IMAGE SEGMENTATION BASED ON HIERARCHICAL MAPPING

Apimun Junda and Orachat Chitsobhuk

Department of Computer Engineering, Faculty of Engineering
King Mongkut's Institute of Technology Ladkrabang,
Ladkrabang, Bangkok 10520, Thailand
e-mail: {s5061051, kcoracha}@kmitl.ac.th

ABSTRACT

In this paper, an efficient image segmentation technique is presented, which combines an image segmentation algorithm with pyramidal approach to form a scale space representation. First, the coarsest level of pyramidal image is quantized to coarse color space. Then, the segmentation is achieved using JSEG algorithm followed by region merging for further refinement. Finally, hierarchical mapping is performed to determine region boundaries in a coarse-to-fine manner using combined global and local features until the final segmentation is accomplished. The multi-resolution approach of the proposed algorithm not only offers a significant reduction in computational cost, but also helps reducing the over-segmentation problem of traditional region growing and watershed techniques. The experimental results show good segmentation performance over a variety of images, also great reduction in the amount of processing time.

1. INTRODUCTION

In computer vision applications, image segmentation constitutes a crucial initial step before performing high-level task including content-based image retrieval, image analysis and recognition. Several segmentation approaches are proposed including morphological watershed and region merge [1, 2], region growing [3, 4], and edge flow and curve evolution [5, 6]. No algorithm could work well on every type of images, which still makes segmentation a challenging problem. The flexible strategies are necessary but time consuming especially when the technique is applied to high-resolution image. Therefore, another solution is to use multi-resolution approach applied to the existing segmentation algorithm. The application of a pyramid proves to be not only more computationally efficient, but also more effective for segmentation of the image. The ideas of segmentation algorithm using multi-scale representation have been proposed [2, 7, 8]. However, all of the above algorithms using only global feature could not produce accurate edge locations since global feature does not well represent information at the boundary of the region. Therefore, local information should be integrated with global feature in order to successfully locate the actual edges of the objects.

The proposed algorithm uses region growing technique as the segmentation kernel at the coarsest level of the image pyramid, since it exploits spatial information and guarantees the formation of closed connected regions. In this case, we use JSEG [9] to accomplish as our region growing method, resulting in over-segmentation. Then, region merge technique is performed at the final segmentation kernel process to obtain the segmented image. The result of this process is passed to the hierarchical mapping step. To evaluate the segmentation results, the quantitative evaluation methods are suggested [10, 11]. In our experiment, we adopt one of the most widely used methods to measure our segmented results compared to other segmentation techniques.

2. THE OVERVIEW OF THE PROPOSED SEGMENTATION TECHNIQUE

We adopt the following five steps for the proposed segmentation technique. The schematic of proposed technique is shown in Fig. 1.

2.1. Hierarchical Image Preparation

At the beginning, the coarser scales of pyramidal images are created from the original image based on bilinear interpolation to down sampling by factor of two. In the pyramid, every element in level *i* has four "children" in level *i*-1 and one "parent" in level *i*+1. From our experience with a large number of different image dataset in a variety of image categories, the smallest image size at the coarsest level should not be less than 10% the original size in order not to loss the important detail of the interested objects.

0-7803-8554-3/04/$20.00 ©2004 IEEE.

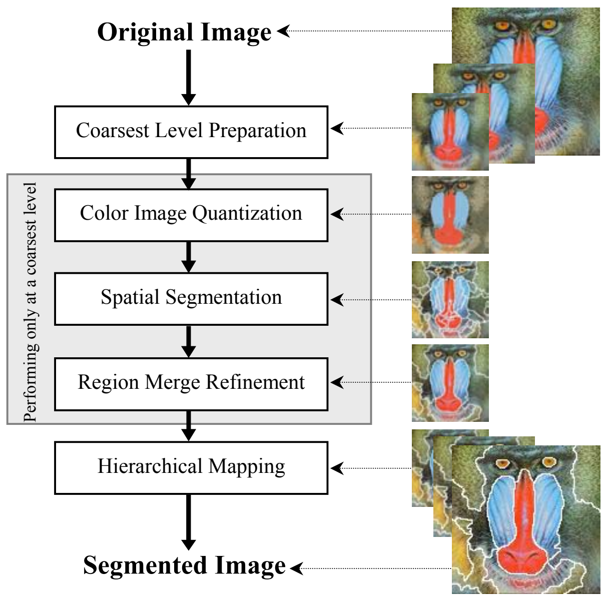


Figure 1. Schematic of proposed segmentation technique.

