Journal of Zhejiang University SCIENCE A
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

A novel texture clustering method based on shift invariant DWT and locality preserving projection*

Rui XING†, San-yuan ZHANG, Le-qing ZHU
(*School of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China*)
†E-mail: xingrui@zju.edu.cn

Received Mar. 2, 2008; Revision accepted June 11, 2008; Crosschecked Dec. 26, 2008

Abstract: We propose a novel texture clustering method. A classical type of (approximate) shift invariant discrete wavelet transform (DWT), dual tree DWT, is used to decompose texture images. Multiple signatures are generated from the obtained high-frequency bands. A locality preserving approach is applied subsequently to project data from high-dimensional space to low-dimensional space. Shift invariant DWT can represent image texture information efficiently in combination with a histogram signature, and the local geometrical structure of the dataset is preserved well during clustering. Experimental results show that the proposed method remarkably outperforms traditional ones.

Key words: Shift invariant DWT, Texture signature, Local preserving clustering, Dimension reduction, *k*-means
doi:10.1631/jzus.A0820145 **Document code:** A **CLC number:** TP391

INTRODUCTION

As an important feature, texture plays a critical role in image analysis such as in content based image retrieval (CBIR) (Kokare *et al*., 2005), object detection (Jeong and Nedevschi, 2005), and remote sensing image analysis (Kandaswamy *et al*., 2005). Discrete wavelet transform (DWT) is a commonly used tool to analyze image textures. Since small shifts in the input signal can result in large differences in DWT coefficients of different scales, the same two patterns with small spatial shifts will produce distant feature vectors. To overcome this, shift invariant DWT is introduced as a powerful tool in image texture analysis. Olkkonen and Olkkonen (2007) proposed a general framework for shift invariant DWT based on two parallel wavelet transforms, where the wavelet forms a Hilbert transform pair. Many methods showing how to design shift invariant DWT exist in the literature. Kingsbury (2001) proposed a dual tree DWT method, where the real and imaginary parts of the complex wavelet coefficients are approximately a Hilbert transform pair. Subsequently, Selesnick (2002) formalized the notion of a Hilbert pair and derived the condition in a pair of filter banks that yield equivalent wavelet functions that are a Hilbert transform of each other. By following the work of Kingsbury and Selesnick, Gopinath (2003) presented the phaselet transform, an approximately shift invariant redundant dyadic wavelet transform. Recently, Tay (2006) proposed a new class of filter bank, where the filters have even-length and are used to match a given odd-length filter bank. Therefore the equivalent wavelet functions of both filter banks are an approximate Hilbert transform of each other.

So far a lot of clustering methods have been proposed. However, an image space is typically of high dimensionality, which makes clustering on original image data a difficult task. This is regarded as the curse of dimensionality. To ease the problem, raw

* Project supported by the Hi-Tech Research and Development Program (863) of China (Nos. 2007AA01Z311 and 2007AA04Z1A5), the National Basic Research Program (973) of China (No. 2009CB320804), the National Research Foundation for the Doctoral Program of Higher Education of China (No. 20060335114), and the Science and Technology Program of Zhejiang Province, China (No. 2007C21006)

data are usually projected from high-dimensional space onto low-dimensional space so that clustering can be done in the lower one. Traditional methods used to reduce dimensionality are principal component analysis (PCA) (Jolliffe, 1989) and linear discriminant analysis (LDA) (Duda *et al*., 2000).

In this paper we propose a novel clustering method on image texture. First dual tree DWT, a typical type of shift invariant DWT, is employed on texture images, which generates texture signatures from the decomposition coefficients. Then a locality preserving projection is used to explore the low-dimensional locality structure lying in high-dimensional space. In order to keep a good locality relationship, the similarity matrix is constructed based on the nearest neighbor algorithm. Our experimental results show that the proposed method succeeds well in clustering texture images and performs better than traditional methods.

SHIFT INVARIANT DWT

Recent studies show that a human interprets an image in many different directions and frequency channels, which is similar to the multiple-resolution property of a wavelet. For this reason, the wavelet transform has been widely applied in image processing such as in texture analysis and texture synthesis. However, because of the shift variance, even a small shift in image texture would greatly perturb the transform and produce very different coefficients.

