*Technical Note*

# Unsupervised Rough Segmentation of Natural Images Containing Man-Made Objects

Xiaoyan Dai,† Yukinori Suzuki† and Junji Maeda†

We present a method of unsupervised rough segmentation for natural images containing man-made objects that uses a proposed texture feature extraction and a hierarchical segmentation algorithm. We use Statistical Geometrical Features (*SGF*) as texture descriptors and propose to obtain uniform texture features and good boundary features to incorporate anisotropic diffusion. The segmentation is performed by using three processes: hierarchical splitting, agglomerative merging and pixelwise classification. Experimental results demonstrate the effectiveness of the proposed technique in obtaining rough segmentation.

## 1. Introduction

Unsupervised rough segmentation of natural images containing man-made objects is a challenging problem to obtain segmentation that represents each main object or meaningful part of an object by one region respectively without paying much attention to region interiors. The segmentation should possess these necessary properties: Segmented regions should be uniform, region interiors should not contain a large number of small holes and boundaries of each region should be spatially accurate[1]. Therefore, an effective segmentation method based on a set of texture features having good discriminating capability is essential.

Among the existing unsupervised texture segmentation methods[2)~4)], Ojala and Pietikäinen present a hierarchical segmentation algorithm[5] that uses local binary pattern and contrast ($LBP/C$) features as texture measures and performs segmentation in hierarchical splitting, agglomerative merging and pixelwise classification. Though the method performs well for the segmentation of texture mosaics, it is not suitable for the segmentation of natural images containing man-made objects. Such natural images usually contain small pictorial texture patterns and strong boundaries, which make it necessary to set a small minimum block for segmentation to avoid the errors occurring during each stage of processing. However, it is difficult to use a minimum block size smaller than 16 due to the unstable histogram distribution of $LBP/C$.

This paper presents unsupervised rough segmentation of natural images containing man-made objects based on the algorithm by Ojala and Pietikäinen. We propose to use the Statistical Geometrical Features ($SGF$)[6] extracted from an original image and those from a preprocessed image because the $SGF$ can easily discriminate various types of textures. Though it is desirable to use a small window for the $SGF$ to obtain good boundary features, it is difficult to obtain uniform texture features using a small window. We propose to incorporate an edge-preserving smoothing preprocess using anisotropic diffusion[7] to reduce the statistical variation and to obtain smooth texture features when calculated with a small window. We then apply a feature reduction technique[8] to reduce the dimension of the features. The adoption of the $SGF$ makes it possible to set a small minimum block of 4 in the segmentation algorithm, which solves the problem of the algorithm by Ojala and Pietikäinen.

## 2. Texture Feature Extraction

### 2.1 Statistical Geometrical Features

The extraction of the $SGF$ starts by thresholding a textured image into a number of binary planes. Then for each binary image, the number of connected 1-valued regions and that of 0-valued regions give two geometrical measures and two irregularity measures. Each of these four measures is further characterized by the maximum value, the average value, the sample mean and the sample standard deviation, allowing 16 feature measures for a textured image to be obtained. A sliding overlapping window is used for calculating the $SGF$ of each pixel of a textured image.

Same as all other texture descriptors, the $SGF$ suffer from the dilemma of the window

† Muroran Institute of Technology

size: a large window is desirable to obtain uniform texture features, but a small window is necessary to obtain accurate boundary features. Since a small window is preferable for the segmentation of natural images containing man-made objects, we decided to use a small window and to solve the dilemma by employing an image-smoothing preprocess.

### 2.2 Proposed Feature Extraction

We incorporate an anisotropic diffusion preprocess to perform image smoothing without blurring boundaries. Though we prefer to use a small window for the *SGF* calculation to obtain good boundary features, it will cause the high statistical variation of texture features. Therefore, we utilize this preprocess to reduce the statistical variation of texture features when calculated with a small window.

