ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 12, December 2016

Reconstruction of 3D Realistic Models from Multiple Images – A Survey

Swarna Priya RM[1], S. L. Aarthy[2], R Sujatha[3], Srinivas Koppu[4], Dr. Madhusudhana Rao[5]

Assistant Professor (Senior), School of Information Technology and Engineering, VIT University, Vellore, Tamilnadu, India [1, 2, 3]

Assistant Professor (Selection Grade), School of Information Technology and Engineering, VIT University, Vellore, Tamilnadu, India[4]

Lecturer, Department of IT, Higher college of Technology, Muscat, Oman [5]

ABSTRACT: The task by which very fast and accurate 3D models of a scene are generated has variety of applications in the field of computer vision which includes robotics, virtual and augmented reality, and entertainment. So far, the computer vision scientific community has provided innovative reconstruction algorithms that exploit variant types of equipment. This paper is a survey of the recent methods that are able to generate complete and dense 3D models using various equipment are given. This also includes the methodology for generating 3D models from multiple images.
.

KEYWORDS:3D Modelling, Laser Scan, Image Reconstruction, Realistic Models.

I.INTRODUCTION

The fast innovative world requires many innovative ideas for visualizing and projecting the already available data in an attractive manner. For attracting the end user as well as for analyzing the data in a still deeper way, the three dimensional (3D) information is required by many applications and in many industries. The 3D information helps in making decision while analyzing the environment either static or dynamic. The various applications where the three dimensional information is much required are in surveillance, robotics, movie industries, medical domain, etc. In these applications, the third dimension whichever is visualized or derived from the two dimensional (2D) data helps in analyzing the depth information about the environment which is focused. Apart from these the 3D information plays a major role in 3D image processing, digital photography, games, multimedia, 3D visualization and augmented reality.

There are various methods existing for modelling 3D models form the available 2D data. These methods can be broadly classified as per the input data required, the details of the final model and the time required for processing them. Hence, based on the requirements of the user, the approach used can be decided which can be any of the automated, semi-automated or manual image based approaches. The automated methods are used when the requirement is to have a better visualization and attraction. Here the entire process is automated and is visualized in a simple manner and is widely used for real-time 3D recording. The semi-automated methods help to give a better accuracy and could be used in documentation applications.

This paper discusses the various methods available for 3D model construction from 2D images and which will give a realistic feel to the end user.

II. RELATED WORK

In the previous years, the reconstruction of three dimensional object modelling used the image matching technique. The method multi-view stereo is based on this image matching technique which recovers the three dimensional geometry by using the relationships between the images and then applying surface fitting and triangulation over the estimated cloud of points [18], [19], [20]. The major disadvantage of multi-view stereo is that at the process of finding the effective correspondence the image view has to be very close together and then the correspondences has to be found in the image.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 12, December 2016

The final partial 3D clouds of points must be then triangulated to get the fused model. Then the other drawback is that the occlusion between the various viewpoints are not addressed. To overcome these drawbacks, the model based multi-view stereo approaches were introduced [21]. These techniques measure the deviation between the actual scene of the image and the reconstructed 3D model. These overcome the correspondence issue as well as the occlusion problem. But these model-based approaches require an initial 3D model. To overcome the drawbacks of both these methods, 3D scene spaces are proposed [22], [23], [24]. These use the volumetric details of an object and hence they are known as scene modelling methods. These techniques eliminate the correspondence problem. The occlusion problem is also overcome by the use of voxels.

The popular rendering methods are plenoptic modelling [25], light field rendering [26] and the lumigraph [27] as they accept only images and then reconstruct the 3D scene directly from the images and do not use the 3D geometry.

III. RECONSTRUCTION USING LASER SCANNER TECHNOLOGY

Three dimensional models were reconstructed in the previous years using laser scanner technology. The major advantages of the earlier method are

1. High speed
2. Increased Accuracy
3. The resultant resolution of the model which is reconstructed.

The other advantages are

1. Can be used for a range of various size starting from centimetres to metres.
2. Can be used for large-sized field of view.

Due to these advantages, these methodologies are well suited for developing the 3D models for archaeological purposes.
The appearance for reconstructing a scene using a scanner is as given in the figure 1.

