International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

3D Imaging using Graph Cuts

Swati Arora[1], Navpreet Singh[2]

[1]Global Institute of Management and Emerging Technologies, Punjab Technical University, India

[2]Global Institute of Management and Emerging Technologies, Punjab Technical University, India

Abstract: ***3D imaging has become one of the peak development segments in the field of biomedical image analysis, computer vision, radar imaging, scene reconstruction, human computer interaction systems and robotics. Different layers of image can be known through disparity maps as closer objects will have elevated disparities whereas objects at background will have comparatively poorer disparities. Based upon disparity map, diverse techniques can be used for extrication of foreground from background and a correct 3D perception can be created. The approach taken here is based on graph cut techniques.***

Keywords: Disparity map, stereoscopic image, s-t minimum cut graph, median filter, max flow**.**

1. Introduction

Creating three dimensional images from two dimensional data with proper brightness and contrast is a hard limitation in several areas like computer vision, biomedical image analysis, radar imaging, scene reconstruction, human computer interaction systems and robotics [7]. 3D adds the third dimension of depth, which can be perceived by human vision. Generation of 3D scenes from 2D contents is an alternative and attractive solution that overcomes the current discrepancy and fills up the lack of 3D image contents. The most common method is to obtain the corresponding depth information from the existing 2D image using which the 3D image is generated based on human visual principles. The key point of the conversion algorithm is how to get more accurate depth information of a 2D image in a fast manner and how to optimize the algorithm that generates the depth map. A 3D display takes profit of this phenomenon, creating two slightly different images of scenes and then resenting them to the individual eyes. By an appropriate disparity and calibration of parameters, a correct 3D perception can be realized.

The human vision system is a natural made perfect system of 3D with two eyes is apart in a fixed distance. Stereo vision therefore tries to replicate the ability of the human brain to conclude depth from a scene and consequently uses the same principle. Stereo pair is set of images taken from different viewpoints (cameras). Figure 1 shows a set of stereo images.

Figure 1: Pair of stereo images

Stereoscopy develops a false impression of three dimensional depths from given two dimensional images. The 3D information can be obtained from a pair of images, also acknowledged as a stereo pair, by estimating the relative depth of points in the scene. These estimates are represented in a stereo disparity map, which is constructed by matching corresponding points in the stereo pair [6].

A disparity map is a 2D function that gives the depth (with respect to the viewpoint) of an object point as a function of the image coordinates. Usually, it is represented as a gray level image with the intensity of each pixel registering its depth.. Fig.2 shows the depth map of a 2D stereo pair. The length of the horizontal displacement vector is generally referred to as disparity, and a pixel's disparity is inversely proportional to the pixel's distance from the cameras. By means of this principle, the human brain converts the disparity information into a three-dimensional impression of the world.

Figure 2: Disparity map of a stereo pair.

Even though the process seems to be very simple and the stereo task is undyingly solved by the human visual system without us even noticing the effort, the same problem turns out to become very tricky when it needs to be solved by a computer.

The stereo matching methods can be generally divided into two categories, local and global ones. There are several local correspondence algorithms; we can compare these algorithms in terms of both performance and efficiency. The global algorithms first make explicit smoothness assumptions of the disparity map and then calculate globally optimized matching by minimizing energy function. Graph Cuts algorithm have attracted much attention due to their good performance. In this global methods are carried out with Mean-Shift color segmentation algorithm.

2. Related Work

There is a significant amount of literature on the segmentation based conversion to 3D. We therefore focus our summary on a few techniques that we consider as important.

Viral H. Borisagar et al. (2011) proposed a new algorithm [1] and evaluate disparity map results for Mean shift, Hill climbing, Otsu and Graph-based color segmentation techniques. He also presents a novel segment-based stereo matching algorithm for disparity map generation which is computationally inexpensive. Structured lighting [10] offers another method for depth reconstruction. There are some algorithms that can perform depth reconstruction from single images in very specific settings. Daniel Scharstein et al. [2] presented taxonomy of dense, two-frame stereo methods designed to assess the different components and design decisions made in individual stereo algorithms. This algorithm particularly emphasis on stereo methods that (1) operate on two frames under known camera geometry, and (2) produce a dense disparity map, i.e., a disparity estimate at each pixel. Any vision algorithm, explicitly or implicitly, makes assumptions about the physical world and the image formation process. Common assumptions are Lambertian surfaces, i.e., surfaces whose appearance does not vary with viewpoint. Some algorithms also model specific kinds of camera noise, or differences in gain or bias. In this some limitation related to local and global methods have been discussed. Michels et al. [11] used supervised learning to estimate 1D distance to obstacles, for the application of autonomously driving a small car.

