MVA '92 IAPR Workshop on Machine Vision Applications Dec. 7-9,1992, Tokyo

TEXTURE CLASSIFICATION BY WAVELET PACKET SIGNATURES

Andrew Laine and Jian Fan

Center for Computer Vision and Visualization
Computer and Information Sciences Department
Computer Science and Engineering Building, Room 301
University of Florida
Gainesville, FL 32611-2024, USA

ABSTRACT

In this paper, we investigated six distinct partitions of wavelet packets selected from a complete decomposition and measured their performance in terms of sensitivity and selectivity for the classification of twenty-five natural textures. Both energy and entropy metrics were computed for each wavelet packet node and included in distinct features for representation. Each subset of wavelet packet nodes reflected a specific scale and orientation sensitivity. Wavelet packet representations for twenty-five natural textures were classified without error by a simple two-layer network classifier.

Texture signatures computed from wavelet packet energy were more reliable than entropy signatures computed from the same wavelet packet basis. A longer analyzing function (D_{20}) was shown to be more efficient in representation and discrimination than a similar function of shorter length (D_6). Energy representations computed from the standard set of wavelet nodes alone (17 features) were sufficient for errorless classification of the twenty-five textures included in our study. The reliability exhibited by texture signatures based on wavelet packets suggest the potential to realize robust classification and subtle discrimination.

Index Terms–Feature extraction, texture analysis, texture classification, wavelet transform, wavelet packet, neural networks.

1 Introduction.

Texture is an important characteristic for the analysis of many types of images including natural scenes, remotely sensed data and biomedical imaging modalities. The perception of texture is believed to play an important role in the human visual system for recognition and interpretation. Despite the lack of a complete and formal definition of texture, a large number of approaches for texture classification have been suggested [1-7].

Previous methods of analysis for accomplishing texture classification may be roughly divided into three categories : statistical, structural and spectral. [10][11][12][13]. Experimental evidence on human and mammalian vision support some sort of spatial-frequency (scale) analysis that maximizes the simultaneous localization of energy in both spatial and frequency domains.Along this line, several important studies have been reported recently.[16] [17][18] [19][20][7][8][9]. In this paper, we present a novel method of texture classification by multiresolution representations from wavelet analysis.

Wavelet theory provides a precise and unified framework for spatial-scale analysis. Carter [21] first reported texture classification results using Morlet and Mexican hat wavelets. He achieved 98 percent accuracy on 6 types of natural textures. However, these wavelets were not orthonormal, and the Mexican hat wavelet lacked direction selectivity. In this paper, these drawbacks are overcome by using orthonormal and compactly supported wavelets. The advantages are twofold. First, since the representation features at each scale are obtained by decomposing a signal (image) onto an orthonormal basis [22], correlation between scales is avoided. Second, orientation selectivity is built into the two-dimensional orthonormal wavelets included in our study. Experimentally, these capabilities demonstrated increased sensitivity and selectivity for reliable texture discrimination.

In this paper, we introduce a methodology for identifying texture representations based on wavelet packet analysis [23][24]. Wavelet packets are a generalization of orthonormal and compactly supported wavelets [27] [22]. In this paper we show that such an analyses provides a powerful method for accomplishing robust texture classification. The efficacy of the technique is demonstrated by classification without error of *25 natural textures*.

2 Wavelet packet signatures.

Wavelet packet analysis, pioneered by R. Coifman *et al* [23][24] has been successfully used for data compression [25] [26]. Wavelet packets may be described by a collection of functions $\{W_j(x)|j \in Z^+\}$ obtained from

$$2^{\frac{p-1}{2}} W_{2n}(2^{p-1}x - l) = \sum_m h_{m-2l} 2^{\frac{p}{2}} W_n(2^p x - m) \quad (1)$$

$$2^{\frac{p-1}{2}} W_{2n+1}(2^{p-1}x - l) = \sum_m g_{m-2l} 2^{\frac{p}{2}} W_n(2^p x - m) \quad (2)$$

where p is a scale index, l is a translation index, $W_0(x) = \phi(x)$, $W_1(x) = \psi(x)$, $\phi(x)$ is a scaling function and $\psi(x)$ is a basic wavelet [27] [22]. The discrete filters h_k and g_k are quadrature mirror filters [28][27] [22].

