Image Segmentation using Weighted K-Means Clustering on RGB-D Data

Mehmet Hakan DURAK

Faculty of Engineering, Department of Electrical-Electronics Engineering, Gazi University, Ankara, Turkey

mhdurak@gazi.edu.tr

Abstract. Image segmentation is a key technology of computer vision and image processing which partition an image into segments rely on the image parameters such as color, gray level, texture, motion or depth. In this paper, we present the approach of image segmentation using weighted k-means clustering on RGB-D data. K-means clustering algorithm is an unsupervised algorithm and it is used to segment the specific region from the other region. Weighted k-means clustering, also, is based on the weighting of the parameters used. In the proposed method, we tried to segmentation on RGB-D data using weighted k-means clustering by changing the weights of parameters such as the Depth (D), Lightness (L), a and b for the color opponent dimensions and the surface normals. Ground truth images have been used to compare the results. Method is open to development with different parameters and weights.

Keywords: image segmentation, weighted k-means clustering, RGB-D, Kinect

1 INTRODUCTION

With increasing advances in computer and imaging technology, human-specific skills are examined, the research on transferring these skills to computerized systems has gained speed [4]. We locate everything we perceive in our everyday life and distinguish easily with meaningful fragments. For example, we can detect where the door is and then we can find the doorknob. Segmentation which is process of separating scenery into non-overlapping meaningful region we do it quickly without notice at every moment has become the main topic of computer vision science.

Image segmentations are performed by finding boundaries between objects in the image. In traditional methods, features were extracted from the color images obtained from RGB camera. Because color segmentation which is a one of the traditional method is performed by finding the same color regions in the image, it is inadequate in partitioning of different object which is closest color. Segmentation using the discontinuity in the depth data is easier than the one using color information and finds the different object with same color [10]. However, due to the low resolution and noise in the depth data, it does not give the right segmentation result alone [1]. With the popularity of low-cost and easy-to-use RGB-D cameras such as Microsoft Kinect, it has become easier to get both color and depth information of the same scene. The availability of color and depth information gives a new insight into segmentation algorithms and provides better partition performance [2,3,9]. Local surface information such as surface normal used in many applications in the field of computer vision also provides descriptive information about the region. However, normal related to the first derivative causes incorrect results in depth data because of noise [12]. All of these methods such as RGB, depth discontinuity and normal used to extract features have been used all together to eliminate the disadvantages separately.

ISSN 2162-3228 Vol 8, No 6, Dec 2016

There is no only one method for image segmentation since each image has a different characteristic and each method gives good results for different images [5,6]. One of the simplest and most popular methods used is the K-means clustering algorithm [7]. The main advantage of this approach is that it is very fast and is easily applied in large data sets [8]. With the K-means, the data is divided to independent and homogeneous clusters into k numbers and each cluster is characterized by a center point. However, since the K-means algorithm is sensitive to initialization, it may result in undesired segmentations [7]. Color and depth information is given to the K-means algorithm. It is a disadvantage that RGB color information is not very sensitive to light values [9]. To get rid of this, it is preferred to convert the RGB color information to the CIE-Lab color space, which is similar to the color perception scheme of the human eye.

Most of the existing K-means algorithms are equally weighted against all features. However, because different features in each image come into prominence, the weights of these features also affect the desired segmentation performance [11

2 PROPOSED ALGORITHM

First of all, the proposed method loads the RGB-D data and L^*, a, b Color Space transformation is applied to RGB data. Surface normals (N_x, N_y, N_z) are obtained from Depth data. All of these data such as N_x, N_y, N_z , Depth, a and b are used for create a feature matrix. After this, the feature matrix is weighted by weight matrix for some features to be foreground. Finally, with applying K-Means Clustering, different objects or surfaces are separated from each other.

The flowchart of our algorithm is showed in Fig. 1.

