A GRAPH-BASED VISION SYSTEM FOR AUTOMATIC OBJECT DETECTION IN OUTDOOR SCENES

RASTI, J[avad]; MONADJEMI, S[ayyed] A[mirhassan] & VAFAEI, A[bbas]

Abstract: In this paper, the characteristics of the human perception system as well as the image features are exploited in a dual-resolution vision system for segmentation/object detection in outdoor scenes. The texture details are deliberately removed and similar color shades are combined together in a low-resolution version of the image, to reduce the excess image information. Using a color clustering algorithm, the color regions of the low-resolution image are found. Then, a weighted graph is constructed, whose nodes contain the detailed features of the regions, derived from the high-resolution image. The weight of the edge between two nodes, Nodes' Merging Potential (NMP), denotes the advantage of merging them together to construct the fundamental image regions. This graph is then pruned regarding the NMP values, so that the main segments are developed and then identified. The proposed algorithm has shown high speed and accuracy for segmentation/object detection in outdoor scenes.
Keywords: vision system, outdoor scenes, object detection, graph-based segmentation

1. INTRODUCTION

Image processing and machine vision techniques have been vastly used in automatic systems (Gonzalez & Woods, 2008). An important related field is designing the environmental machines, to name but a few, wearable computers (Mayol-Cuevas, 2004) and mobile robots as studied by (Bailey, 2002), (Rufus Blas et al., 2008) and (Kim et al., 2007). These automatic systems operate based on object detection in outdoor scenes. For example, consider a wearable computer which is designed to help blind or other visually impaired people (Everingham et al. 2003). This computer should assist the impaired human to find the route out of the home and pass it with no crashes. This is possible using a scene analysis system which can detect and classify the outdoor objects. The environmental robots localization and navigation are other applications who exploit outdoor scenes analysis.

Regarding object detection and scene analysis, image segmentation is one of the most important preliminary stages, which is a determinant factor in the ultimate success of the analysis process (Gonzalez & Woods, 2008). The goal of segmentation is to divide the image into basic disjoint quasi-homogeneous partitions, whose properties can be employed by object detection procedures.

Although there is an extensive literature on image segmentation, the typical algorithms are not such promising in the case of outdoor images. It is due to the special characteristics of such images, like major texture details and luminance adverse effects. Hence, the traditional methods should be customized as suggested in (Kim et al., 2007) and (Wangenheim et. al, 2008).

According to (Bosch et al., 2007), an effective solution is to combine the segmentation and object detection concepts. In that research, the basic features (color, texture, and location) of typical outdoor objects are determined through a voting/training system, and used to develop a database. To analyze an arbitrary input image, it is divided into equal sections and then compared to the database to locate the seeds of the region growing procedure for segmentation. Some unknown regions will remain at the end of the region growing process, which will be merged to the adjacent objects via a Region Belief Fusion technique. This algorithm is not such straightforward, and is accomplished with much processing overhead. In addition, the segmentation is performed based on object detection here, while the reverse process can use some other features, like the shape and the size of segments as well as edge information.

In (Makrogiannis et al, 2005), a Region Adjacency Graph (RAG) is investigated whose nodes are the atomic constructive regions of the image, detected by the Watershed algorithm. Then, the built RAG is pruned considering the color similarity among nodes, so that the main parts of the image are depicted. The Watershed method degrades the process efficiency, since it introduces plenty of primary regions (nodes) in the outdoor images, considering the texture details and luminance effects. Moreover, the regions are so tiny that only their color can be used for graph pruning.

The characteristics of the human perception system can be a reasonable model for artificial object detection and scene analysis. For example, when an unconscious human comes to, the objects are totally blurred and only the colors of some main objects are distinctive. Then, the objects gradually become more obvious in shape and texture, while some new small objects are exposed. This process can motivate novel segmentation algorithms towards object detection, as we customized the *k*-means algorithm for color clustering in outdoor images through a multi-resolution pyramid in (Rasti et al., 2010).

Combining the characteristics of the human perception system with the outdoor objects features in a Region Adjacency Graph, can be profitable for segmentation/object detection in outdoor conditions. In this paper, a dual-resolution humanlike procedure based on the RAG is proposed, which employs various perceptual properties of the objects for segmentation and object detection. The algorithm is described in the next section, and the conclusion and future work are presented in section 3.

2. METHODOLOGY

As mentioned earlier, the segmentation in outdoor scenes is not such straightforward. Figure 1 shows how diverse textures in the greensward and various blue shades in the sky have confused the segmentation algorithm.

(Makrogiannis et al, 2005) start the segmentation from spatial space, thus introduce plenty of primary nodes.

Fig. 1. (a) An outdoor image, and (b) the 6-colored version

On the other hand, some important features of the regions, such as texture and shape, cannot be considered in the segmentation process due to the tiny size of the regions. In this paper, we propose a dual-resolution procedure in the feature space motivated by the human perception system that shows a credible quality for outdoor images segmentation.

This algorithm comprises the following steps:

(1) The texture details and diverse color shades are suppressed using a 5×5 averaging smoothing filter. Thus, the color of the objects becomes the most important property, as what an unconscious human observes while coming to.

(2) Using the MVQ algorithm (Wan et al., 1990), the image colors are clustered into N groups (here 20 clusters), which in turn make the "color regions" of the image, as can be seen in Figure 2.a. Note that every "color cluster" contains some "color regions" which are scattered all over the image. Here, we focus on each region (not cluster) as a probable object.

(3) Each region forms a node in a RAG structure, containing the features such as mean hue and brightness, texture, and region boundary/size, derived from the original (high-resolution) image. The weight of the edge between two nodes (regions) shows the Nodes' Merging Potential (NMP), which is determined based on the nodes' similarity in color, brightness and texture, as well as the relative size and common boundaries of the regions.

