Filter banks for texture segmentation

Trygve Randen and John Håkon Husøy
Høgskolesenteret i Rogaland
Institutt for elektroteknikk og databehandling
P.O. Box 2557 Ullandhaug, N-4004 Stavanger, Norway

Abstract

Several approaches to multichannel filtering for texture classification and segmentation with Gabor filters have been proposed. The rationale presented for the use of the Gabor filters is their relation to models for the early vision of mammals as well as their joint optimum resolution in time and frequency. In this work we present a critical evaluation of the Gabor filters as opposed to filter banks used in image coding – in both full rate and critically sampled realizations. In the critically sampled case, tremendous computational savings can be realized. We further evaluate the commonly used octave band decomposition versus alternative decompositions.

We conclude that for a texture segmentation task it is possible to use a wider range of filters than just the Gabor class of filters, it is possible to use alternative decompositions and, most important, it is possible to use sub-sampled filters.

Introduction

In several recent papers the successful application of multichannel filtering for texture segmentation is reported [1, 2, 3, 4, 5, 6, 7]. The channel filters utilized are Gabor filters [8] designed to respond to different spatial frequencies. The rationale presented for the use of the Gabor filters is their relation to models for the early vision of mammals as well as their joint optimum resolution in time and frequency [8, 9]. A critical evaluation of the use of Gabor filters as opposed to alternative filter choices is not presented in the works referred to above.

In this paper, a broader point of view to the task of selecting filters for the filter banks is adopted. Rather than postulating a restricted class of filters such as the Gabor filters, we examine the performance of multichannel segmentation schemes based on more general classes of filters including the Gabor filters. It is shown that, for the present application, performances comparable to those obtained by the use of Gabor filters are obtained by other filters that can be implemented at a lower computational cost.

It has been common to decompose the frequency band by the octave band decomposition [1, 2, 3, 5, 6, 7] (often referred to as the wavelet decomposition). This choice is motivated by the similarities between the octave band decomposition and the early vision of mammals. Octave band decomposition implies a finer frequency resolution in lower than in higher frequency regions. The work by Chang and Kuo [10], however, indicate that texture features are most prevalent in intermediate frequency bands. Thus, the octave band decomposition is not optimal. We present a broad range of subband decompositions, and thereby show that alternative decompositions may, in some cases, be superior to the octave band decomposition.

Most approaches to multichannel texture segmentation apply filter banks in which the filtered images representing various spatial frequencies each *have the same number of image samples as the original image to be segmented*. Thus, after filtering, the amount of data is increased by a factor corresponding to the number of frequency channels in the filter bank. It is well known that sub-sampling

of the various frequency bands, in such a way that the total number of signal samples equals the number of signal samples in the input image, can be performed without losing any information as long as the filters are designed appropriately [11]. We examine what kind of segmentation can be obtained by operating on sub-sampled frequency channels when using critically sampled filter banks allowing perfect reconstruction of the input signal. By using this approach, tremendous computational savings are be realized.

The main results of our research are: 1) Other filters than the Gabor class is useful in a texture segmentation context. 2) Other subband decompositions than the octave band decomposition may be better suited for texture segmentation. 3) By applying sub-sampled perfect reconstruction filter banks, good segmentation performance can be achieved at a significantly lower computational cost than what is the case for non-sub-sampled Gabor schemes.

Multichannel filtering for texture segmentation

Experiments done on cats and apes, which have a visual system similar to that of humans [12], show that these animals decompose the visual images into several filtered images, with variations over a small interval in frequency and orientation [13]. The objective of multichannel filtering, and the subsequent local energy function, is to transform the edges between the textures into detectable discontinuities [14].

Figure 1 shows the general experimental setup. We will, in the current and following subsections, describe the elements of the experimental setup. We will go further into the details of the various filter banks in the next sections.

