CSVTU International Journal of Biotechnology, Bioinformatics and Biomedical
Vol. 4(2), 67–71, 2019 ISSN: 2455-5762

Semantic Labeling of natural Scene Images Using Color Features

[1]Kyawt Kyawt Htay, [2]G. R. Sinha, [3]Hanni Htun, [4]Nwe Ni Kyaw

Myanmar Institute of Information Technology, Mandalay Myanmar- 05053
Email: kyawt_kyawt_htay@miit.edu.mm(1), gr_sinha@miit.edu.mm(2), hanni_htun@miit.edu.mm(3), nwe_ni_kyaw@miit.edu.mm(4)

Received August 16, 2019; received in revised form August 27 2019; accepted August 28, 2019; Available online August 2019

Research Article

Abstract

Scene image classification systems firstly need to locate the objects, and then classify the whole image. The color feature is important to describe the properties of an image surface. The paper presents a framework for scene images to label local regions using color features. The paper uses maker-controlled watershed algorithm to segment the input image into regions. This paper uses the segmented regions as a basic input unit, and then extracts Color Histogram (CH) and Color Moment (CM) features in HSV space. This system performs labeling using 3-layer Feed Forward Neural Network (FFNN) classifier. The system tests accuracy on public Microsoft Research Cambridge (MSRC) 9-class dataset.

Keywords: Scene classification, Color features, Color moment, Color histogram, Semantic concepts

1. Introduction

Semantic labeling is a difficult problem for both objects recognition and image processing applications. There most well-known object types are well-defined and natural scene objects in image recognition and labeling applications. Labeling and recognizing objects in natural scene images is more complicated than objects of well-defined. The well-defined objects are found in research of automated target recognition and computer vision applications. The natural objects are mostly found in scene analysis and remote sensing applications. Early labeling systems worked on individual pixels. The early works have drawbacks such as complexity increases depend on having an individual pixel. There are many difficult problems in recognizing objects of natural scene images. The multiclass image labeling methods try to classify all pixels and concurrent multi-class object recognition in an image. The early methods work on pixel level and utilize super pixels as a starting point. The author Fulkerson, et al. (2009) detected each super pixel and regularizes the classifier by aggregating histograms in the neighborhood of each super pixel. In the proposed model by Vogel & Schiele (2004) partitioning the input image into 10x10 grid blocks and then label into a class of predefined semantic classes using SVM classifier. The low-level features extract from regions are presented using HSI color histogram of 84 bins, edge direction histogram of 72 bins and texture features. The classification performance is 71.7% using color features for region level. The performance dropped to 65.7% for gray scaled regions. This paper is shown in section 2, the previous works of region labelling in scene images. Section 3 presents about MCWS segmentation algorithm that uses for region dataset creation. Section 4 describes about HSV colour space and feature extraction methods. Finally, discuss about the proposed method and conclusion will be given in the last sections.

2. Related Works

The old fashion approaches extract color and texture properties from the whole images, and then use as low level features. In the complex datasets, the approaches are not enough to have detailed idea about the image content. Using only low-level features, intermediate or high-level features generally give satisfactory results on limited datasets. The idea of using a combination of the different feature types together makes more reliable. In an article by Verbeek & Triggs (2008) uses Naive Bayes model that omits four neighbors of couplings, each label of patch depends only on the observation functions. For each label, parameter estimation reduces counting visual word frequencies. This model returns 67.1%accuracy on the use of MSRC 9-class dataset. Therefore, the system uses isolated features for semantic labeling is not sufficient to get higher accuracy. The proposed approach by

