# Regions Labeling in Outdoor Scene Images

Kyawt Kyawt Htay and Nyein Aye

**Abstract** Outdoor scene analysis is a complex problem for both image processing and pattern recognition domains. This paper proposes an approach for labeling regions in outdoor scene images. The basic idea of this approach is to label local image regions into semantic objects such as tree, sky and road etc. There are four phases in the approach: segmentation, feature extraction, region labeling and merging. Firstly, modified Marker-Controlled Watershed (MCWS) algorithm proposes for segmented regions generation. And then, color feature extracted from segmented regions are given as input to 3-layer Artificial Neural Network (ANN) classifier for labeling. Finally, region merging is performed if the two regions are adjacent with the same color values. The proposed method is test on our real scene image dataset which are collected from our university campus.

**Keywords** Region labeling · Segmentation · ANN classifier · Region-merging

## 1 Introduction

Scene understanding is a fundamental problem in robotics and perception environment. Most robotic scene understanding work has focused on outdoor scenes captured with laser scanners, where technologies have matured and enabled high-impact applications such as mapping and autonomous driving [1]. Autonomous robots capture visual information of scenes and train a classifier, predict the taversability of real scene images with the trained classifier. The promising classification results provide the basis for obstacles avoiding and path planning. Gi-Yeul Sung et al. [2] presented a terrain classification method based upon color and texture features,

K.K. Htay(✉)
University of Computer Studies, Mandalay, Myanmar
e-mail: kyawtkyawthtay2014@gmail.com

N. Aye
Head of the Department of Hardware, University of Computer Studies, Mandalay, Myanmar
e-mail: nyeinaye@gmail.com




T.T. Zin et al. (eds.), *Genetic and Evolutionary Computing*,
Advances in Intelligent Systems and Computing 387,
DOI: 10.1007/978-3-319-23204-1_26

they extracted texture features with discrete wavelet transform coefficients mean and energy values. Furthermore, they adopted spatial coordinates where a terrain is located in the image as additional features. Color histograms are a popular method of representing the image [3]. Moreover, image segmentation is one of the most complex stages in image analysis. It becomes essential for subsequent image description and recognition tasks. The problem consists of partitioning an image into its constituent regions, objects or labels [4]. The level of division depends on the specific problem being solved. This partition is accomplished in such a way that the pixels belonging to homogeneous regions, regarding to one or more features (i.e., brightness, texture or color), share the same label, and regions of pixels with significantly different features have different labels. There must be considered four objectives for developing an efficient generalized segmentation algorithm [5]: continuous closed contours, non-over segmentation, independence of threshold setting and short computation time. In particular, the over-segmentation error, which occurs when a single semantic object is divided into several regions, is a tendency of some segmentation methods like watershed [6]. Therefore, some subsequent region merging process is needed to improve the segmentation results.

## 2 Related Works

Many scene classification systems needs to be sure which objects are in the image before classifying the scene. The common method of doing this is to segment the objects in the image and then identify the low-level features and semantic concepts. There are two methods to segment images, one is block-based and one is region-based [7] [8]. The block-based method simply segments the image into several rectangles. The region-based method also segments the image, but the objects are more meaningful to human eye. The objective of a scene understanding system consists of recognizing and localizing the significant imaged objects in the scene and identifying the relevant object relationships. Consequently, a system must perform segmentation, region characterization and labeling processes. Gao-Hua, Iaeng and Zhao Chun-xia, Zhang Hao-feng [9], proposed a scene classification method. First, texture features are extracted in gray channel. Then, color moments are computed in all three color channels of RGB color space. The parameters of GMM are estimated by training these labeled samples. Finally, new scene images are classified by the trained GMM model. Vogel and Schiele [10] modeled the semantic content of an image and used this model to classify local image regions into semantic concept classes. Model-based systems rely upon the configuration of the scene components. Luo and Stephen [11] proposed a model-based approach to detecting sky. This approach consists of color classification, region extraction and physics motivated sky signature validation.

In this paper, our contribution is modified Marker-Controlled Watershed (MCWS) algorithm is proposed and region merging is performed on classifier output image. Therefore, our approach will reduce computational complexity and time than others pixel-based and clustering-based segmentation approaches. Our work will focus on the problem of image regions labeling to classify every region of a given image into one of several pre-defined classes.

