Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®

Pattern Recognition Letters 25 (2004) 391–398

Pattern Recognition Letters

www.elsevier.com/locate/patrec

Cosine-modulated wavelet based texture features for content-based image retrieval

Manesh Kokare *, B.N. Chatterji, P.K. Biswas

Electronics and Electrical Communication Engineering Department, Indian Institute of Technology, Kharagpur, West Bengal 721 302, India

Received 30 April 2003; received in revised form 27 October 2003

Abstract

Feature extraction is one of the most important tasks for efficient and accurate image retrieval purpose. In this paper we have presented a Cosine-modulated wavelet transform based technique for extraction of texture features. The major advantages of Cosine-modulated wavelet transform are less implementation complexity, good filter quality, and ease in imposing the regularity conditions. Texture features are obtained by computing the energy, standard deviation and their combination on each subband of the decomposed image. To check the retrieval performance, texture database of 1856 textures is created from Brodatz album. Retrieval efficiency and accuracy using Cosine-modulated wavelet based features is found to be superior to other existing methods.
© 2003 Elsevier B.V. All rights reserved.

Keywords: Cosine-modulated wavelet; Content-based image retrieval; Image database; Query image; Texture analysis

1. Introduction

The advancement of information technology demands communication and sharing of visual information globally. This is made possible due to the development of global computer network in the form of World Wide Web. Storage of visual information or image database has wide applications such as entertainment, film and video archieves, finger print or face identification, geographical information systems, remote sensing, medicine, surveillance etc. Retrieval of visual information from a large image database is one of the most challenging tasks in content-based image retrieval (CBIR). Comprehensive and extensive literature survey on content-based image retrieval is presented by Rui et al. (1999), Smeulders et al. (2000) and Kokare et al. (2002a). Content-based image retrieval lies at the crossroads of multiple disciplines such as database, artificial intelligence, image processing, statistics, computer vision, high performance computing, and human–computer intelligent interaction. Efficient and accurate image retrieval necessitates development of image description techniques, which will not only describe an image uniquely but also should be computationally very efficient.

* Corresponding author.
E-mail addresses: mbk@ece.iitkgp.ernet.in (M. Kokare), bnc@ece.iitkgp.ernet.in (B.N. Chatterji), pkb@ece.iitkgp.ernet.in (P.K. Biswas).

0167-8655/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2003.11.008

An image is described by a set of features. These features represent various image characteristics such as colour, texture, shape etc. Importance of texture feature is due to its presence in many real as well as synthetic data. As tigers and cheetahs have same colours but different texture patterns, so using colour feature alone cannot clearly distinguish between them. This phenomenon gives clear justification for texture features to be used in content-based image retrieval along with colour and shape. Texture describes the content of many real world images: for example, clouds, trees, bricks, hair, fabric etc. all of which have textural characteristics.

The main texture features currently used are derived from either Gabor wavelets or the conventional discrete wavelet transform. There is evidence that images are decomposed into a collection of band pass subimages by the simple visual cortical cells to form features for pattern recognition. Daugman (1980) reported that Gabor filters are suitable for such decomposition because impulse response of Gabor filter is similar to that of mammalian cortical cells. Manjunath and Ma (1996) reported Gabor wavelet based texture image retrieval results using four scales and six orientations. For constructing feature vector they used mean and standard deviation of the magnitude of the Gabor transform coefficients, resulting in a feature vector of size 24×2. Though Manjunath and Ma (1996) had done extensive experiments on a large set of textured images and shown that retrieval performance is better using Gabor filters than using conventional orthogonal wavelets, but the computational effort and storage requirement cause major problems. The basic requirement in content-based image retrieval for online application is that image feature extraction method should be computationally efficient with high retrieval accuracy and should require less storage space. Recent development in wavelet theory has provided a promising alternative through multichannel filter banks that have several potential advantages over Gabor filters namely,

(i) Wavelet filters cover exactly the complete frequency domain.
(ii) To facilitate computation fast algorithms are readily available.