Htay & Aye (2014) performed semantic labeling based on region based approach for outdoor scene images. The approach uses Modified Marker-Control Watershed (MCWS)algorithm to get more accurate and compact regions. The paper uses GLCM texture features vector as input to label regions. The authors Htay & Aye (2015) proposed a method for regions labeling and also proposed region merging approach. The paper uses color histogram descriptor in RGB color space. And then, the extracted features given as input to 3-layer FFNN classifier. The proposed method by Htay & Aye (2018) used over-segmented regions as a basic unit. The proposed method produced segmented regions with meaningful labels: aero-plane, cow, bicycle, building, face, car, sky, vegetation and unknown. The method produced labeling accuracy of 68.2% for color features alone and combines features with gray level co-occurrence matrix is 72.4%. The method used the combine features for labeling regions observed most of the vegetation classes. The method has difficulty in labeling the entire body of aero-plane and face, as some parts are mis-categorized as vegetation and building. The system performance evaluated on the public MSRC 9-class dataset. The paper Campbell, et al. (1996) describes the objects classification technique on350 urban and rural scenes images of Bristol Image database. The classification is performed on 11 objects classes. The paper performs pixel-based and region-based classification in two separate cases using ANN classifiers. In the case pixel-based classification based on pixels, each pixel is classified by using color and Gabor texture features rely on frequency. The authors found that intensity feature alone obtains 73.4%recognition rate. This accuracy improves 85.0% adding color features and also improves 87.1% combine with texture feature. In the case of classification based on regions, firstly the images are segmented using k-means segmentation technique. In the case of region-based, contextual and shape features are used. The overall recognition accuracy 82.9% is obtained by using over 3000 regions.

3. Segmentation Technique

The watershed transform is an image segmentation technique based on morphological operations Kasmir Raja, et al. (2009). The boundaries of image regions can determine depend on watershed lines. The watershed transform computes ridge lines and catchment basins for regional boundaries. There are many noise and other local irregularities exist in the traditional watershed transform technique, that leads to over segmentation problem. Many researchers proposed marker controlled watershed segmentation, hierarchical and multi-scale segmentation algorithms to reduce the over segmentation problems.

3.1 Marker Controlled Watershed Segmentation

The segmentation algorithm is very fast to generate segmented regions than other segmentation algorithms like Ncut and Mean shift. The Marker-controlled Watershed Segmentation (MCWS) algorithm is used with 3x3 Prewitt edge mask to produce segmented regions. The morphological dilation and erosion operations are basic in the segmentation. The morphological opening-closing-by-reconstruction and Otsu's thresholding methods are used to compute foreground markers and background markers. This algorithm gives 75% segmentation accuracy on the use of MSRC 9-class dataset. The processing time for each image is approximately 0.7 second. This is very fast and computationally efficient to produce segmented regions. This paper doesn’t use under segmented images as input images.

Table 1 Segmentation Result

Total Images	Over Segmented Images	Under Segmented Images	Segmentation (Time)	Correct Segmentation
240	180	60	~0.7Sec	75%

The table shows the correct segmentation percentage of the MCWS algorithm. The segmentation algorithm produces under segmented regions for some images.

4. Feature Extraction

Feature extraction is a measurement which describes the significant characteristics of an object. The effective combination of color features and the selection of a suitable features are very important for improving classification accuracy. The choice of color space and color quantization schema are key issue in color feature extraction process. There are many color space models RGB, HSV and HSI. Among them, we need to first determine the color space for describing a region by its color features. The selection of best color space relies on the special needs of applications.

4.1 Color Model

Choosing the suitable color model is important to produce the most accurate results. This paper uses HSV color space for color features extraction. The HSV model interested on the visual properties of an image pixels. The HSV color model is a different view of RGB cube Monika Deswal1 & Neetu Sharma (2012). This paper uses (16,4,4) quantization schema for HSV color space.