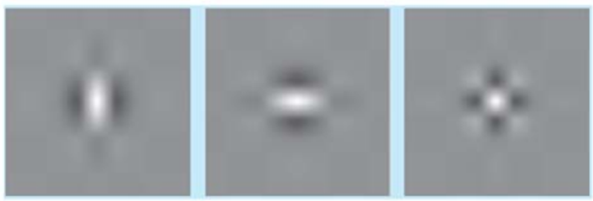
To overcome the problem, a shift invariant DWT was proposed and exploited to extract image texture features. In this study we used one classical shift invariant DWT, namely dual tree DWT, proposed by Kingsbury (2001). Dual tree DWT is performed based on a real wavelet transform instead of using complex coefficient filters. Decomposition is applied to two parallel DWT trees. By elaborately designing the filters of the two trees, the samplings are evenly distributed in the space, which ensures the approximate shift invariance. After dual tree DWT, an image is decomposed into six sub-bands in six different directions. Though the transform does not use complex coefficient filters, the output of the two wavelet trees can be considered as the real and imaginary parts of a complex wavelet transform. Compared with conventional DWT, dual tree DWT works well because of approximate shift invariance as well as good directional selectivity, high computational efficiency and low redundancy.

A 1D dual tree DWT uses a pair of filters on input data. It contains two parallel 1D discrete wavelet transform trees (tree A and tree B). Let $\psi_h(t)$ and $\psi_g(t)$ be the wavelets of tree A and tree B, respectively. The transform can be written as follows:

$$\psi(t)=\psi_h(t)+j\psi_g(t), \tag{1}$$

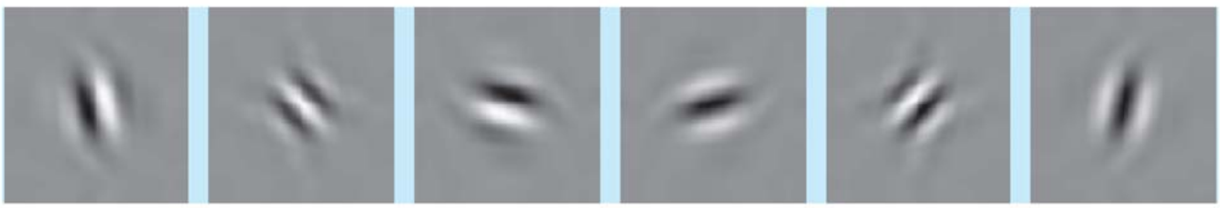
where $\psi_g(t)$ is approximately the Hilbert transform of $\psi_h(t)$, i.e., $\psi_g(t)\approx H\{\psi_h(t)\}$.

As shown in Fig.1, a 2D DWT is oriented in three directions (vertical, horizontal, and diagonal). Dual tree DWT, as can be seen in Fig.2, is oriented in six directions (±15°, ±45°, ±75°). Therefore, 2D dual tree DWT shows more directional selectivity than 2D DWT. In other words, 2D dual tree DWT produces more directional texture information than 2D DWT.



(a) (b) (c)

Fig.1 Impulse responses in three different directions of a 2D DWT. (a) Vertical; (b) Horizontal; (c) Diagonal



(a) (b) (c) (d) (e) (f)

Fig.2 Impulse responses in six different directions of a 2D dual tree DWT
(a) −75°; (b) −45°; (c) −15°; (d) 15°; (e) 45°; (f) 75°

TEXTURE SIGNATURE

After the wavelet decomposition is performed, two feature parameters, energy (E_k) and standard deviation (σ_k), can be obtained:

$$E_k = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \left| W_k(i, j) \right|, \tag{2}$$

$$\sigma_k = \left[\frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} (W_k(i, j) - \mu_k)^2 \right]^{\frac{1}{2}}, \tag{3}$$

where $W_k(i, j)$ is the kth wavelet-decomposed sub-band, $M \times N$ is the size of the wavelet-decomposed sub-band, and μ_k is the mean value of the kth sub-band.

In addition, a histogram signature (van de Wouwer *et al*., 1999) is adopted to improve the texture discriminative ability. The detail histograms of natural textured images can be modeled by a family of exponentials:

$$h(u) = K \exp[-(|u| / \alpha)^{\beta}], \tag{4}$$

where $h(u)$ denotes the histogram of the wavelet detail coefficients, α and β are the histogram signatures of the detailed texture images, and K is a normalization constant to ensure that $\int h(u) \mathrm{d}u = 1$.