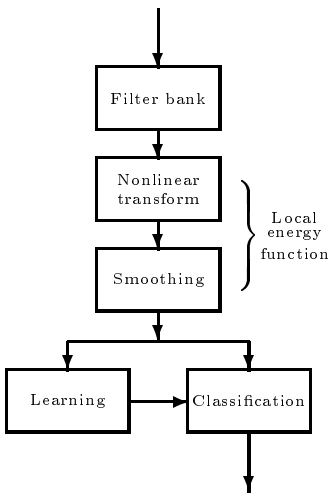

Figure 1: Block diagram for the basic experimental setup

Local energy function

The filter bank is in essence a set of bandpass filters with frequency and orientation selective properties. The local energy function is supposed to transform areas where the band pass frequency components are strong into a high constant gray-level and areas where it is weak into a low constant gray-level.

We found the local energy function proposed by Farrokhnia [1, 2, 3] to be appropriate for our experiments. Thus we use

$$e_k = |\Psi(r_k(x,y))|, \qquad (1)$$

where

$$\Psi(t) = \tanh(\alpha t),$$

as the nonlinear transform, where $r_k(x,y)$ is the image corresponding to frequency channel no. k and (x,y) denote spatial coordinates. The parameter α gives the saturation characteristic of this function. By experience, $\alpha = 0.25$ [1, 2, 3].

The nonlinear transform is succeeded by a Gaussian low-pass (smoothing) filter of the form

$$H_G(u,v) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(u^2+v^2)}, \qquad (2)$$

where

$$\sigma = \frac{\sqrt{2}}{2}\frac{1}{f_0}, \tag{3}$$

and f_0 is the absolute value of the center frequency from the band-pass filter. The center frequencies are normalized to the interval $[-1/2, 1/2]$. Gaussian filters have the property of giving accurate localization of the texture boundaries [1].

Classification

The output from the local energy function is a set of images, one image per filter in the filter bank. These images are the feature images that will form the basis for the classification. We use these images to form feature vectors. Each feature image corresponds to one dimension in the feature vectors. Classification is the task of assigning these feature vectors to, hopefully, the right class.

We use the "Learning Vector Quantizing (LVQ)" supervised classifier by Kohonen [15, 16]. As is common in other work [1, 2, 3], we use a subset of 6% of the evaluation data-set as the training data-set. Tests verified that a subset of 6% is appropriate [17].

The Gabor filter approach

In this section we will review three different approaches to using the Gabor filter in texture segmentation. The Gabor filter is a modulated Gaussian filter [8, 9].

A fixed filter structure

Farrokhnia [1] proposes a bank of Gabor filters with even symmetry and an octave band decomposition. The decomposition is determined by the size of the input image, and is therefore fixed with respect to the characteristics of the textures in the input image.

We have tested the segmentation scheme on several texture mosaics. One representative

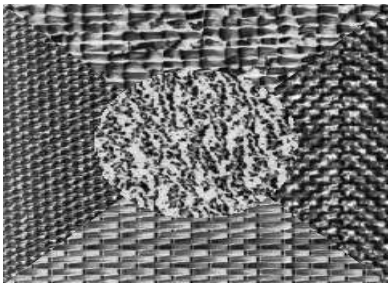

Figure 2: The 'Nat-5b' texture mosaic (D24, D77, D55, D84 and D17).

image is 'Nat-5b', with natural textures from Brodatz [18], see figure 2. As the segmentation quality measure we use the percentage of misclassified pixels. The method presented here gives good segmentation results, 95.1% correct segmentation. A drawback with the method is its high computational complexity. An image of i.e. 256×256 pixels will give rise to 28 feature images of the same size as the input image. Farrokhnia [1, 2, 3] presents a way of reducing the number of feature images, but the computational complexity is still high.

Optimal parameters

The method presented in the previous section applies a large fixed set of filters, and has therefore a high computational complexity. If we can find *one* filter which leads to good segmentation, we will get large savings in computational complexity.