4.2 Color Moment

The color moment descriptor is scaling and rotation invariant. The color distribution can be represented by its moments. The authors Hui Yu, et al. (2002), calculated the moments for each RGB or HSV channels. The three moments are average color, variance and skewness respectively. The three color moments are computed by using the following formulas:

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

$$\sigma_i = \sqrt{\tfrac{1}{n}\sum_{j=1}^{n}(f_{ij}-\mu_i)^2} \quad (2)$$

$$S_i = \sqrt[s]{\tfrac{1}{n}\sum_{j=1}^{n}(f_{ij}-\mu_i)^3} \quad (3)$$

This paper uses the first two moments of HSV color space because the low order moments contain most of the color distribution information. Where symbol n represents the total pixels of an image, i and j for color channel and the j^{th} pixel value of the image.

4.3 Color Histogram

Color histogram describes color distribution of an image with an image istranslation, rotation and scale invariant. This paper uses 16x4x4=256 features in HSV space to represent the color histogram. Color histogram. The disadvantage of color histogram is very sensitive to noise as reported by Mangijao Singh & Hemachandran (2012). The color histogram of an image is represented as a vector:

H={H(0),H(1),…,H(i),…,H(n)}(4)

Where the symbol i for the color bin, H(i) represents the pixels of color in the image and the number of total histogram bins represent with the symbol n

5. Region Merging

Region merging is a technique that is uses for postprocessing step to merge adjacent regions. The RAG represents region neighborhood relations in the segmented image mentioned by Sonka, M. et al. (1999). The graph G={V,E,W} is weighted undirected which each node represents a region of the over segmented image. The region merging approach have two predicates:

- For first predicate, the region pair need to have minimum edge.
- For second predicate, the two regions must have color homogeneity distribution.

For the first predicate, the adjacent regions which have minimum edge is obtained using the graph G. For the second predicate, 3-layer (FFNN) classifier output images use to obtain color homogeneity distribution. The region pair satisfies these two predicates are actually merged. The two adjacent regions which are not satisfy the above two predicates are ignored.

6. Results and Discussion

The system involves segmentation, features extraction and labeling steps to produce regions with semantic labels.

Figure 1 MSRC 9-Class Dataset

The above figure 1 shows the original and its ground truth images from MSRC 9-class dataset. The above images obtain from Microsoft Research Cambridge (MSRC) 9-class dataset as reported by Criminisi (2004).

6.1 Region Dataset

This system uses MCWS algorithm to produce regions. The regions are used as input for feature extraction and labelling steps. This paper extracts color features in HSV space using the color moment and color histogram descriptors. For the color moment, the first two moments mean and standard deviation (2x3=6) are used as a feature vector. For the color histogram, (16, 4, 4) quantization scheme and

16x4x4=256 features are used as a feature vector for labelling step. The extracted color features from each region are given as a feature set to 3-layer FFNN classifier. The region dataset is used in the training phase of 3-layer FFNN classifier. This paper aims to compare labelling accuracies using the two features.

The segmentation algorithm significantly reduces over segmentation problems than a traditional watershed transform algorithm. The above figure shows the segmented images of MCWS segmentation algorithm. The algorithm produces segmented regions to create as a region dataset.

Figure 2 Segmentation Result

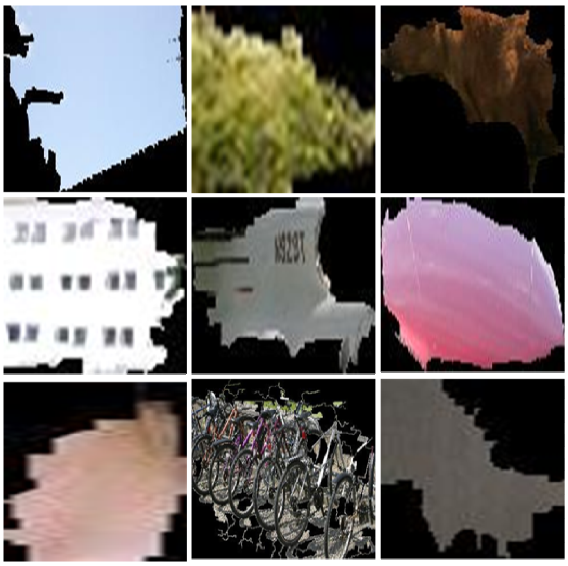

Figure 3 Region Dataset

The above figure 3 shows the result of segmentation algorithm. The output regions are collected as a region dataset as shown in the figure. The first row represents sky, vegetation, cow regions. The second row represents building, aeroplane, car regions. The third row shows face, bicycle, unknown regions.

