Unsupervised Texture Segmentation: Comparison of Texture Features

AHSAN AHMAD URSANI*, ABDUL WAHEED UMRANI*, AND FAHIM AZIZ UMRANI**

ABSTRACT

Texture is an important image-content that has been utilized for different machine intelligent tasks, like those in machine vision and remote sensing, which identify objects of interest by segmenting the image texture. This paper aims at comparing texture features based on DFT (Discrete Fourier Transform) with ones based on Gabor wavelets for unsupervised image segmentation. The comparison is realized theoretically, analytically, as well as empirically. Images of natural scenes from a standard image database have been taken as test images. Analytical comparison shows that the DFT-based features are computationally less expensive than those based on Gabor wavelets. Empirical results show that the performance of the texture features based on DFT is comparable to those based on Gabor wavelets.

Key Words: Texture Segmentation, DFT-Based Texture Features, Gabor Wavelets, K-Means Clustering.

1.   INTRODUCTION

Although one can find many definitions of texture in the contemporary literature, it has no definition that is uncontroversial and agreed upon unanimously by the image processing research community. Researchers have devised several definitions of texture comprising from a single sentence to several sentences. Nonetheless, one can describe texture as a pattern of gray level changes in an image or a local region therein [1]. This property of image, called texture, is an important image characteristic and an essential part of the computer vision, CBIR (Content-Based Image Retrieval) systems and the remote sensing alike. The texture processing, including texture segmentation, recognition, classification, synthesis, and shape from texture, is now already a mature field of research. Texture segmentation and classification is performed on the medical images for diagnostics [2], on natural images for computer vision [3], and on the remote sensing images [4] for land-cover classification besides several other applications.

The process of texture-based image segmentation consists of texture features extraction followed by clustering the feature space. If the segmentation result is granular with unacceptably high levels of salt and pepper effect [5], then the segmentation is refined by a post-processing like region growing. However, this paper avoids salt and pepper effect by pre-processing the feature images with Gaussian averaging. The experimental work presented herein concerns with the comparison of the two texture features,

* Associate Professor, and ** Lecturer,
Department of Telecommunication Engineering, Mehran University of Engineering & Technology, Jamshoro.

each based on Gabor wavelets and DFT, respectively, for image segmentation. The following paragraphs explain image texture, and the feature description of image texture. There are several kinds of texture descriptors, which are generally [6] categorised as based on:

(i)   Signal Processing Methods

(ii)   Statistical Methods

(iii)   Model-Based Methods

(iv)   Geometrical Methods

The approach in the first category considers the spatial frequencies involved in the textures and analyses them in the spatial frequency domain using Gabor wavelets, discrete wavelet transforms, curvelet transforms and ridgelet transforms, etc. These methods are very popular in texture analysis for computer vision, remote sensing [7] and other applications.

The second category includes the texture features that come from histogram-like representations such as GLCM (Gray Level Co-Occurrence Matrix), local binary patterns, HLAC (Higher Order Local Auto-Correlation), and the LFH (Local Fourier Histograms). These features are known for their ability of describing the textures with local neighbourhood properties [6].

The third category, called model-based methods, includes MRF (Markov Random Fields) and fractals. The methods based on MRF consider that each pixel depends statistically on the rest of the image, whereas the fractals are a geometrical measure of self-similarity of a texture at different scales.

The fourth category includes the texture descriptors based on geometric features. This type of texture description finds applications in texture synthesis and texture mapping for 3D rendering of the surfaces [8,9].

The two methods categorised in the third category of model-based approaches can be re-categorised. Those based on MRF can be considered as belonging to the category of statistical methods, because MRF is a statistical entity. Furthermore, methods based on fractals can be categorised as belonging to the geometrical methods, since these texture features are based on the principles of fractal geometry, despite the fact that it is non-Euclidean geometry. Hence there remain three categories. Out of the three, only first two, i.e. "Signal Processing" and "Statistical" methods have been successfully used for texture analysis purposes. Therefore, methods from these two categories are compared herein.