Studies on successful application of wavelet theory on texture analysis mainly use the multiresolution signal decomposition developed by Mallat (1989). He used quadrature mirror filters to relate information at different scales of decomposition of the embedded subspace representation. The work of Chang and Kuo (1993) indicates that the texture features are more prevalent in the intermediate frequency band. Smith and Chang (1996) used mean and the variance of discrete wavelet transform coefficients to develop fully automated content-based image retrieval system called *VisualSEEk*.

A drawback of standard wavelets is that they are not suitable for the analysis of high-frequency signals with relatively narrow bandwidth. Kokare et al. (2002b) used the decomposition scheme based on M-band wavelets, which yields improved retrieval performance. Unlike the standard wavelet decomposition, which gives a logarithmic frequency resolution, the M-band decomposition gives a mixture of a logarithmic and linear frequency resolution. Further as an additional advantage, M-band wavelet decomposition yields a large number of subbands, which improves the retrieval accuracy. One of the drawbacks with M-band wavelet in content-based image retrieval is that computational complexity increases and hence retrieval time with number of bands. Gopinath and Burrus (1991) introduced Cosine-modulated class of multiplicity M wavelet tight frames (WTF's). In these WTF's, the scaling function uniquely determines the wavelets. This is in contrast to general multiplicity M case, where one has to, for any given application, design the scaling function and the wavelets. Hsin (2000) used a modulated wavelet transform approach for texture segmentation and reported that texture segmentation performance can be improved with this approach. Guillemot and Onno (1994) had used Cosine-modulated wavelet for image compression. They have presented procedure for designing Cosine-modulated wavelets for arbitrary length filters. This procedure allows obtaining filters with high stopband attenuation even in the presence of additional regularity constraints. Their results show that these filter solution provide good performance in image compression. The advantages of the Cosine-modulated wavelet are their low design and implementation complexities, good filter quality, and ease in imposing the regularity conditions, which yields improved retrieval performance both in terms of accuracy and retrieval time.

The main contributions of this paper are summarized as follows. First, in this paper we have presented novel texture features for content-based image retrieval using Cosine-modulated wavelet transform. Second, our approach of using the Canberra distance metric for similarity measurement improves the retrieval performance from 57.16% to 74.78% compared with the traditional Euclidean distance metric (where same features were used but Euclidean distance metric is used for similarity measurement). This shows that good performance in retrieval comes not just from a good set of features but also together with the use of suitable similarity measurement, which supports our approach. Another advantage of proposed method is that the retrieval time required is 6.69 times less than the Gabor based method, which is very important in CBIR. Third, a detailed comparison of the retrieval performance with standard Daubechies wavelet and Gabor wavelet method proposed by Manjunath and Ma (1996) is presented. The result indicates that retrieval performance of proposed method is superior to standard Daubechies wavelet and Gabor wavelet both in terms of accuracy and retrieval time. For large-scale evaluation our retrieval results are checked on large database of 1856 images.

This paper is organized as follows. In Section 2, Cosine-modulated wavelet for image retrieval is discussed in brief. The proposed image retrieval procedure is given in Section 3. Experimental results are given in Section 4, which is followed by the conclusion.

2. Cosine-modulated wavelet for content-based image retrieval

Fig. 1 shows an M-channel filter bank with analysis filters h_i and synthesis filters g_i. Filter bank is said to be *perfect reconstruction* if $y(n) = x(n)$. A *perfect reconstruction* filter bank is unitary if $g_i(n) = h_i(-n)$. Vaidyanathan (1992) reported that unitary (FIR) filter banks are practically important since they can be completely parameterized and efficiently implemented.