Dunn et.al. [14] introduces a method for finding the optimum parameters to the Gabor filter for a given pair of textures. He assumes a general Gabor filter, as opposed to the even symmetric filter in the previous section, but doesn't present any segmentation experiments. We have taken the idea one step further, and implemented the method with full segmentation. We have also generalized the

method to images with more than two textures, in which case we apply one filter for each texture pair. We refer to Randen [17] for a further description of the method.

For 'Nat-5b' the computational savings are ten versus 28 filters, although we see an improvement in the accuracy. Experiments [17, 19] showed that images with fewer textures had more significant computational savings, but poorer segmentation results than the previous method. We further found this approach suitable only for a subset of our test images.

The method has even more weaknesses. 1) The computation of the optimal Gabor filter parameters is computationally demanding, so if it's impossible to perform this computation off-line, then much of the savings in computational complexity will be lost. 2) In order to find the optimal parameters, we have to know the textures in advance, and therefore an unsupervised system is out of the question. In contrast, the previous method becomes unsupervised just by replacing the supervised classifier by an unsupervised classifier. 3) The number of filters increases rapidly with the number of textures in a image, two textures implies one filter, five textures implies ten filters and eight textures implies 28 filter, etc.

In conclusion, this method may be an accurate and faster alternative to the previous method if the circumstances are right, but for a large class of images, the method is poor or unusable.

Alternative subband decompositions

In the fixed filter structure approach we used the octave band decomposition. The papers that suggest using the octave band decomposition usually gives similarities with the human visual system as the argument for using this decomposition. In this section we will evaluate the octave band decomposition against some other decompositions while still using the Gabor filter.

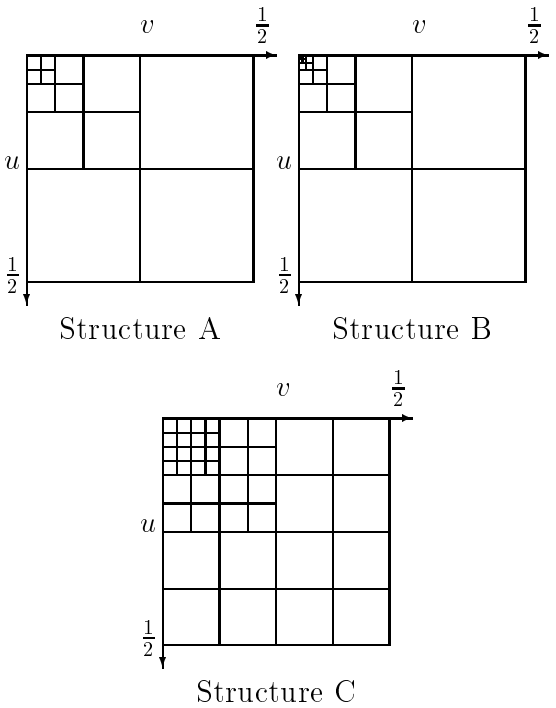

Figure 3: Subband decompositions, structures A and B are octave band decompositions, where A has four and B has seven levels.

Some useful subband decompositions are given in figure 3. In these experiments we have chosen to use the even symmetric Gabor filter. The center frequency, f_0, is set to the absolute value of the frequency in the center of the pass-band and θ is set to the corresponding angle.

The decomposition we have proposed here will only make it possible to specify θ from 0° to 90°. To maintain selectiveness of angles from 90° to 180°, as in the two previous methods, we apply the filter whose frequency response is the sum of the frequency responses for θ and $\theta + 90^\circ$.

The experimental results from this filtering approach, 95.7%, 94.4% and 95.5% for structure A, B and C respectively, shows that non-octave band decompositions are viable alternatives to octave band decompositions. The problem, however, remains the same for all de-

compositions: to achieve a good segmentation quality we need a large number of subbands, and this leads to high computational complexity.

New filter banks for texture segmentation

The problem with the approaches presented above is that they have a high computational complexity or that they show poor classification for some textures. We know that filtering through a bank of N filters gives rise to N feature images, each of the same size as the input image – we have an N-fold increase in the amount of data. We have, however, no increase in the amount of information, thus we have *redundancy*. To achieve correct classification we should in principle not need a redundant representation. Is it possible to encode the images without redundancy, without any critical loss in segmentation quality?