The model parameters α, β and K can be computed as follows. Given

$$c_1 = \int |u|\, h(u) \mathrm{d}u, \tag{5}$$

$$c_2 = \int |u|^2\, h(u) \mathrm{d}u, \tag{6}$$

inserting Eq.(4) into Eqs.(5) and (6) using the normalization condition, we have

$$\beta = H^{-1}(c_1^2 / c_2), \tag{7}$$

$$\alpha = c_1 \frac{\Gamma(1/\beta)}{\Gamma(2/\beta)} = c_1 \exp[\ln \Gamma(1/\beta) - \ln \Gamma(2/\beta)], \tag{8}$$

where

$$H(x) = \frac{\Gamma^2(2/x)}{\Gamma(3/x)\Gamma(1/x)}, \tag{9}$$

$$\Gamma(x) = \int_0^{\infty} \mathrm{e}^{-t} t^{x-1} \mathrm{d}t. \tag{10}$$

A fast algorithm is presented to compute the natural logarithm of Gamma function in Eq.(8) (Cody and Hillstrom, 1967).

In this study the histogram signatures α and β are combined with energy and standard deviation features for clustering, which are demonstrated in the next section.

LOCALITY PRESERVING PROJECTION

Generally, image data are of high dimensionality. This makes dimension reduction a necessary task before clustering. The graph based clustering (Shi and Malik, 2000; O′Callaghan and Bull, 2005) has received a lot of attention since it explores the local similarity between data points. In this work we use the locality preserving projection (LPP) (He and Niyogi, 2003; Cai *et al*., 2005; He *et al*., 2005) to reduce the high dimensionality.

Suppose that we have a set of data points $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, …, $\boldsymbol{x}_n$, and the similarity matrix $\boldsymbol{S}$ in which the element s_{ij} denotes the similarity between point $\boldsymbol{x}_i$ and point $\boldsymbol{x}_j$. Thus we can use graph $G=(\boldsymbol{V}, \boldsymbol{E})$ to represent the dataset. Vertex $\boldsymbol{v}_i$ corresponds to data point $\boldsymbol{x}_i$. Two points are connected if the similarity between them is larger than zero. The weight on the connecting edge is the similarity s_{ij}. The cluster problem can be solved by finding a minimum cut over the graph.

The local preserving clustering is stated as follows:

Given graph $G=(\boldsymbol{V}, \boldsymbol{E})$ and its vertex set $\boldsymbol{V}=\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$, the similarity matrix is $\boldsymbol{S}=(s_{ij})_{n\times n}$. Because G is an undirected graph, $s_{ij}=s_{ji}$. The degree of vertex $\boldsymbol{v}_i$ is

$$d_i = \sum_{j=1}^{n} w_{ij}. \tag{11}$$

Degree matrix $\boldsymbol{D}$ is a diagonal matrix whose elements are d_1, d_2, …, d_n. Define the Laplacian matrix of G as $\boldsymbol{L}=\boldsymbol{D}-\boldsymbol{W}$, where $\boldsymbol{W}$ is the adjacency matrix constructed by $\boldsymbol{S}$.

Input: dataset $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, …, $\boldsymbol{x}_n$, similarity matrix $\boldsymbol{S} \in \mathbb{R}^{n\times n}$, and clustering number k.

Step 1: Construct adjacency matrix $\boldsymbol{W}$ with $\boldsymbol{S}$, degree matrix $\boldsymbol{D}$ and Laplacian matrix $\boldsymbol{L}$.

Step 2: Remove the weighted mean from each data point $\boldsymbol{x}_i$ and obtain $\hat{\boldsymbol{x}}_i$:

$$\hat{\boldsymbol{x}}_i = \boldsymbol{x}_i - \overline{\boldsymbol{x}}, \ \overline{\boldsymbol{x}} = \sum_i \boldsymbol{x}_i D_{ii} \Big/ \sum_i D_{ii}. \quad (12)$$

Step 3: Project the vector into the singular value decomposition (SVD) sub-space by removing zero singular values:

$$\hat{\boldsymbol{X}} = \boldsymbol{U\Sigma V}^{\mathrm{T}}. \quad (13)$$

where $\hat{\boldsymbol{X}} = [\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \dots, \hat{\boldsymbol{x}}_n]$. Let $\boldsymbol{W_V}$ be $\boldsymbol{U}$. The SVD projection is

$$\tilde{\boldsymbol{x}} = \boldsymbol{W}_{\boldsymbol{V}}^{\mathrm{T}} \hat{\boldsymbol{x}}. \quad (14)$$

The data obtained in this step are $\tilde{\boldsymbol{X}} = [\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_2, \dots, \tilde{\boldsymbol{x}}_n]$.