We can show [23] that such wavelet packets are orthonormal in $L^2(R)$ and serve as bases similar to sinusoid functions in Fourier analysis. Furthermore, wavelet packets are well localized in both time and frequency and thus provide an attractive alternative to pure frequency (Fourier) analysis.

The inverse relationship between wavelet packets of different scales can be shown by,

$$\begin{aligned} 2^{\frac{p}{2}} W_n(2^p x - k) &= \sum_l h_{k-2l} 2^{\frac{p-1}{2}} W_{2n}(2^{p-1}x - l) \\ &+ \sum_l g_{k-2l} 2^{\frac{p-1}{2}} W_{2n+1}(2^{p-1}x - l). \end{aligned} \quad (3)$$

Analogous to Fourier methods, any function $f(x) \in L^2(R)$ can be decomposed onto a wavelet packet basis. The coefficients of this decomposition are simply the inner products of $f(x)$ with distinct wavelet packets. For example, coefficients from the inner product $\langle f(x), W_n(2^p x - k)\rangle$ indicate the intensity of this component in $f(x)$. An approximation of an original function $f(x)$ using wavelet packet W_n at scale 2^p can be written as,

$$A_n^{2^p} f(x) = \sum_k S_{n,k}^p \, 2^{\frac{p}{2}} \, W_n(2^p x - k) \quad (4)$$

where

$$S_{n,k}^p = 2^{\frac{p}{2}} \int_{-\infty}^{\infty} f(x) \, W_n(2^p x - k) dx \quad (5)$$

and $\overline{x}$ denotes the complex conjugate of x. Next, we show how wavelet packets may be computed efficiently.

From equation (3) , we have

$$S_{n,k}^p = \sum_l h_{k-2l} \, S_{2n,l}^{p-1} + \sum_l g_{k-2l} \, S_{2n+1,l}^{p-1} \, . \quad (6)$$

Using equations (1) and (2), coefficients at coarser scales are calculated by

$$S_{2n,l}^{p-1} = \sum_m h_{m-2l} \, S_{n,m}^p \quad (7)$$

$$S_{2n+1,l}^{p-1} = \sum_m g_{m-2l} \, S_{n,m}^p \quad (8)$$

Note that for standard wavelet decompositions [22], only two wavelet packets W_0 and W_1 are used. In this case the index n is restricted to $n = 0$, and only S_0^p are decomposed from equations (7) and (8).

For discrete signals, we treat the original discrete signal as the set of wavelet packet coefficients for the first scale ($p = 0$), and then apply the technique described above.

The basis functions are obtained by translation and scale change. They remain well localized in both time (spatial) and frequency domains and thus represent scale and spatial information. Therefore, a complete tree presents the distribution of a signal within a scale space continuum. Note that the total number of coefficients in a complete tree decomposition is exactly equal to the number of points (pixels) in an original signal.

Energy distributions within transform spaces have been applied in Fourier analysis. Since wavelet packets form orthogonal bases, their decompositions also preserve energy. It is easy to show that

$$\sum_k (S_{n,k}^p)^2 = \sum_l (S_{2n,l}^{p-1})^2 + \sum_l (S_{2n+1,l}^{p-1})^2 . \quad (9)$$

Therefore, if we define an energy measure as $E_n^p = \sum_k (S_{n,k}^p)^2$, then $E_n^p = E_{2n}^{p-1} + E_{2n+1}^{p-1}$.

Our stratagy was to first compute the energy associated with each wavelet packet node. We hypothesized that the energy pattern distributed in scale space shall provide unique information, and support a representation (signature) for classification. Thus, a signature was a vector consisting of a set of *energy features*. In the next section, we shall demonstrate that such signatures provide a powerful and efficient means to accomplish signal classification.

An alternative measure is *entropy*, defined by, $H(x) = -\sum_k |x_k|^2 \log |x_k|^2$. This measure was previously proposed in [10] as a feature for texture analysis, and has also been used in [29] to identify a "best basis" for building wavelet packet libraries for signal compression. In this paper, we compare the entropy and energy measures described above for their performance in texture discrimination.