2.2. Color Image Quantization

The original image, typically 24-bit true color in general natural scenes, contains thousands of colors, which makes it difficult to process directly by spatial algorithm. Accordingly, image colors can be more coarsely quantized using vector quantization (VQ) [12] and a modified general Lloyd algorithm (GLA) [13]. During quantizing process, first, a peer group filtering algorithm [14] is applied to remove noise without edge blurring. As a result, given perceptual weight values indicating smoothness of the local areas is assigned to each pixel. Second, a GLA is used in VQ to quantize the pixel colors. The update centroid condition is modified to incorporate the pixel weights and is defined as in [13]. The centroids are shifted towards points with higher weights. The initializing GLA clusters are estimated by splitting algorithm. In the last step of VQ, generally, there are some close cluster centroids tend to be merged by an agglomerative clustering algorithm [15], such that the minimum distance between two centroids satisfies a specified threshold. Finally, each pixel in the quantized color image is assigned their corresponding color class labels instead of its new quantized value. From this task, the image of labels is called a class-map.

2.3. Spatial Segmentation Algorithm

In this task, we select the JSEG algorithm as our region growing technique to segment uniformly distributed, cluttered and textured regions. The spatial segmentation algorithm is performed on the class-map image to further segment on remaining texture regions. The value of each point in the class-map can be viewed as a spatial data point located in a 2-D plane. Let us first consider the measure J defined as follows. Let Z be the set of all N data points, let $z = f(x,y)$, $z \in Z$, suppose C is the number of color cluster, Z_i, i=1, … , C. Let m be the mean of the entire image and m_i be the mean of cluster Z_i, let

$$S_T = \sum_{z \in Z} \| z - m \|^2 \tag{1}$$

and

$$S_W = \sum_{i=1}^{C} S_i = \sum_{i=1}^{C} \sum_{z \in Z_i} \| z - m_i \|^2 \tag{2}$$

The measure J is calculated as

$$J = S_B / S_W = (S_T - S_W) / S_W \tag{3}$$

The high value of estimated J indicates that the classes are more separated from each other. On the other hand, a small value of J occurred, if all color classes are uniformly distributed over the entire image. The idea is similar to Fisher's multi-class linear discriminant [15], but for arbitrary nonlinear class distributions.

After calculating J-values, the region growing is applied on J-image. First, seeds are determined by the following simple heuristics: Let μ_J and σ_J be mean and standard deviation of the local J values in the region, let T_J be the seed determination threshold

$$T_J = \mu_J + a\sigma_J \tag{4}$$

a is chosen from several present values that will result in the most number of seeds. As a simple condition, pixels with local J values less than T_J and the size larger than the minimum size associated with the selected window size are considered as a seed. For each seed, we eradicate holes in the seeds and grow seeds by connected pixels having the value less than the average of J to form growing areas.

2.4. Region Merging Refinements

Region growing usually results in over-segmentation. These regions tend to be merged based on their feature similarity. A simple strategy like an agglomerative method [15] is applied to merge the regions. The Euclidean distance measure is used to define the different between two neighboring feature histograms. Let H_i and H_j be the histogram of region i^{th} and j^{th} respectively. The distance $D(i,j)$ between H_i and H_j is obtained by:

$$D(i,j) = \| H_i - H_j \| \tag{5}$$

Two neighboring regions are selected to merge with the minimum distance below a threshold. The process continues until every pair of regions has the distance greater than the maximum distance threshold. In the end of this process, the final segmentation results are emerged in the coarsest level of hierarchical images.

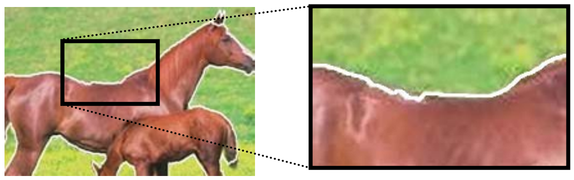
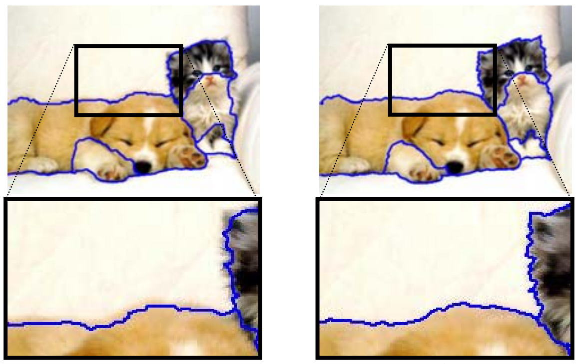


Figure 2. Example mapping results using only global features with the region contains a few color variance.



(a) (b)

Figure 3. Example mapping results on the region with large variance (a) using only global features, (b) using combined local and global features.