For critically sampled filter banks with perfect reconstruction, as used in image coding, see for example Husøy [20], the problem with redundancy is solved. Is it possible to achieve segmentation accuracy comparable to that of the Gabor filters, when using such filter banks? And if so, is it possible to use the filter outputs in a non-redundant encoding and achieve acceptable segmentation? Before we go further into the experiments, we will start by a short description of critically sampled filter banks.

Critically sampled filter banks

Figure 4 shows a two-level tree-structured critically sampled filter bank. The symbol $\downarrow L$ represents *sub-sampling* by L (also called *down-sampling* or *decimation*), which is the process of discarding all but every L'th sample. A critically sampled filter bank has *perfect reconstruction* if it is possible to reconstruct the input from the sub-sampled outputs. There exists a large (in fact infinite) class of filter banks with perfect reconstruction. We will restrict our further discussion to *separable* perfect reconstruction filter banks.

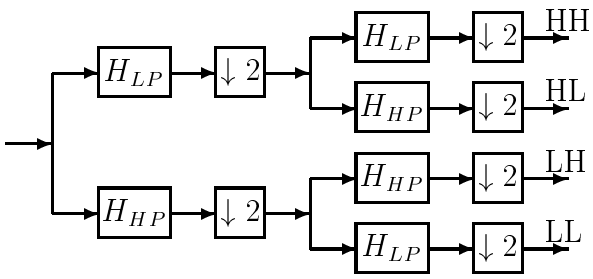

Figure 4: Tree-structured filter bank

Note that it is possible to implement the arithmetic operation of the FIR and some of the IIR filter banks to get the sub-sampling *within* the filter, and thus achieve big savings in computational complexity [21]. The theory of filter banks is covered in detail in [11, 21].

We will use two types of perfect reconstruction filters, *finite impulse response* (FIR) and *infinite impulse response* (IIR). The IIR filters are implemented with *all-pass* filters. This implementation is attractive due to the computational simplicity [20].

Multirate texture segmentation

We have previously presented an experimental setup for multichannel filtering for texture segmentation with Gabor filters. To be able to use sub-sampled filter banks successfully, we need to make a modification to the smoothing filter in the local energy function. All other parts, except from the filter bank, remain the same.

When we decimate the signal by a factor N, a signal with period M and frequency $f = 1/M$ will get the period M/N, and the frequency $N/M = Nf$. Thus, the smoothing filter designed for a given full rate signal will no longer be designed for the same signal

Structure	'F_2_1_09'	'f32d'
A	86.9%	90.5%
B	90.8%	92.5%
C	96.4%	95.5%

Table 1: Full rate segmentation of 'Nat-5b' with various subband decompositions.

after sub-sampling. We remember that σ in equation (3) was inversely proportional to the frequency f_0, therefore we use the equation

$$\sigma = \frac{\sqrt{2}}{2}\frac{1}{N f_0},$$

where N is the sub-sampling factor and f_0 is the absolute value of the center frequency of the filter bank filter.

Experimental results

For the purpose of studying the effect of sub-sampling with respect to segmentation accuracy, we performed tests with both full rate and sub-sampled filter banks. We tested segmentation using the perfect reconstruction filter banks with several different filters, FIR filters from Johnston [22] and IIR filters from Husøy [20]. We performed experiments on a large number of combinations of filters, images and subband decompositions, and found the FIR filter 'f32d' and the IIR filter 'F_2_1_09' to be quite representative of the different classes of filters [17]. These filters are formerly used in texture *classification* [23]. Our experiments revealed that filter structure C, figure 3, was better than the octave band decompositions, see table 1. Other subband decompositions may lead to better segmentation results, but of the decompositions we tested, structure C was a good compromise between segmentation accuracy and computational complexity [17].