Table 2 shows the labelling accuracies on the MSRC 9-class dataset of another technique. The authors found that color feature alone is not enough for accurate labelling. The labelling accuracy can be improved by combining with another feature.

Table 2 Pixel-wise Labeling Accuracy

Method	Accuracy
Verbeek & Triggs (2008)	67.1%
Htay & Aye (2018) with CH in RGB space	68.2%
Htay & Aye (2018) with CH+GLCM	72.4%

7. Conclusions

The paper proposes a framework to test classification accuracy using color moment features alone and combine the features with color histogram features in HSV space. There are 6 features of color moment and256 features of color histogram are given to the 3-layer FFNN classifier. The proposed approach produces nine semantic labels aero-plane, cow, bicycle, building, face, car, sky, vegetation and unknown. The paper presents a continuous research and also intends to get higher accuracy for semantic labeling of regions in natural scene images. The results obtained by our method as well as a comparison of these results with those of the existing methods will be detailed in future.

Conflicts of interest

There are no conflicts of interest.

References

[1] Fulkerson, B., Vedaldi, A., & Soatto, S. (2009, September). Class segmentation and object localization with super pixel neighborhoods. In 2009 IEEE 12th international conference on computer vision (pp. 670-677). IEEE.

[2] Vogel, J., & Schiele, B. (2004, July). Natural scene retrieval based on a semantic modeling step. In International Conference on Image and Video Retrieval (pp. 207-215). Springer, Berlin, Heidelberg.

[3] Triggs, B., & Verbeek, J. J. (2008). Scene segmentation with CRFs learned from partially labeled images. In *Advances in neural information processing systems* (pp. 1553-1560).

[4] Htay, K. K., & Aye, N. (2014). Semantic concepts classification on outdoor scene images based on region-based approach. International Journal of Future Computer and Communication, 3(6), 427.

[5] Htay, K. K., & Aye, N. (2016). Regions Labeling in Outdoor Scene Images. In Genetic and Evolutionary Computing (pp. 259-268). Springer, Cham.

[6] Htay, K. K., & Aye, N. (2018, October). Region Labeling in Natural Scene Images. In 2018 IEEE 7th Global Conference on Consumer Electronics (GCCE) (pp. 781-782). IEEE.

[7] Campbell, N. W., Mackeown, W. P., Thomas, B. T., & Troscianko, T. (1996, June). The automatic classification of outdoor images. In *International Conference on Engineering Applications of Neural Networks* (pp. 339-342).

[8] Raja, S. K., Khadir, A., Shaik, A., & Ahamed, S. R. (2009). "Moving Toward Region Based Image Segmentation Techniques: A Study", Journal of Theoretical and Applied Information Technology (JATIT).

[9] Deswal, M., & Sharma, N. (2014). "A Fast HSV Image Color and Texture Detection and Image Conversion Algorithm", International Journal of Science and Research (IJSR), *3*(6).

[10] Yu, H., Li, M., Zhang, H. J., & Feng, J. (2002, September). "Color Texture Moments for Content Based Image Retrieval", Proceeding IEEE Internal Conference on Image Processing, (Vol. 3, pp. 929-932). IEEE.

[11] Singh, S. M., & Hemachandran, K. (2012). Image retrieval based on the combination of color histogram and color moment. International Journal of Computer Applications, 58(3).

[12] Sonka, M., Hlavac, V., & Boyle, R. (1999). Image processing, analysis, and machine vision second edition. International Thomson, 2.

[13] Criminisi, A. 2004 "Microsoft Research Cambridge Object Recognition Image Database", Version-1.0.