2.5. Hierarchical Mapping

In this process, the exacted edges of the interested region on the next finer adjacent level is evaluated by using clustering of combined local and global features at the particular region boundary. Mean and variance of each region is calculated and used for clustering condition. Two regions separated by an edge are designated to be two cluster centroids. The thick edges mapped from the previous level are viewed as the unsegmented part. The mapping process is done by selecting a pixel next to cluster i and j. Then, we calculate the distance D_i and D_j between the pixel value and features of cluster i and j. The pixel is considered to be connected to the cluster i when the condition of $D_i<D_j$ is true. This process is continued until there is no adjacent pixel to cluster i. The unsegmented pixels are combined to cluster j and then the results are mapped to the next finer level until the final segmentation is obtained at the finest level of pyramidal image.

As shown in Fig. 2 No effect occurs, if a region contains only homogeneous color, the basic global feature such as mean value is sufficient to handle the problem. Unfortunately, several natural scenes are rich in combined color and texture. It is difficult to identify edge locations at the region boundaries containing color-texture patterns as shown in Fig. 3(a). Therefore, just the basic global features are inadequate. In the proposed technique, local features weighted together with the global features are used to efficiently improve the previous limitation as shown in Fig. 3(b).

3. EXPERIMENTAL RESULTS

To demonstrate the effectiveness of the proposed technique, the original true color images, nearly ten thousand images from the Corel stock photo gallery with 24 bits per pixel, are used. In our experiments, segmentation performance and quality are evaluated, which are explained in the following sections.

3.1. Performance Evaluation

For comparison, Table 1 lists the segmentation time of the proposed technique versus that of the JSEG approach. The number after the letter H denotes the number of levels in the image hierarchy and the percentage indicates the size of the image at the coarsest level compared to the original one. The experiments were performed on a single-CPU 1.5 GHz Pentium PC with 256 of RAM.

Table 1. Segmentation Time (second)

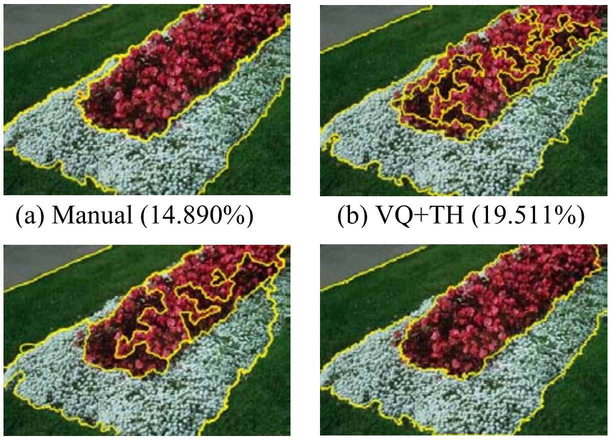
Technique \ Image Size	1600x1200 (pixels)	800x600 (pixels)	512x512 (pixels)	384x256 (pixels)
JSEG	2082.98	98.01	52.62	7.89
H2-50%	116.25	15.12	7.93	1.37
H3-25%	18.11	2.38	1.14	0.24
H4-12.5%	3.69	0.63	0.35	0.18

From Table 1, we can see that the higher resolution of the original image, the lower the processing time of the proposed technique compared to that of JSEG. Especially on the image size 1600x1200, JSEG algorithm requires almost 35 minutes while using the proposed technique spends only few minutes or less.

3.2. Quantitative Evaluation of Segmentation Results

For quantitative evaluation, the popular method proposed by Bosotti et al. [11] has been selected to measure the quality of our segmentation algorithm. The result of the visual evaluation method is in the form of error percentage indicating how many color errors of each region in the segmented image. The evaluation results show whether the segmentation results are good or not.

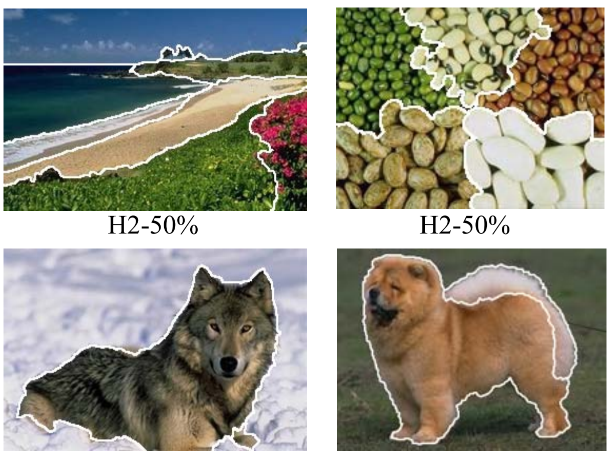
Fig. 4 shows the examples of segmentation results from the proposed algorithm compared to those of other segmentation techniques. From the results, we can see that percent of errors from the proposed algorithm is less than those from other segmentation techniques but very close to human manual segmentation. It can be seen that the results obtained from the proposed technique represents better regions than those obtained by the other selected algorithms. The segmentation results of four different images are also presented in Fig. 5.