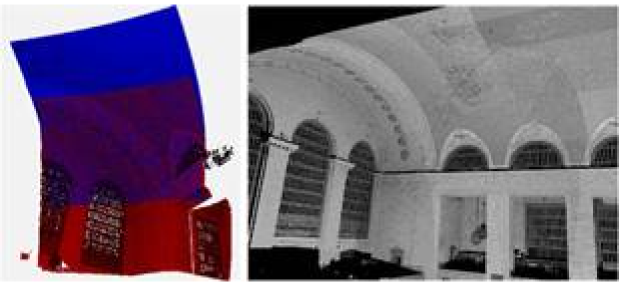

Figure 1. Reconstruction of a scene using Scanner

The steps include

1. Acquisition of scene:
   Here the scene which has to be reconstructed in 3D has to be captured. Hence a range scanner equipment is placed in an appropriate position to acquire number of views of the scene to cover the entire scenario.
2. Registration of the Scene:
   The scene which is acquired via the equipment would be in different alignments based on the projection line. Hence the image has to be converted to a common co-ordinate system which is done in this stage.
3. Refinement:
   In this stage initially noise is removed. Then the duplicacy in information is reduced or eliminated.
4. Construction:
   The 3D model is reconstructed using the points identified, features detected and by including the missed third dimension of the data.

The registration procedure is divided into two steps namely,

1. First level initial registration which gives the alignment details and helps in transformation.
2. Fine registration which further accurately aligns the image scene to the co-ordinate system.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 12, December 2016

Initial Registration

Various methodologies have been proposed by various researchers for initial level of registration of the scanned image. Most of the methods first extract the geometric features using some basic feature extraction algorithm and then proceed by matching them using a distance metric and then register the same to a co-ordinate system.
This section discusses about the various methodologies proposed and used by different authors. Stamos et al [1] designed a system for registering the scanned image scene by integrating the automated 3D to 3D registration method along with the geometric features. Chen and Stamos [2] proposed a method named 3D to 3D feature based registration where the features are extracted linearly and are registered with respect to one co-ordinate system. But these linear features were inadequate for accurate alignment of the scene. Hence, another author Jiang et al [3] proposed a methodology for identifying angular invariant feature which helped in a reliable selection of the points. This methodology was robust to scaling and rotation.

Later developments were towards identifying 3D objects for helping the registration process. Among these developments, Johnson and Hebert [4] method holds a significant position since this formed the base for free-form surface registrations. The major gap in this methodology was that the resultant features were ambiguous and was hence computationally inefficient since further filtering was required at various stages. Mian et al [5] proposed a scheme for 3D free form representation using third order tensors for registering the surface area. This algorithm was robust to changes in resolution and noise. Makadia et al [6] tried to further enhance the registration accuracy by using the Extended Gaussian Images for alignment. Fourier Alignment was then used by Kang and Ikeuchi [7].

When the scanned images vary by intensity, then the registration method used has to be including the intensity features also. Wyngaerd and Gool [8] developed a method where texture features were used and the feature points were identified using spherical spheres. This methodology is well suited for surfaces which have poor geometry and are rich in texture. Smith et al [9] presented a registration system which identified 2D SIFT features [10] and matched them. Then further these features were projected to 3D using RANSAC [11]. But these methods perform well only when the scanned scene is rich in texture and is useless if not.

Fine Registration

Fine registration is performed based on the estimation provided by the initial registration. This phase is used only for the purpose of optimizing the already estimated values and aligning still further. The most popularly used methodology for 3D data fine registration is the Iterative Closest Point (ICP) [12]. The other popularly used method is Maximum Likelihood Method which was proposed by Sharp et al [13] that uses the missing data and do not require the ICP refinement procedure.

Though the range scanners help in getting high accurate results, they are not cost effective. Their huge size, the high requirement of power and the way it handles data are the other major drawbacks of this system. This lead to the emergence of multi-view stereo methods.

III. MULTIVIEW STEREO METHODOLOGES

This section deals with the various multi-view stereo methods which are popularly used for 3D reconstruction is discussed. This provides a full 3D Model of the scene from multiple views of the image scene. These methods are broadly classified as

1. Methods reconstructing the Visual Hull of the object.
2. Methods for recovering the photo-hull of the object.
3. Algorithms for minimizing the integral part of the cost function for a surface shape.

Visual Hull Reconstruction

The very first methodology is based on Silhouette information for generation of cones which intersect and can be visualized. These cones can be used for representing the object in 3D. These methodologies are widely used since they are simple and efficient. Matsuyama [14] proposed a method using parallel pipeline process for reconstruction of 3D objects using the video of the scene. Franco and Boyer proposed a framework where the 3D reconstruction was performed by the fusion of multiple view of Silhouette information. An author Lai and Yilmaz eliminated the

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 12, December 2016

drawbacks of the dense features and errors due to camera calibration using an algorithm. He used the concept of projective geometry in the images which are taken in multiple views. The shape of the object was reconstructed using a hypothetical plane which slices the object volume and hence estimates the relation of the images with the projective geometry.