Moreover, they give rise to orthonormal wavelet bases for $L^2(R)$. A unitary filter bank where the lowpass filter satisfies the additional linear constraints given in Eq. (1) gives rise to wavelet tight frames.

$$\sum_{k=0}^{N-1} h_0(k) = \sqrt{M} \tag{1}$$

where $h_0(k)$ is lowpass filter of length N, and the number of channels are M.

This filter is the unitary scaling vector, and the remaining filters in the filter bank are the unitary wavelet vectors. The scaling and wavelet vectors determine the scaling function, $\psi_0(t)$, and the $(M-1)$ wavelets, $\psi_i(t)$, are defined by

$$\psi_i(t) = \sqrt{M} \sum_k h_i(k) \psi_0(Mt - k)$$
$$i \in \{0, \ldots, M-1\} \tag{2}$$

The $(M-1)$ wavelets $\psi_i(t)$, $i \in \{0, \ldots, M-1\}$, their translates and dilates by powers of M form

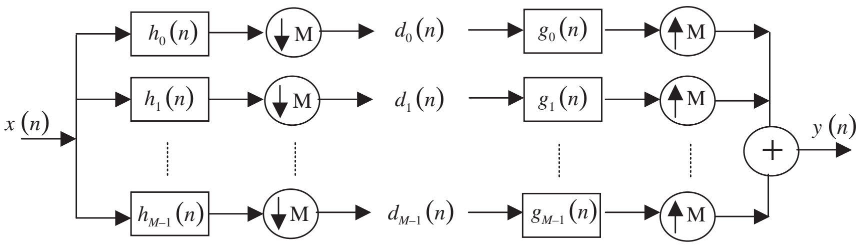

Fig. 1. M-channel filter bank.

a wavelet tight frame for $L^2(R)$ as reported by Gopinath and Burrus (1991) and Veterli and Herley (1992). For every function $f(t) \in L^2(R)$ one has

$$f(t) = \sum_{i=1}^{M-1} \sum_{j,k} \langle f, \psi_{i,j,k}(t) \rangle \psi_{i,j,k}(t) \tag{3}$$

where $\langle \cdot \rangle$ is an inner product and $\psi_{i,j,k}(t) = M^{j/2}\psi_i(M^j t - k)$.

A scaling vector is said to be K regular if its Z transform is of the form

$$H_0(z) = (1 + z^{-1} + \cdots + z^{-(M-1)})^K P(z) \tag{4}$$

for maximal possible K, and $P(z)$ is a polynomial in z^{-1}. Steffen et al. (1993) had shown that the minimal length K-regular scaling vectors are generically of length $N = MK$ and it can be constructed for all multiplicity M. The corresponding wavelet tight frames are called K-regular wavelet tight frames.

Modulated filter banks are special class of filter banks where the analysis and synthesis filters are obtained by modulation of prototype filters. Koilpillai and Vaidyanathan (1992) had reported that Cosine-modulated FIR filter banks are the special class of unitary filter banks, where the analysis filters $h_i(n)$ are all Cosine-modulates of a low pass linear-phase prototype filter $g(n)$. The fundamental idea behind Cosine-modulated filter banks is the following: In an M-channel filter bank, the analysis and synthesis filters are meant to approximate ideal Mth band filters, which are shown in Fig. 2. The passband of these filters occupy adjacent frequency channels that are $\frac{\pi}{M}$ apart. Given a real, prototype filter $g(n)$ with passband in $[-\frac{\pi}{2M}, \frac{\pi}{2M}]$, if it is modulated by $\cos((2i+1)\frac{\pi}{2M}n + \varepsilon_i)$, (where ε_i is arbitrary phase), has a passband equal to the desired band for the ith filter in Fig. 2. This technique gives rise to modulated filter banks (Eqs. (5) and (6)).

