ROTATION INVARIANT TEXTURE CLASSIFICATION USING BAMBERGER PYRAMIDS

José Gerardo Rosiles
Elec. and Comp. Eng. Dept.
Univ. of Texas at El Paso
El Paso, TX, 79968-0523
grosiles@utep.edu

Mark J. T. Smith
School of Elec. and Comp. Eng.
Purdue University
West Lafayette, IN 47907-2035
mjts@purdue.edu

Russell M. Mersereau
School of Elec. and Comp. Eng.
Georgia Institute of Tech.
Atlanta, GA,30332-0250
rmm@ece.gatech.edu

ABSTRACT

We study the application of the Bamberger directional filter bank to the problem of rotation invariant texture classification. We explore the use of purely directional decompositions and the use of polar-separable Bamberger pyramids. We obtain comparable classification performance to Gabor-based methods using a smaller feature set.

1. INTRODUCTION

In the real world, texture-based image analysis algorithms and image database retrieval systems deal with textures that have some type of distortion with respect to training samples and database content. One commonly found distortion is rotation. Our visual system recognizes rotation easily, but it is a non-trivial problem in image analysis. Hence it is of interest to obtain feature sets that are rotation invariant (RI).

Extraction of RI feature sets for texture classification has been addressed in a few works. One of the early approaches was the use of RI random field models like circular AR models [1] and Gaussian-Markov models [2]. A second approach based on multichannel filtering (i.e spatial filtering) is particularly well suited to deal with rotation invariance. Greenspan, et al., [3] used a steerable pyramid and DFT encoding to obtain RI features. Haley and Manjunath [4] used a 2-D analytical polar Gabor representation to form a set of RI macrofeatures based on local magnitude, phase, and autocorrelation features. As far as wavelet-based approaches, Hill, et al. [5] used the DT-CWT and DFT encoding to generate RI features.

In our previous work [6, 7] we have studied the use of the Bamberger Directional Filter Bank (BDFB) for texture segmentation and analysis. In this paper, we exploit the orientation selectivity of the BDFB for RI texture analysis, and introduce the use of Bamberger pyramids as an alternative to Gabor filters.

In the following section we review the theory of the BDFB and the implementation of Bamberger Pyramids. In section 3, we introduce DFT-encoding and its use with BDFBs. In section 4, we present RI classification results using a purely directional approach. We introduce the use of Bamberger pyramids for RI classification in section 5. We finish the paper doing a comparison with Gabor-based schemes, and discuss extensions for future work.

2. THE BAMBERGER DIRECTIONAL FILTER BANK

Next we review the implementation of the BDFB. For a thorough discussion we refer the reader to [8, 9]. The BDFB partitions the frequency plane into a set of complementary wedge-shaped passbands as shown in Figure 4-(a). A tree-structured 2^S-band BDFB has been commonly employed. This is illustrated in Figure 1 where we show an analysis section for $S = 3$. Each stage in the tree structure uses Fan Filter Banks (FFBs) as a building block. Hence, the implementation of the BDFB is reduced to the design of FFBs. The FFB is a 2-D two-channel filter bank with complementary fan-shaped support filters. For alias-free maximal decimation, each channel is downsampled using a quincunx resampling matrix $\mathbf{M}$ as shown in Figure 2.

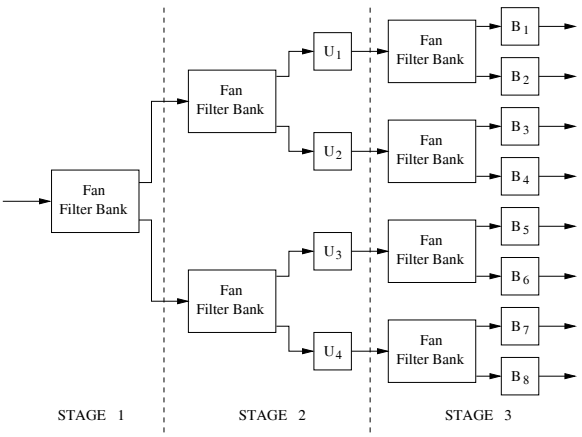


Fig. 1. Analysis stage for a tree-structured BDFB

To implement perfect reconstruction FFBs we use the three-stage ladder structure proposed by Ansari, et al., [10]. The analysis stage is presented in Figure 2. The functions $\beta_i(z)$ are 1-D FIR filters of length L_i with all-pass magnitude responses. In this paper we let $\beta_1(z) = \beta_2(z) = \beta_3(z) = \beta(z)$. Hence, the 2-D filter design problem reduces to the design of the prototype $\beta(z)$. We note that the 2-D filtering is performed efficiently in the polyphase domain as separable operations. The synthesis structure is straightforward to derive using the properties of ladder structures.

