FUSION TECHNIQUE FOR SCENE SEGMENTATION

[1]*PURVI A. PATEL. ,* [2]*KUNAL ACHARYA*

[1] **Student of (Communication System Engineering) Master of Engineering, Department of Electronics and Communication , L.D. College of Engineering, Ahmedabad**

[2] **Asst.Professor, Department of Electronics and Communication, L.D. College of Engineering, Ahmedabad**

purvipatel369@gmail.com

acharyakunal1@gmail.com

ABSTRACT : *The need for scene segmentation with high accuracy is increasing day by day in the area of computer vision and image processing. We introduce various approaches for designing algorithms for scene segmentation with high reliability in this thesis.*

This report attempts to introduce various segmentation techniques and focuses on different stereo matching algorithms for geometry information, which is mainly used for 3D video, game controlling, etc. Depth estimation plays a crucial role for extracting geometry information from stereo images. Depth estimation and a complete study of different algorithms as well as software based approach for depth estimation is also presented. It is known that no segmentation technique is equally good for all scene and all techniques are not good for a particular type of scene. So here thought of fusing geometry information and color information for better scene segmentation is proposed as only geometry information is not sufficient for scene segmentation.

KEY WORDS *: Scene Segmentation, Geometry information, 3D, Fusion.*

1. Introduction

Image processing is rapidly growing area of computer science. Its development has been fueled by scientific advances in digital imaging, computer processors and mass storage devices. Digital image processing is apprehensive primarily with extracting useful information from images. Ideally, this is done by computers, with little or no human involvement. Image processing algorithms may be placed at three levels. At the lowest level are those techniques which deal directly with the raw, possibly noisy pixel values, with denoising and edge detection being good examples. In the middle, algorithms which make use of low level results for further means, such as segmentation and edge linking. At the highest level are those methods which attempt to extract semantic meaning from the information provided by the lower levels, for example, handwriting recognition. A segmentation could be used for object recognition, occlusion boundary estimation within motion or stereo systems, image compression, image editing, or image database look-up.

Image segmentation is a very challenging task. Huge amount of study activity is going on image segmentation. There are many different segmentation techniques based on different tools. Method based on graph theory and in particular on the formulation of the segmentation problem in terms of a graph-cut problem[2]. Another group of methods are based on clustering algorithms, e.g. the method of exploits the mean shift clustering algorithm for image segmentation.[1] There are also methods based on region merging, level sets, watershed transforms, spectral methods and many other different techniques. Despite huge amount of time is washed-out on this issue, it remains unsolved. One of the main reasons is that the information content of the image is not always sufficient to properly recognize the object in the framed scenes. Human eye can differentiate object and background with same color of scene, but same thing computer cannot do easily. To do this, geometry information of scene is necessary. Many approaches are proposed to obtain geometry information from real world scene, each one with pros and cons. Now a days market offers more expensive and precise active methods such as structured light systems and laser scanners. Also time-of-flight range cameras(e.g., Mesa Imaging SR4K)[4] and structured-light cameras(e.g., Microsoft Kinect)[5] are gaining popularity. Unstructured scene reconstruction tools like Microsoft Photosynth[6] can also provide the geometrical representation of a scene from a collection of pictures taken from random positions. Stereo vision systems provide estimates of the 3D geometry of a framed scene from two or more views of it. In this field there has been sizable amount of research and current methods are able to give dense and reliable depth information.

The main objective of this paper is implementation and accuracy evaluation of different segmentation techniques. There is no general segmentation technique for every scenes. Each segmentation techniques has is own advantages and disadvantages. In order to achieve accurate result of segmentation of scene depth map and color information must be correct.

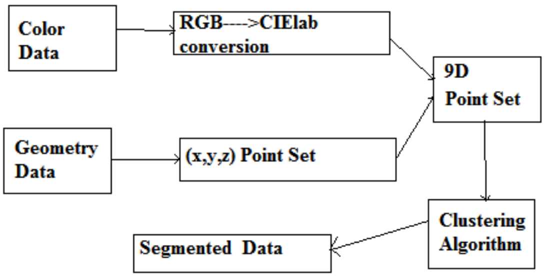


Figure 1. Architecture of the proposed segmentation scheme.

2. Representation of Geometry Information

Depth estimation refers to the set of techniques and algorithms aiming to achieve a representation of the spatial formation of a scene. In other expressions, to obtain a measure of the distance of, each point of the seen scene, geometry information is required. Depth information for image may be acquired in several ways. Many different approaches exist for the estimation of the 3D scene geometry. They can be roughly divided into active methods, that project some form of light over the scene, like la ser scanners, structured light systems (including the recently released Microsoft's Kinect) or Time-Of-Flight cameras. Passive methods instead do not use any form of projection and usually rely only on a set of pictures framing the scene. In this class binocular stereo vision systems are the most common approach due to the simplicity of the setup and to the low cost. The choice of the best suitable system depends upon the trade-off among the system cost, speed and required accuracy[]. There are stereo matching algorithms for disparity depth map, which gives geometry information.

