SEMANTIC CONCEPTS CLASSIFICATION ON OUTDOOR SCENE IMAGES

Kyawt Kyawt Htay
Associate Professor Dr. Nyein Aye
Computer University, Mandalay, Republic of the Union of Myanmar.

ABSTRACT

Outdoor scene analysis is a complex problem for both image processing and pattern recognition domains. A number of different approaches have been used for recognizing different objects in such scenes. There are two methods of segmenting images to look for objects in an image. One is block-based and one is region-based. Region-based method can provide some useful information about objects even though segmentation may not be perfect. In this paper, propose a method for semantic concepts classification on outdoor scene images. The basic idea of this approach is to classify local image regions into semantic concept classes such as tree, sky and road etc. There are three phases in this approach: segmentation phase, features extraction phase and classification phase. In segmentation phase, Modified Marker-Control Watershed algorithm (MCWS) is used. Second, color feature vectors are extracted from segmented regions by color moments in RGB space. Finally, classification is performed by 3-layers Artificial Neural Network (ANN). The propose method is applied on real scene images dataset.

KEYWORDS

Color moments, Marker Control Watershed, Outdoor Scene

INTRODUCTION

The semantic information of an image carries the meaning of that image. It is trivial for the human eye to extract semantic information from photos. However, for a computer, it is difficult to identify the semantic features of high-level images in photos. Therefore, if a computer can be made to correctly identify the semantic features of objects in photos, it will enhance the image identification rate. Towards this goal, segmentation of an image into regions has been used in recent years. Some researchers believe that a segmentation of images into regions can provide more semantic information than the usual global image features. Scene classification has become a popular research topic in recent years. However, even though much research has been done before, classifying photos into semantic types of scene (e.g., portrait, landscape) is still a difficult problem. Image classification can be achieved either by only using low-level features, or by integrating the low-level and high-level features [1]. Color histograms are a popular method of representing the image [2]. In the system of Town and Sinclair [3], neural networks are trained to classify previously segmented image regions into one of eleven semantic classes such as brick, cloud, fur or sand. The image regions are represented by color and texture features and images are retrieved using visual features. Vogel and Schiele [4] modeled the semantic content of an image and used this model to classify local image regions into semantic concept classes. Mode-based systems rely upon the configu-ration of the scene components. Luo and Stephen [5] proposed a model-based approach to detecting sky. This approach consists of color classification, region extraction and physics motivated sky signature validation.

RELATED WORK

Image segmentation is a preliminary and critical step in segment-based image analysis. The best segmentation result is used in image classification. Scene classify-cation system 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 [6], [7]. 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 the 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 relation-ships. Consequently, a system must perform segmentation, region characterization and labelling processes. Gao-Hua, Iaeng and Zhao Chun-xia, Zhang Hao-feng [8], 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. In the system of Town and Sinclair [9] neural networks are trained to classify previously segmented image regions into one of eleven semantic classes such as brick, cloud, fur or sand. Serrano et al. [10] proposed a set of low-dimensional, computationally efficient low-level features that are extracted from LST color space and wavelet texture features. In reference [11], a K-means segmentation technique was used to provide closed-region segmentations automatically. The same classification technique, when used on these regions, was shown to give a classification accuracy of 81.4% per image by area for the identical 11 class problem. For such a complex task, this classification rate is very impressive given the fully automatic nature of the segmentation and labeling steps.

In this paper region-based image segmentation approach is used for semantic concepts classification on outdoor scene images. As a first step Modified Marker-Control Watershed algorithm (MCWS) is used to solve the problem of segmented regions generation. Therefore, we will focus our work in the problem of image regions labeling to classify every region of a given image into one of several predefined classes.

PROPOSED SYSTEM

There are three mainly phases in this proposed system: segmentation, feature extraction and classification.

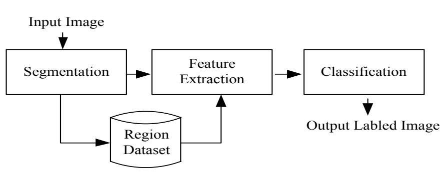

Figure 1. Overview of Proposed System.

