

# Review on 3d Reconstruction Using Multi View Stereo

<sup>1</sup>Anupriya K, <sup>2</sup>Rinju Ravindran, <sup>3</sup>Swathi Ravindran, <sup>4</sup>Jerrin Yomas

MTech Scholar, <sup>4</sup>Asst. Professor, ECE

<sup>1, 2, 3, 4</sup>Vimal Jyothi Engineering, College, Chemperi,  
Kannur, Kerala, India

## Abstract-

**3** D reconstruction is one of the areas in which innovations are to be brought out for every single element. The method of reconstruction includes mainly 3 steps. Collection of a sequence of images of the object using a single camera and finding out the relation between or checking the match between the same points on the object in two different images using triangulation method and cost function are the first two steps. After finding the relation in mapping, at the final step the images combined together to generate a 3D model of the object on the display. This paper reviews about various reconstruction algorithms as well as the fundamental properties of multi view stereo. Properties include mainly the pixel characteristics and image quality. And the reconstruction methods mainly deal with the classification as active and passive methods.

**Key words-Reconstruction, Triangulation, Cost function, Multi view stereo, Active and Passive methods.**

## I. INTRODUCTION

Computer graphics and animation is a developing area in the current scenario it have got a wide popularity. Since 3D images are used for animation a large no .of reconstruction methods are used now a days. 3D reconstruction mainly deals with the generation of a character or object in the display screen by using a large set of 2D images. These images are obtained by taking photos of the object from variety of known camera view points. So many algorithms were developed for the 3D reconstruction of which each was concentrated in overcoming the difficulties of the previous algorithm.

In earlier times the image acquisition was done using laser scanners where the object is divided into small points. Using the algorithm each point is reconstructed using the same algorithm and at the final stage of the algorithm all the different elements are combined together to get the single desired object. Even though the laser reconstruction was efficient the nature of laser scanners limits its usefulness in the outdoor reconstructions. Thus further researches were done for developing a better algorithm. So by later times another method of sequence acquisition is used where an RGB camera is placed on a rotating table. That single camera is used to take the images of the object and can be used to analyze the characteristics of the object from different angles of projections. Image acquisition by the use of camera is more efficient since we can adjust the quality and other characteristics of the camera as per the object and time.

## II. RELATED WORKS

At the beginning of the studies of 3D reconstruction methods mainly there were four classes of algorithms. First category includes the formation of cost function of 3D volume which will give the correct characteristic of object. Voxel colouring algorithm and graph cut algorithm come under this section. The second type is variational method in which iterative evolution of surface is done in order to reduce error probability. Third category is related to the multiple depth maps ensuring the consistent 3D scene. Combining all the three categories the forth class is called a hybrid algorithm, which is the efficient reconstruction algorithm.

The researcher Sylvain Paris developed a reconstruction algorithm in which images are collected and each pixel is mapped to a ray in space. It used the method of triangulation where a feature is mapped in to at least 2 views. For this method similar points are identified and a scalar function is generated. If the value of the function is low it is a good match and if it's a high value it shows a poor match.

At a later time Patrick Labatut and Jean Philippe developed a new algorithm where in the image collection is done without any conditions. It generates a quasi dense 3D point cloud of object by the method of triangulation. Using this Nader Salman and Mariette Yvinec introduced the depth map representation that is used for large scale outdoor scene reconstructions. And thus it's now used for the reconstruction of current states of art.

## III. FUNDAMENTAL PROPERTIES

The scene representation is an important property which is used to distinguish between different reconstructions algorithms. There are various ways to represent geometry of an object or a scene. It includes voxels, level-sets, polygon meshes, or depth maps. There can be single representation or variety of representations that can be used in the reconstruction methods. Voxels is the method of representing the geometry on a 3D volume or grid as a discrete occupancy function. Level sets are created by describing the geometry on a 3D volume as a function encoding distance to the closest surface. 3D grids are simple and uniform way of approximating any surface. Another efficient method of

representation is by the use of polygon meshes. It makes use of connected plane facets to represent a surface. Depth is the 3D component that is missing from the given image. The calculation of this depth is actually a difficult and challenging process.