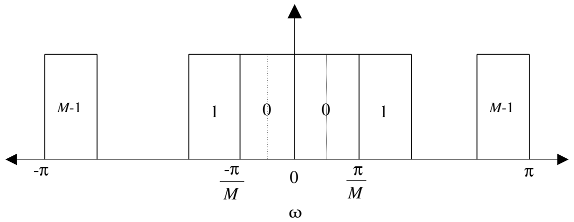

Fig. 2. Ideal frequency responses in M-channel filter bank.

$$h_i(n) = h(n)\cos\left(\frac{\pi}{2M}(2i+1)n + \varepsilon_i\right) \tag{5}$$

and

$$g_i(n) = g(n)\cos\left(\frac{\pi}{2M}(2i+1)n + \gamma_i\right) \tag{6}$$

where ε_i and γ_i are phase factors. Several choices of ε_i and γ_i have been reported by Koilpillai and Vaidyanathan (1992) and Malvar (1990). In this work we have used filter coefficients designed by Gopinath and Burrus (1995) by assuming $g(n) = h(n)$ with following phase factor.

$$h_i(n) = c_{i,n}g(n) \tag{7}$$

where $g(n)$ is an even-symmetric prototype filter of length $N = 2Mm$ for some nonnegative integer m and

$$C_{i,n} = \cos\left(\frac{\pi}{2M}(2i+1)\left(n - \frac{N-1}{2}\right) + \tau_i\right) \tag{8}$$

The phase factor τ_i can be taken to be $(-1)^i\frac{\pi}{4}$. In the filtering stage we make use of filter coefficients for $M = 2$ to decompose the texture image in to four channels, corresponding to different direction and resolutions. After decomposing image with wavelet transform we get horizontal, vertical and diagonal information. Hsin (2000) has reported that diagonal filter gives strong response to textures with orientations at or close to ±45°, the wavelet based features of similar textures with symmetric orientations are almost indistinguishable. Pattichis et al. (1997) reported that a large class of natural textures can be modeled as a quasiperiodic pattern and represented by modulated function, which motivates us to use Cosine-modulated wavelet for extracting texture features for content-based image retrieval.

3. Retrieval procedure for texture images

In this section texture image database used for experimental purpose, feature database creation and image retrieval method are discussed.

3.1. Texture image database

The texture database used in our experiment consists of 1856 texture images. For creating this database 116 different textures classes are used. We have used 108 textures from Brodatz (1966) texture photographic album, seven textures from USC database and one artificial texture. Size of each texture image is 512×512. Each 512×512 image is divided into sixteen 128×128 nonoverlapping subimages, thus creating a database of 1856 texture images.

3.2. Feature database creation

Each image from the database was analyzed using standard Daubechies wavelet and Cosine-modulated wavelet filter banks. The analysis was performed up to third level (4×3 = 12 subbands) of the wavelet decomposition. For constructing the feature vector feature parameters such as energy, standard deviation and combinations of both were computed separately on each subband and are stored in vector form. The basic assumption of this approach is that the energy distribution in the frequency domain identifies a texture. Besides providing acceptable retrieval performance from large texture, this approach is partly supported by physiological studies of the visual cortex as reported by Hubel and Wiesel (1962) and Daugman (1980). The energy and standard deviation of decomposed subbands are computed as follows:

$$\text{energy} = E_k = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} |W_{ij}| \tag{9}$$

$$\text{standard deviation} = \sigma_k = \left[\frac{1}{M \times N} \sum_{i=1}^{N} \sum_{j=1}^{M} (W_{ij} - \mu_{ij})^2 \right]^{\frac{1}{2}} \tag{10}$$

where $W(i,j)$ is the wavelet-decomposed subband, $M \times N$ is the size of wavelet-decomposed subband, k is the number of subbands ($k = 12$ for three levels), and μ_{ij} is the subband mean value.