Methods from the first category are usually computationally expensive [10], since these come from large filter banks, but these are moderately insensitive to noise. On the other hand, those from the second category are computationally less expensive and have comparatively much higher descriptive power than the former, but are highly sensitive to the image noise [11]. Varma, M., et. al. [10] observes that the texture features based on statistical methods and extracted from neighbourhoods as small as 3x3 pixels can yield as good as or even better results than the features extracted from large filter banks, since texture information is contained in the local characteristics of the immediate pixel neighbourhood. Ursani, A.A., et. al. [12-13] found that the DFT-based features perform better than the features based on Gabor wavelets in the application of texture retrieval. This paper presents a comparison of features based on Gabor wavelets, and texture features based on DFT in the application of texture segmentation.

2.   THE GABOR WAVELETS

The studies on human vision system analyze the perceived image in the spatial frequency domain. For this reason, the Gabor wavelets are popular among the image processing and computer vision experts who believe that the Gabor wavelets are best at replicating human vision system [14]. Although there are several works reporting the use of variations of Gabor wavelets, we present empirical results

with the wavelets first presented in [15]. These wavelets are designed quite carefully to avoid wavelets causing redundant information. The mother Gabor wavelet is a function in two spatial dimensions defined as:

$$\psi(x,y,W) = \frac{1}{2\varpi_x \sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \exp\left(j2\pi W\right) \tag{1}$$

where σ_x and σ_y are functions of the scales (m), and W is the modulation frequency. The child Gabor wavelets are variations of the mother Gabor wavelets with various orientations and scales defined as:

$$\psi_{m,n}(x,y,W_m) = a^{-m} \psi(\tilde{x}, \tilde{y}, W_m) \tag{2}$$

where m is the scale, $\tilde{x}$ and $\tilde{y}$ are the rotated coordinates given as:

$$\tilde{x} = a^{-m}\left(\left(x - \frac{S}{2}\right)\cos\theta_n + \left(y - \frac{T}{2}\right)\sin\theta_n\right) \tag{3}$$

and

$$\tilde{y} = a^{-m}\left(\left(x - \frac{S}{2}\right)\sin\theta_n + \left(y - \frac{T}{2}\right)\cos\theta_n\right) \tag{4}$$

where S and T specify the template size in no. of columns and rows, respectively, and θ_n is defined as:

$$\theta_n = \frac{n\pi}{N} \tag{5}$$

where N is the total number of orientations, n varies from 0-N-1 and specifies the orientation of the wavelet. In this equation, a>1 and is given by:

$$a = \left(\frac{U_h}{U_l}\right)^{\frac{1}{M-1}} \tag{6}$$

where M is the number of scales, U_l and U_h are the lower and upper centre frequencies, whose values are proposed in [15-16] as 0.05 and 0.4 respectively. The modulation frequency Wm is given as:

$$W_m = a_m U_l \tag{7}$$

where m is scale of the wavelet and varies from 0 to M-1. These values of m produce the wavelets having the frequencies of 0.0500, 0.0841, 0.1414, 0.2378, and 0.4000 cycles per pixel. The σ_x and σ_y of the children wavelets are scale dependent and are given by:

$$\sigma_{x,m} = \frac{(a+1)\sqrt{2\ln 2}}{2\pi a^m (a-1)U_l} \tag{8}$$

$$\sigma_{y,m} = \frac{1}{2\pi \tan\left(\frac{\pi}{2N}\right)\sqrt{\frac{U_h^2}{2\ln 2} - \left(\frac{1}{2\varpi_{x,m}}\right)^2}} \tag{9}$$

These scales ensure that the wavelets cause minimal redundancy. The mother wavelet of Equation (1) comprises of a real and an imaginary part. For computational simplicity, we used only the real part. Therefore, the mother Gabor wavelet becomes:

$$\psi(x,y,W) = \frac{1}{2\varpi_x \sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \cos\left(2\pi W\right) \tag{10}$$

Extraction of Texture Features using Gabor Wavelets: Thirty 30 child Gabor wavelets (templates) were generated using five scales, i.e. M=5, and six orientations, i.e. N=6. Each of these templates is then convolved with the image as:

$$G_m(x,y) = \sum_{s=-S/2}^{S/2} \sum_{t=-T/2}^{T/2} I(x-s, y-t)\, \psi_{m,n}^*(s,t) \tag{11}$$

where I is the image to be segmented, and * represents operation of complex conjugation. The convolution yields feature images G_{mn}, measuring (X+S-1)×(Y+T-1) pixels, where X and Y are the dimensions of the image in terms of no. of columns and rows, respectively. Before using these feature images for clustering, S/2 columns from top as well as bottom and T/2 rows from each side are removed from the feature images, so as to have the same dimensions as

of the actual image to be segmented.