Ishikawa used segmentation for the reconstruction but this methodology lead to various segmentation errors which directly affected the visual hull. To overcome this error caused by segmentation, Grauman used Bayesian Approach. This approach could be used over the already proposed segmentation approach and could reduce the error which was created in the already generated 3D model. But the other drawbacks of this method is that the knowledge about the objects is to be known well in advance to get a better accuracy. Hence these kinds of approaches though help in generating 3D models from dynamic scenes, there are lots of errors occurring during the information extraction.

A modelling system for generating a 3D Model using the Silhouette information was designed by Franco et al [15] under an environment which is completely controlled. The system which was designed for acquiring images around the object and to identify the camera parameters with respect to these views had a camera which was fixed on a rotatory table which is shown in the figure 2.

Figure 2. System for Image Acquisition

Space Carving Reconstruction

This is the next level of methodology for reconstructing the 3D model which takes the photometric consistency from the input images and produce the photo-hull which forms the basis for reconstruction. The initial object detail which is required for further object reconstruction is constructed by the space carving reconstruction methods. Kutulakos and Seitz [16] proposed an algorithm where an arbitrary camera is placed and a sweeping plane was repeatedly used throughout the captured scene and the photo consistency of the voxels on a particular plane at a time is tested. But the major gap was that it was hard to avoid the removal of the voxels. That is, in some cases, the voxels are considered to be errors and are removed by the algorithm, but this leads to inconsistency in the 3D model which is reconstructed. To overcome this drawback, a space carving methodology based on probability was proposed by Bhotika et al [17]. But still there are stagnant drawbacks in the Space Carving Approaches which are as follows:

1. The initial Space Carving Algorithm which was proposed had disadvantage of decisions taken in removal of noise.
2. The threshold which was fixed as global was also an issue.
3. The voxel based date representation lead to erroneous model.

IV.3D RECONSTRUCTION FROM MULTIPLE IMAGES

The various challenges to be dealt in this type of reconstruction are as follows.

1. The images and photographs would be taken by different cameras and with different imaging conditions.
2. The ordering of the images to produce an accurate reconstruction also plays a major role.
3. Scaling of the parameters for various algorithms.

The steps involved in this is shown in the figure 3.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 12, December 2016

Figure 3. Reconstruction from Multiple Images

Sorting of Photos

The photos which are collected from various sources of a particular place or area or scene is usually unordered. The first and foremost step is to organize and order the images. The SIFT algorithm which is a feature detector and extractor could be used to compare the images and order them and this forms the input for the next level of the process.

Feature Detection and Matching

The steps in this approach are

- The feature points in each image out of the multiple images is determined using SIFT.
- The feature points which are identified as the key points are matched using the nearest neighbour for a pair of images.
- If the pair of images are matched and if there are lesser than 20 matching key points, then the pair is discarded.
- Instead of SIFT, the SURF algorithm could be used for detecting the features in each image and the Brute force technique for matching the features.
- The other variation is to use the optical flow technique for matching the features identified. This is a faster method and the number of matching points retrieved are higher compared to the other two approaches but it is assumed that the images are acquired by the same camera and with same illumination.

Structure from Motion

The approach uses the set of camera parameters and 3D location of each track for 3D reconstruction. Since the images are taken by different cameras, the parameters cannot be estimated at a stretch and hence an incremental approach is used. Initially, the parameters are estimated for a single pair of image. The parameters ae initialized for a new camera image using the Direct Linear Transform. From the parameters identified, the tracks are optimized. Then, the Iterative Closest Point Algorithm, which triangulate the points and create a 3D point is used.

Multiview Stereo

The Multiview Stereo Algorithms help to generate the 3D reconstruction of the object or the scene from the 3D points generated. A consistency function is used for testing whether the 3D point is consistent with the input image is used. The consistency function can be based on colour or texture of the image.

V.CONCLUSION

This paper presents an overview of the various algorithms and methodologies available for reconstruction of dense full models of 3D. The survey is broad as the methods which exploit the various types of equipment are discussed. However, except for the equipment-based categorization a further classification for each category is given based on the particular techniques that methods use to recover the 3D shape of objects.

REFERENCES

1. Stamos, I., Liu, L., Chao, C., Wolberg, G., Yu, G., Zokai, S., “Integrating Automated Range Registration with Multiview Geometry for the Photorealistic Modeling of Large-Scale Scenes”, IJCV, Springer, 78(2/3), 237–260, (2008).
2. Chen, C., Stamos, I., “Semi-automatic range to range registration: a feature-based method”, in the 5th international conference on 3-D digital imaging and modeling, 254-261, (2005).
3. Jiang, J., Cheng, J., Chen, X., “Registration for 3-D point cloud using angular invariant feature”, Neuro computing, Elsevier, 72, 3839–3844 (2009).
4. Johnson, A., Hebert, M., “Surface registration by matching oriented points,” In International Conference on Recent Advances in 3-D Imaging and Modelling, 121-128, (1997).