The extension into 2–D signals is straight forward by using a special class of separable 2–D wavelet packets. In this case, the energy preserving equation is specified by the sum

$$E_{n,m}^p = E_{2n,2m}^{p-1} + E_{2n,2m+1}^{p-1} + E_{2n+1,2m}^{p-1} + E_{2n+1,2m+1}^{p-1}. \quad (10)$$

Orientation selectivity is embedded in the tensor product of the lowpass filter h and highpass filter g, and therefore energy distributions are captured in all orientations.

3 Methodology.

3.1 Texture selection and sampling.

Twenty-three distinct natural textures were selected from the Brodatz album [30] and two additional textures from public archive. The complete set of twenty-five textures is shown in Figure 1. Each selected texture was digitized and stored as a 512×512, 8bit/pixel digital image. Each sample texture was then broken down into 128×128 sub-samples. Our selection criteria was such that each selected texture pattern maintained a certain degree of spatial periodicity within its 128×128 sample size.

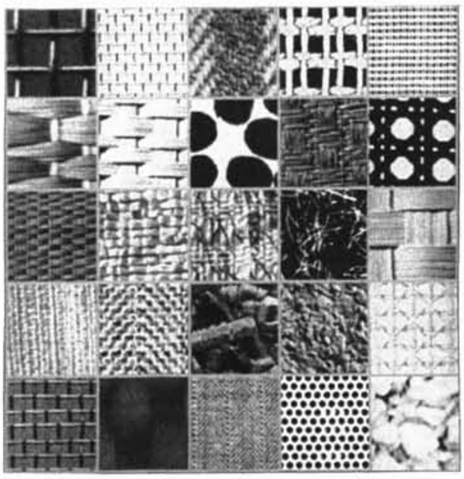

Figure 1: Twenty-five natural textures.
row1: D1,D6,D11,D20,D21;
row2: D56,D65,D75,D82,D101;
row3: D55,D84,D81,D110,D64;
row4: D78,D17,D5,D4,D52;
row5: D14,D8,D16,O1,O2.

A two-layer neural network and a minimum-distance classifier were used to accomplish supervised classification. To obtain a large amount of data for training the classifiers, we adapted a method of overlapped sampling. We extracted 64 sub-samples of size 128×128 (pixels) from each original 512×512 sample texture.

3.2 Patitions of wavelet packet space.

A complete set of wavelet packets were computed for each 128 x 128 subsample. Discrete filters D_6 and D_{20} were obtained from Daubechies [27]. Due to downsampling at each decomposition step, the size of each subsample was reduced by a factor of four. Thus, a subsample at level four, consisted of 64 (8 x 8) coefficients. Recall that each parent node had four children. Therefore, a complete five-level decomposition (levels 0 through 4) had exactly 341 wavelet packet nodes.

We measured the energy (and entropy) contributed by each wavelet packet node, and treated its real value as a distinct feature element. Thus, in our representation, the maximum number of features encoded for a sample texture consisted of a vector of 341 real values. However, we investigated the classification performance of each signature using distinct subsets of wavelet packet nodes. Each subset consisted of nodes reflecting a certain scale and orientation sensitivity. We considered six distinct partitions of wavelet packets selected from a complete decomposition tree (full recursion) and measured their performance in terms of sensitivity and selectivity for the classification of all twenty-five natural texures. Below, we identify the six partitions of wavelet packet nodes that provided distinct bases for feature representation:

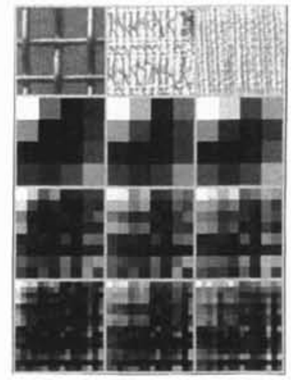

Figure 2: Log energy maps for wavelet packet nodes of levels 2,3 and 4.

1. *Complete set of wavelet packet nodes.* Each texture subsample was represented by a vector of 341 features.
2. *Standard wavelets.* As mentioned earlier, nodes of the standard wavelet decomposition (Mallat [22]) are a subset of a complete wavelet packet decomposition. In this case, the four leftmost nodes at each level of a complete wavelet packet decomposition tree were selected. Thus, each subsample was represented by only 17 features.
3. *Levels 1, 2 and 3.* In this case, each texture sample was represented by exactly 84 features. Energy features computed from wavelet packet nodes of levels 0 and 4 were discarded.
4. *Levels 2 and 3.* Each texture sample was represented by exactly 80 features.
5. *Level 3.* Each texture sample was represented by exactly 64 features.
6. *Level 4.* Each texture sample was represented by exactly 256 features. In this case, energy features were from a single coarse scale.

For each of the six representations above, we computed feature vectors for each of the 64 subsamples of each texture. Our database consisted of 42 samples for classification training and 22 samples for testing. Therefore in total, our study processed 1050 (42×25) sample signatures in training and 550 (22×25) sample signatures in testing classification performance.

In Figure 2, we show three texture samples, and the energy signatures corresponding to the wavelet packet of levels 2, 3 and 4 (from top to bottom).The energy map was obtained by first computing the logarithm for each feature value, then globally scaling the values within a 0-255 range. Therefore, Figure 2 shows a normalized energy distribution for each signature.

3.3 Discrimination using a simple minimum-distance classifier.

To decide the efficacy of wavelet packet signatures for texture classification, the performance of a simple minimum-distance classifier was evaluated. The single prototype minimum-distance classifier [33] was based on the assumption that each pattern class ω_k is representable by a prototype pattern $\mathbf{Z}_k$ (class center). The minimum-distance classifier assigned a pattern $\mathbf{X}$ of unknown classification to the class $\omega_{\mathbf{k}}$, if the distance D_k between $\mathbf{X}$ and $\mathbf{Z}_k$ was minimum among all possible class prototypes $\omega_j \neq \omega_k$. In addition, more sophisticated classifier, described in the next section, was evaluated.

3.4 Discrimination using a neural network classifier.

We examined the classification performance of each wavelet packet representation for several network topologies. We used a two-layer back propagation network [31] with a conjugate gradient function for error correction [32]. For each topology, the number of input nodes was matched to the dimension (number of values) of each wavelet packet representation. All network topologies had 25 output nodes, the total number of distinct textures targeted.

By using wavelet packet representations, we reduced the number of bits required for each original texture pattern by a factor of 240 ($\frac{128\times128\times8}{17\times32}$). Thus we were able to reduce the number of input nodes (bandwidth) of a neural network by a factor of 960 ($\frac{128\times128}{17}$).

4 Results and discussion.

In Table 1 we show the performance of three parameters included in our study: (a) Energy versus entropy based signature metrics, (b) a minimal (standard) and maximum (complete) number of wavelet packet nodes for representation, (c) D_{20} (long) versus D_6 (short) analyzing functions. When using complete wavelet packet representations, perfect classification for the twenty-five textures was observed regardless of the signature (energy versus entropy) or analyzing wavelet (D_6 versus D_{20}) included. However, when texture signatures were computed from a minimal (standard) number of wavelet packet nodes, perfect classification was observed only for the D_{20} analyzing wavelet. This demonstrates that (a) longer analyzing functions are more efficient for discriminating salient textural features (b) perfect classification is achievable by a minimal representation of energy (based on 17 wavelet packet nodes). For the textures included in our investigation, we observed that signatures computed from energy performed slightly better than entropy based representations.

In Table 2, we show the classification performance of energy signatures computed from redundant and complete sets of wavelet packet representations. Signatures computed from wavelet packets of level 3 alone (64 features) and level 4 alone (256 features) yielded perfect classification for all twenty-five textures. However, texture signatures computed from levels 1,2 and 3 (redundant representations) resulted

Table 1: Classification results comparing the performance of two signature metrics, number of wavelet packet nodes for representation and two analyzing functions.

Sig.	Selected Wavelet Packets	N	Analy. Func.	Num. of Err.	% Correct
E	Comp.	341	D_{20}	0	100
			D_6	0	100
	Stand.	17	D_{20}	0	100
			D_6	4	99.3
H	Comp.	341	D_{20}	0	100
			D_6	0	100
	Stand.	17	D_{20}	1	99.8
			D_6	5	99.1

E: Energy, H: Entropy, N: Number of Features

Table 2: Classification results for energy signatures computed from redundant and complete wavelet packet representations, using a D_{20} analyzing function

Selected Wavelet Packets	Number of Features	Number of Errors	% Correct
Level 1,2,3	84	1	99.8
Level 2,3	80	0	100
Level 3 only	64	0	100
Level 4 only	256	0	100

in a classification error. This suggests that redundancy may increase uncertainty (degrees of freedom) for the classifier employed in our study.

Table 3 compares the performance of network topologies for each (D_6 and D_{20}) analyzing function. Energy signatures for twenty-five textures were computed from a standard decomposition (17 wavelet packet nodes) and trained for classification. In the case of the D_6 analyzing function, errors were observed for all five configurations. Perfect classification was observed only for the D_{20} analyzing function when the network consisted of exactly three hidden nodes. Note that fewer training epochs were required for networks consisting of more hidden nodes. In general, signatures computed from D_6 analyzing functions required significantly more training time than signatures obtained from D_{20} analyzing functions. This demonstrates that longer analyzing functions provide a more efficient representation for texture discrimination.

The simple minimum-distance classifier using wavelet signatures from level 3 alone was able to discriminate 550 sample patterns (22 samples/texture $\times$ 25 textures) with 96 percent (not shown above) accuracy! This result comfirmed that texture signatures computed from wavelet packet energies alone, are highly efficient representations for texture classi-

Table 3: Network training times and topologies for classification using D_6 and D_{20} analyzing functions.

	Hidd. Nodes	Training Epoches	Train. Time (sec.)	Num. of Err.	% Correct
D_6	1	225	50740	1	99.8
	2	226	53770	3	99.5
	3	102	24828	4	99.3
	5	32	8300	8	98.5
D_{20}	1	153	35231	2	99.6
	2	125	29857	1	99.8
	3	21	5203	0	100
	5	33	8822	1	99.8

fication.

5 Summary and Conclusions.

Wavelet packet representations for twenty-five natural textures were classified without error by a simple two-layer network classifier. A longer analyzing function was shown to be more efficient in representation and discrimination than a similar function of shorter length. Experimentally, we observed that a neural network classifier performed best in terms of accuracy and minimal training time when configured with exactly three hidden units.

Surprizingly, our results showed that energy representations computed from the standard set of wavelet nodes alone were sufficient for errorless classification. However, finer discrimination may be more strongly supported by additional subsets of wavelet packets. We suggest that identifying an "optimal" set of wavelet packets for texture representation may depend on the aggregate of textural features targeted for classification. Thus, similar textures consisting of variations mostly at finer scales may be best discriminated by representations computed from wavelet packets of higher levels.

With respect to the 128 x 128 pixels/sample sizes included in this study, we suggest that representations computed from level three (64 feature values) have sufficient selectivity and sensitivity for autonomous texture classification. Texture signatures computed from wavelet packet energy performed better than entropy signatures computed from the same wavelet packet nodes.

References

[1] R.M.Haralick. "statistical and structural approaches to texture". *Proc. IEEE*, vol. 67, pp. 786–804, 1979.

[2] L.S.Davis. "image texture analysis techniques-a survey," in Digital Image Processing(Proc. NATO Advanced Study Inst.). J.C.Simon and R.M.Haralick, Eds., Bonas, France, June 23-July 4, 1980. Dordrecht, The Netherlands: D.Reidel, pp. 189-201

[3] R.Bajcsy. "computer description of textured surfaces," in *Proc. 3rd Inl. Joint Conf. Artificial Intell.,* Aug. 1973, 572-579

[4] C.H.Chen. "a study of texture classification using spectral features," in Proc. IEEE *6th Inl. Conf. Pattern Recongnition*, Munich, Germany, Oct. 19-22, 1982, pp. 1074-1077

[5] V.L.Vickers *et al* "a maximum likelihood approach to texture classification" *IEEE Trans.* vol. PAMI-4, pp. 61–68, 1982.

[6] R.Wilson *et al. Image segmentation and uncertainty.* Research Studies Press Ltd, 1988.