We have previously shown that the octave band decomposition is approximately equal to other decompositions in terms of performance for the Gabor filters, but we see from table 1 that the octave band decomposition is not suited for this set of filters. Can we find any perfect reconstruction filters for which the octave band decomposition is suited? The octave band decomposition is often associated with the wavelet class of filters. We therefore found it natural to test some wavelet filters with octave band decomposition, and we tested the set of Daubechies [24] orthogonal wavelet filters. We only tested these filters in the full rate setup. The best octave band decomposed Daubechies filter, 'dau4' with structure B, got a classification accuracy of approximately 88% for 'Nat-5b', which is far poorer than the approximately 96% of the 'F_2_1_09' filter.

We have now shown that perfect reconstruction filters are suited to texture segmentation. From these full rate experiments, we see that the perfect reconstruction filters are comparable to the Gabor filter with respect to segmentation accuracy, when we use *full rate* perfect reconstruction filters. But by using full rate filters we don't achieve the desired savings in computational complexity. Is it possible to use critically sampled filter banks and still achieve good segmentation? Table 2 and figure 5 shows the segmentation results.

We see from the tables 1 and 2 and figure 5 that the segmentation accuracy in the sub-sampled case is a little bit degraded from the results in the full rate case, but that we obtain high quality segmentation even with sub-sampled perfect reconstruction filters. We especially note that the FIR filter has no more than 0.6% degradation from the full rate case to the critically sampled case. We performed several tests, that verified our findings [17].

We see that the segmentation based on critical sampling has a stair-case effect near the edges. Further experiments show, as suspected, that the stair-case effect is reduced if we use a lower sub-sampling rate, at the cost

Filter	'Nat-5b'	'S-IO'	'D17-D55'
'F_2_1_09'	91.7%	85.0%	97.0%
'f32d'	94.9%	84.3%	97.4%

Table 2: Segmentation from critically sampled filters with filter structure C.

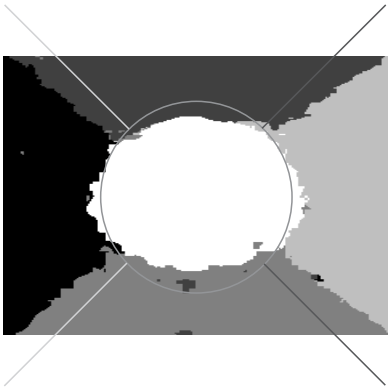

Figure 6: Segmentation result from four times critical sampling with filter structure C, filter 'F_2_1_09' and texture 'Nat-5b'.