Step 4: Solve the following eigen-problem:

$$\tilde{\boldsymbol{X}}\boldsymbol{L}\tilde{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{a} = \lambda \tilde{\boldsymbol{X}}\boldsymbol{D}\tilde{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{a}. \quad (15)$$

Sort the eigenvalue of Eq.(15) in ascending order and obtain the corresponding eigenvector $\boldsymbol{W_L}=[a_1, a_2, \dots, a_k]$. The dimensionality-reduced representation of the data point is

$$\boldsymbol{y}_i = (\boldsymbol{W_V W_L})^{\mathrm{T}} \hat{\boldsymbol{x}}_i. \quad (16)$$

Step 5: Cluster the dataset $\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_n$ into $C_1, C_2, \dots, C_k$ via k-means (McQueen, 1967).

Output: cluster set $A_1, A_2, \dots, A_k$, where $A_i=\{j | \boldsymbol{y}_j \in C_i\}$.

THE PROPOSED METHOD

Our proposed method is performed as follows:

Step 1: Preprocess and normalize images.

Step 2: Apply 2D dual tree DWT on the image, using a near-symmetric 13-, 19-tap filters for the first level and Q-shift 14-tap for the rest.

Step 3: Generate texture signatures from the wavelet coefficients.

Step 4: Compute the Canberra distance between images according to the obtained texture features. Construct the weighted adjacency matrix $\boldsymbol{W}$, where $W_{ij}=\exp[-\|\boldsymbol{x}_i-\boldsymbol{x}_j\|^2/(2\alpha^2)]$. In our case we set $\alpha=0.2$.

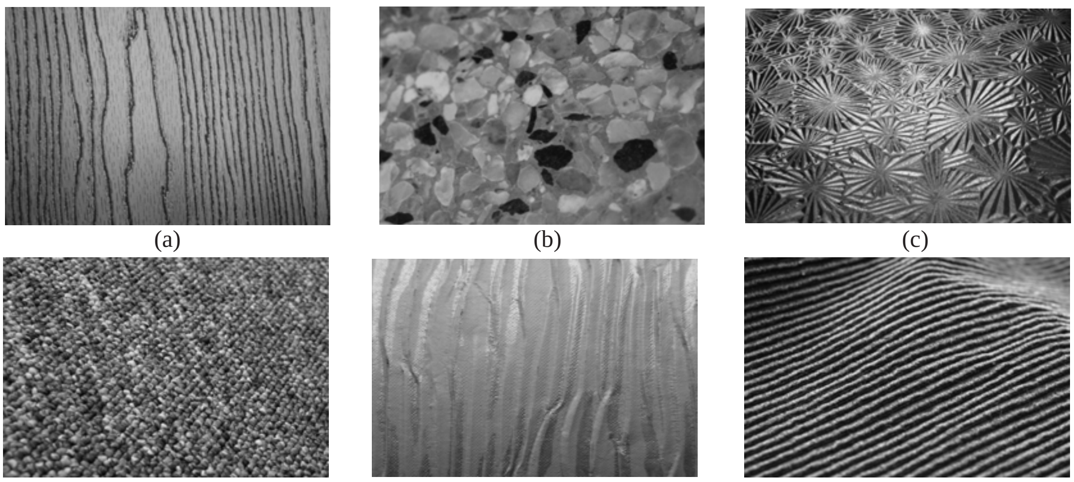
Step 5: Change $\boldsymbol{W}$ based on the p-nearest neighbor algorithm (Funkunaga and Navendra, 1975). In our experiment we set $p=10$.

Step 6: Perform dimension reduction using locality preserving projection to obtain representative data $\boldsymbol{Y}=[\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_n]$.

Step 7: Cluster $\boldsymbol{Y}$ using the k-means algorithm.

EXPERIMENTAL RESULTS

In our experiment we compared the proposed method with traditional PCA+k-means and direct k-means methods. The texture image data were downloaded from the Ponce texture database of UIUC (Lazebnik *et al*., 2005), which has 25 kinds of textures and each kind has 40 images, as shown in Fig.3.



