International Journal of Engineering Research and General Science Volume 3, Issue 5, September-October, 2015
ISSN 2091-2730

RECONSTRUCTION OF 3D MODEL FROM 2D SURVEILLANCE IMAGES

Noel Vincent[1], Shiny Mathew[1], Shilu Mathew[2] and Ishtiaq Qadri[3]*

[1]Department of multimedia technology, Karunya University, Coimbatore, India

[2]Center of Excellence in Genomic Medicine, King Abdul Aziz University, Jeddah, Saudi Arabia

[4]King Fahad Medical Research Center, King Abdul Aziz University, Jeddah, Saudi Arabia

*Corresponding author
E-mail: ishtiaq80262@yahoo.com

Abstract — This paper presents a comprehensive scheme for 3D modeling of objects from series of images that perform multi-view calibration and dense 3D point cloud computation. The proposed methodology is induced by 3D model construction which involves Image matching that finds common local sub-image between two pictures ,on which Structure from Motion (SFM) estimates the relative camera position from anchor points computed in the previous step and Multiple View Stereovision (MVS) which calculates a dense representation of the 3D model (Dense point cloud). The position of a camera can also be computed from correspondences between 3D points and corresponding projections in the image plane such that the Pi configuration minimizes the re-projection errors between the rays passing through optical camera center to 3D points and the 2d image plane coordinates. The modern techniques of SFM and Image Based modeling have provided simple and accurate automation to record and construct 3 dimensional data.

Index Terms—3d model, Photographs, Three-dimensional, 3D Reconstruction, Image matching.

I. INTRODUCTION

Getting three dimensional (3D) models of scenes from pictures has been an enduring exploration theme in photogrammetric and in computer vision. Numerous applications exist, for example, laser (ground), Lidar (Aerial), organized light etc. But they have their own advantages and disadvantages [1, 2]. The advancement of computers is such that today even personal computers can show complex 3D models. Computer recreations are situated in 3D universes and the utilization of 3D models and situations is not currently a typical practice in the Internet. But there are difficulties in generating such 3D models is the when we take into consideration of the cost.

3D recreation has been an enduring examination enthusiasm for some scientists and researchers. Distinctive systems have been proposed with variable camera situating approaches based upon their application.

In this paper we propose a 3D innovation in picture coordinating which has been setup with novel building design in putting the camera position for observation in a successful way to catch most extreme extent with least turn. The goal is to make a 3D model from image sequence by a surveillance camera for modeling an ordinary object or by using the new architecture in setting up the camera position to reconstruct a 3D model for the use of surveillance. This paper proposes a methodology of 3D reconstruction of scenes on surveillance using image acquisition and processing the images through Patch Multi View Stereo for effective results. The methodology is applicable for indoor and outdoor surveillance.

II. RELATED WORKS

The 3D model development or remaking of a scene caught from distinctive perspectives by the arrangement of 2D pictures is a noteworthy issue being considered for quite a long time in the region of computer vision. There are a vast number of problems already published on addressing them. Here some of the problems 1) the amount of input; 2) the complexity of the information given; 3) the representation of the output and 4) 3d model construction of a human model [3].

The quantity of inputs given will give a related yield. The multi-view camera adjustment remains a testing undertaking. This multi-view adjustment is critical since it is the person who focus the scene's exactness development and the nature of the 3D dense model.

Late research gave a proficient item, for example, ARC 3D [4] and 123D catch [5] display however the issue is that they utilize the cloud as capacity and computational stage, so the pictures are figured and all the control is not given to the client. In any case, some are free, yet at the same time they utilize cloud for calculation.

Since the developing of OpenSource structure has begun to perform multi-view alignment and dense 3D point cloud calculation [6]. Here the paper proposes a new camera positioning architecture and also a free developing software to make a 3D digital model in ease. The only problem is that of the computational speed, which depends on the user’s computer. The processing is fast for small scenes or small images, but computation takes time for large scenes or large images [7] [8].