(a) Manual (14.890%) (b) VQ+TH (19.511%)

(c) JSEG (26.700%) (d) H3-25% (15.108%)

Figure 4. Example of segmentation results evaluation, (a) manual, (b) combining VQ and Thresholding, (c) JSEG, (d) the proposed technique.



H2-50% H2-50%

H3-25% H2-50%

Figure 5. The segmental results of four different images.

From the experiments, we can see that the proposed technique not only requires less amount of processing time, but also produces the segmentation results subjectively close to human partitioning of these images.

4. CONCLUSIONS AND FUTURE WORK

This paper presents a fast solution to the image segmentation problem by using pyramidal approach. The technique consists of two main parts, spatial segmentation kernel and hierarchical mapping.

The main contributions of the proposed technique are its abilities to preserve the interested regions while increase its time efficiency. The experimental results show that the proposed technique is able to reduce the amount of processing time thus obtain good perceptual quality of the segmented results on a wide range of image categories. However, some limitations are found for the proposed strategy. One case is that several control parameters are needed to be tuned. Another case is when an error occurred during the mapping process, these errors become propagation effect. Therefore, additional strategy should be integrated to the proposed algorithm to reduce possible propagation errors between adjacent levels. Moreover, the appropriate parameter selection should be automatically defined on individual image. Automatic selection of the appropriate minimum size of the coarsest level should also be specified in order to obtain the excellent performance.

5. REFERENCES

[1] K. Haris, S.N. Estradiadis, N. Maglaveras and A.K. Katsaggelos, “Hybrid image segmentation using watersheds and fast region merging”, *IEEE Trans. on Image Processing*, vol.7, no. 12, pp. 1684-1699, Dec. 1998.

[2] A.S. Wright and S.T. Acton, “Watershed pyramids for edge detection”, *Proc. of IEEE ICIP*, Santa Barbara, CA, Oct. 1997.

[3] S.C. Zhu and A. Yuille, “Region Competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentatoin”, *PAMI*, vol. 18, no. 9, pp. 884-900, Sep. 1996.

[4] R. Adams and L. Bischof, “Seeded region growing,” *IEEE Trans. On PAMI*, vol. 16, pp. 641–647, 1994.

[5] B. Sumengen, B.S. Manjunath and C. Kenney, “Image Segmentation using Curve Evolution and Flow Fields”, *Proc. of IEEE ICIP*, Rochester, NY, USA, Sep. 2002.

[6] W.Y. Ma and B.S. Manjunath, “EdgeFlow: A technique for boundary detection and segmentation”, *IEEE Trans. on Image Processing*, vol. 9, no. 8, pp. 1375-88, Aug. 2000.

[7] Jianzhuang Liu and Weizin Xie, “Pyramid Segmentation of Color Image Using Fuzzy C-Means Clustering Alogorithm”, *IEEE TENCON*, 1993.

[8] M.R. Rezaee, P.M.J. van der Zwet, B.P.F. Lelieveldt, R.J. van der Geest and J.H.C. Reiber, “A Multiresolution Image segmentation Technique Based on Pyramidal Segmentation and Fuzzy Clustering”, *IEEE Trans. Image Processing*, vol. 9, No. 7, pp. 1238–1248, 2000

[9] Y. Deng and B.S. Manjunath, “Unsupervised Segmentation of Color-Texture Regions in Image and Video”, *IEEE Trans. on PAMI*, vol. 23, no. 8, pp. 800-810, Aug. 2001.

[10] J. Liu and Y.H. Yang, “Multiresolution Color Image Segmentation”, *IEEE Trans. on PAMI*, vol. 16, no. 7, pp. 689-700, 1994.

[11] M. Borsotti, P. Campadelli and R. Schettini, “Quantitative Evaluation of Color Image Segmentation Results”, *Pattern Recognition letters*, vol. 19, no. 8, pp. 741-48, 1998.

[12] R.M. Gray, “Vector Quantization”, *IEEE Acoustics, Speech and Signal Processing Society Magazine*, pp. 4-29, April 1984.

[13] A. Gersho and R.M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publisher, Boston, USA, 1992.

[14] Y. Deng, C. Kenney, M.S. Moore and B.S. Manjunath, “Peer group filtering and perceptual color image quantization”, *Proc. of IEEE ISCAS'99*, Orlando, FL, USA, vol. 4, pp. 21-24, June 1999.

[15] R.O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York, 1970.