Photo-consistency measure is another fundamental property that describes a reconstruction algorithm. The similarity between images can be studied by comparing pixels of one image with the pixels of other image. This method helps in acquiring the correlation between input images and thus provides visual compatibility of the reconstruction methods. There are 2 classes of photo consistency measures: one which uses scene space and other includes image space. In case of scene space, we are interested in considering pixels or geometries which are projected towards the input images. By doing this we find the correlation of the projected parts and the input image. Another approach takes into account two images at a time and verifies the cross correlation between the two images. In the case of image space, we are predicting a different view using the available scene geometry. This creates a prediction error between the measured and predicted images. Surface integration is carried out for finding scene space error whereas integration is done over the set of images of a scene to obtain the prediction error.

Next important property is visibility models. It is a factor which is studied for obtaining the proper photo-consistency. There are various methods for approaching the visibility models namely geometric, quasi-geometric, outlier based approaches, etc. The shape of the scene and the image formation process are explicitly modelled using the geometric approach. It determines the visibility of scene structures in the images. Appropriate geometric reasoning is carried out for performing quasi geometric approach also. In order to get an accurate relationship between neighbouring pixels, a rough estimate of the required surface is obtained. But in the third category of outliers, the explicit geometric reasoning is avoided.

Shape priors are used to shape the reconstruction algorithm with our desired characteristics. According to the technique used for reconstruction, we prefer either minimal surfaces with small overall area or maximal surfaces. These shape priors can be also used to smoothen the surfaces.

It is very important to choose the proper reconstruction algorithm. The two main categories of 3D object reconstruction are active and passive methods. If active triangulation methods are undertaken to obtain the 3D data of an object, we call it as active reconstruction method. It make use of mechanical or radiometric steps to interfere with the reconstructed object. These processes are highly expensive and needs the help of efficient and skilled persons. The lack of scanners with proper recording devices is also another problem in this regard. The passive devices does not require such devices and hence is simple for evaluation purposes. It does not interfere with the reconstruction objects. Here the 3D structure is obtained by the use of sensors that evaluates the radiance reflected from the surface of the object. In this, a digital RGB camera is used to capture the images at required viewpoints. These 2D images are collected and the corresponding 3D image is extracted as the output. It is cost effective compared to active methods.

In addition to the reconstruction algorithms, there are various initialization requirements for all methods. It includes the use of bounding box or volume, the segmentation methods, etc.

#### **IV. RECONSTRUCTION METHODS**

Multi-view stereo algorithms take the information on scene that has to be reconstructed and the geometric extend of the object as the input. For the initial estimation of the scenes some algorithms require background as well as foreground information so a visual hull can be created. While some of the algorithms require only space carving variants. For evaluating the efficiency of the algorithm many set of multi-view images are required.

##### **A. Reconstruction Algorithm**

The steps that are performed for obtaining the 3D view of an object depend upon the reconstruction method being used. But the reconstruction of a 3 D scene from several calibrated images has mainly three steps.

##### **1) Merging, Filtering and smoothing the track**

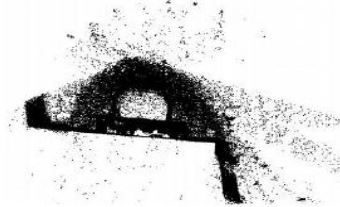
To the algorithm provided by stereo vision algorithm, the figure 1 is taken as the input. Set of eleven images were taken from different viewpoints and the results were taken from it. The input track will store the 3-D point location as well as the camera location from which the image was captured .There should be a line of sight between the 3-D position of the track captured by camera and the camera location. The extracted tracks are then Pre-processed by merging, filtering and smoothing.

The first pre-processing step is merging those tracks which are close in their position. 3-D position of track is considered and its incremental Delaunay triangulation is built. The 3-D position of each track has to be added and nearest neighbor to this 3-D track is requested .So based on the distance between these two 3-D tracks, one of the two cases will arise. If the two 3-D tracks are close then they are merged to a single one and the list of cameras will be updated with the union of the cameras of the two tracks being merged. Otherwise both the tracks are preserved in the final set of tracks.

Tracks are characteristic points from the image. The 3-D location of the tracks should spread over the objects. Thus the tracks are Isolated and are said to be outliers. These outliers are filtered out based on the criteria called line of sight. So the 3-D locations form small angles and the imprecise intersection computation will make the 3-D points outliers .Outliers can be detected based on two criteria Distance to neighbours, which eliminates the far away tracks from highly populated regions and the cone angle, to eliminate tracks that can only be observed from few camera locations. Set of images taken from video sequence can also be filtered using this criterion.

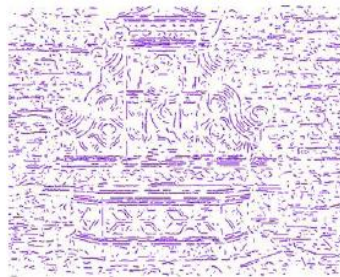


(a)



(b)

Fig 1. (a) one image out of eleven; (b) top view of the corresponding point cloud



(c)

Fig 2. Both the image and the point cloud are used to detect the constraints.

Smoothing is done on the remaining tracks which are the final pre-processing step. The algorithm used is smoothing algorithm based upon jet-fitting and re-projection

## 2) *Triangle soup extraction*

The pre-processed tracks and the input image are combined to form an image plane triangulation depth map. To form a triangle soup, where on the actual surface of the scene will the triangles of the image lie, the triangles of each depth map is lifted into 3-D space that is further filtered so that the visibility violations and photo-consistency constraints can be minimized. The triangulation soup may contain big triangles, i.e. the triangles whose circumradii are larger than the scene. Such big triangles can be filtered using photo-consistency criterion. Photo-consistency is determined based on two criteria. They are the Sum of the Squared Difference (SSD) and the Normalized Cross Correlation (NCC) between the pixels inside the triangle that is projected in the image. NCC is less sensitive to the variation in the illumination between the views. Visibility criteria is that the line of sight between the camera location and the 3-D position of the track seen using the camera does not intersect any object in the scene. In 2-D reconstruction of the image the straight and curved edge in the image is important keys. These information has to be included in the depth maps. Edge corresponds to a contour in 2-D image and is an occluding boundary in 3-D image. Edge detection can also be done using gradient based method using Canny algorithm. The properties of this algorithm are with aliasing high positional accuracy can be achieved, sensitive to high frequency data and data can be reduced. Only those images having the angle between the normal to the 3-D triangle and the line of sight through the centre of the image is small is mainly used.

## 3) *Reconstruction*

Delaunay refinement surface mesh generation algorithm is used in the final step. Algorithm need to know the surface that has to be meshed. Given a line segment, the intersection of the line segment with the surface is checked. If it intersects then it returns the intersection point. Oracle required for meshing algorithm is implemented by using triangle as an approximation of the surface. The intersection with the triangle is computed by the oracle and the algorithm queries are answered. Reconstruction steps are applied on the extracted triangle soup and the Delaunay refinement parameters are altered.

## V. CONCLUSION

In this paper, we have described the steps in obtaining the 3D view of real objects from a sequence of images. This paper presented the fundamental properties of multi-view stereo algorithms. This work explains the importance of multiple views for 3D reconstruction. Paper also describes the algorithms used at the beginning for the study of 3-D reconstruction.

**REFERENCES**

- [1] Steven M. Seitz, Brian Curless, James Diebel, Daniel Scharstein, Richard Szeliski: *“A Comparison and Evaluation of multi view stereo reconstruction algorithms”*.
- [2] Nader Salman, Mariette Yvinec :*“Surface Reconstruction from Multi-View Stereo of Large-Scale Outdoor Scenes”*, The International Journal of Virtual Reality, 2010, 5(3):1-6
- [3] Carlos Hernandez Esteban, Francis Schmitt: *“Multi-stereo 3D object reconstruction”*, Ecole Nationale Sup'erieure des Tele Communications Signal and Image Processing Department
- [4] Conglin Huang :*“3d Reconstruction Using Multiview Imaging System”*. University of Kentucky Master's Theses, 2009
- [5] Sylvain Paris :*“Methods for 3D Reconstruction from Multiple Images”*.,MIT CSAIL
- [6] Carlos Hernandez Esteban Francis Schmitt: *“Multi-Stereo 3D Object Reconstruction”*: Ecole Nationale Sup'erieure des T'el'ecommunications Signal and Image Processing Department 46, rue Barrault 75634 Paris Cedex 13