Understanding of deep information in a scene, is great significance in the field of machine vision, robotics and image analysis. In this Chapter, an explicit analysis of the existing stereo matching methods, is presented. The presented algorithms are discussed in terms of speed, accuracy, coverage, time consumption, and disparity range. Towards the direction of real-time operation, the development of stereo matching algorithms, suitable for efficient hardware implementation is highly desirable.
Stereo matching by computing correlation is a basic technique for obtaining dense depth map from images. Correlation based matching typically produces dense depth maps by calculating the disparity at each pixel within a neighborhood. The depth estimation world is a quite complex research field, where many techniques and setups have been proposed. This is achieved by taking a square window of certain size around the pixel of interest in the reference image and finding the homologous pixel within the window in the target image, while moving along the corresponding scan line. The goal is to find the corresponding (correlated) pixel within a certain disparity range d (d E [0,….,d_{max}]) that minimizes the associated error and maximizes the similarity.
In brief, the matching process involves computation of the similarity measure for each disparity value, followed by an aggregation and optimization step. Since these steps consume a lot of processing power, there are significant speed-performance advantages to be had in optimizing the matching algorithm. The images can be matched by taking either left image as the reference (left-to-right matching, also known as direct matching) or right image as the reference (right-to-left matching, also known as reverse matching).

A. Sum of Absolute Differences(SAD)

It is one of the simplest of similarity measures which is calculated by subtracting pixel within square neighborhood between reference image Il and target image Ir followed by the aggregation of absolute differences within the square window, and optimization with the winner-take-all strategy. If both images exactly match, the result will be zero.

$$\sum_{(i,j)\in w} |(I_1(i,j) - I_2(x+i, y+j|))$$

B. Sum of squared Differences (SSD)

It is one of the algorithm in which differences are squared and aggregated within square window and optimization with winner-take-all strategy. Its computational complexity is higher than SAD algorithm.

$$\sum_{(i,j)\in w} (I_1(i,j) - I_2(x+i, y+j))^2$$

C. Normalized Cross Correlation (NCC)

It is more complex than SAD and SSD algorithms. Due to its complexity execution time is higher than SAD and SSD algorithms.

$$\frac{\sum_{(i,j)\in w}(I_1(i,j).I_2(x+i,y+j))}{\sqrt[2]{\left(\left(\sum_{(i,j)\subset w} I_1{}^2(i,j)\right).\left(\sum_{(i,j)\subset w} I_2{}^2(i+x,j+y)\right)\right)}}$$

D. Proposed Algorithm For Geometry information

Fusion of edge detection based method and sum of squared difference gives higher accuracy than above stereo matching algorithms.

- In this algorithm Edge of right and left side view image is calculated.
- According to disparity sum of absolute difference is calculated

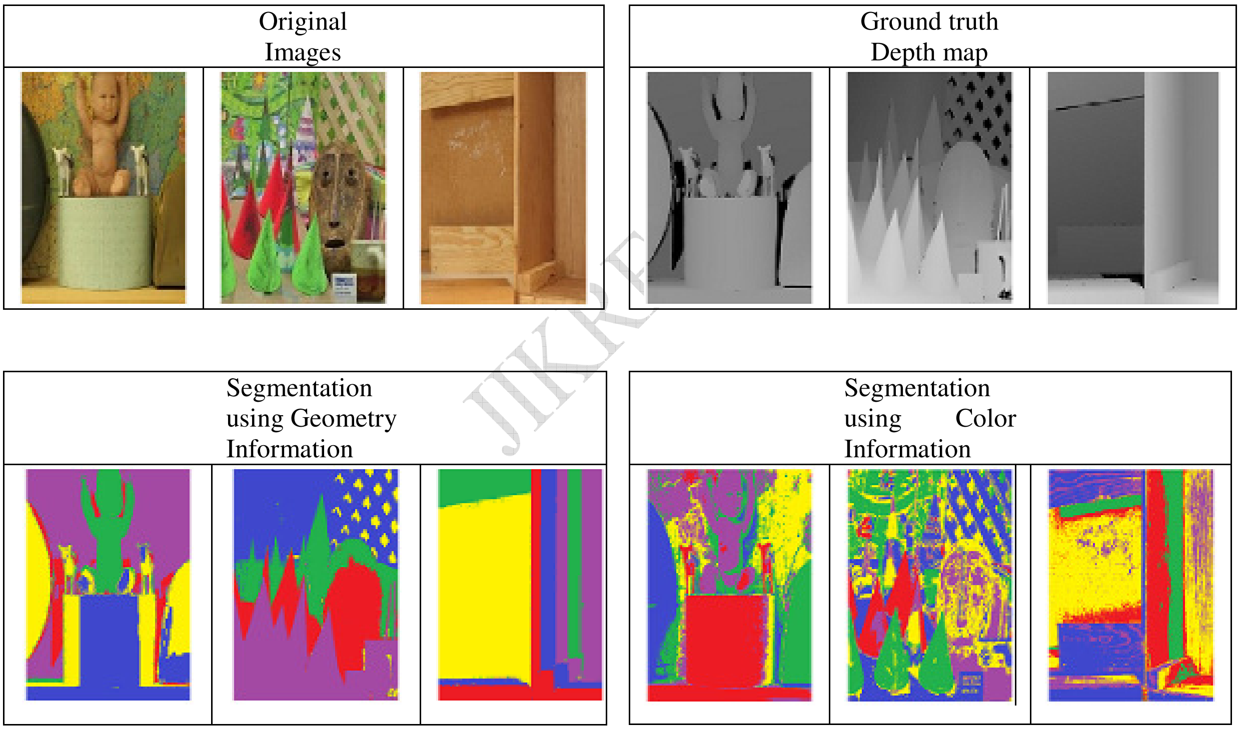
- For each pixel Find minimum difference and assign its disparity level to that pixel and make disparity depth map.
- Then smoothing of disparity depth map is done using filter process.
- Then fusion of disparity depth map using SSD algorithm and Edge detection based Algorithm.