A feature vector is now constructed using E_k and σ_k as feature components. Length of feature vector will be equal to (*No. of subbands*×*No. of feature parameters used in combination*) elements. Resulting feature vectors are as follows:

Using only energy feature measure

$$\bar{f}_E = [\, E_1 \quad E_2 \quad \cdots \quad E_k \,] \tag{11}$$

Using only standard deviation feature measure

$$\bar{f}_\sigma = [\, \sigma_1 \quad \sigma_2 \quad \cdots \quad \sigma_k \,] \tag{12}$$

Using combination of standard deviation and energy feature measure

$$\bar{f}_{\sigma\mu} = [\, \sigma_1 \quad \sigma_2 \quad \cdots \quad \sigma_k \quad E_1 \quad E_2 \quad \cdots \quad E_k \,] \tag{13}$$

For creation of feature database above procedure is repeated for all the images of the image database and these feature vectors are stored in feature database.

3.3. Image retrieval method

A query image is any one of the images from image database. This query image is processed to compute the feature vector as in Section 3.2. Traditional Euclidean distance metric and Canberra distance metrics are used to compute the similarity or match value for given pair of images. If x and y are two d-dimensional feature vectors of database image and query image respectively, then these distance metrics are defined as:

The Euclidean or L2 metric is

$$d_{\mathrm{E}}(x,y) = \sqrt{\sum_{i=1}^{d} (x_i - y_i)^2} \tag{14}$$

Euclidean distance is not always the best metric. The fact that the distances in each dimension are squared before summation, places great emphasis on those features for which the dissimilarity is large. Hence it is necessary to normalize the individual feature components before finding the distance between two images. This has been taken care of in Canberra distance metric, which motivates us to use Canberra distance metric as dissimilarity measure. This is also supported by our experimental results given Table 1. Canberra distance is given by

Table 1
Average retrieval rate of all the database images and retrieval time of query image

Feature	Standard Daubechies wavelet			Cosine-modulated wavelet			Gabor wavelet method proposed by Manjunath and Ma (1996)	
	Retrieval accuracy		Retrieval time (s)	Retrieval accuracy		Retrieval time (s)	Retrieval accuracy (%)	Retrieval time (s)
	Euclidean distance (%)	Canberra distance (%)		Euclidean distance (%)	Canberra distance (%)			
Energy	41.76	70.91	0.488	40.19	70.85	0.488	69.83	3.429
Standard deviation	49.95	68.64	0.516	58.99	67.24	0.516	59.64	3.528
Energy + standard deviation	48.55	71.71	0.529	57.16	**74.78**	0.529	74.32	3.54

$$\mathrm{Canb}(x,y) = \sum_{i=1}^{d} \frac{|x_i - y_i|}{|x_i| + |y_i|} \tag{15}$$

In Eq. (15) the numerator signifies the difference and denominator normalizes the difference. Thus distance values will never exceed one, being equal to one whenever either of the attributes is zero. Thus it would seem to be a good expression to use, which avoids scaling effect. It is obvious that the distance of an image from itself is zero. The distances are stored in increasing order and the closest sets of patterns are retrieved. In the ideal case all the top 16 retrievals are from the same large image. The performance is measured in terms of the average retrieval rate, which is defined as the average percentage number of patterns belonging to the same image as the query pattern in the top 16 matches.

4. Experimental results

Table 1 provides a detailed comparison of average retrieval accuracy for entire database and retrieval time for query image using standard Daubechies wavelet, and Cosine-modulated wavelet with feature parameters such as energy, standard deviation and combination of both. Comparison of retrieval performance using traditional Euclidean distance metric and Canberra distance metric is also presented. These results are also compared with Gabor wavelet based method proposed by Manjunath and Ma (1996). The proposed retrieval system has been implemented using MATLAB on Pentium III, 866 MHz machine. It is observed that concerning retrieval accuracy the combination of energy and standard deviation as feature measure gives the best performance and Canberra distance metric outperforms the traditional Euclidean distance metric. Table 1 shows that average retrieval performance using Cosine-modulated wavelet is (74.78%) better than standard Daubechies wavelet (71.71%) and is marginally better than that in case of Gabor wavelet method (74.32%) proposed by Manjunath and Ma (1996). The retrieval time with Cosine-modulated wavelet greatly outperforms (6.69 times less) Gabor wavelet based method. The reason behind getting good retrieval accuracy and less retrieval time is that Cosine-modulated wavelets transform has less implementation complexity, good filter quality, and ease in imposing the regularity conditions.