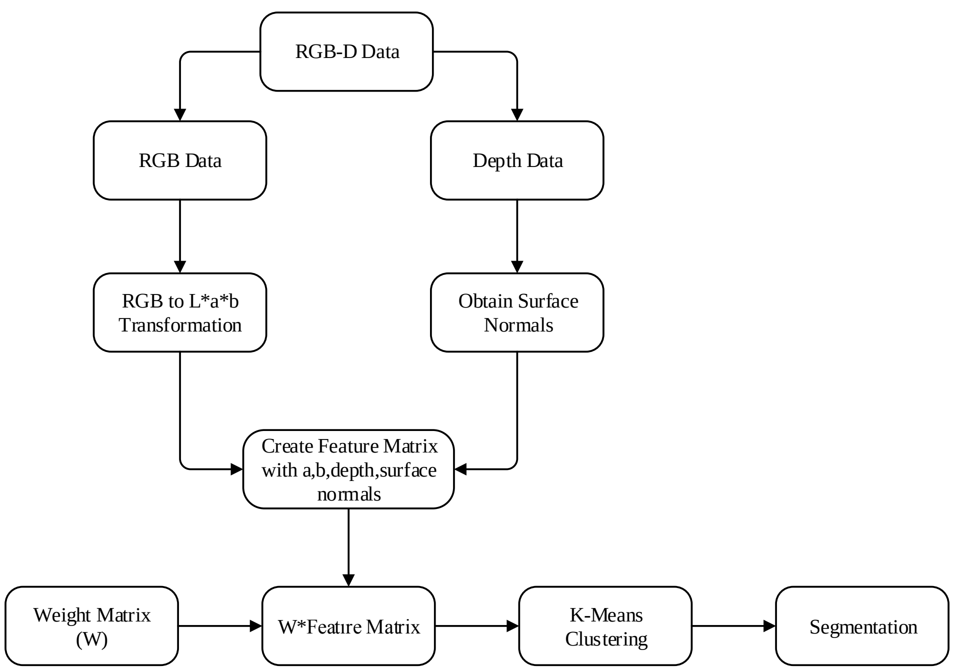


Fig. 1. The flowchart of the proposed algorithm

In the following sections, details about the algorithm are given.

ISSN 2162-3228 Vol 8, No 6, Dec 2016

2.1 RGB to Lab transformation

The Lab color space includes all perceivable colors and is designed to approximate human vision. This color space is organized in a cube form as shown in the Fig. 2. The highest value of the L is 100, which means excellent reflecting diffuser, the lowest value of the L is zero, which means zero. The a and b axes haven't such numerical limits, but positive and negative a are red and green, respectively. And positive and negative b are yellow and blue, respectively. [9]

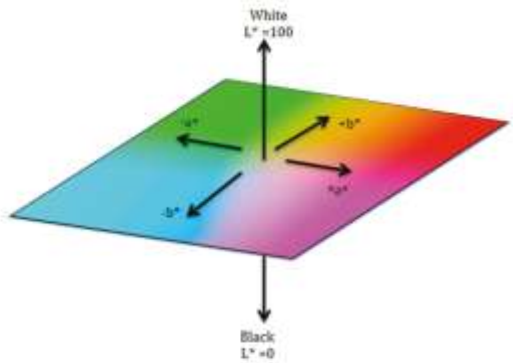


Fig. 2. Lab color space [9]

The coordinates are calculated from a linear transformation of the RGB space components as shown equation 1.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \tag{1}$$

In equation 1, the X and Z components are normalized spectral ponderation curves based on statistical experiments with human observers, the Y component describes the illumination.

The color coordinates of every pixel in the Lab are obtained from non-linear transformation applied on the XYZ coordinates using equations 2, 3, 4 and 5.

$$L^* = 116\left[f\left(\frac{Y}{Y_w}\right)\right] - 16 \tag{2}$$

$$a = 500\left[f\left(\frac{X}{X_w}\right) - f\left(\frac{Y}{Y_w}\right)\right] \tag{3}$$

$$b = 200\left[f\left(\frac{Y}{Y_w}\right) - f\left(\frac{Z}{Z_w}\right)\right] \tag{4}$$

$$f(t) = \begin{cases} t^{\frac{1}{3}} & if\ t > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 t + \frac{4}{29} & if\ t \le \left(\frac{6}{29}\right)^3 \end{cases} \tag{5}$$

X_w, Y_w, and Z_w are tristimulus of XYZ values with reference to the "white spot".

$$\begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} = \begin{bmatrix} 0.9504 \\ 1.000 \\ 1.0887 \end{bmatrix} \tag{6}$$

With applying to this transformation, a and b are used for create the feature matrix. These parameters are useful for segmentation algorithms. After this process, Feature matrix is created with a, b, depth and surface normals.

