 presents the average retrieval rate for the correct patterns when top n matches are considered. This results are also contained in Table 1. Note that using ML we obtained the best average retrieval.

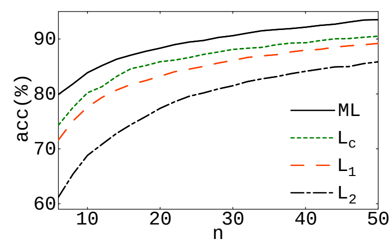


Figure 4. Average retrieval rate using QMF-wavelets

Top	5	10	25	50
L_2	62.43	68.86	78.83	85.14
L_1	72.36	76.34	81.41	89.62
L_c	76.32	79.15	83.67	90.18
ML	80.06	83.58	88.66	94.24

Table 1. Comparison of retrieval performance using QMF-wavelets for different metrics

3.2. Similarity Noise for Gabor Wavelet Transform

A Gabor wavelet transform (GWT) enables us to obtain image representations which are locally normalized in intensity and decomposed in spatial frequency and orientation. It thus provides a mechanism for obtaining (1) invariance under intensity transformations, (2) selectivity in scale by providing a pyramid representation and (3) it permits investigation of the local oriented features. In this paper, for the non-orthogonal Gabor wavelet transform we used 4 scales (S=4) and 6 orientations/scale ($K = 6$).

The mean and the variance of the energy distribution of the transform coefficients for each subband at each decomposition level were used to construct a 48 ($6\times4\times2$) dimensional feature vector. We calculated the similarity noise distribution for both features and fitted them with the model distributions. As seen from Table 2, the Cauchy distribution was the best match for the measured noise distribution. The Exponential was a better match than the Gaussian.

Feature	Gauss	Exponential	Cauchy
Mean	0.186	0.128	0.114
Variance	0.049	0.035	0.027

Table 2. The approximation error for the noise distribution using GWT

Figure 5 presents the average retrieval rate when different metrics were used. Note that L_c had better retrieval rate than L_1 and L_2. *ML* provided the best results.

In summary, L_c performed better than the analytic distance measures, and the *ML* metric performed best overall. Note that the results obtained with GWT were superior to the ones obtained using QMF-wavelet transform.

4. Discussion and Conclusions

This research is differentiated from the previous works in texture classification in that we had investigated the role of the underlying noise distribution and corresponding metric in the paradigm of maximum likelihood. Our experiments on both the noise distribution and the retrieval rates from using a particular distortion measure provided strong evidence of the maximum likelihood theory.

In the maximum likelihood paradigm, it is provable that the Gaussian distribution results in the L_2 metric, the Exponential distribution results in the L_1 metric, and the Cauchy distribution results in the L_c metric. By linking the distributions with the metrics, we can directly show why a particular metric would outperform another metric. Specifically, the metric which will have the best retrieval rate should be the metric whose distribution best matches the real noise distribution from the test set.

Here in this paper, we have found that the noise distribution is modeled better by the Cauchy distribution than the Exponential or Gaussian distributions. Consequently, among the analytic distortion measures, L_c consistently had a better retrieval rate than L_1 or L_2.

Given that the modeling of the real noise distribution is linked with the retrieval rate, the next logical question is, "What is the retrieval rate when we directly model the real noise distribution?" It was also validated that the highest retrieval rate occurs when we use an approximate, quantized model for the real noise distribution. The corresponding distortion measure clearly outperformed the rest of the distortion measures.

Minimizing the *ML* metric is optimal with respect to maximizing the likelihood of the difference between feature elements when the real noise distribution is representative. Therefore, the breaking points occur when there is no ground truth, or when the ground truth is not representative.

Thus, the primary contribution of this paper is the theoretical link between the noise distribution and the corresponding distortion measure. In this context, we showed that prevalent Gaussian distribution assumption is often invalid and proposed the Cauchy metric as an alternative for both the L_1 and L_2 metrics. Furthermore, we provided a method for deriving an optimal distortion measure from the real noise distribution, which experimentally provided consistently improved results over the other metrics.

References

[1] J. Beck, A. Sutter, and R. Ivry. Spatial frequency channels and perceptual grouping in texture segregation. *CVGIP*, 37, 1987.

[2] T. Chang and C. Kuo. Texture analysis and classificatrion with tree-structured wavelet transform. *IEEE Trans. Image Proc.*, 2(4):429–441, 1993.

[3] J. Daugman. Entropy reduction and decorrelation in visual coding by oriented neural receptive fields. *IEEE Trans. on Biomedical Engineering*, 36(1), 1989.

[4] P. Huber. *Robust Statistic*. Wiley, 1981.

[5] A. Kundu and J.-L. Chen. Texture classification using QMF bank-based subband decomposition. *CVGIP: Graphical Models and Image Processing*, 54(5), 1992.

[6] W. Ma and B. Manjunath. Texture features and learning similarity. *CVPR*, 1996.

[7] Y. Ma and B. Manjunath. A comparison of wavelet transform features for texture image annotation. *ICIP*, 1995.

[8] P. Ohanian and R. Dubes. Performance evaluation for four classes of textural features. *Pattern Recognition*, 25:819–833, 1992.

[9] T. Ojala, M. Pietikainen, and D. Harwood. A comparative study of texture measures with classification based on feature distribution. *Pattern Recognition*, 29:51–59, 1996.

[10] T. Reed and J. Du Buf. A review of recent texture segmentation and feature extraction techniques. *CVGIP*, 57(3):359–373, 1993.

[11] T. Reed and H. Wechsler. Segmentation of textured images and gestalt organization using spatial/spatial-frequency representations. *IEEE Trans. on PAMI*, 12(1):1–12, 1990.

[12] J. Smith and S.-F. Chang. Transform features for texture classification and discrimination in large image databases. *ICIP*, 1994.

[13] L. Van Gool, P. Dewaele, and A. Oosterlinck. Texture analysis. *CVGIP*, 29(3):336–357, 1985.