(4) The maximum NMP denotes the couple of nodes which must be merged together. So they develop a new node, and the graph information is updated.

(5) Step 4 is repeated until a minimum total weight within the graph is met. Figure 2.b shows the final combination of the regions.

(6) Finally, the objects corresponding to the developed segments are identified using the object models suggested in (Bosch et al., 2007). Other features like the edge shape and object size, are also employed for more accuracy. The labels assigned in Figure 2.b are the outputs of this step. The mentioned Region Belief Fusion also has been used to uniform the segments and reduce the unknown parts.

Several experimental results have shown the speed and accuracy of this method. However, the features used to identify the objects are database-dependent and must be re-calculated for each new data set.

3. CONCLUSION AND FUTURE WORK

General segmentation/object detection algorithms are not such promising in the case of outdoor scenes. In this paper, a dual-resolution segmentation procedure motivated by the human perception system was suggested. The texture details and diverse color shades are neglected using the low-resolution image, so that the color becomes the most important feature of the objects. Then, an over-segmented image is created in the feature space using a color clustering procedure. The developed regions form the nodes of a RAG, whose features are derived from the original (high-resolution) image. This dual-resolution scheme prevents the method to be confused by the texture details from scratch, while considering them in a suitable stage. Finally, the developed segments are assigned to the relevant objects using pre-defined models. This algorithm shows a credible quality in the case of outdoor scenes.

The accuracy of the suggested procedure can be improved using an optimized feature set for determining the NMP values. For example, we used the GLCM matrices for texture analysis as (Bosch et al., 2007) have proposed. The other methods, especially the mixture of color and texture features (Freixenet et.al 2004) can be helpful here. Meanwhile, involving the high-level features of the image can increase the whole process quality both in segmentation and object detection. For instance, the node(s) representing the car tires or windows are to be merged with the node(s) corresponding to the relevant car chassis.

(Makrogiannis et al, 2005) have investigated some other graph pruning algorithms (e.g. SST) that comparing their speed and accuracy with our merging scheme can be informative.

(a)   (b)
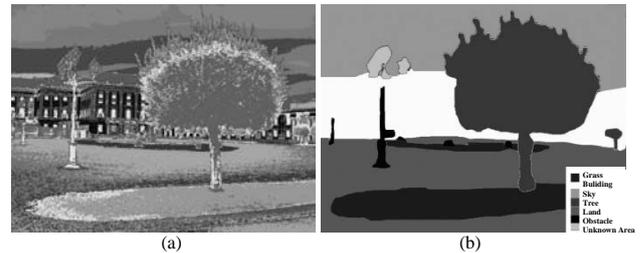
Fig. 2. a) The over-segmented and b) the labeled image

4. REFERENCES

Bailey, T. (2002) *Mobile Robot Localization and Mapping in Extensive Outdoor Environments*, PhD dissertation, Univ. of Sydney, Australia

Bosch A.; Mun~oz, X. & Freixenet J. (2007), Segmentation and description of natural outdoor scenes, *Image and Vision Computing*, Volume 25, Issue 5, pp. 727-740, ISSN: 0262-8856

Everingham, M.R.; Thomas, B. & Troscianko T. (2003) A wearable mobility aid for low vision using scene classification in a Markov random field model framework. *International Journal of Human Computer Interaction*, *special issue on mediated reality*, vol. 15 (2), pp. 231-244, ISSN: 1044-7318

Freixenet, J.; Muñoz, X.; Martí, J. & Lladó, X. (2004) Colour Texture Segmentation by Region-Boundary Cooperation, in Proc. 8th European Conference on Computer Vision, Vol. II, pp. 250-261, Prague, Czech Republic

Gonzalez, R. C. & Woods, R. E. (2008) *Digital Image Processing*, 3rd ed. New York: Pearson Education, ISBN: 9780131687288

Kim, D.N.; Trinh, H.H. & Jo, K. H. (2007) Regions Segmentation using Multiple Cues for Robot Navigation on Outdoor Environment, *International Conference on Control, Automation and Systems*, pp. 1768-1772, Seole, Korea

Makrogiannis, S.; Economou, G.; Fotopoulos, S. & Bourbakis, N.G. (2005) Segmentation of color images using multiscale clustering and graph theoretic region synthesis, *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans,* , Volume: 35 Issue:2, pp. 224 – 238, ISSN: 1083-4427

Mayol-Cuevas, W. (2004) *Wearable Visual Robots*, PhD dissertation, Univ. of Oxford, England

Rasti, J.; Monadjemi, A.H. & Vafaei, A. (2010) A Hierarchical Clustering Method With Gradual Elimination For Color Reduction and Preliminary Segmentation in Outdoor Scenes, *in Proc. 6th Iranian conference on Machine Vision and Image Processing (MVIP)*, Isfahan, Iran

Rufus Blas, M.; Agrawal M.; Sundaresan A. & Konolige K. (2008) Fast Color/Texture Segmentation for Outdoor Robots, *Procceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4078-4085, Nice, France

Wan, S. J.; Prusinkiewicz, P. & Wong, S. K. M. (1990) Variance based color image quantization for frame buffer display, *Color Research and Application 15-1*, pp. 52–58

Wangenheim, A.; Bertoldi R.F.; Abdala D. D.; Richter M.M.; Priese L. & Schmitt F. (2008) Fast two-step segmentation of natural color scenes using hierarchical region-growing and a color-gradient, *Journal of the Brazilian Computer Society*, vol.14, n.4, pp. 29-40, ISSN 0104-6500