3. DFT-BASED TEXTURE SIGNATURES

Zhou, F., et. al. [17] proposed extracting texture features using DFT in one dimension. Coefficients of the 1D (1-Dimensional) DFT of the 8-pixel sequence around each image pixel provide the texture signatures that are useful in texture segmentation, texture recognition, and retrieval. Cutter, G.R., et. al. [18] notes that the DFT-based features are an accurate description of the local spatial distribution of the pixel grey-levels, and uses features extracted from local DFT from 3 3-pixel neighbourhoods for remote sensing application. Ursani, A.A., et. al. [11] compares the features based on DFT and Gabor wavelets for texture retrieval; and finds the former more efficient and computationally less expensive than the latter.

Extraction of DFT based features: Moving 3 3-pixel window across a texture image, 1D DFT of the digital numbers (pixel values) of sequence of 8 pixels, P_0 through P_7, in the spatial domain, shown in Fig. 1 is computed as:

$$X_k = \sum_{n=0}^{7} x_n \exp\left(-\frac{\pi i}{4} k\right)$$

(12)

where $0 \leq k \leq 7$, Xk represents the kth Fourier coefficient, x_0 through x_7 correspond to digital numbers of pixels P_0 through P_7. The symbol x, hereafter, represents the sequence of eight pixel values x_0 through x_7.

From the computed DFT, the absolute values of the first five coefficients, i.e. $|X_0|$ through $|X_4|$ are used for the texture description. It should be noted that the phases are not used because these are sensitive to image rotation. Also, the magnitudes $|X_7|$ through $|X_5|$ are replica of the magnitudes $|X_1|$ through $|X_3|$, respectively. Altogether, the computed coefficients are normalised to take values from 0-255.

In Equation (12), replacing k with 0, 1, 2, 3, and 4 yields the DFT coefficient X_0, X_1, X_2, X_3, and X_4, respectively. X_0 and X_4 are real, whereas others are complex. It is easy to show that the eight templates of Table 1 yield the real and imaginary values of these DFT coefficients if convolved with the 9-pixel neighbourhood of Fig. 1.

4. THEORETICAL COMPARISON

This section presents a theoretical comparison of the two texture descriptors under study, with respect to their descriptive power and the computational complexity.

4.1 Descriptive Power and Applicability
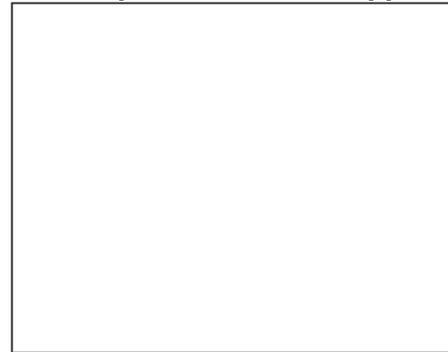
FIG. 1. THE 9-PIXEL NEIGHBOURHOOD IN THE SPATIAL DOMAIN

TABLE 1: THE TEMPLATES TO EXTRACT THE DFT-BASED TEXTURE SIGNATURES

$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 & 1 \\ -\sqrt{2} & 0 & \sqrt{2} \\ -1 & 0 & 1 \end{bmatrix} \times \frac{1}{\sqrt{2}}$	$\begin{bmatrix} -1 & -\sqrt{2} & -1 \\ 0 & 0 & 0 \\ 1 & \sqrt{2} & 1 \end{bmatrix} \times \frac{1}{\sqrt{2}}$
X_0	X_4	X_1 (real)	X_1 (imaginary)
$\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & -1 \\ -\sqrt{2} & 0 & \sqrt{2} \\ 1 & 0 & -1 \end{bmatrix} \times \frac{1}{\sqrt{2}}$	$\begin{bmatrix} -1 & \sqrt{2} & -1 \\ 0 & 0 & 0 \\ 1 & -\sqrt{2} & 1 \end{bmatrix} \times \frac{1}{\sqrt{2}}$
X_2 (real)	X_2 (imaginary)	X_3 (real)	X_3 (imaginary)