Fig. 3 shows retrieval performance of standard Daubechies wavelet, Cosine-modulated wavelet, and Gabor wavelet according to the number of top matches considered. From that figure it is clear that the retrieval performance of Cosine-modulated wavelet is superior to standard Daubechies wavelet and Gabor wavelet. If the top 116 (6% of the database) retrievals are considered the performance increases up to 94.77%, 91.65% and 92.375% using Cosine-modulated wavelet, standard Daubechies wavelet and Gabor wavelet respectively.

Retrieved top 20 similar images from the database of 1856 images using Cosine-modulated wavelet for a sample query image are shown in Fig. 4.

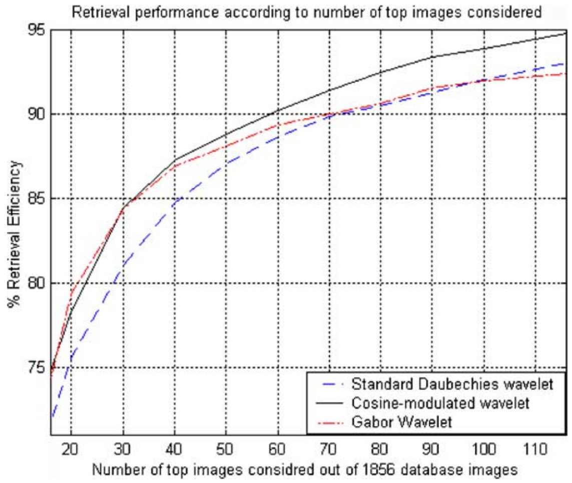

Fig. 3. Average retrieval rate according to no. of top images considered.

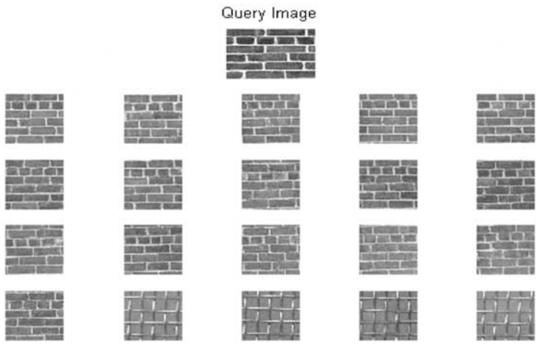

Fig. 4. Retrieved top 20 similar images from the database of 1856 images using Cosine-modulated wavelet.

5. Conclusion

Two main requirements of content-based image retrieval are that it should have high retrieval accuracy and less computational complexity. We have presented a Cosine-modulated wavelet technique for content-based image retrieval. Large texture database of 1856 images is used to check the retrieval performance. All the database images were decomposed using standard Daubechies wavelet, and Cosine-modulated wavelet with pyramidal representation. Features were computed on the decomposed subbands. A Canberra distance metric is used to discriminate 116 different textures. A detailed comparison of the retrieval performance using feature measures such as standard deviation, energy and combinations of both is presented. Amongst all these feature parameters we found that combination of standard deviation and energy gives the best retrieval performance. The result indicates that retrieval performance of Cosine-modulated wavelet is superior to standard Daubechies wavelet and Gabor wavelet both in terms of accuracy and retrieval time. Previous work reported by Manjunath and Ma (1996) indicates that Gabor feature gives the best retrieval performance (74.32%) than other multiresolution method at the cost of high computational complexity and hence requires large retrieval time. Here the proposed method provides a strong alternative with less computational complexity and high retrieval accuracy (74.78%) than Gabor features. Retrieval time required for Cosine-modulated wavelet is 6.69 times less than Gabor wavelet. This makes the proposed method a strong candidate, which achieves both the important requirements for texture feature extraction in content-based image retrieval.