Firstly, there are three steps in image segmentation phase: pre-processing, watershed transform and post-processing.

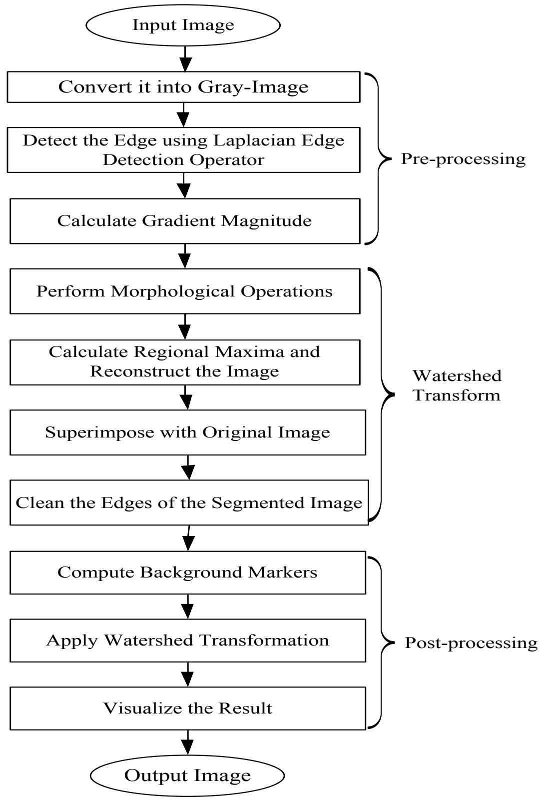

Figure 2. Flow Chart for Segmentation Phase: Modified Marker-Control Watershed Algorithm (MCWS).

SEGMENTATIONAND CLASSI-FICATION

In the pre-processing step, scene image is given as input and filter the image by using Laplacian operator instead of Sobel. Marker controlled watershed segmentation:

Now mark the background objects with the help of marker. A variety of procedures could be applied here to find the foreground markers which must be connected blobs of pixels inside each of the foreground objects. Here the method that is used is morphological techniques called opening by reconstruction and closing by reconstruction to clean up the image. These operations will create flat maxima inside each object that can be located using imregionalmax. Now compute background markers. Now you need to mark the background. In the cleaned-up image, the dark pixels belong to the background, so you could use a thresholding operation to mark the background objects. Now compute the watershed transform of the segmentation function. The function imimposem in can be used to modify an image so that it has regional minima only in certain desired locations. Here you can use imimposem in to modify the gradient magnitude image so that its only regional minima occur at foreground and background marker pixels. Finally compute the watershed transform of the modified segmentation function.

With the segmentation, objects of interest from image are extracted. Various techniques discovered till now for segmentation, here watershed algorithm is used. Watershed is also based on morphology. It is a region based algorithm having low computational complexity and high efficiency. It provides complete division of the image. Besides, all these advantages, it has a major drawback; it suffers from over-segmentation. Due to this, image content is distorted completely. So, some modifications are required to remove the problem of over-segmentation. In this paper, propose a pre-processing step in Marker-Control Watershed (MCWS) algorithm which actually reduces the number of segments produced by watershed algorithm.

A. Marker Control Watershed Transformation

The advantages of watershed transformation are that it is simple, instinctive knowledge, and can be parallelized. 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 proposed. 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. Segmentation using the watershed transforms works well if you can identify, or "mark," foreground objects and background locations. Marker-controlled watershed segmentation follows this basic procedure:

1. Compute a segmentation function. This is an image whose dark regions are the objects you are trying to segment.
2. Compute foreground markers. These are connected blobs of pixels within each of the objects.
3. Compute background markers. These are pixels that are not part of any object.
4. Modify the segmentation function so that it only has minima at the foreground and background marker locations.
5. Compute the watershed transform of the modified segmentation function.

Figure 3. Segmented Regions Dataset for Road, Sky, Tree and Grass.

In this system, classification phase: 3-layers ANN is used to classify previously segmented image regions into one of four predefined five semantic classes sky, grass, road, tree and unknown.