Almost all the applications of the texture analysis require rotation invariance of the texture descriptors. The texture invariance is inherently present in the DFT based texture signatures, whose values are dependent on the DFT of the digital numbers of a pixel sequence around a central pixel. Since magnitudes of the DFT coefficients are insensitive to the rotation or shift. Whereas rotation invariance is not inherently present in the texture descriptors based on Gabor wavelets, because each Gabor wavelet is sensitive to a single direction. For the applications of texture recognition or image retrieval, a feature vector is constructed comprising statistics of a complete query image with uniform texture. A means of rotation invariance was introduced in such a vector [16]. Since the rotation invariance is possible only after computing the feature vector representing a region with a homogeneous texture, it cannot be applied in case of image segmentation, where an image comprises several textures and the rotation invariance is required on the pixels level. Keeping in view the closeness of Gabor wavelets with the human vision system, computer vision community has extensively used Gabor wavelets for the extraction of texture features, but [18] observes that the human vision is deceptive for some applications such as remote sensing. In such applications, texture descriptors are required that are more capable of making objective evaluation rather than humanlike assessment that is subjective in nature.

4.2 Computational Cost Analysis

The computational complexity of performing convolution [19] is:

$$\text{Complexity} = O(S^2) \tag{13}$$

Where S is the size of the template convolved with the image. Since one generally performs this convolution with several templates, say with n templates, on an image for extracting multiple features, this complexity also becomes a function of the n. Therefore, the complexity becomes:

$$\text{Complexity} = O(S^2 n) \tag{14}$$

In case of the DFT-based texture signatures, the size of the template, i.e. S=3, and the no. of templates is n=8. Therefore, the computational complexity of the process of extracting DFT-based signatures is:

$$\text{Complexity} = O(3^2 \text{x}8) = O(72) \tag{15}$$

Similarly, in case of the features based on Gabor wavelets, size of template is 17x17-pixels, and the no. of templates used is 30. Hence, the computational complexity of extracting these features becomes:

$$\text{Complexity} = O(17^2 \text{x}30) = O(289) = O(72^{1.325}) = 1.325 \text{ x } O(72) \tag{16}$$

Equations (15-16) show that the computational complexity of the texture signatures based on Gabor wavelets is 1.325 times more than that of the DFT-based texture signatures.

5. CLUSTERING IN THE FEATURE SPACE

After extracting texture features as described in the preceding sections, k-means algorithm as implemented in CLUSTR 3.0 [20] was used to segment the images in the feature space. The distance measure used was the norm-2 distance, i.e. Euclidean distance.

Although k-means is a fast converging and efficient data clustering algorithm, but it works in the feature domain instead of spatial domain. Clustering the target images in the feature space causes salt and pepper effect in the spatial domain. Therefore, before clustering the feature images both, resulting from Gabor wavelets as well as from DFT, were averaged using a Gaussian filter of 9x9 pixels described in Equation (17).

$$g_{xy} = \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma^2} + \frac{y^2}{\sigma^2}\right)\right] \tag{17}$$

where x is the column coordinate, y is the row coordinate, and ? is the standard deviation selected to be one fifth of the filter dimension. On one hand, filtering with larger filters reduces the salt and pepper effect but on the

other hand not only reduces the discriminatory ability of the texture features as well as accuracy of the border localization. The smaller the filter size, more severe is the salt and pepper effect in the segmented image. The values selected for the filter dimension and ? were found to be the optimal values.

6. RESULTS

Table 2 shows results on 10 images from BSD (Berkley Segmentation Database) [21] and their segmentations using texture features based on DFT and Gabor wavelets respectively. The second column from the left shows the number of segments the image is divided into.