## 3 Watershed Transform

The advantages of watershed transformation are that it is simple, instinctive knowledge, and can be parallelized [12]. The main drawback of this method is the over-segmentation due to the presence of many local minima. To decrease the effect of severe over-segmentation, marker-controlled watershed transformation is used. Separating touching objects in an image is one of the more difficult image processing operations. The watershed transform is often applied to this problem. The watershed transform finds "catchment basins" and "watershed ridge lines" in an image by treating it as a surface where light pixels are high and dark pixels are low.


(a) (b)

**Fig. 1** (a) Original Image (b) Result of Standard Watershed Algorithm

The above described figure 1. (a) and (b) show original outdoor scene image and segmented result of standard watershed algorithm. The standard watershed algorithm produces 23833 segmented regions. Therefore, we need to reduce the number of segmented regions by using marker controlled approach.

### 3.1 *Modified Marker-Controlled Watershed (MCWS) Algorithm*

This paper proposes modified Marker-Controlled Watershed (MCWS) algorithm for segmented regions generation. There are three steps in the propose algorithm: pre-processing, watershed transform and post-processing. In the proposed modified MCWS algorithm, a variety of procedures could be applied to find the foreground markers which must be connected blobs of pixels inside each of the foreground objects. The method that is used to find foreground markers is morphological techniques called "opening-by-reconstruction" and "closing-by-reconstruction" to "clean" up image. These operations will create flat maxima inside each object that can be located using imregionalmax. Background markers are computed in the cleaned-up image by using a thresholding operation. And then compute watershed transform of the segmentation function.

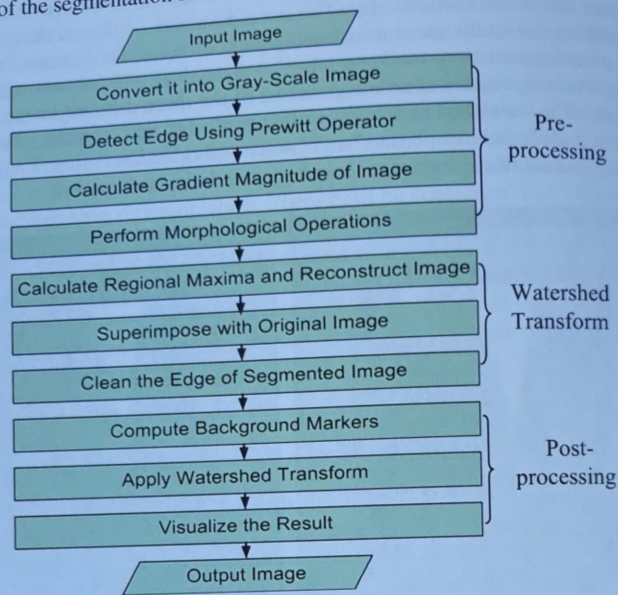
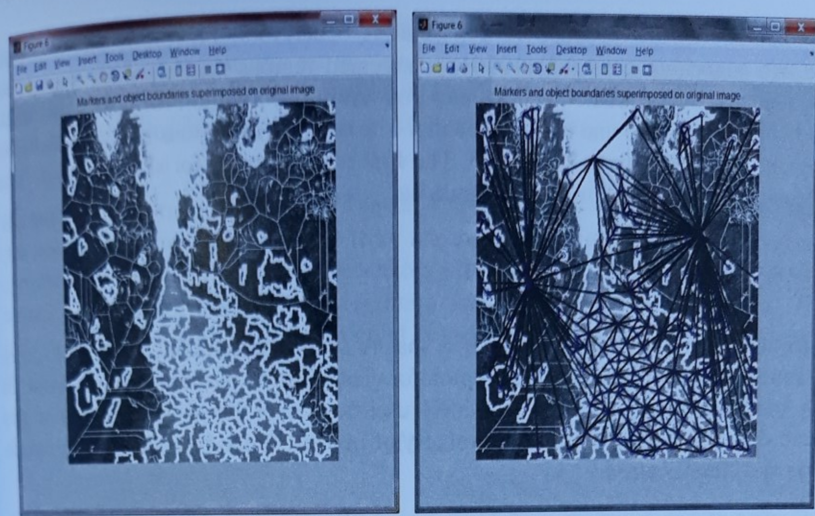