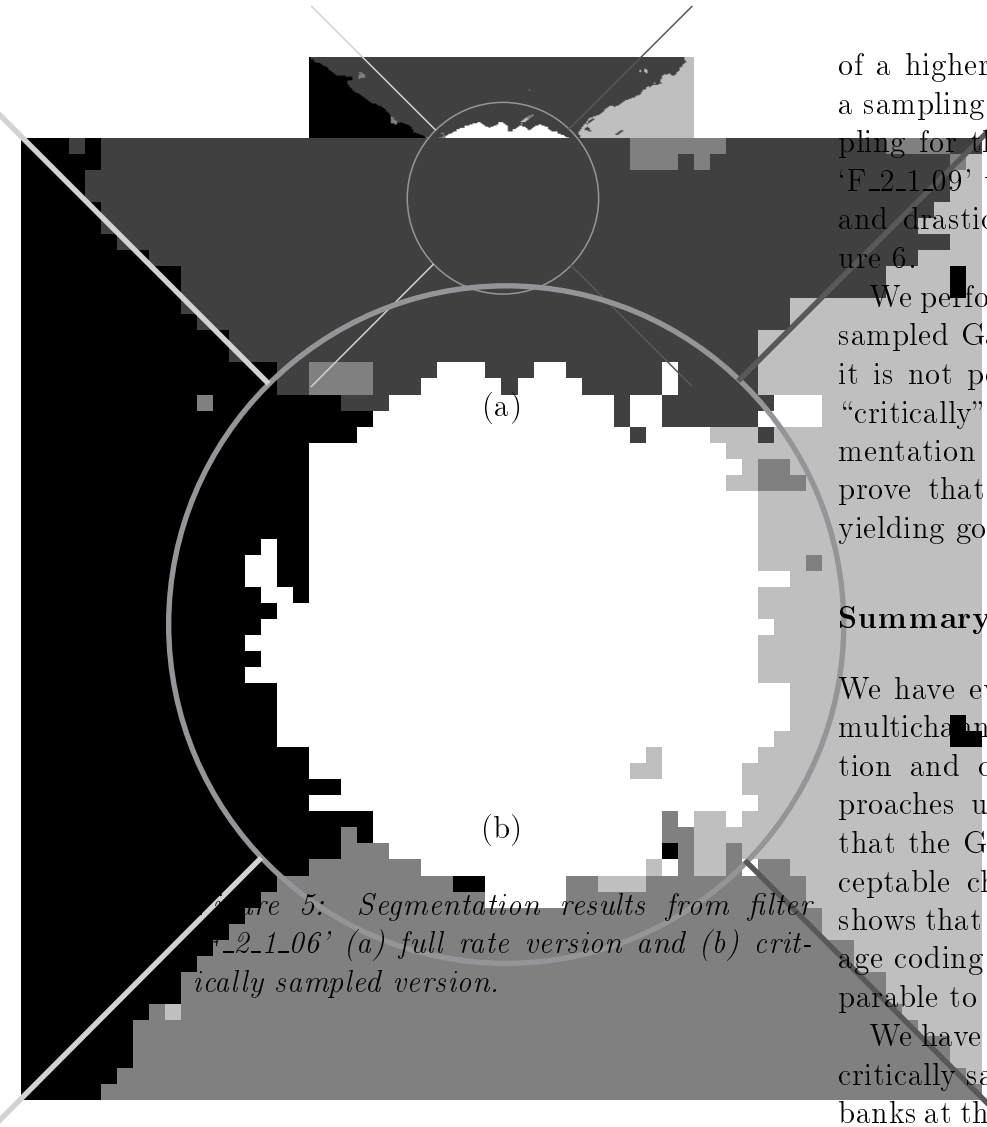

Figure 5: Segmentation results from filter 'F_2_1_06' (a) full rate version and (b) critically sampled version.

of a higher computational complexity. With a sampling rate of four times the critical sampling for the texture 'Nat-5b' with the filter 'F_2_1_09' we got 95.6% correct classification and drastically reduced stair-case effect, figure 6.

We performed the same kind of tests on subsampled Gabor filters. These tests show that it is not possible to sample the Gabor filter "critically" without severely reducing the segmentation quality. This, however, does not prove that non-perfect reconstruction filters yielding good segmentation doesn't exist.

Summary and conclusions

We have evaluated a large class of filters for multichannel filtering for texture segmentation and compared them against some approaches using the Gabor filter, and shown that the Gabor filter is far from the only acceptable choice of filters. Our experiments shows that several filter banks known from image coding may provide results that are comparable to those of the Gabor filter.

We have also shown that it is possible to use critically sampled perfect reconstruction filter banks at the cost of only a modest degradation

in segmentation quality. In this way, tremendous savings in computational complexity are realized.

We have further shown that other subband decompositions than the octave band decomposition may be viable alternatives. We found a subband decomposition which splits the band pass-regions, that for the Gabor filter provided approximately equal performance to the octave band decomposition, and for the perfect reconstruction filter banks provided better performance.

Therefore we have shown that the assumptions made on the human visual system not necessarily are the only appropriate model for a synthetic texture segmentation scheme. In fact, our experiments shows that other approaches may have some considerable advantages.

References

[1] F. Farrokhnia, *Multi-channel filtering techniques for texture segmentation and surface quality inspection.* PhD thesis, Michigan State University, 1990.