Perona and Malik[7] proposed an edge-preserving smoothing algorithm to diffuse an image $I(x, y, 0)$ with time by the following anisotropic diffusion equation:

$$\frac{\partial I(x,y,t)}{\partial t} = c(x,y,t)\triangle I(x,y,t) + \nabla c(x,y,t)\nabla I(x,y,t), \quad (1)$$

where $\triangle$ and $\nabla$ are the Laplacian and the gradient operators, respectively, and $c(x, y, t)$ denotes the diffusion conductance coefficient which is proposed to be the function of the magnitude of the intensity gradient. Since this preprocess maintains the boundaries well and decreases the statistical variation of the textured regions, we adopt it to obtain uniform texture features from the preprocessed image even with a small window.

Though the features composed of the *SGF* calculated from the original image and those from the preprocessed image retain a large amount of information for segmentation, the 32-dimensional features in total will unavoidably bring about some redundancy and inaccuracy. Therefore, we adopt the Karhunen-Loève Transformation $(KLT)$[8] for feature reduction as follows:

$$S_t t = \lambda t, \quad (2)$$

where $S_t$ is the total scatter matrix of $N$-dimensional *SGF* feature vectors of all pixels, $\lambda$ is the eigenvalue matrix and $t$ is the eigenvector matrix. The feature vector with a reduced dimension of each pixel $V$ is next obtained by

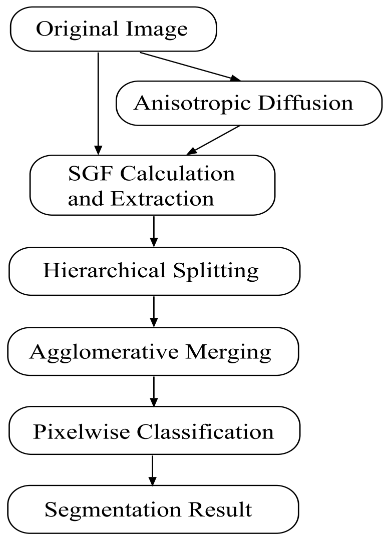


**Fig. 1** Block diagram of the segmentation procedure.

$$V = t^T U, \quad (3)$$

where $U$ is the $N$-dimensional feature vector and $t^T$ are the eigenvector matrix corresponding to the $M$ largest eigenvalues. The $M$-dimensional features are then used for segmentation. We can reduce the dimension of features from $N = 32$ to $M = 4$ since the cumulative contribution of 4 largest eigenvalues reaches 96%.

## 3. Segmentation Algorithm

The hierarchical segmentation procedure is shown in **Fig. 1**. First, the anisotropic diffusion preprocess is applied to obtain an edge-preserving smoothed image. Second, the *SGF* of the original image and those of the smoothed image are obtained using a $3 \times 3$ window and then feature reduction is performed to extract features for segmentation. Next, the segmentation is executed. In the following, we will demonstrate the progress of the segmentation algorithm on a $256 \times 256$ natural scene composed of the sky, mountains covered with snow, mountains with sunshine, mountains without sunshine (a dark region) and a desert as shown in **Fig. 2** (a).

Hierarchical splitting in Fig. 1 is performed to divide the image into regions of different size. The image is first divided into rectangular blocks of size $S_{max}$. Then the uniformity of its four subblocks based on the Euclidean distance is tested. If the ratio of the largest distance to the smallest one is higher than a predetermined

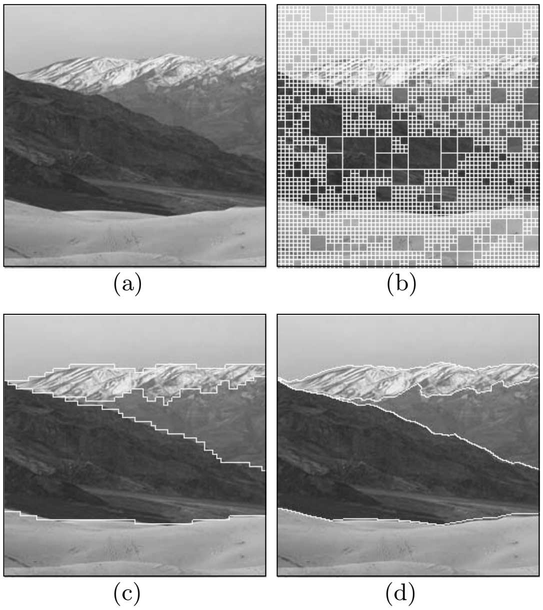

**Fig. 2** Segmentation of a natural image. (a) original image, (b) result of hierarchical splitting, (c) result of agglomerative merging, (d) result of pixelwise classification.