FEATURES EXTRACTION

The type of features to be extracted from an image depends on the nature of the scene classification task. Semantic concepts classification on outdoor scene images based on color and texture features have been addressed by many researchers. In the present work, we deal with the scene images primarily containing natural regions. Although not sufficient, low-level features such as color and texture contain good representation power for the region classification of natural scenes.

A. Color

Color is an important component of the natural scene classes. However, the color-based features suffer from the problem of color constancy. For natural scenes, we argue that given enough variations in the training data set, we can capture the class color distribution in varying illumination conditions. In addition, in outdoor natural scenes, the problem of artificial illuminants is not as serious as that of changes in brightness. To extract the color features, we need to represent the color in a suitable space. Any color distribution can be characterized by its moments. The first (mean), the second (Variance) and the third (skewness) color moments have been proved to be efficient in representing color distribution of images [11].

The first three moments are defined as:

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} f_{ij} \quad (1)$$

$$\sigma_i = \sqrt{\tfrac{1}{N}\sum_{j=1}^{N}\left(f_{ij} - \mu_{ij}\right)^2} \quad (2)$$

$$s_i = \sqrt[s]{\tfrac{1}{N}\sum_{j=1}^{N}\left(f_{ij} - \mu_{ij}\right)^3} \quad (3)$$

In this system, the first two moments is used for color feature vectors extraction in RGB space. It is sufficient for our analysis. Where f_{ij} is the value of the i^{th} color component of the image pixel j, and N is the number of pixels in the image. Since only 6 (two moments for each of the three color components) numbers are used to represent the color content of each image, color moments are a very compact representation compared to other color features.

RESULTS AND DISCUSSION

The proposed system is applied on our real scene images dataset by various edge detection filters (Disk, Average, Motion, Gaussian, Laplacian, LOG, Prewitt, and Sobel). Laplacian operator is used for our proposed Modified Marker-Control Watershed (MCWS) Algorithm. For each filter: three different parameters are calculated i.e. ENTROPY, MSE, and PSNR. On the basis of these values, final result and conclusion has been drawn.

A. Entropy

$$Entropy = \sum_i p_j \log_2 p_j \quad (4)$$

B. Mean Square Error (MSE)

It considers the quantity of the removed noise. The mean square error (MSE) is defined as:

$$MSE = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\left(f(m,n) - f'(m,n)\right)^2 \quad (5)$$

Value of MSE should be low for an efficient filter [12].

C. Peak Signal to Noise Ratio (PSNR)

PSNR is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. PSNR is usually expressed in terms of the logarithmic decibel scale. A higher PSNR generally indicates that the reconstruction is of higher quality. PSNR is most easily defined by the Mean Squared Error (MSE). PSNR can be defined as:

$$PSNR = 10 * \log_{10} \left[\frac{255^2 * M * N}{\sum\sum (x(i,j) - y(i,j))^2} \right] \quad (6)$$

Value of PSNR should be high for an efficient filter [12].

Table 1. Comparision of Various Filter Methods.

Filter	ENTROPY	PSNR	MSE
Disk	1.8327	20.4992	579.6473
Average	1.8493	20.8531	534.2779
Motion	1.8318	20.4668	583.984
Gaussian	1.7833	21.3256	479.2084
LOG	1.9797	19.5412	722.709
Laplacian	2.1467	19.7896	701.2922
Prewitt	1.9873	19.725	692.7596
Sobel	1.9854	19.6782	701.6556

On the basis of table, it is observed that LOG, Laplacian, Prewitt and Sobel filters are giving better performance results for denoising and segmentation according to Entropy values. Disk, Average and Motion filters have lower performance than other filters. Laplacian filter has highest in Entropy and Gaussian filter has lowest in Entropy, MSE and highest in PSNR values. According to table, Laplacian filter has higher in Entropy value than Sobel filter in existing marker-control watershed algorithm. Therefore, in this system Laplacian filter is used for denoising and segmentation. It is suitable for our analysis.