References

Brodatz, P., 1966. Textures: A Photographic Album for Artists & Designers. Dover, New York.

Chang, T., Kuo, C.C., 1993. Texture analysis and classification with tree-structured wavelet transform. IEEE Trans. Image Process. 2 (4), 429–441.

Daugman, J.G., 1980. Two-dimensional spectral analysis of cortical receptive field profiles. Vision Res. 20, 847–856.

Gopinath, R.A., Burrus, C.S., 1991. Wavelets and filter banks. In: Chui, C.K. (Ed.), Wavelets: A Tutorial in Theory and Applications. Academic, San Diego, CA, pp. 603–654. Tech. Report CML TR91-20, Rice University.

Gopinath, R.A, Burrus, C.S., 1995. On Cosine-modulated wavelet orthonormal bases. IEEE Trans. Image Process. 4 (2), 162–176.

Guillemot, C., Onno, P., 1994. Cosine-modulated wavelets: New results on design of arbitrary length filters and optimization for image compression. In: Proc. Internat. Conf. on Image Processing 1, Austin, TX, USA, pp. 820–824.

Hsin, H.C., 2000. Texture segmentation using modulated wavelet transform. IEEE Trans. Image Process. 9 (7), 1299–1302.

Hubel, D.H., Wiesel, T.N., 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. 160, 106–154.

Koilpillai, R.D., Vaidyanathan, P.P., 1992. Cosine-modulated FIR filter bank satisfying perfect reconstruction. IEEE Trans. Signal Process., 770–783.

Kokare, M., Chatterji, B.N., Biswas, P.K., 2002a. A survey on current content-based image retrieval methods. IETE J. Res. 48 (3&4), 261–271.

Kokare, M., Chatterji, B.N., Biswas, P.K., 2002b. *M*-Band wavelet based texture features for content-based image retrieval. In: Proc. Indian Conf. on Vision, Graphics Image Processing (ICVGIP 2002), Ahemadabad, India, pp. 173–178.

Mallat, S., 1989. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans. Pattern Anal. Machine Intell. 11 (7), 674–693.

Malvar, H.S., 1990. Modulated QMF filter banks with perfect reconstruction. Electron. Lett. 26, 906–907.

Manjunath, B.S., Ma, W.Y., 1996. Texture features for browsing and retrieval of image data. IEEE Trans. Pattern Anal. Machine Intell. 8 (8), 837–842.

Pattichis, M.S., Christodoulou, C., Pattichis, C.S., Bovik, A.C., 1997. Non-stationary texture segmentation using an AM-FM model. In: Proc. Internat. Conf. on Neural Networks 3, Houston, TX, pp. 1552–1557.

Rui, Y., Huang, T.S., Chang, S.F., 1999. Image retrieval: Current techniques, promising directions and open issues. J. Visual Comm. Image Representation 10, 39–62.

Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain, R., 2000. Content-based image retrieval at the end of the early years. IEEE Trans. Pattern Anal. Machine Intell. 22 (12), 1349–1380.

Smith, J.R., Chang, S.F., 1996. VisualSEEk: a fully automated content based image query system. In: Proc. ACM Internat. Conf. on Multimedia, Boston, MA, pp. 87–98.

Steffen, P., Heller, P., Gopinath, R.A., Burrus, C.S., 1993. Theory of regular *M*-band wavelet bases. IEEE Trans. Signal Process. 41 (12), 3497–3510.

Vaidyanathan, P.P., 1992. Multirate Systems and Filter Banks. Prentice Hall, Englewood Cliffs, NJ.

Veterli, M., Herley, C., 1992. Wavelets and filter banks: Theory and design. IEEE Trans. Acoust. Speech Signal Process. 40 (9), 2207–2232.