**Fig. 2** Flow chart of proposed modified MCWS algorithm


**Fig. 3** (a) Foreground Marker by using Morphological Operations (b) Background Marker by using Thresholding Method




(a) (b)

**Fig. 4** (a) Result of Modified MCWS Algorithm, (b) Region Adjacency Graph

The figure 4 (a) shows the result of our proposed modified MCWS algorithm. Our proposed modified MCWS algorithm can reduce over segmentation problem of standard watershed algorithm by using marker controlled. The number of

**Fig. 5** Regions Dataset

segmented regions produce is 215 regions. Therefore, the proposed modified MCWS algorithm can reduce 23618 segmented regions than standard watershed algorithm. The segmented regions can give useful information for later steps such as feature extraction and region-labeling. The segmented regions are collected as a region dataset is shown in figure 5. The first row represent for sky and road. The second row represent for tree and grass etc.

## 4 Feature Extraction

An important step in classification is that of feature extraction. Here, 768 color features are extracted from each region for classification purposes. Color is one of the most effective, simplest and widely used low level visual features. Color has been thought of as an important property of objects. It can be used for segmentation and classification.

### 4.1 *Color Histogram*

In the analysis of color images, the description of image regions has mainly been performed using color histograms. The color histogram is obtained by counting the number of times each color occurs in the image array. Color histogram H for a given image is defined as a vector

$$\mathbf{H} = \{H[0], H[1], \ldots, H[n]\} \tag{1}$$

Where i represent the color in color histogram and H[i] represent the number of pixels of color i in the image, and N is the number of bins used in color histogram. To extract color features, we need to represent the color in a suitable space. This system uses RGB color space to extract color features. The number of bins used in color histogram is 256+256+256=768 bins for RGB channels. Therefore, the 768 feature sets are given as input to 3-layer ANN classifier for regions labeling.

## 5 Regions Labeling

### 5.1 *Classifier*

In this system, 3-layer Artificial Neural Network (ANN) classifier is used to classify local image regions. Using the feature set described above, a neural network with 768 inputs and 5 label outputs has been trained and optimized. Conjugate gradient descent optimization was used and the optimal network architecture was found to have 24 nodes in the hidden layer.

**Training.** In the training step, weighting parameters are computed through off-line training by supervised learning using error back-propagation. Once an input pattern is applied to the network, the weights are adjusted iteratively by the back-propagation method, which minimizes the mean square error between the actual and desired output. We set the learning rate to 0.3 and, the training process is iterated until the mean square error is smaller than 0.1. In order to train the network, 768 color features are used in RGB color spaces.

**Running.** In the running step, the 3-layer ANN classifier is performed by loading the network parameters computed from the training results. The classified result can then be computed using feature vectors extracted from an input image. Neural network inputs are made up of weighting parameters and feature vectors. The output of the network is the classified region type to which the input region belongs.

### 5.2 *Region Adjacency Graph (RAG)*

Region adjacency graph (RAG) which is a usual data structure used to represent region neighborhood relations in a segmented image [13]. RAG is a weighted undirected graph G={V,E,W}, where each node represents a region of the over segmented image and each edge is a symmetric dissimilarity function between adjacent regions.

**Region Merging.** Region merging is a post-processing technique that merges adjacent regions with the same color values. In the proposed region merging approach, there are two predicates.

- The first predicate is minimum pair for each region.
- The second one is color homogeneity distribution.

From RAG, minimum edge pairs for each region are discovered. Color homogeneity distribution can be obtained from 3-layer ANN classifier output. Only if the region pair can satisfy these two predicates, the region pair is actually merged. If not, this region pair is ignored.

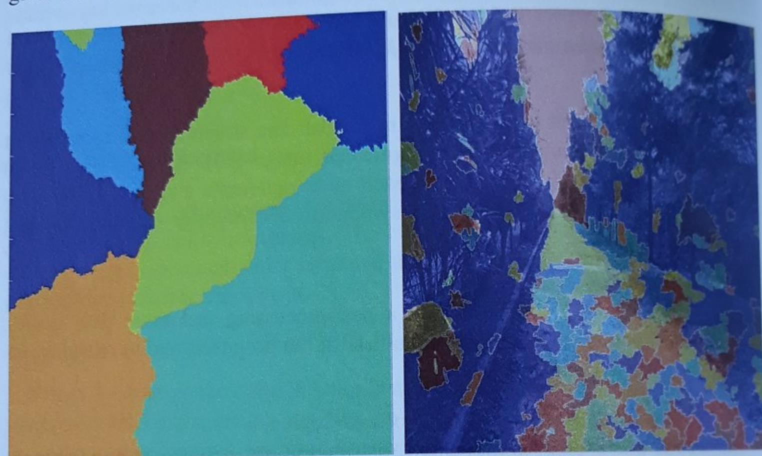
## 6 Results and Discussion

The proposed method is applied on our real scene images dataset was collected from our university campus. All of the images were taken during daytime to give a realistic view of the real environment. Environmental conditions during capture were dry and good atmospheric visibility. All of the images collected are stored in JPEG (.jpg) format with 24-bit color depth and the size of image is resize into 512x512 pixels. The 768 color features extracted for every region are presented to 3-layer ANN classifier and produces five output labels, which include sky, road, building and tree etc. And then, to reduce the number of labeled regions, region merging is performed if the two regions are adjacent with the same color values.

As a limitation, our system cannot accept under segmented images. However, the system can accept over segmented image. So the segmented image must not be under segmented image for later feature extraction and region-labeling steps.

Some segmentation methods like N-cut at which is a well-known segmentation method but like all methods it only works on specific datasets. The problem of the N-cut algorithm is to determine the best N value which is the number of segments in an image. Determining a fixed N-value is not very reliable so N-cut gives satisfactory result only for specific images with specific N-values.

When our proposed modified MCWS is compared with N-cut segmentation, the N-cut segmentation gives under segmented image. The proposed method can gives satisfy result than N-cut and traditional watershed algorithm.



(a) (b)

**Fig. 6** (a) Result of N-cut Segmentation and (b) Result of our proposed modified MCWS

The above figure 6. (a) Shows the result of N-cut segmentation on the input scene image with N-value is 8. The result image faces the problem of under segmentation for the input N-value. N-cut segmentation is not used in the system because it produces under segmented image. In (b) shows the result using our proposed modified MCWS algorithm. The proposed modified MCWS algorithm gives suitable over segmented image for later feature extraction and region-labeling steps. Therefore our system uses the proposed modified MCWS algorithm for segmentation.

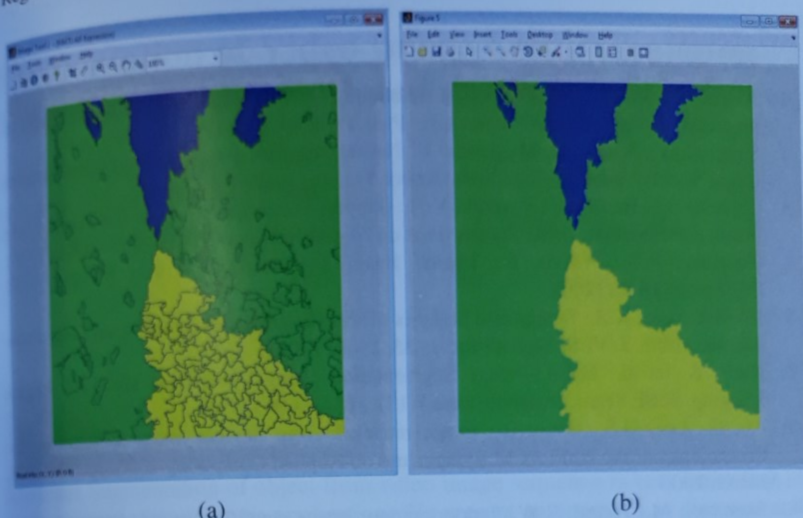



**Fig. 7** (a) 3-layer ANN Classifier Output Image (b) Result of Region Merging Image

Sky Road Tree

The above described image (a) and (b) show the result of 3-layer ANN classifier output and the result of after region-merging. The blue color represents for sky, yellow color represent for road and green color represent for tree etc. After region merging, we can reduce the number of labeled regions 215 into 7 regions. Therefore, our proposed regions labeling method can produce meaningful regions efficiently.