[7] M.Unser. *Sum and difference histograms for texture classification. IEEE Trans.*, vol. PAMI-8, pp. 118–125, 1986.

[8] M.Unser. *Local linear transforms for texture measurements. Sig. Processing*, vol. 11, pp. 61–79, July 1986.

[9] M.Unser *et al Multiresolution feature extraction and selection for texture segmentation. IEEE Trans.* vol. PAMI-11, pp. 717–728, 1989.

[10] R.M.Haralick *et al.* "textural features for image classification". *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, pp.610–621, 1973.

[11] J.S.Weszka *et al.* "comparative study of texture measures for terrain classification". *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, pp. 269–285, 1976.

[12] R.L.Kashyap *et al.* "a model-based method for rotation invariant texture classification". *IEEE Trans.* vol. PAMI-8, pp. 472–481, 1986.

[13] F.S.Cohen *et al.* "classification of rotated and scaled textured images using Gaussian Markov random field models". *IEEE Trans.* vol. PAMI-13, pp. 192–202, 1991.

[14] B.Julesz *et al.* "visual discrimination of textures with identical third-order statistics". *Biol Cybern*, vol. 31, pp. 137-140, 1978

[15] A.Gagalowicz. "a new method for texture fields synthesis: some applications to the study of human vision". *IEEE Trans.*, vol. PAMI-3, pp. 520-533, 1982

[16] M.R.Turner. "texture discrimination by Gabor functions". *Biol. Cybern.*, vol. 55, pp. 71–82, 1986.

[17] A.C.Bovik *et al.* "multichannel texture analysis using localized spatial filters". *IEEE Trans.* vol. PAMI-12, pp. 55–73, 1990.

[18] A.K.Jain *et al.* "unsupervised texture segmentation using Gabor filters". *Pattern Recognition* vol.24,No. 12, pp. 1167–1186, 1991.

[19] T.R.Reed *et al.* "segmentation of textured images and gestalt organization using spatial/spatial-frequency representations. *IEEE Trans.* vol. PAMI-12, pp. 1–12, 1990.

[20] J.R.Bergen *et al.* "computational modeling of visual texture segregation". in "computational models of visual processing". M.S.Landy and J.A.Movshon, Ed., *The MIT Press* , pp. 472–481, 1991.

[21] P.H.Carter. "texture discrimination using wavelets". In *SPIE Applications of digital image processing XIV*, volume vol. 1567, pp. 432-438, 1991.

[22] S.Mallat. "a theory for multiresolution signal decomposition : the wavelet representation". *IEEE Trans.* vol. PAMI-11, pp. 674–693, 1989.

[23] R.R.Coifman and Y.Meyer. "orthonormal wave packet bases". preprint, Yale Univ.,New Haven, CT, Aug., 1989.

[24] R.R.Coifman *et al.* "best-adapted wavelet packet bases". preprint, Yale Univ.,New Haven, CT, Feb. 1990.

[25] M.V.Wickerhauser. "acoustical signal compression with wavelet packets". preprint, Yale Univ.,New Haven, CT, Aug., 1989.

[26] R.R.Coifman *et al.* "entropy-based algorithms for best basis selection". *IEEE Trans. Infor. Theory*, vol. 38, pp. 713–718, 1992.

[27] I.Daubechies. "orthonormal bases of compactly supported wavelets". *Commun. Pure Appl. Math.*,, vol. XLI, pp. 909–996, 1988.

[28] R.E.Crochiere *et al.* *Multirate digital signal processing.* Prentice-Hall, 1983.

[29] R.R.Coifman *et al.* "signal processing and compression with wave packets". preprint, Yale Univ.,New Haven, CT, 1990

[30] P.Brodatz. *Textures-a photographic album for artists and designers.* Dover, 1966.

[31] D.E.Rumelhart *et al.* *Parallel distributed processing.* MIT Press, 1986.

[32] B.L.Kalman. "super linear learning in back propagation neural nets". Wucs-90–21, Washington Univ., St. Louis, MO, 1990.

[33] J.T.Tou *et al.* *Pattern recognition principles* Addison-Wesley, 1974, second printing 1977