III. METHODOLOGY

IV. *3D from Images*

Building a 3D model from pictures comprises in recoup 3D camera positions identified with pictures and 3D positions of specific substance of the pictures. It is finished by distinguishing comparative substance between N views and tackles 3D geometry issues [9]. Client info comprises of a picture gathering and camera parameters. The processed yield is an arrangement of 3D camera positions and 3D point in a thick cloud [10]. 3D points Fig. 1.

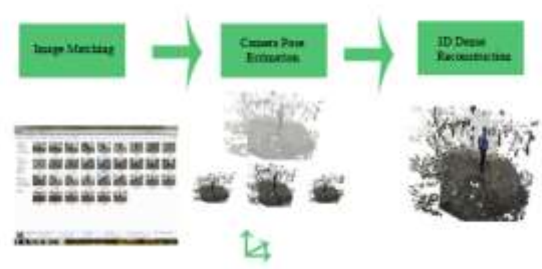

Fig. 1: Structure from Motion/Image-Based Modeling's standard work process.

3D from pictures is a dynamic exploration space that depends on Computer Vision and all the more particularly, Image recovery/coordinating, Structure from Motion (SFM) and Multiple View Stereovision (MVS) [11].

• Picture coordinating discovers normal neighborhood sub-picture between two photos [14].

• Structure from Motion appraises the relative camera position from stay focuses figured at the past step [18].

• Multiple View Stereovisions appraise a dense representation of the 3D model (Dense point cloud).

B. Proposed Architecture Camera Positioning

To make a 3D model of an object for surveillance, image acquisition is performed by a camera which is setup and the images are taken. The new camera architecture involves in placing three cameras exactly 120 degree apart from each other, where each camera can rotate from 0 degree to 60 degree to left taking five images and again to the right using 60 degrees taking five images and these images are sent for processing.

The same process is followed by other cameras, thus sequentially taking the series of image covering most of the area of interest for surveillance 3D modelling.Fig.2.

Fig. 2. Proposed Architecture Camera Positioning

C. Camera Pose Estimation

The camera posture estimation finds the camera positions by taking care of the relative estimation issues [13]. The posture estimation by assessing an inflexible to movement between the cameras a turn R and an interpretation T. Fig.3. If the object is not exactly aligned or very far away from the camera which has been placed the images being captured won't be producing the exact 3D model. The images won't be of exact focus hence the image feature extraction won't be possible.

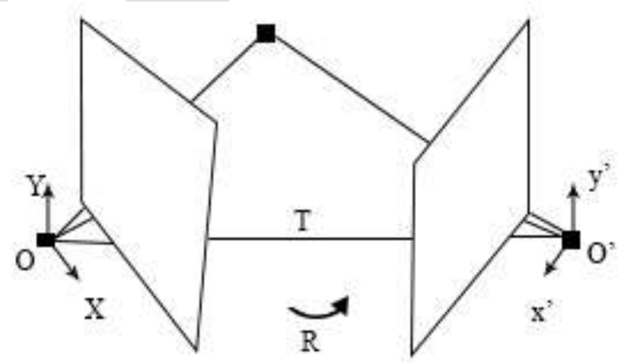

Fig.3. Camera Position Estimation

D. Feature Detection

The feature detection is done by the SFM (Structure from motion) from each and every images which has been given as an input. For every image I and next image J, F (I) is the image feature of the first image I.

For each feature $f_i \in$ F (I) and each corresponding feature $f_i' \in$ F (J) gets the nearest neighbor $f_{nn} \in$F (J) by

$$f_{nn} = \arg\min_{f' \in F(J)} ||fi = f'_i||_2$$

D. Image Matching

Picture coordinating recognizes the pictures that can be utilized to figure the relative introduction of the cameras and in this way to ascertain the pictures' geometry. This procedure of picture coordinating is done in 3 stages:

1. