3rd International Conference on Electrical & Computer Engineering
ICECE 2004, 28-30 December 2004, Dhaka, Bangladesh

COLOR OBJECT SEGMENTATION FOR HELPER ROBOTS

Md. Altab Hossain, Rahmadi Kurnia, Akio Nakamura and Yoshinori Kuno
Department of Information and Computer Sciences, Saitama University,
255 Shimo-Okubo, Sakura-ku, Saitama-shi 338-8570, Japan
Email :{hossain, kurnia, nakamura, kuno}@cv.ics.saitama-u.ac.jp

ABSTRACT

We are developing a helper robot that carries out tasks ordered by the users through speech. Users sometimes may order for an object by specifying its color. The robot needs a vision system to recognize the object appearing in the orders. Specific color object segmentation plays the most important role in recognizing such object. This paper presents the segmentation technique of a specific color object(s) in different illumination conditions by combining the Mean shift algorithm and the HSI color space. The Mean Shift algorithm has been used first for image pre-processing. This reduces the number of colors in the image and divides it into several regions. Then the Hue, Saturation and Intensity components of HSI color space have been used for merging regions of specific color object(s). Finally, the Median filter has been used for smoothing the image and the region growing algorithm has been used to eliminate small regions as image post processing. The results show the effectiveness of the proposed segmentation technique for specific color object(s) in different illumination conditions.

1. INTRODUCTION

Helper robots or service robots in welfare domain have attracted much attention of researchers for the coming aged society [1, 2]. Multimodal interfaces [3, 4, 5] are considered good interface means for such robots. Thus, we are developing a helper robot that carries out tasks ordered by the user through voice and/or gestures [6, 7, 8]. In addition to gesture recognition, such robots need to have vision systems that can recognize the objects mentioned in speech.

It is, however, difficult to realize vision systems that can work in various conditions. Image segmentation is one of the most difficult but essential tasks in computer vision [9, 10]. Unlike other vision tasks such as parametric model estimation, fundamental; matrix estimation, optical flow calculation etc., there is no widely accepted model or analytical solution for image segmentation. Probably, there is no "one true" segmentation acceptable to all different people and under different psychophysical conditions [11] There has been proposed many segmentation methods [12]. These methods can be classified into [9]:

(1) Histogram thresholding
(2) Clustering
(3) Region growing
(4) Edge-based
(5) Physical-model-based
(6) Fuzzy approaches and
(7) Neural network approaches.

The appearance of the objects in an image depends on the combined effects of object's 3D geometry, pose in the scene, surface properties, and illumination conditions. All of these properties contribute to the imaging process and can result in significantly different appearances of the same object in an image. Especially illumination conditions can have a dramatic effect on the appearance of an object. Several segmentation/object recognition methods in the past ignored variations in illumination, or performed simple normalization procedures, or used some "illumination invariant" features in the image like edges. Adini et al. [13] have demonstrated for faces that variations in illumination cause larger variations in the image than different subjects seen under the same illumination conditions. Wang and Suter proposed a segmentation method introducing both global information and local homogeneity into mean shift algorithm [14]. They have used cyclic property of the hue component in the mean shift algorithm.

In this paper, we propose an illumination insensitive segmentation method by first applying the mean shift algorithm for reducing the colors and regions in the image and then applying the Hue-Saturation-Intensity components of HSI color space for merging the homogenous regions of a specific color object.

2. THE MEAN SHIFT ALGORITHM

Since Fukunaga and Hosterler [15] introduce the mean shift algorithm, the mean shift algorithm has been extensively exploited and applied in low-level computer vision tasks [10] for its ease and efficiency. The mean shift algorithm is based on kernel density estimation.

2.1 Mean Shift Algorithm and Kernel Density Estimation.

Let {Xi}i=1,…,n be a set of n data points in a d-dimensional Euclidian space R^d, the multivariate kernel density estimator with kernel K and window radius (band-width) h is defined as follows ([16]):

$$\hat{f}(x) = \frac{1}{nh^d}\sum_{i=1}^{n} k\left(\frac{x - X_i}{h}\right) \quad (1)$$

The kernel function should satisfy some conditions.

The Epanechinkov kernel [16] is one of the optimum kernels, which yields the minimum mean integrated square error (MISE):

$$k_e(x) = \begin{cases} \frac{1}{2}c_d^{-1}(d+2)(1-x^T x) & \text{If } x^T x < 1 \\ 0 & \text{Otherwise} \end{cases} \quad (2)$$

Where c_d is the volume of the unit d-dimensional sphere, e.g., $c_1=2$, $c_2=\pi$, $c_3=4\pi/3$.

Thus, the density gradient estimate of the Epanechinkov kernel can be written as:

$$\hat{\nabla} f(x) \equiv \nabla \hat{f}(x) = \frac{1}{nh_d}\sum_{i=1}^{n} \nabla k\left(\frac{x - X_i}{h}\right) \quad (3)$$

Equation (3) can be rewritten as:

$$\hat{\nabla} f(x) = \frac{n_x}{n(h^d c_d)} \frac{d+2}{h^2}\left(\frac{1}{n_x}\sum_{X_i \in S_h(x)} (X_i - x)\right) \quad (4)$$

Where the region $S_h(x)$ is a hypersphere of the radius h, having the volume $h^d c_d$, centered at x, and containing n_x data points.

The mean shift vector $M_h(x)$ is defined as:

$$M_h(x) \equiv \frac{1}{n_x}\sum_{X_i \in S_h(x)} (X_i - x) = \frac{1}{n_x}\sum_{X_i \in S_h(x)} X_i - x \quad (5)$$

From equations (4) and (5), we get:

$$M_x(x) \equiv \frac{h^2}{d+2} \frac{\hat{\nabla} f(x)}{\hat{f}(x)} \quad (6)$$

Equation (6) firstly appeared in [15]. The mean shift is an unsupervised nonparametric estimator of density gradient and the mean shift vector is the difference between the local and the center of the window.

3. HSI COLOR SPACE

The human color perception is conveniently represented by the following set of color features: H (Hue), S (Saturation), I (Intensity). Hue is the specification of intrinsic color, i.e., the color aspect of a visual impression. Saturation describes "how pure the color is". The last component of the HSI triple is a measure of "how bright the color is". The HSI color space can be described geometrically as in Figure 1.

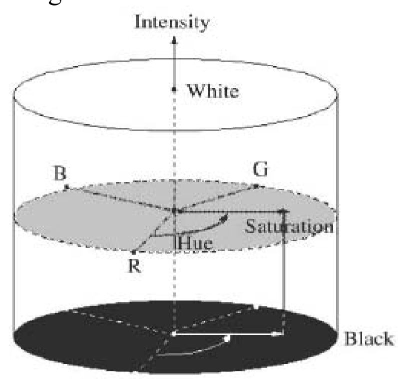


Fig. 1 HSI color Space [9].

The RGB/HSI transformation is non-linear and in "classical" version is defined as:

$$H = \arccos\left(\frac{0.5((R-G)+(R-B))}{\sqrt{(R-G)^2 + (R-B)(G-B)}}\right) \quad (7)$$

If B>G then: H :=360^0-H

$$S = 1 - 3\frac{\min\{R, G, B\}}{R+G+B} \quad (8)$$

$$I = \frac{(R+G+B)}{3} \quad (9)$$

Generally, hue is considered as an angle between a reference line and the color point in RGB space. The range of the hue value is from 0^0 to 360^0, for example, blue is 240^0, yellow is 60^0, green is 120^0, and magenta is 300^0.

The advantages of HSI over RGB are [9]:

- Hue is invariant to certain types of highlights, shading, and shadows.
- The segmentation is performed on only one dimension (H) and results of segmentation have fewer segments than using RGB.

4. SEGMENTATION METHOD

Although the mean shift algorithm and the HSI color space can be separately used for color image segmentation, they surely fail to segment image when illumination condition will change. In this paper, we use the mean shift algorithm as an image preprocessing tool to reduce region and number of colors used and then use the HSI color space to merge regions for segmentation of a single colored object in different illumination condition.

Our proposed method consists of the following parts:

- Apply the mean shift algorithm into a real image to reduce colors and divide it into several regions.
- Merge the regions of a specific color based on H, S, I components of the HSI color space.
- Filter the result using the median filter.
- Eliminate the small regions using the region growing algorithm.

The details of the proposed method are:

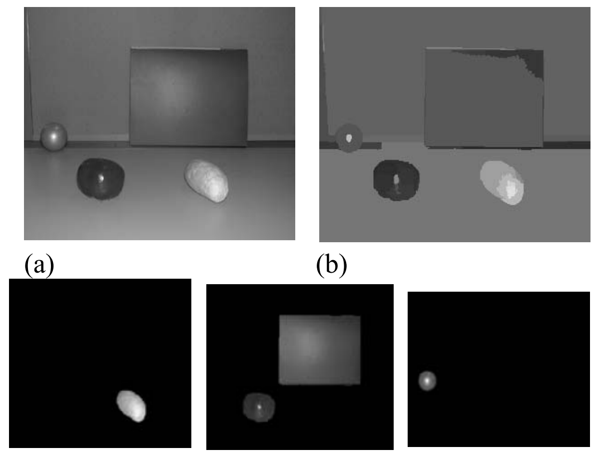
The input image is first analyzed using the mean shift algorithm. The image may contain many colors and various regions. The algorithm reduces the number of colors and regions. Thus, the output of the mean shift algorithm is several regions with a few numbers of colors than the input image.

However, this region does not imply that each region comes from a single object. The mean shift algorithm may divide even a single colored object into several regions with more than one color. To remove this ambiguity, we use the Hue, Saturation and Intensity components of the HSI color space to merge the homogeneous regions which likely come from a single object. For transformation of the image from RGB to HSI space, we use equations (7), (8) and (9). For homogeneous regions we use threshold values for each component of HSI. The threshold values are selected dynamically based on the illumination condition of the image and thereby efficiently segment out specific color object in different illumination conditions.

Then we use the median filter as image post-processing. This may help to smooth the image boundary and also helps to reduce the unwanted regions. Finally, we also use the region growing procedure as post-processing procedure to avoid over segmentation or remove small highlight from objects. This procedure removes all regions smaller than a given area by analyzing their 4-connected neighborhood and thus segmentation of a specific color object is finished.

5. EXPERIMENTS ON COLOR IMAGE

We test our color image segmentation method on natural images. Figures 3 and 4 show the segmentation results.



(a) (b) (c) (d) (e)

Fig. 3 (a) Original Image (used colors 39399), (b) Image after applying mean shift algorithm (used colors 55), (c) Segmented Yellow object, (d) Segmented Red objects and (e) Segmented Green object.

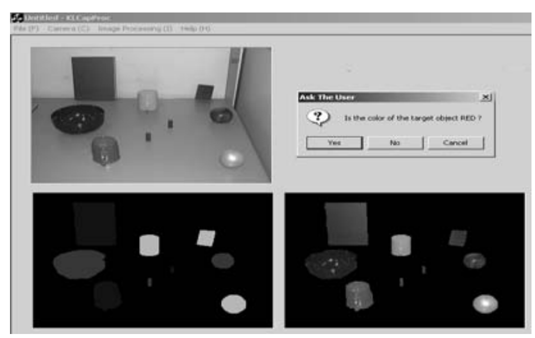


Fig. 4 (a) Original image (top left); Color segmentation (bottom left); Foreground objects (bottom right).