After the second stage, it is necessary to resample each subband using unimodular matrices $\mathbf{U}_i$. This operation rearranges the sampling lattice without downsampling, such that FFBs can be used throughout the BDFB structure [8, 9]. Finally, the "backsampling" matrices $\mathbf{B}_i$ are used to rearranged the directional subbands over a rectangular lattice [9].

The non-ideal realization of wedge-shaped responses is par-

0-7803-9332-5/05/$20.00 ©2005 IEEE

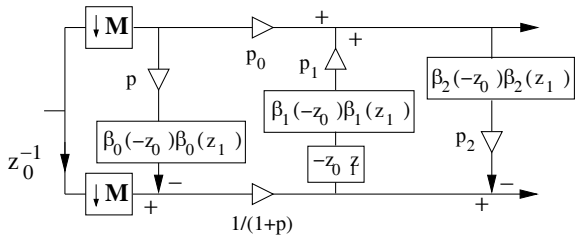


Fig. 2. FFB implementation with a three-stage ladder structure.

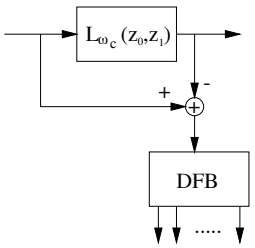


Fig. 3. Lowpass-highpass analysis structure.

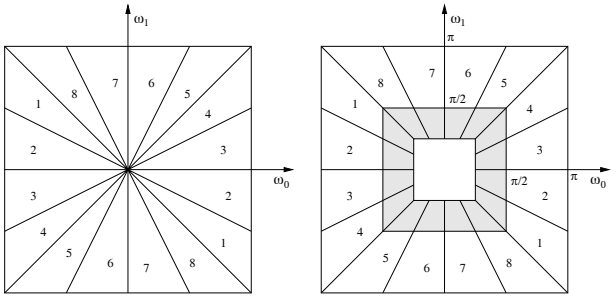
ticularly manifested around the origin. For natural images, the subbands present uneven distribution of the low frequency information. Since directional information resides mostly in the mid-frequency to high-frequency range, the simplest approach to deal with this issue is too use the "lowpass-highpass" decomposition presented in Figure 3. The image is first filtered with a lowpass filter $L_{\omega_c}(z_0, z_1)$ with cutoff frequency ω_c. The highpass component is given by the difference between the input and the lowpass image. Finally, the highpass component is processed with the BDFB.

Finally we note that the BDFB is a maximally decimated representation. Compared to other directional filter banks (e.g., Gabor functions), memory savings are substantial.

2.1. Bamberger Pyramids

Polar-separable pyramids (PSPs) allow us to extract information across different resolutions and orientations. The decomposition is obtained by combining a radial pyramid and an angular filter bank in sequence. This allows independent control of the radial and angular selectivity. In [11, 12] we proposed a family of PSPs where the BDFB is used as the angular component, and the radial component can be chosen according to the application. We identify this family of PSPs as *Bamberger Pyramids* (BPs). In particular, we presented Bamberger pyramids using the undecimated BDFB and the Laplacian pyramid in [11]. A possible partitioning of the frequency plane is presented in Figure 4-(b).

In this work, we use the BP configuration shown in Figure 5. Here, the image is decomposed into J *undecimated* pyramid levels. Except for the lowpass channel, each of the J-1 bandpass channels is decomposed by the maximally-decimated N-band BDFB.



(a)Bamberger DFB (b) Bamberger Pyramid

Fig. 4. Comparison of the frequency plane partitioning between the BDFB and a Bamberger pyramid.

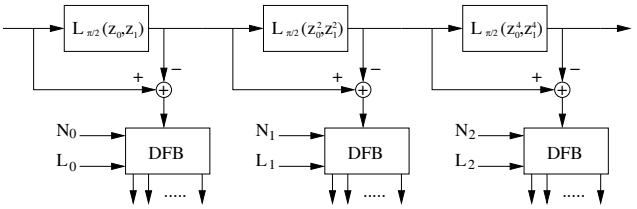


Fig. 5. Polar-separable Bamberger Pyramid using the undecimated Gaussian pyramid and the BDFB. The BDFB is applied to the difference image between contiguous lowpass images

3. EXTRACTION OF ROTATION INVARIANT FEATURES USING DFT ENCODING

In this section, we describe the DFT-encoding step used in various works [3, 5]. More specifically, we relate this method to the extraction of RI features using the Bamberger DFB.