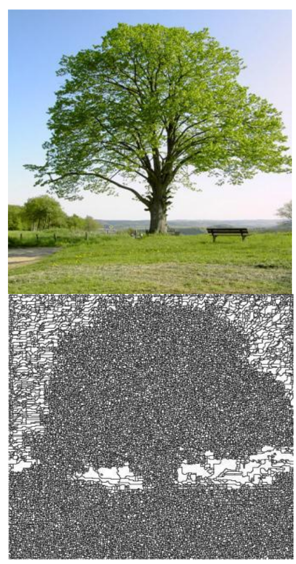

Figure 4. (a) Original Image (b) Gradient Image by Standard Watershed Algorithm.

Figure 5. (a) Marker-Control Watershed Algorithm, (b) Modified Marker-Control Watershed Algorithm.

CONCLUSION

Semantically meaningful labels to local image regions that can be used for subsequent retrieval step. In this paper, propose a method for semantic concepts classification based on region-based approach. First, input image is segmented by Modified Marker-Control Watershed algorithm (MCWS) in order to get more accurate compact regions. And then, color feature is extracted. Finally, classification is performed by 3-layers ANN to label the regions produced by the segmentation process. For future work, the proposed method can be tested on various segmentation algorithms and low-level features. The classifications results are varying depend on the use of segmentation algorithm and low-level feature.

ACKNOWLEDGMENT

We would like to thank my supervisor, Dr. Nyein Aye, at University of Computer Studies, Mandalay for providing excellent supporting and the feedback from the discussions of our research topic.

BIBLIOGRAPHY

[1]J. Luo and A. Savakis, “Indoor vs outdoor classification of consumer photographs using low-level and semantic features,” Proceedings of the 2001 International Conference On Image Processing (ICIP 01), Thessaloniki, Greece, vol. 2, pp. 745-748, Oct. 2001.

[2] O. Chapelle, P. Haffner, and V. N. Vapnik, “Support vector machines for histogram-based image classification,” IEEE transactions on neural networks, vol. 10, issue 5, pp. 1055-1064, Sep. 1999.

[3] C. Town and D. Sinclair. Content based image retrieval using semantic visual categories. Tech. Rep. 2000.14, AT&T Laboratories Cambridge, 2000.

[4] J. Vogel and B. Schiele, “Semantic modeling of natural scenes for content-based

image retrieval," International Journal of Computer Vision, 2004.

[5] J. Luo and S. P. Etz, "A physical model-based approach to detecting sky in photographic images,"IEEE Transactions of Image Processing, vol. 11, issue 3, pp. 201-212, Mar. 2002.

[6] C. Ko, H. S. Lee, and H. Byun, "Image retrieval using flexible image subblocks," Proceedings of the 2000 ACM symposium on Applied computing 2000, pp.574-578, March 2000.

[7] M. Szummer and R.W. Picard, "Indoor-outdoor image classification," Proceedings of IEEE International Workshop on Content-based Access of Image and Video Databases, Bombay, India, pp. 42-51, 1998.

[8] GAO Hua Member, IAENG and ZHAO Chun-xia, ZHANG Hao-feng "Visual Features Fusion for Scene Images Classification " Proceedings of the International Multi-Conference of Engineers and Computer Scientists 2012 Vol II IMECS 2012 , March 14-16 2012 ,Hong Kong.

[9] B. L. Saux and G. Amato, "Image classifiers for scene analysis," International Conference on Computer Vision and Graphics 2004.

[10] N. Serrano, A. Savakis, and J. Luo, "Improved scene classification using efficient low-level features and semantic cues," Pattern Recognition 37(9), pp.1773-1784, Sep. 2004.

[11] N. W. Campbell, W. J. Mackeown, B. T. Thomas and T. Troscianko, Automatic interpretation of outdoor scenes, British Machine Vision Con.(., Birmingham, U.K. (September 1995).

[12] Ng, H. P. (2008). Medical - Image Segmentation Using Water-shed Segmentation with Texture-Based Region Merging. 30th Annual International IEEE EMBS Conference (pp. 4039-4042). Vancouver, Canada: IEEE.