Yichen WEI et al. [3] proposed region based progressive stereo matching algorithm in which reliable regions are firstly identified and matched using GCPs. Remaining regions are matched progressively in a growing-like process using a global best first strategy based on a cost function that integrates disparity smoothness and visibility constraints and an ambiguity measure that is defined to be the ratio of the best and second best costs. Dong Yang et al. [4] described a new segment-based dense stereo matching algorithm. Firstly, the reference view and matching view are over segmented using mean-shift segmentation method. A new region-based approach [4] is proposed to obtain the initial disparity maps of the two views. Then, the unreliable matching points are filtered out by left-right consistency checking technique. An improved greedy search algorithm [4] is applied to propagate the reliable disparity to the segments which don't have reliable disparity. Finally, the disparity map in coarse regions is refined. Skbastien Roy et al. [5] describes a new algorithm for solving the IC'-camera stereo correspondence problem by transforming it into a maximum-flow problem. Once solved, the minimum-cut associated to the maximum-flow yields a disparity surface for the whole image at once. This global approach to stereo analysis provides .a more accurate and coherent depth map than the traditional line-by-line stereo. Moreover, the optimality of the depth surface is guaranteed and can be shown to be a generalization of the dynamic programming approach that is widely used in standard stereo. Results show improved depth estimation as well as better handling of depth discontinuities.

3. Proposed Work

The intention is to execute tool to practice Foreground object segmentation(i.e. to separate foreground from background) from stereo pair images using s-t minimum graph cut algorithm. Michels et al. [12] used supervised learning to estimate 1D distance to obstacles, for the application of autonomously driving a small car. The results obtained from graph cut method will be analyzed. The following hierarchy shows the proposed procedure for implementing the conversion. First of all the stereo pair is loaded and correspondence of left and right image is obtained. According to the correspondence obtained, the disparity map is formed. Median filtering is applied to even the disparity map. Foreground is separated from the background using the disparity map and 3D illusion is formed. Figure 3 shows the hierarchy of the proposed algorithm.

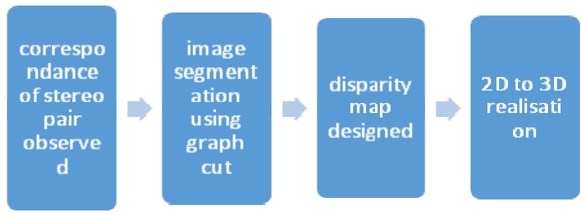

Figure 3: Proposed algorithm of the work done

The approach followed is a hybrid approach based on minimum graph cuts.

To enforce two-dimensional smoothness, those approaches convert the stereo correspondence task into a maximum flow/minimum cut problem. The maximum flow problem is explained in figure 4.

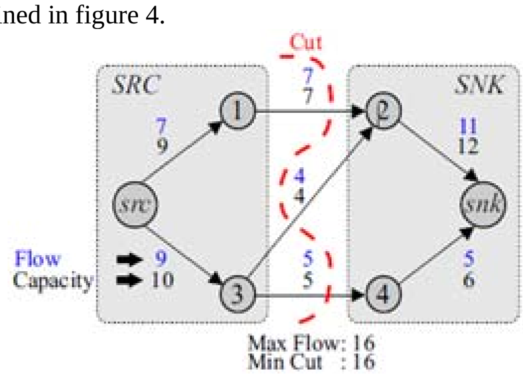

Figure 4: Maximum flow/minimum cut approach

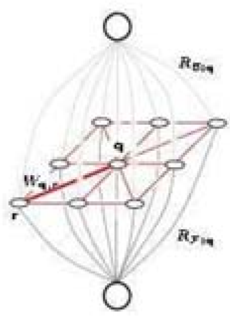

Figure 5: Graph representing a 3 by 3 image

A graph cut is process of partitioning a directed or undirected graph into disjoint sets. The concept of optimality of such cuts is usually introduced by associating energy to each cut. Graph cuts have been successfully applied to stereo, image restoration and texture synthesis and image segmentation. General approach to construct an undirected graph from an image is shown in figure 5.

Basically each node is viewed as node in a graph, edges are formed between nodes and weights corresponding to how alike two pixels are. In order to reduce the number of edges in a graph, only pixels with smaller and predetermined neighborhood N of each other are considered.

The proposed algorithm consists of three key parts, (i) Motion/edge detection and image segmentation (ii) Depth estimation and (iii) Shift algorithm

3.1 Image Segmentation

Segmentation based matching algorithm is used that divide one or sometimes both images into non-overlapping regions of homogeneous colour. Instead of computing a disparity for each individual pixel, those techniques assign a single disparity value (or model) to a complete segment.

3.2 Color Segmentation

In principle, any algorithm that divides the reference image into regions of homogeneous color can be used for the proposed stereo algorithm. The current implementation uses a mean-shift-based segmentation algorithm that incorporates edge information. [8]Pixels belonging to the same segment are assigned the same color. To derive the desired plane models, we first compute an initial disparity map and use the computed disparity values to fit the plane for each segment.

3.3 Median Filtering

Median filtering is a nonlinear operation often used in image processing to reduce "salt and pepper" noise. A median filter is more effective than convolution when the goal is to simultaneously reduce noise and preserve edges. In median filtering, a window slides along the image and median intensity of pixels within the window becomes output intensity of pixel being processed. Figure 4 shows the result of Median filtering on an image.

Figure 6: Median filtering of an image

3.4 Calculate Disparity Map

We compute an initial disparity map using a local window-based method that exploits the results of the image segmentation and operates on different window sizes. We benefit from the image segmentation by exploiting the assumption of smoothly varying disparities inside a segment. Stereo correspondence or disparity is conventionally determined based on matching windows of pixels. Every time disparity have some value, if disparity is zero it means a depth is zero because disparity is inversely propositional to depth.

4. Experimental Setup and Results

For implementation MATLAB, a high-performance language for technical computing is used. First, the reference image is segmented using a technique called Mean Shift Segmentation. This is a clustering algorithm that "over-segments" the image. The result is a very 'blocky' version of the original image. Then, for each segment, we look at the associated pixel disparities. In this implementation, we assign each segment to have the median disparity of all the pixels within that segment.

As bright colors represent closer objects (foreground), a range of bright colors is selected for foreground. I chose red color so i computed red objects, all objects which are not green, all objects which are not blue as shown in figure 7. Smaller objects which are less than 100 pixels are removed. For computed red mask, combine red, green and blue components to obtain foreground objects.

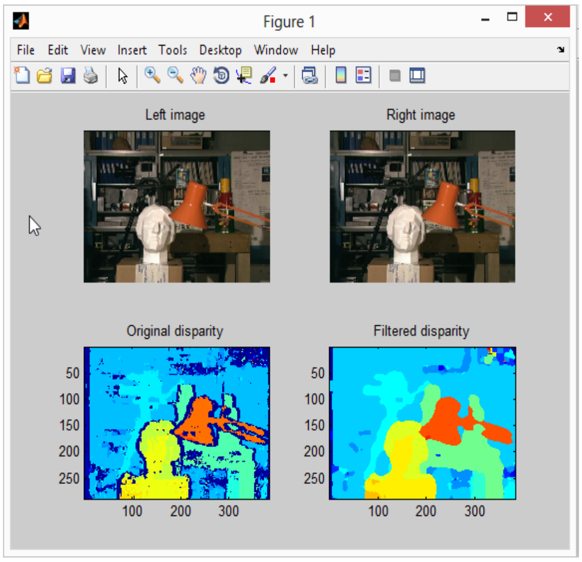

Figure 7: Original and filtered disparity maps of stereo pair images.

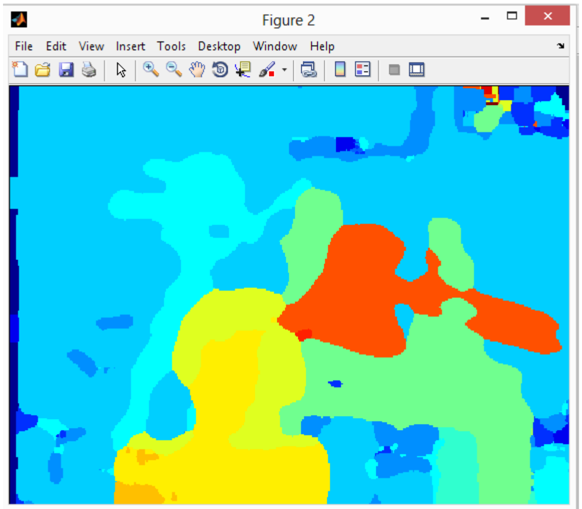

Figure 8: Colored disparity segmented image

Figure 9 shows the total time the algorithm is taking to convert the 2D image into coloured segmented disparity.

Profile Summary

Generated 09-Sep-2014 17:09:27 using cpu time.

Function Name	Calls	Total Time	Self Time*	Total Time Plot (dark band = self time)
FirstClass	1	7.161 s	0.099 s	
ColorDetection	1	2.776 s	0.489 s	
stereo	1	1.712 s	0.013 s	
stereo>slide_images	2	1.617 s	0.571 s	
saveas	1	1.484 s	0.032 s	
print	1	1.371 s	0.550 s	
imshow	18	1.172 s	0.109 s	
strel>MakeDiskStrel	1	0.865 s	0.339 s	
print>LocalPrint	1	0.633 s	0.010 s	
newplot	48	0.577 s	0.022 s	
newplot>ObserveAxesNextPlot	48	0.554 s	0.003 s	
cla	43	0.551 s	0.003 s	
graphics\private\clo	43	0.548 s	0.111 s	
strel>strel.strel	10	0.504 s	0.003 s	
stereo>shift_image	161	0.489 s	0.489 s	

Figure 9: Tabular form of total time consumed by the process

5. Conclusion and Future Scope

The approach is hybrid approach that aims at minimum time consumption that most of the existing approaches are not able to fabricate and is supposed to give a clearer 3D view of 2D stereo pair of images. It describes how prior information can be brought into a graph cut framework through the use of terminal node weights and learning techniques. Also the noise content is tried to remove. In near future by improvising the stereovision system we can implement it for real time applications such Face Recognition, 3D Face Reconstruction and 3D Terrain Mapping.

References

[1] Viral H. Borisagar and Mukesh A. Zaveri "A Novel Segment-based Stereo Matching Algorithm for Disparity Map Generation" International Conference on Computer and Software Modeling IPCSIT 2011.

[2] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms", International Journal of Computer Vision, vol. 47, no. 1-3, pp. 7–42, 2002.

[3] Yichen WEI, Long QUAN, "Region-based progressive stereo matching", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04), 2004.

[4] Dong Yang, Zhengguang Xu Lixin Zhang, "A Segment-Based Dense Stereo Matching Algorithm," in Second International Symposium on Information Science and Engineering, pp. 574-578.

[5] Skbastien Roy and Ingemar J. Cox, "A Maximum-Flow Formulation of the N-camera Stereo Correspondence Problem", Sixth International Conference on Computer Vision, 1998.

[6] M. Z. Brown and D. Burschka and G. D. Hager, "Advances in Computational Stereo", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003.

[7] 2D TO 3D CONVERSION FOR IMAGES USING HOUGH TRANSFORM-Sneha. K1 T. Sheela.

[8] Carlo Dal Mutto and Pietro Zanuttigh and Guido M. Cortelazzo, "Scene Segmentation Assisted by Stereo Vision" International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission, 2011.

[9] Manaf A. Mahammed "Object Distance Measurement by Stereo VISION" International Journal of Science and Applied Information Technology (IJSAIT)Vol.2 , No.2, Pages : 05-08(2013) Special Issue of ICCTE 2013 -Held during 11-12March, 2013 in Hotel Crowne Plaza, Dubai.

[10] Scharstein, D. and Szeliski, R. (2003) High-accuracy stereo depth maps using structured light, Computer vision and pattern recognition (CVPR), Vol.1, Pp.195-202.

[11] Michels, J., Saxena, A. and Ng, A. Y. (2005) High speed obstacle avoidance using monocular vision and reinforcement learning, 22nd international conference on machine learning (ICML), Pp. 593-600.

[12] Yuri Boykov, Olga Veksler, and Ramin Zabih. Fast approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23:1222–1238, 2001.

Author Profile

Swati Arora is currently pursuing her M.Tech. in Electronics and Communication from Global Institute of Management and Emerging Technologies affiliated to Punjab Technical University. She received her B.Tech. degree in Electronics and Communication from IET Bhaddal in 2011. Her area of interest is Image Processing.