2.2 Weighted K-Means Clustering

K-means clustering is popular unsupervised learning algorithm and this algorithm is used to several important clustering problems successfully. In K-means clustering, it divides a collection of data into a k number group of data. Thus, it separates a given set of data into k number of different cluster. This algorithm, firstly, calculates the k centroid and then it takes each point to the cluster which has nearest centroid from respective data point. In this process, Euclidean distance is used for define the distance of the nearest centroid. Euclidean distance is calculated by $d = \|p(x,y) - c_k\|$ where k is number of cluster, c_k is the cluster centers and $p(x,y)$ is the input pixel in the image which is size of $x \times y$. When the labeling is done, new centroid of each cluster is calculated by $c_k = \frac{1}{k}\sum_{y \in c_k}\sum_{x \in c_k} p(x,y)$. And then new Euclidean distance which have minimized is calculated between each center and each data point and assigns the points in the cluster. Each cluster in the partition is identified by its member objects and by its centroid. Since it minimizes the sum of distances from each object to its cluster centroid iteratively, K-means clustering is an iterative algorithm. [7]

In this paper, K-means clustering is applied to feature matrix after weighting. Thus, with multiplying weight matrix, it increases some feature's influence while reduces other feature's influence.

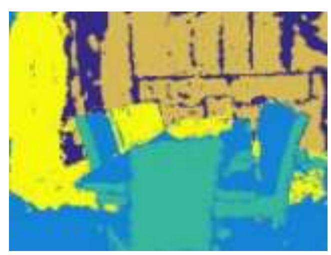
3 EXPERİMENTS

NYU Depth Dataset V2 [13] is used for the analysis. This dataset has 1449 RGB-D data of 464 scenes from 3 cities. Matlab is used to implement the proposed algorithm. The results compared with ground truth images. Two experiments were carried out to develop the method. First experiment is about the cluster number. The success of segmentation was examined by changing the number of clusters. Second experiment is about weight of the depth data. We examined the effect of segmentation by changing weight of this feature.

3.1 Effect of cluster number

(a)

(b)



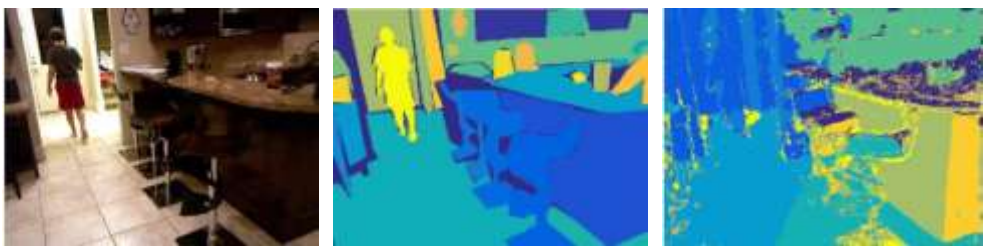
(c) (cluster number=5)

(d) (cluster number=10) (e) (cluster number=20) (f)(cluster number=30)

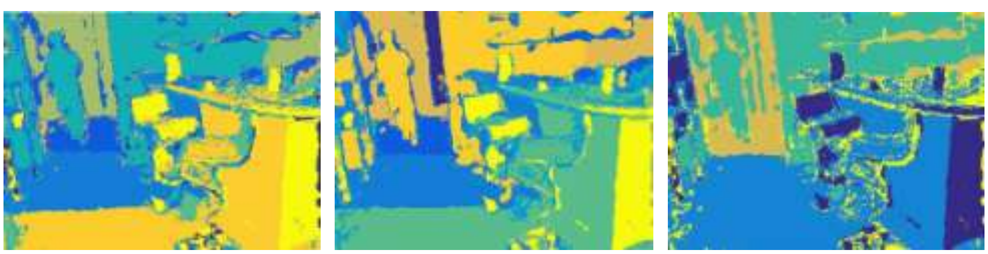
Fig. 3. (a) RGB color image (b) Ground Truth image (c, d, e, f) Proposed method with different cluster number

We analyze the effect of cluster number by performing many tests in dataset mentioned. As it is seen, cluster number is very important parameter for our algorithm. When we increase the number of clusters, the processing load increases and the segmentation of the objects is not as desired. Also, when we reduce the number of clusters too much, the processing load is reduced, but the segmentation of the objects is not as desired. In the experiments performed for this RGB-D data, the number of clusters of 10 can be interpreted as a good result.