Define the 2-D rotation operator as $\mathcal{R}_\theta(\cdot)$ which performs of a rigid rotation by an angle θ. We start by letting $x(t_0, t_1)$ be a continuous time texture. Rotation by some θ gives

$$x'(t_0, t_1) = \mathcal{R}_\theta(x(t_0, t_1)).$$

Let $X(\Omega_0, \Omega_1)$ and $X'(\Omega_0, \Omega_1)$ be the Fourier transforms of $x(t_0, t_1)$ and $x'(t_0, t_1)$ respectively. It can be shown that

$$X'(\Omega_0, \Omega_1) = \mathcal{R}_\theta(X(\Omega_0, \Omega_1)).$$

A rotation in the (continuous) spatial domain generates a similar rotation in the (continuous) frequency domain.

When texture images are digitized, the above relationship does not hold in general. However, we assume for practical purposes that

$$x'[n_0, n_1] \approx \mathcal{R}_\theta(x[n_0, n_1])$$

and

$$X'[\omega_0, \omega_1] \approx \mathcal{R}_\theta(X[\omega_0, \omega_1]),$$

where $X[\omega_0, \omega_1]$ is the DFT of $x[n_0, n_1]$. As the texture is rotated, texture energy is redistributed on the frequency plane.

As presented in [7], the energy estimates of the BDFB subbands provide an excellent feature set. However, rotation redistributes energy across subbands, inducing changes in the feature

vector. As the rotation angle increases between two textures of the same class, the feature vectors become dissimilar. Hence it becomes necessary to compensate for the undesired rotation. A way to generate RI features is to do the so-called DFT-encoding.

Suppose that we use a BDFB with a large number of subbands N. (In the limit, a subband will become a diametral slice of the frequency plane.) Hence, we have an N-dimensional feature vector

$$\mathbf{f} = [\ f_0 \quad f_1 \quad f_2 \quad \dots \quad f_{N-1}\ ]^T,$$

where f_i is an energy measure of the i^{th} subband. We can map $\mathbf{f}$ to a discrete 1-D signal $f[n]$ for $n = 0, 1, \dots, N-1$.

Next, we define the circular shift operator for $f[n]$ as $S_\ell(f) = f((n-\ell) \text{mod} N)$. Suppose we have a texture $x[n_0, n_1]$ with a feature vector described by the 1-D sequence $f[n]$. Then we can show that

$$\mathcal{R}_\theta(x[n_0, n_1]) \Rightarrow S_\ell(f[n]). \tag{1}$$

In words, rotation of a texture implies a circular shift on the feature vector.

In reality, the circular shift on the feature vector is not exact since we are working with a discrete sequence of energy measures while θ is a continuous variable. As N gets smaller (e.g., the case of eight- and four-band BDFBs), the rotation-to-shift relationship in Equation 1 gets coarser and could become inaccurate.

The DFT-encoding step generates RI features by removing the circular shift from $f[n]$. This scheme consists of first taking the N-point DFT $F[k]$ of the feature sequence $f[n]$. Next, the magnitude coefficients $|F[k]|$ are taken to generate RI features. This in effect removes the phase (i.e., the shift) from the feature vector. The RI feature vector is given by

$$\mathbf{F} = [\ |F[0]| \quad |F[1]| \quad \dots \quad |F[N/2+1]|\ ]^T. \tag{2}$$

Since the DFT magnitude is an even function, the new RI feature vectors have only $\frac{N}{2}+1$ elements. Therefore, the generation of RI features could affect classification performance. On the other hand, this has the positive effect of reducing classifier complexity.

4. RI TEXTURE CLASSIFICATION USING THE BAMBERGER DFB

We use the same data set tested by Haley and Manjunath in [4]. This image set is available over the internet [13], and consists of 13 textures. Each texture is scanned and digitized at rotations of 0^o, 30^o, 60^o, 90^o, 120^o, and 150^o. Each image is 512×512 pixels, and is subdivided into 16 non-overlapping 128×128 images which are used for training or testing.

Our RI classification system is set-up as follows. In order to avoid uneven distribution of the lowpass energy across subbands, we use the lowpass-highpass structure described in Section 2. For these tests, the ladder filter $\beta(z)$ was designed with the Parks-McClellan program with $L = 12$. As features we compute subband energies using

$$e_{1,i} = \frac{1}{N_1 N_2} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} |y_i(n_1, n_2)|, \tag{3}$$

where $y_i(n_1, n_2)$ represents the ith subband. Finally, we perform DFT-encoding on the resulting feature vector.