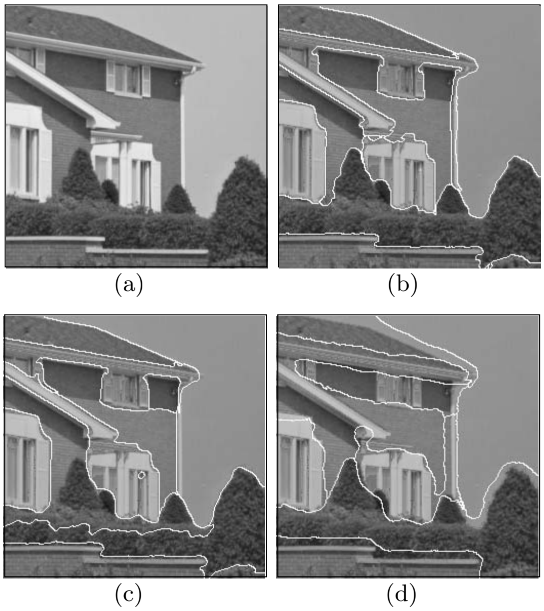

**Fig. 3** Segmentation results. (a) original image, (b) result of the proposed method, (c) result using the features extracted from the original image alone, (d) result of the conventional algorithm.

threshold $X$, the block is regarded as nonuniform and is divided into four subblocks. The process is performed on all the subblocks until the minimum block size $S_{min}$ is reached. The result is shown in Fig. 2 (b). Here $S_{max} = 64$, $S_{min} = 4$ and $X = 2$.

Agglomerative merging in Fig. 1 is then executed to merge most similar adjacent regions globally. The process will merge two adjacent segments with the largest merger importance among all possible mergers for each merge until the ratio of the merger importance for the current best merger to the smallest merger importance of all preceding mergers exceeds a threshold $Y$. $Y$ can be interpreted as the scale of distinct texture regions we want to discriminate. We set $Y = 2$ to obtain rough segmentation shown in Fig. 2 (c).

Pixelwise classification in Fig. 1 is finally performed to improve boundary localization. If an image pixel is on the boundary of different textures, a discrete disc is placed on it and the Euclidean distance between the disc and its neighboring regions is calculated individually. The pixel will be relabeled if its label is different from that of the most similar neighboring region. The process is continued until no pixels are relabeled. The final segmentation shown in Fig. 2 (d) represents that the approach well divides the natural image into five homogeneous regions.

## 4. Experimental Results

The parameters of the proposed method are $S_{max} = 64$, $S_{min} = 4$, $X = 2$ and those of the conventional algorithm by Ojala and Pietikäinen are $S_{max} = 64$, $S_{min} = 16$ and $X = 1.2$. We set $Y$ differently according to different images to obtain rough segmentation.

**Figure 3** (a) is a $256 \times 256$ natural image composed of the sky, trees and a man-made house. Figure 3 (b) is the segmentation result of the proposed method. The result shows rough segmentation in which the trees are well maintained as a complete region, the segmented regions are uniform and the boundaries of each region, such as windows, roofs and walls, are spatially accurate. Figure 3 (c) is the segmentation result using the features extracted from the original image alone. The figure shows that the trees are divided into several parts, some important boundaries of the trees are lost and some shadows are kept as a separate region. The comparison verifies the effectiveness of the proposed feature extraction. Figure 3 (d) is the result of the conventional segmentation algorithm. Though each region approximates