In Image 1, the shirt of the subject is segmented better by the Gabor wavelets than by the DFT. On the other hand, the background is better segmented by the DFT. In Image 2, DFT segmented image shows more precise localization of the cluster borders than the Gabor wavelets do. On the other hand, DFT segmented image shows misclassification on the bottom-left of the image. In Images 3, 6, 7, and 9 also, the border localization is better with DFT but the clusters are more granular than with Gabor wavelets. In Image 4, the Gabor wavelets results in more homogeneous clusters than DFT do, but Gabor-wavelet-based segmentation confuses the bottom-right of the image with the fish. In case of Image 5, though DFT results in much better segmentation than Gabor wavelets do, nevertheless it is far from perfect. Gabor wavelets segmentation confuses the bird's neck with the background. In case of Image 8, DFT features succeed in segmenting the clusters correctly, but the background cluster is granular. On the other hand, Gabor wavelets result in homogenous but confused clusters. It can be seen that part of the subject's shirt is confused with the background. In the last image, DFT features succeed in segmenting the zebras from the background only marginally, whereas, the Gabor wavelets fail miserably.

The granularity in the segmentation results with the DFT features indicate that the feature extraction method lacks a means of consolidating the DFT values over a spatial neighbourhood. The confusing clusters in the Gabor wavelet segmentations in the Images 4, 5, 8, and 10 show that descriptive power of the Gabor wavelets is poorer than the DFT. The larger size of the Gabor templates result in spatially more homogenous clusters but also cause coarser border localization.

7. CONCLUSIONS

The results show that the DFT-based features perform slightly better in some cases and little worse in others, than the features based on Gabor wavelets. More or less, the two perform the same. DFT-based features produce spatially less homogenous clusters than the features based on Gabor wavelets do. On the other hand, DFT-based features provide better border localization than the features based on Gabor wavelets do. However, it should be noted that the DFT based texture features are computationally much less expensive than those based on Gabor wavelets.

These results on texture segmentation are in contrast with the results published for texture retrieval, where DFT-based texture features demonstrated considerably better than those based on Gabor wavelets. This is because the DFT values were consolidated over total area of a test image in case of image retrieval, where as in case of texture segmentation, each pixel is individually represented by a set of DFT values that are averaged over only 9x9 pixel neighbourhoods. Averaging over larger neighbourhoods not only tends to cause less accurate border localization, but also affects the discrimination ability of the texture features. Therefore, an effective way of consolidating the DFT values over a spatial region needs to be found out that may retain their discrimination ability, while maintaining accuracy of border localization.

ACKNOWLEDGEMENTS

REFERENCES

TABLE 2. THE RESULTS OF TEXTURE SEGMENTATION USING TEXTURE FEATURES BASED ON DFT AND GABOR WAVELETS

No.	BSD Image No	Number of Segements	Image	Segmentation	
				DFT	Gabor Wavelets
1.	198054	3			
2.	86016	2			
3.	58060	3			
4.	164074	3			
5.	66075	2			
6.	163085	3			
7.	134052	2			
8.	198023	3			
9	15088	2			
10.	253027	2			
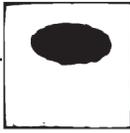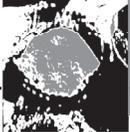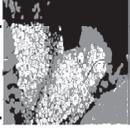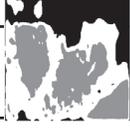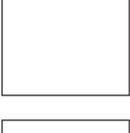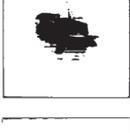
[1] Karu, K., Anil, K.J., and Ruud, M.B., "Is There Any Texture in the Image?", Pattern Recognition, Volume 29, No. 9, pp. 1437-1446, 1996.

[2] Olowoyeye, A., Tuceryan, M., and Fang S., "Medical Volume Segmentation Using Bank of Gabor Filters", SAC, Honolulu, Hawaii, USA, March 8-12, 2009.