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 12, December 2016

5. Mian, A., Bennamoun, M., Owens, R., “A Novel Representation and Feature Matching Algorithm for Automatic Pairwise Registration of Range Images, IJCV, Springer, 66(1), 19–40, (2006).
6. Makadia, A., Patterson, A., Daniilidis, K., “Fully automatic registration of 3d point clouds, In CVPR, (2006).
7. Kang, S., Ikeuchi, K., “The complex EGI: A new representation for 3-D pose determination”, TPAMI, 15(7), 707-721, (1993).
8. Wyngaerd, V., Gool, V., “Combining texture and shape for automatic crude patch registration”, In Proc. Fourth Int. Conf. on 3D IM, pages, 179–186, (2003).
9. Smith, E., King, B., Stewart, C., Radke, R., “Registration of Combined Range Intensity Scans: Initialization through Verification”, Computer Vision and Image Understanding, 110, 226-244 (2008).
10. Lowe, D.G., “Distinctive Image Features from Scale-Invariant Keypoints”, International Journal of Computer Vision, 60, 91-110 (2004).
11. Fischler, M., Bolles, R., “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography”, Graphics and Image Processing, 24(6), 381-395, (1981).
12. Besl, P., McKay, N., “A method for registration of 3-d shapes”, IEEE Trans. Pattern Anal. Mach. Intell, 14(2), 239-256, (1992).
13. Sharp, G., Lee, S., Wehe, D., “Maximum-Likelihood Registration of Range Images with Missing Data”, TPAMI, 30(1), 120–130, (2008).
14. Matsuyama, T., Wu, X., Takai, T., Wada, T., “Real-Time Dynamic 3-D Object Shape Reconstruction and High-Fidelity Texture Mapping for 3-D Video”, In IEEE Transactions on Circuits and Systems for Video Technology, 14(3), 357–369, (2004).
15. Franco, J., Boyer, E.,”Fusion of multi-view silhouettecues using a space occupancy grid”, In ICCV, 2, 1747-1753, (2005).
16. Kutulakos, K., Seitz, S., “A Theory of Shape by Space Carving”, In International Journal of Computer Vision, 38(3), 199–218, (2000).
17. Bhotika, R., Fleet, D., Kutulakos, K, “A Probabilistic Theory of Occupancy and Emptiness”, In Proc. ECCV, 3, 112–132, (2002).
18. J. Y. Zheng, “Acquiring 3-d models from a sequence of contours,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.16, no.2, pp.163–178, February 1994.
19. T. Kanade, M. Okutomi, and T. Nakahara, “A multiple-baseline method,” in DARPA Image Understanding Workshop, 1992, pp. 409–426.
20. T. Kanade, S. B. Kang, J. Webb, and C. L. Zitnick, “A multi base line stereo system with active illumination and real-time image acquisition,” in International Conference on Computer Vision, 1995, pp. 88–93.
21. P. E. Debevec, C. J. Taylor, and J. Malik, “Modeling and rendering architecture from photographs: A hybrid geometry and image-based approach,” in SIGGRAPH’96, Computer Graphics Proceedings, Annual Conference Series, August 1996, pp.11–20.
22. S. M. Seitz and C. R. Dyer, “Photo realistic scene reconstruction by voxel coloring,” International Journal of Computer Vision, vol. 35, no. 2, pp. 151–173, 1999.
23. K. N. Kutulakos and S. M. Seitz, “A theory of shape by space carving,” International Journal of Computer Vision, vol. 38, no. 3, pp.1999–218, 2000.
24. W. B. Culbertson, T. Malzbender, and G. Slabaugh, “Generalized voxel coloring,” in International Conference on Computer Vision, September 1999, pp. 100–115.
25. L. Mc Millan and G. Bishop, “Plenoptic modeling: An image-based rendering system”, Proc. SIGGRAPH’95, pp.39-46, 1995.
26. M. Levoy, P. Hanrahan : Light field Rendering. Proceedings SIGGRAPH’96, pp 31–42, ACM Press, New York, 1996.
27. S. Gortler, R. Grzeszczuk, R. Szeliski, M. F. Cohen : The Lumigraph Proceedings SIGGRAPH’96, pp 43–54, ACM Press, New York, 1996.