4. Representation of Color Information

Color data require a 3-D vector, in order to account for the R, G, and B color components and another 3-D vector is required for geometry information in order to describe coordinates of a point with respect to a given reference system (such a reference system can be obtained from the calibration data and the depth-maps produced by many acquisition systems). First of all, geometry and color information need to be unified in a meaningful way. They choose to represent the color values in a perceptually uniform space in order to give a perceptual significance to the Euclidean distance between colors. This helps keeping consistent with the perceived color difference the distances used in the clustering process. Uniform color space ensures that the distances in each of the three color components are comparable, thus simplifying the clustering of the 3-D vector associated to color information. The CIELab space was selected for color representation.

After Collecting geometry information and color information, segmentation is required. For segmentation ,there are number of algorithms (i.e- K-means clustering, mean shift clustering, spectral clustering etc.)

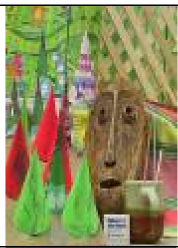
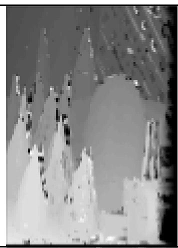
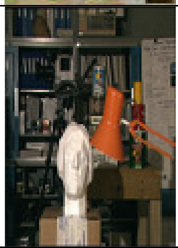
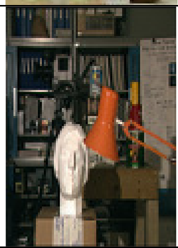
5. Scene Segmentation based on color information and geometry information



From above results we can conclude that different objects with same color can not be separated using segmentation based on color information ,same as Objects with different color but same depth can not be separated using segmentation based on geometry information. So fusion of both geometry and color information is required.

6.Fusion of Geometry and color information for scene segmentation

Here are some results of Scene segmentation using geometry and color information using different clustering techniques.

	Left Image	Right Image	Depth Map using proposed algorithm	Scene Segmentation using Fuzzy C-means clustering	Scene Segmentation using K-means clustering
1 Cones					
2 Lamp					
3 Venus					

7. Conclusion

In this paper, we proposed a new approach for geometry information using edge delectation based algorithm and Sum of squared difference algorithm. K- means clustering algorithm gives fast output as compare to fuzzy C-means clustering. Among stereo matching algorithms, SSD algorithm is more complex then SAD algorithm. SSD algorithm gives better disparity map. Time taken by SSD algorithm is less than other algorithms.

8. References

[1] D. Comaniciu and P. Meer. Mean shift: a robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* 24(5):603 –619, may. 2002.

[2] Pedro F. Felzenszwalb and Daniel P. Huttenlocher. Efficient graph-based image segmentation. *Int. J. Comput. Vision,* 59(2):167–181, 2004.

[3] Carlo Dal Mutto, Pietro Zanuttigh and Guido M. Cortelazzo. Fusion of Geometry and Color Information for Scene Segmentation*, IEEE journal of selected topics in signal processing, vol. 6, no. 5,September 2012*

[4] Mesa Imaging, [Online]. Available: http://www.mesa-imaging.ch

[5] Microsft Kinect, [Online]. Available: http://www.xbox.com/en-US/ kinect

[6] Microsoft Photosynth, [Online]. Available: http://photosynth.net/

[7] Barıs Baykant ALAGÖZ. Obtaining Depth Maps From Color Images By Region Based
8Stereo Matching Algorithms, *OncuBilim Algorithm And Systems Labs.* Vol.08, Art.No:04,(2008)

[8] Pablo Revuelta Sanz, Belén Ruiz Mezcua and José M. Sánchez Pena, Depth Estimation, Available: http://dx.doi.org/10.5772/45904

[9] Middleburey, http://vision.middlebury.edu/stereo