3.2 Effect of weight of depth data



(a) (b) (c)



(d) (e) (f)

Fig. 4. (a) RGB color image (b) Ground Truth image (c, d, e, f) proposed algorithm with weight of depth :1,10,20,50

Experiments on the weight of the depth data, we saw that the segmentation process is more successful by increasing the weight of depth. But when we increase too much, the process load increases and the performance decreases.

4 CONCLUSIONS

In this paper, we propose a weighted K-means clustering method for image segmentation. A fast method with k-means is targeted because the process load is low. The effect of several parameters has been examined. Chose the cluster numbers is very important for segmentation as desired. And it has been seen that the effect of depth data is more on these. Algorithm is a basic and is open to developments.

References

[1] Cruz, L., Lucio, D., Velho, L., (2012*). Kinect ant RGBD Images Challenges and Applications.* Graphics, Patterns and Images Tutorials (SIBGRAPI-T),(pp. 36-49)

[2] Mutto, C. D., Zanuttigh, P., Cortelazzo, G. M., (2010). *Scene Segmentation by Color and Depth Information and its Applications.* Streaming Day.

[3] Dahan, M. J., Chen, N., Shamir, A., Cohen-Or, D., (2012) *Combining color and depth for enhanced image segmentation*. The Visual Computer, Volume 28, Issue 12, (pp.1181–1193)

[4] Zhu, H., Meng, F., Cai, J., Lu, S., (2016) *Beyond pixels A comprehensive survey from bottom-up to semantic*. Journal of Visual Communication and Image Representation 34, (pp. 12-27)

[5] Khan, W., (2013) *Image Segmentation Techniques A Survey*. Journal of Image and Graphics (pp. 166-170)

[6] Zhu, H., Meng, F., Cai, J., & Lu, S. (2016). *Beyond pixels: A comprehensive survey from bottom-up to semantic image segmentation and cosegmentation*. Journal of Visual Communication and Image Representation, 34, 12-27.

[7] Dhanachandra, N., Manglem, K., & Chanu, Y. J. (2015). *Image Segmentation Using K-means Clustering Algorithm and Subtractive Clustering Algorithm*. Procedia Computer Science, 54, 764-771.

[8] Coates, A., & Ng, A. Y. (2012). *Learning feature representations with k-means*. In Neural Networks: Tricks of the Trade (pp. 561-580). Springer Berlin Heidelberg.

[9] Hernandez-Lopez, J. J., Quintanilla-Olvera, A. L., López-Ramírez, J. L., Rangel-Butanda, F. J., Ibarra-Manzano, M. A., & Almanza-Ojeda, D. L. (2012*). Detecting objects using color and depth segmentation with Kinect sensor*. Procedia Technology, 3, 196-204.

[10] Liu, H., Philipose, M., & Sun, M. T. (2014). *Automatic objects segmentation with RGB-D cameras*. Journal of Visual Communication and Image Representation, 25(4), 709-718.

[11] Walid Ayech, M., & Ziou, D. (2015). *Automated feature weighting and random pixel sampling in k-means clustering for terahertz image segmentation*. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 35-40).

[12] Liu, Y., & Xiong, Y. (2008). *Automatic segmentation of unorganized noisy point clouds based on the Gaussian map*. Computer-Aided Design, 40(5), 576-594.

[13] Silberman, N., Hoiem, D., Kohli, P., & Fergus, R. (2012, October). *Indoor segmentation and support inference from RGBD images*. In European Conference on Computer Vision (pp. 746-760). Springer Berlin Heidelberg.

Mehmet Hakan Durak is a research assistant and PhD student at the Gazi University. He received his BS and MS degrees in electrical-electronics engineering from the Atatürk University and Eskişehir Osmangazi University in 2011 and 2014, respectively. His current research interests include speech and image processing and telecommunications.