## 7 Conclusion

In this paper, modified Marker-Controlled Watershed (MCWS) algorithm is proposed for segmented regions generation. The proposed modified MCWS algorithm can produce suitable over segmented image than N-cut and standard watershed algorithm. The segmented image is used for later feature extraction and region-labeling steps. Color histogram is used in RGB space to extract color features. The extracted 768 color features are given as input to 3-layer ANN classifier to label segmented regions. In the classifier output image, many labeled regions are adjacent with the same color values. Therefore, region merging is performed to produce meaningful labeled regions. For future work, the proposed method can be test with various segmentation algorithms and low-level features that are used for region labeling.

## References


1. Xiong, X., Munoz, D., Bagnell, J., Hebert, M.: 3-D scene analysis via sequenced predictions over points and regions. In: Proc. of ICRA (2011)
2. Sung, G.-Y., Kwak, D.-M., Lyou, J.: Neural Network Based Terrain Classification Using Wavelet Features. Intelligent Robot System **59**, 269–281 (2010)
3. Chapelle, O., Haffner, P., Vapnik, V.N.: Support vector machines for histogram-based image classification. IEEE Transactions on Neural Networks **10**(5), 1055–1064 (1999)
4. Gonzalez, R.C., Woods, R.: Digital Image Processing, 2nd edn. Prentice Hall, Englewood Cliffs (2002)
5. Ho, S.Y., Lee, K.Z.: Design and analysis of an efficient evolutionary image segmentation algorithm. J. VLSI Signal Process **35**, 29–42 (2003)
6. Haris, K., et al.: Hybrid Image Segmentation Using Watersheds and Fast Region Merging. IEEE Trans. Image Process **7**(12), 1684–1699 (1998)
7. Ko, C., Lee, H.S., Byun, H.: Image retrieval using flexible image sub-blocks. In: Proceedings of the 2000 ACM Symposium on Applied Computing 2000, pp. 574–578, March 2000
8. Szummer, M., Picard, R.W.: Indoor-outdoor image classification. In: Proceedings of IEEE International Workshop on Content-Based Access of Image and Video Databases, Bombay, India, pp. 42–51 (1998)
9. Gao, H., IAENG, Zhao C.-X., Zhang H.-F.: Visual features fusion for scene images classification. In: Proceedings of the International Multi-Conference of Engineers and Computer Scientists, IMECS 2012, Hong Kong, vol. II, March 14–16, 2012
10. Vogel, J., Schiele, B.: Semantic modeling of natural scenes for content-based image retrieval. International Journal of Computer Vision (2004)
11. Luo, J., Etz, S.P.: A physical model-based approach to detecting sky in photographic images. IEEE Transactions of Image Processing **11**(3), 201–212 (2002)
12. Beucher, S., Meyer, F.: The morphological approach to segmentation: the watershed transform. In: Dougherty, E.R. (ed.) Mathematical Morphology in Image Processing, vol. 12, pp. 433–481. Marcel Dekker, New York (1993)
13. Sonka, M., et al.: Image processing, analysis and machine vision, 2nd edn. PWS (1999)


Segmentation
Detection and I

Mie Mie Tin and Mie M

**Abstract** Segmentation
many aspects of multin
segment foreground ob
any frame and compute
pare each frame of bloc
automatically segment
terest. We apply our m
better results than the s

**Keywords** Video ima
Block processing

1 Introductio

Image processing has
robot vision, airport s
ical image analysis fo
technique to separate
image. Segmenting o
processing and imag
and challenging task
objects in terms of co

M.M. Tin
University of Compute
e-mail: Miemietin198

M.M. Khin(✉)
Myanmar Institute of
e-mail: Miemie.khin9@


T.T. Zin et al. (eds.), Ge
Advances in Intelligent S
DOI: 10.1007/978-3-3