We chose the Bayes distance rule for classification. We assume that the probability density functions of the feature vectors have a multivariate Gaussian distribution with mean vectors and covariance matrices $(\mathbf{m}_k, \mathbf{C}_k)$, for $k = 1, 2, \dots, 13$. Class parameters were obtained using maximum likelihood estimators. The classification of a feature vector $\mathbf{F}$ is done by assigning it to the class with minimum distance value

$$d_k(\mathbf{F}) = (\mathbf{F} - \mathbf{m}_k)^T \mathbf{C}_k^{-1} (\mathbf{F} - \mathbf{m}_k) + \log(\det(\mathbf{C}_k)). \tag{4}$$

Each feature vector was tested using the *leave-one-out* method.

Table 1. DFB RI correct classification as a function of ω_c.

ω_c	0	$\pi/16$	$\pi/8$	$\pi/4$
N=4	50.64%	66.11%	66.99%	61.62%
N=8	50.32%	72.04%	72.12%	68.35%

Table 2. BDFB RI correct classification as a function of ω_c using a combination of three energy features.

ω_c	0	$\pi/16$	$\pi/8$	$\pi/4$
N=8	68.67%	90.14%	91.35%	87.66%

Our first results are shown in Table 1 for different values of N and ω_c. The best result for this system is 72.12% correct classification. There are some possible reasons for this poor performance. First, the sample size (128×128) might not be large enough to capture representative subband energies. Second, feature vector dimension is reduced after DFT-encoding. Finally, the degree of directional selectivity might be too coarse, so the rotation-shift relationship from Equation (1) is not captured adequately. To improve classification, we increase the number of features using the following subband energy measures

$$e_{2,i} = \frac{1}{N_1 N_2 - 1} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} y_i(n_1, n_2)^2, \tag{5}$$

and

$$e_{3,i} = \frac{1}{0.6745} \text{MED}(|y_i(n_1, n_2)|). \tag{6}$$

In this case, the feature extraction vector is formed by three energy measures e_1, e_2, e_3 for each subband. The features are grouped by energy measure and are DFT encoded separately. The resulting features are concatenated into a single feature vector of dimension $3(\frac{N}{2}+1)$. We show the result for an eight-band BDFB in Table 2. Performance is improved significantly with the best result at 91.35% correct classification for $\omega_c = \pi/8$.

5. RI TEXTURE CLASSIFICATION USING BAMBERGER PYRAMIDS

As we have discussed, the BDFB does not differentiate between coarse or fine information along a particular direction. Hence, adding a multiresolution component could improve the classification performance of our system. In this section, we introduce the use of BPs for RI texture classification. In our experiments, we use the BP described in section 2.

We group the subband energies for each resolution level to form an N-dimensional feature vector $\mathbf{f}^{(j)}$ for $j = 1, 2, ..., J-1$. Then we form RI vectors $\mathbf{F}^{(j)}$ using DFT-encoding. The feature

Table 3. Bamberger pyramid RI correct classification for $N = 4$.

	J=2	J=3	J=4	J=5
L=4	49.60%	78.13%	91.59%	94.95%
L=12	49.28%	84.05%	95.35%	97.28%
L=18	49.52%	84.13%	95.19%	97.04%

Table 4. Bamberger pyramid RI correct classification for $N = 8$.

	J=2	J=3	J=4	J=5
L=4	47.92%	81.81%	92.95%	93.99%
L=12	55.93%	92.39%	96.88%	96.71%
L=18	54.81%	92.16%	96.96%	96.88%

vectors $\mathbf{F}^{(j)}$ are concatenated to a single feature vector given by

$$\mathbf{F} = \left[\begin{array}{ccccc} \mathbf{F}^{(1)^T} & \vdots & \mathbf{F}^{(2)^T} & \cdots & \mathbf{F}^{(J-1)^T} \end{array} \right]^T. \quad (7)$$

The resulting feature vector has dimension $(J-1) \times (\frac{N}{2}+1)$.

In Tables 3 and 4 we show classification results using (3) as a feature. We show the performance progression as J is increased. Beyond $J = 5$ there was no improvement in classification. The tables show the advantage of decomposing a texture at different resolutions and directions. We draw the following observations from these tables.