[3] Gadelmawla, E.S., Eladawi, A.E., Abouelatta, O.B., and Elewa, I.M., "Application of Computer Vision for the Prediction of Cutting Conditions in Milling Operations", Proceedings of IMechE, Journal of Engineering Manufacture, Volume 223, Part-B, 2009.

[4] Kartikeyan, B., Sarkar, A., and Majumder K.L., "A Segmentation Approach to Classification of Remote Sensing Imagery", International Journal of Remote Sensing, Volume 19, No. 9, pp. 1695-1709, 1998.

[5] Gong, Y., Shu, N., Li, J., Lin, L., and Li, X., "A New Conception of Image Texture and Remote Sensing Image Segmentation Based on Markov Random Field", Geo-Spatial Information Science, Volume 13, No. 1, pp. 16-23, 2010.

[6] Chen, C.H., Pau, L.F., and Wang, P.S.P., (Editors), "Textural Analysis", The Handbook of Pattern Recognition and Computer Vision, Chapter-2, 2nd Edition, World Scientific Publishing Co., 1998.

[7] Schaale, M., Keller, I., and Fischer, J., "Land Cover Texture Information Extraction from Remote Sensing Image Data", Proceedings of ASPRS-RTI Annual Conference, Washington DC, 2000.

[8] Stamos, I., and Allen, P.K., "Geometry of Texture Recovery of Scenes of Large Scale", Computer Vision and Image Understanding, Volume 88, pp. 94-118, 2002.

[9] Elber, G., "Gemoetric Texture Modelling", IEEE Computer Graphics and Applications, Volume 25, No. 4, pp. 66-76, 2005.

[10] Varma, M., and Zisserman, A., "Texture Classification: Are Filter Banks Necessary?", Conference on Computer Vision and Pattern Recognition, Volume 2, pp. 691-698, 2003.

[11] Ursani, A.A., Kpalma, K., and Ronsin, J., "Texture Features Based on Fourier Transform and Gabor Filters: An Empirical Comparison", International Conference on Machine Vision, pp. 67-72, December 28-29, 2007.

[12] Ursani, A.A., Kpalma, K., and Ronsin, J., "Improved Texture Description with Features Based on Fourier Transform", Wireless Networks, Information Processing and Systems, Volume 20, pp. 19-28, Springer Verlag, 2008.

[13] Ursani, A.A., Kpalma, K., and Ronsin, J., "Texture Features Based on Local Fourier Histogram: Self-Compensation Against Rotation", Journal of Electronic Imaging, Volume 17, No. 3, pp. 1-3, 2008.

[14] Lee, T.S., "Image Representation Using 2D Gabor Wavelets", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 18, No. 10, 1996.

[15] Manjunath, B.S., and Ma, W.Y., "Texture Features for Browsing and Retrieval of Image Data", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 18, No. 8, pp. 837-842, 1996.

[16] Zhang, D., Wong, A., Indrawan, M., and Lu, G., "Content-Based Image Retrieval Using Gabor Texture Features", IEEE International Symposium on Multimedia Information Processing, 2000.

[17] Zhou, F., Feng, F.J., and Shi, Q.Y., "Texture Feature Based on Local Fourier Transform", IEEE Transactions on Image Processing, Volume 2, pp. 610-613, 2001.

[18] Cutter, G.R. Jr., Rzhanov, Y., and Mayer, L.A., "Automated Segmentation of Seafloor Bathymetry from Multibeam Echosounder Data Using Local Fourier Histogram Texture Features", Journal of Experimental Marine Biology and Ecology, pp. 355-370, 2003.

[19] Lu, J., and Yuasa, T., "A New Algorithm for 2D Convolution on Mesh-Connected SIMD Computers", Transactions on Information Processing Society of Japan, Volume 37, No. 12, pp. 2390-2397, Japan, 1996.

[20] Hoon, M.J.L., Imoto, S., Nolan, J., and Miyano, S., "Open Source Clustering Software", Bioinformatics, Volume 20, No. 9, pp. 1453-1454, 2004.

[21] Martin, D., Fowlkes, C., Tal, D., and Malik, J., "A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics", Proceedings of 8th International Conference on Computer Vision, Volume 2, pp. 416-423, 2001.