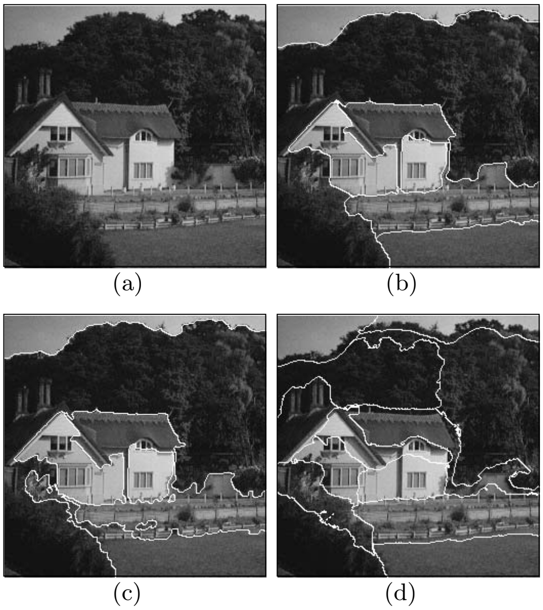

**Fig. 4** Segmentation results. (a) original image, (b) result of the proposed method, (c) result using the features extracted from the original image alone, (d) result of the conventional algorithm.

each object, the boundaries are spatially inaccurate. Though there is no criterion to evaluate rough segmentation, the experiment shows the effectiveness of the proposed method in obtaining rough segmentation corresponding well with the human visual system. We take $Y = 1.5$ in our algorithm and $Y = 1.1$ in the conventional algorithm.

Another $256 \times 256$ natural scene containing a man-made object is shown in **Fig. 4** (a) that consists of the sky, a forest, a house, a yard and grass. The proposed method represents a satisfactory rough segmentation result as shown in Fig. 4 (b) in which the sky, the forest, the yard and the grass are well maintained as separate regions. Figure 4 (c) is the result using the features extracted from the original image alone. There are several holes within the region of the yard and part of the trees in the lower left of the image are regarded as a region of the grass. The result of the conventional segmentation algorithm shown in Fig. 4 (d) fails to maintain accurate boundaries though the yard and the grass are maintained as separate regions. We set $Y = 1.7$ in our algorithm and $Y = 1.2$ in the conventional algorithm.

## 5. Conclusions

In this paper we have proposed unsupervised rough segmentation for natural images containing man-made objects. We proposed a texture feature extraction using the $SGF$ to obtain uniform texture features and good boundary features. The $SGF$ of the original image and those of the edge-preserving smoothed image obtained from an anisotropic diffusion are calculated and used in a hierarchical segmentation algorithm after a feature reduction technique. A small minimum block size of 4 has been used which is suitable for the segmentation of natural images. The experiments have demonstrated that the proposed method provides promising results for rough segmentation of natural images containing man-made objects, which corresponds well with the human visual system.

## References


1) Haralick, R.M. and Shapiro, L.G.: Image Segmentation Technique, *Comput. Vision Graphics Image Process.,* Vol.29, pp.100–132 (1985).
2) Fu, K.S. and Mu, J.K.: A Survey on Image Segmentation, *Pattern Recognition,* Vol.13, No.1, pp.3–16 (1981).
3) Pal, N.R. and Pal, S.K.: A Review on Image Segmentation Techniques, *Pattern Recognition,* Vol.26, No.9, pp.1277–1294 (1993).
4) Reed, T.R. and du Buf, J.M.H.: A Review of Recent Texture Segmentation and Feature Extraction Techniques, *CVGIP: Image Understanding,* Vol.57, pp.359–372 (1993).
5) Ojala, T. and Pietikäinen, M.: Unsupervised Texture Segmentation Using Feature Distributions, *Pattern Recognition,* Vol.32, No.3, pp.477–486 (1999).
6) Chen, Y.Q., Nixon, M.S. and Thomas, D.W.: Statistical Geometrical Features for Texture Classification, *Pattern Recognition,* Vol.28, No.4, pp.537–552 (1995).
7) Perona, P. and Malik, J.: Scale-space and Edge Detection Using Anisotropic Diffusion, *IEEE Trans. Pattern Anal. Mach. Intell.,* Vol.12, No.7, pp.629–639 (1990).
8) Jain, A.K.: *Fundamentals of Digital Image Processing,* Prentice Hall, Englewood Cliffs, NJ (1986).


(Received March 1, 2002)
(Accepted October 16, 2003)