[2] F. Farrokhnia and A. K. Jain, "A multichannel filtering approach to texture segmentation," in *Proc. Computer Vision and Pattern Recognition*, pp. 364–370, IEEE Comp. Soc., 1991.

[3] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognition*, vol. 24, no. 12, pp. 1167–1186, 1991.

[4] A. C. Bovik, "Properties of multichannel texture analysis filters," in *Proc. Int. Conf. Acoust. Speech, Signal Proc.*, (Albuquerque, NM), pp. 2133–2136, IEEE, 1990.

[5] M. R. Turner, "Texture discrimination by Gabor functions," *Biological Cybernetics*, vol. 55, pp. 71–82, 1986.

[6] A. Perry and D. G. Lowe, "Segmentation of textured images," in *Proc. Computer Vision and Pattern Recognition*, pp. 319–325, IEEE Comp. Soc., 1989.

[7] T. N. Tan and A. G. Constantinides, "Texture analysis based on a human visual model," in *Proc. Int. Conf. Acoust. Speech, Signal Proc.*, (Alberquerque, NM), pp. 2137–2140, IEEE, 1990.

[8] D. Gabor, "Theory of communication," *J. Inst Elect. Engr.*, vol. 93, pp. 429–457, 1946.

[9] J. G. Daugman, "Uncertainty relation for resolution in space, spatial-frequency, and orientation optimized by two-dimensional visual cortical filters," *Journal of Optical Society of America*, vol. 2, no. 7, pp. 1160–1169, 1985.

[10] T. Chang and C. C. J. Kuo, "A wavelet transform approach to texture analysis," in *Proc. Int. Conf. Acoust. Speech, Signal Proc.*, vol. 3, pp. 661–664, 1992.

[11] P. P. Vaidyanathan, "Theory and design of M-channel maximally decimated quadrature mirror filters with arbitrary M, having the perfect reconstruction property," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 476–492, Apr. 1987.

[12] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 11, pp. 674–693, July 1989.

[13] F. W. Campbell and J. G. Robson, "Application of fourier analysis to the visibility of gratings," *J. Physiology*, vol. 197, pp. 551–566, 1968.

[14] D. F. Dunn, W. E. Higgins, and J. Wakeley, "Determining Gabor-filter parameters for texture segmentation," in *Intelligent robots and computer vision XI*, vol. 1826, pp. 51–63, 1992.

[15] T. Kohonen, "The self-organizing map," *Proc. IEEE*, vol. 78, pp. 1464 – 1480, Sept. 1990.

[16] T. Kohonen, J. Kangas, J. Laakonen, and K. Torkkola, "LVQ-PAK: The Learning Vector Quantization program package, version 2.1." Internet FTP, 1992.

[17] T. Randen, "Flerkanal filtrering for teksturseg mentering," Master's thesis, Rogaland University Center, 1993 (in Norwegian).

[18] P. Brodatz, *Textures: A photographic album for artists and designers.* NY: Dover, 1966.

[19] T. Randen and J. H. Husøy, "Novel approaches to multichannel filtering for image texture segmentation," in *Proc. SPIE's Visual Communications and Image Processing*, (Boston), pp. 626 – 636, Nov. 1993.

[20] J. H. Husøy, "Low complexity subband coding of still images and video," *Optical Engineering*, vol. 30, pp. 904–911, July 1991.

[21] A. N. Akansu and R. A. Haddad, *Multiresolution Signal Decomposition.* San Diego: Academic Press, 1992.

[22] J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks," in *Proc. Int. Conf. Acoust. Speech, Signal Proc.*, (Denver, CO), pp. 291–294, IEEE, 1980.

[23] J. H. Husøy, T. Randen, and T. O. Gulsrud, "Image texture classification with digital filter banks and transforms," in *Proc. SPIE International symposium on optical instrumentation and applied science: Applications of Digital Image Processing XVI*, (San Diego), pp. 260 – 271, July 1993.

[24] I. Daubechies, *Wavelets.* Philadelphia: S.I.A.M, 1992.