- Classification is improved as a function of the length of $\beta(z)$. For L larger then 12 we did not find a significant improvements.
- With two resolution levels, correct classification jumps from the 54%-56% range to more than 92%.
- The best result is 97.28% correct classification using $J = 5$ and $N = 4$. Hence, higher directional selectivity does not necessarily improve classification.
- For the $N = 8$, the best performance is at 96.96% with $J = 4$.

Using e_2 or e_3 as the energy measure showed a significant loss of performance. For instance, using e_2 with $J = 5$, $N = 4$, and $L = 18$ gives 91.83% correct classification.

6. COMPARISON AND DISCUSSION

We have shown that BPs provide good RI classification results. BPs can be considered as a good alternative to Gabor functions. They present computational and design advantages like separable filtering implementation and excellent directional selectivity.

Haley and Manjunath [4] achieve a 96.8% correct classification on the same data set. We achieve slightly better performance, but with significantly lower feature set complexity. Their feature set is expensive to compute, and consists of a 208-dimensional feature vector. In our case, our feature is simple to compute, and only 12 features per texture sample are required.

We believe that the improvement over the method in [4] is caused by two factors. First, the frequency selectivity for BPs is significantly better than for Gabor functions. Hence, there is less energy leakage among subbands. This effect can also be observed in our results. In Tables 3 and 4, we see that increasing L (i.e., generating sharper filter passbands) improves classification. Second, the downsampling operations in the BDFB could be considered as smoothing operations which remove irregularities among texture samples. As discussed by Wilson and Spann [14], this improves the "classness" of texture features.

As future work, we plan to compare Bamberger pyramids against other multichannel representations. Additionally, we plan to take advantage of the BDFB structure to develop texture classification systems invariant to other types of geometrical distortions like skewing and more general affine transforms.

7. REFERENCES

[1] R. L. Kashyap and A. Khotanzad, "A model-based method for rotation invariant texture classification," *IEEE Trans. PAMI*, vol. 8, pp. 472–481, July 1986.

[2] F. S. Cohen, Z. Fan, and M. A. Patel, "Classification of rotated and scaled textured image using gaussian-markov random field models," *IEEE Trans. PAMI*, vol. 13, pp. 192–202, Feb. 1991.

[3] H. Greenspan, S. Belongie, R. Goodman, and P. Perona, "Rotation invariant texture recognition using a steerable pyramid," in *ICIP'94*, Jerusalem, Israel, Oct. 1994.

[4] G. M. Haley and B. S. Manjunath, "Rotation-invariant texture classification using a complete space-frequency model," *IEEE Trans. Im. Proc.*, vol. 2, no. 8, pp. 255–269, feb 1999.

[5] P. R. Hill, D. R. Bull, and C. N. Canagarajah, "Rotationally invariant texture features using the dual-tree complex wavelet transform," in *ICIP'00*, September 2000.

[6] J. G. Rosiles and M. J. T. Smith, "Texture segmentation using a biorthogonal directional decomposition," in *SCI 2000*, July 2000, Orlando,FL.

[7] J. G. Rosiles and M. J. T. Smith, "Texture classification with a biorthogonal directional filter bank," in *ICASSP'01*, 2001, vol. 3, pp. 1549–1552.

[8] R. H. Bamberger and M. J. T. Smith, "A filter bank for the directional decomposition of images," *IEEE Trans. on Signal Proc.*, vol. 40, no. 4, pp. 882–893, April 1992.

[9] S. Park, M. J. T. Smith, and R. M. Mersereau, "A new directional filter bank for image analysis and classification," in *ICASSP'99*, October 1999, vol. 2, pp. 1286–1290, Chicago, IL.

[10] R. Ansari, Chai W. Kin, and M. Dedovic, "Structure and design of two–channel filter banks derived from a triplet of halfband filters," *IEEE Trans. on Circ. Sys.–II: Anal. and Dig. Sig. Proc.*, vol. 46, no. 12, pp. 1487–1496, December 1999.

[11] J. G. Rosiles and M. J. T. Smith, "A low complexity undecimated directional image decomposition," in *ICIP'03*, 2003, vol. 1, pp. 1049–1052.

[12] J. G. Rosiles, *Image and Texture Analysis Using Biorthogonal Angular Filter Banks*, Ph.D. thesis, Georgia Institute of Technology, July 2004.

[13] "Signal and Image Processing Inst., Univ. Southern California," http://www.sipi.usc.edu.

[14] R. Wilson and M. Spann, "Finite prolate spheroidal sequences and their applications ii: Image feature description and segmentation," *IEEE Trans. PAMI*, vol. 10, no. 2, pp. 193–203, mar 1988.