The image in Figure 3(a) contains 29399 colors, but only four single color foreground objects are visible to us. After applying mean shift algorithm the number of colors are reduced into 55 and the image is divided into several regions (Figure 3(b)). Then the threshold values of H, S and I have been used to merge the regions that are fragmented from a single object. The threshold values are selected dynamically. For example, to segment the yellow object the threshold value of H is selected 35^0 as lower limit and 61^0 as higher limit where the actual yellow is 60^0. Similarly, for red objects the threshold value is 9^0 as lower limit and 360^0 as higher limit (Hue value is ranging from 0^0 to $360^{0.}$ Actual red is 0^0, so the H value of this red object is 360^0-0^0-9^0) and for green object the threshold value is 80^0 as lower limit and 140^0 as higher limit. Similarly all single colored objects are segmented. The segmented image is processed further to divide it into foreground and background for object recognition purposes of our helper robot (Figure 4).

6. CONCLUSION

Robots need vision to carry out their tasks. However, conventional vision systems cannot recognize objects in complex scenes. To make the vision system efficient, we have implemented a specific color object segmentation method based on mean shift algorithm and HSI color space. Experiments show the proposed method achieves promising results for natural single color object segmentation in different illumination conditions. However the system does not work for multi color objects. Future works include multi color object segmentation in different illumination conditions based on this with adding additional attributes of objects.

ACKNOWLEDGMENTS

This work was supported in part by the Ministry of Education, Culture, Sports, Science and Technology under the Grant-in-Aid for Scientific Research (KAKENHI 14350127).

REFERENCES

[1] M. Ehrenmann, R. Zollner, O. Rogalla, and R. Dillmann, "Programming service tasks in household environments by human demonstration," ROMAN 2002, pp.460-467.

[2] M. Hans, B. Graf, R.D. Schraft, "Robotics home assistant Care-O-bot: Past-present-future," ROMAN 2002, pp.380-385.

[3] G. A. Berry, V. Pavlovic, and T. S. Huang, "BattleView: A multimodal HCI research application," Workshop on Perceptual User Interfaces, pp. 67-70, 1998.

[4] I. Poddar, Y. Sethi, E. Ozyildiz, and R. Sharma, "Toward natural gesture/speech HCI: A case study of weather narration," Workshop on Perceptual User Interfaces, pp. 1-6, 1998.

[5] R. Raisamo. "A multimodal user interface for public information kiosks," Workshop on Perceptual User Interfaces, pp. 7-12, 1998.

[6] T. Takahashi, S. Nakanishi, Y. Kuno, and Y. Shirai, "Human-robot interface by verbal and nonverbal communication," IROS 1998, pp.924-929, 1998.

[7] M. Yoshizaki, Y. Kuno, and A.Nakamura, "Mutual assistance between speech and vision for human-robot interface," IROS 2002, pp.1308-1313, 2002.

[8] M. Yoshizaki, A. Nakamura, and Y. Kuno, "Vision-speech system adapting to the user and environment for service robots," IROS2003, CD-ROM, 2003.

[9] H.D. Cheng, et al., "Color Image Segmentation: Advances and Prospects," Pattern Recognition 34, pp. 2259-2281, 2001.

[10]D.Comaniciu, and P. Meer, "Mean Shift: A Robust Approach towards Feature Space A Analysis," IEEE Trans. Pattern Analysis and Machine Intelligence, 2002, 24(5): p. 603-619.

[11]K.S. Fu, J.K. Mui, "A survey on image segmentation," Pattern Recognition 13 (1981) 3-16.

[12]W. Skarbek and A. Koschan, "Color imgae segmentation- a survey," Technical report, Tech. University of Berlin, Germany, October 1994.

[13]Y. Adini, Y. Moses, S. Ullman, "Face recognition: the problem of compensating for changes in illumination direction," PAMI 19 (7) (1997) 721–732.

[14]Hanzi Wang and David Suter, "Color image segmentation using global information and local homogeneity," Proc. VIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Qurselin S. and Adriaansen T. (Eds.), 10-12 Dec. 2003, Sydney.

[15]K. Fukunaga, and L.D. Hostetler, "The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition," IEEE Trans. Info. Theory, 1975, IT-21: p. 32-40.

[16] B.W. Silverman, "Density Estimation for Statistics and Data Analysis," 1986, London: Chapman and Hall.