(a) (b) (c) (d) (e) (f)

Fig.3 Texture samples used in the experiments (all from Ponce collection)

(a) Wood; (b) Floor; (c) Glass; (d) Carpet; (e) Wallpaper; (f) Corduroy

To quantitatively evaluate the results, two metrics, the accuracy (*AC*) and the normalized mutual information (*MI*), are used to measure the clustering performance (Xu *et al*., 2003). *AC* is defined as

$$AC = \frac{1}{n}\sum_{i=1}^{n}\delta(s_i, map(r_i)), \tag{17}$$

where r_i and s_i denote the obtained cluster label and the actual label of image x_i, respectively. $\delta(x, y)=1$ if $x=y$, otherwise 0. $map(r_i)$ is the permutation mapping function (Lovasz and Plummer, 1986) that maps each cluster label r_i to the equivalent label from the data corpus. The best mapping can be found by using the Kuhn-Munkres algorithm.

Another metric is mutual information. Suppose C denotes the actual cluster set and C' denotes the obtained cluster set. Then the mutual information between the two cluster sets is defined as follows:

$$MI(C,C') = \sum_{c\in C, c'\in C'} p(c_i, c'_j)\cdot \log_2 \frac{p(c_i, c'_j)}{p(c_i)\cdot p(c'_j)}, \tag{18}$$

where $p(c_i)$ is the possibility that a randomly selected image belongs to c_i, $p(c_j')$ is the possibility that a randomly selected image belongs to c_j', and $p(c_i, c_j')$ is the possibility that the image belongs to c_i and c_j' at the same time. The normalized mutual information is

$$\overline{MI}(C,C') = \frac{MI(C,C')}{\max(H(C), H(C'))}, \tag{19}$$

where $H(C)$ and $H(C')$ denote the entropies of C and C', respectively.

Table 1 gives the results of the comparison of the three different texture features, namely DWT, dual tree DWT and dual tree DWT signature. As can be seen, dual tree DWT achieves good results in the experiments compared to DWT. Moreover, dual tree DWT signature shows better performance than the other two texture features.

Table 2 shows the comparison of the experimental results based on dual tree DWT signature features. As we can see, the proposed method performs better than direct *k*-means and PCA+*k*-means, in both accuracy and mutual information.

Table 1 Comparison of accuracy and mutual information of the three different texture features

k	Accuracy			Mutual information		
	DWT	Dual tree DWT	Dual tree DWT signature	DWT	Dual tree DWT	Dual tree DWT signature
2	0.863	0.900	0.925	0.726	0.681	0.679
3	0.850	0.892	0.950	0.813	0.854	0.865
4	0.719	0.744	0.756	0.564	0.585	0.597
5	0.690	0.695	0.710	0.541	0.569	0.597
6	0.713	0.746	0.754	0.597	0.623	0.661
Average	0.767	0.795	0.819	0.648	0.662	0.680

Table 2 Comparison of accuracy and mutual information of the proposed algorithm with other high performance clustering algorithms

k	Accuracy			Mutual information		
	k-means	PCA+*k*-means	Our algorithm	*k*-means	PCA+*k*-means	Our algorithm
2	0.887	0.838	0.925	0.579	0.468	0.679
3	0.742	0.925	0.950	0.616	0.823	0.865
4	0.650	0.706	0.756	0.451	0.484	0.597
5	0.635	0.615	0.710	0.512	0.536	0.597
6	0.621	0.750	0.754	0.558	0.633	0.661
Average	0.707	0.767	0.819	0.543	0.589	0.680

CONCLUSION

In this paper, we proposed a new method to cluster texture image using shift invariant DWT and locality preserving projection. Firstly a classical type of approximate shift invariant DWT, dual tree DWT, is employed on texture images. After decomposition, multiple histogram signatures are extracted from the obtained high frequency bands and combined with energy and standard deviation to form image features. Then a weighted adjacency matrix is constructed according to the Canberra distance between each pair of data points. During the construction of the adjacency matrix, we exploit the nearest neighbor algorithm to utilize the local relationship. Locality preserving projection is then applied to reduce the data from high-dimensional space to low-dimensional space, where the traditional *k*-means method is used for clustering.

Since the shift invariant DWT with histogram signatures can extract image texture information efficiently and locality preserving projection keeps well the local geometrical structure of the dataset, the experimental results show that the proposed method achieves good accuracy and outperforms traditional clustering methods in both accuracy and mutual information.

As dual tree DWT is only one of the solutions in the shift invariant DWT family, further research could be carried out on extending the proposed method to other existing shift invariant DWT algorithms for comparison and evaluation.

References

Cai, D., He, X., Han, J., 2005. Document clustering using locality preserving indexing. *IEEE Trans. Knowl. Data Eng.*, **17**(12):1624-1637. [doi:10.1109/TKDE.2005.198]

Cody, W.J., Hillstrom, K.E., 1967. Chebyshev approximations for the natural logarithm of the Gamma function. *Math. Comput.*, **21**(98):198-203. [doi:10.2307/2004160]

Duda, R.O., Hart, P.E., Stork, D.G., 2000. Pattern Classification (2nd Ed.). Wiley Interscience, Hoboken, NJ.

Funkunaga, K., Navendra, P.M., 1975. A branch and bound algorithm for computing *k*-nearest neighbors. *IEEE Trans. Comput.*, **C-24**(7):750-753. [doi:10.1109/T-C.1975.224297]

Gopinath, R.A., 2003. The phaselet transform—an integral redundancy nearly shift-invariant wavelet transform. *IEEE Trans. Signal Processing*, **51**(7):1792-1805. [doi:10.1109/TSP.2003.812833]

He, X., Niyogi, P., 2003. Locality Preserving Projections. Advances in Neural Information Processing Systems 16, Vancouver, Canada, p.153-160.

He, X., Yan, S., Hu, Y., Niyogi, P., Zhang, H., 2005. Face recognition using Laplacianfaces. *IEEE Trans. Patten Anal. Mach. Intell.*, **27**(3):328-340. [doi:10.1109/TPAMI.2005.55]

Jeong, P., Nedevschi, S., 2005. Efficient and robust classification method using combined feature vector for lane detection. *IEEE Trans. Circuits Syst. Video Technol.*, **15**(4):528-537. [doi:10.1109/TCSVT.2005.844453]

Jolliffe, I.T., 1989. Principal Component Analysis. Springer-Verlag, New York.

Kandaswamy, U., Adjeroh, D.A., Lee, M.C., 2005. Efficient texture analysis of SAR imagery. *IEEE Trans. Geosci. Remote Sensing*, **43**(9):2075-2083. [doi:10.1109/TGRS.2005.852768]

Kingsbury, N., 2001. Complex wavelets for shift invariant analysis and filtering of signals. *Appl. Comput. Harmon. Anal.*, **10**(3):234-253. [doi:10.1006/ACHA.2000.0343]

Kokare, M., Biswas, P.K., Chatterji, B.N., 2005. Texture image retrieval using new rotated complex wavelet filters. *IEEE Trans. Syst., Man Cybern.-Part B*, **35**(6):1168-1178. [doi:10.1109/TSMCB.2005.850176]

Lazebnik, S., Schmid, C., Ponce, J., 2005. A sparse texture representation using local affine regions. *IEEE Trans. Pattern Anal. Mach. Intell.*, **27**(8):1265-1278. [doi:10.1109/TPAMI.2005.151]

Lovasz, L., Plummer, M., 1986. Matching Theory. Akademiai Kiado, North Holland, Budapest.

McQueen, J., 1967. Some Methods for Classification and Analysis of Multivariate Observations. Proc. 5th Berkeley Symp. on Mathematical Statistics and Probability, p.281-297.

Olkkonen, H., Olkkonen, J.T., 2007. Half-delay B-spline filter for construction of shift-invariant wavelet transform. *IEEE Trans. Circuits Syst. II: Express Briefs*, **54**(7):611-615. [doi:10.1109/TCSII.2007.895327]

O′Callaghan, R.J., Bull, D.R., 2005. Combined morphological-spectral unsupervised image segmentation. *IEEE Trans. Image Processing*, **14**(1):49-62. [doi:10.1109/TIP.2004.838695]

Selesnick, I.W., 2002. The design of approximate Hilbert transform pairs of wavelet bases. *IEEE Trans. Signal Processing*, **50**(5):1144-1152. [doi:10.1109/78.995070]

Shi, J., Malik, J., 2000. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, **22**(8):888-905. [doi:10.1109/34.868688]

Tay, D.B.H., 2006. ETHFB: A New Class of Even-length Wavelet Filters for Hilbert Pair Design. IEEE Int. Symp. on Circuits and Systems, p.1083-1086. [doi:10.1109/ISCAS.2006.1692777]

van de Wouwer, G., Scheunders, P., van Dyck, D., 1999. Statistical texture characterization from discrete wavelet representation. *IEEE Trans. Image Processing*, **8**(4):592-598. [doi:10.1109/83.753747]

Xu, W., Liu, X., Gong, Y., 2003. Document Clustering Based on Non-negative Matrix Factorization. Proc. Int. Conf. on Research and Development in Information Retrieval, p.267-273. [doi:10.1145/860435.860485]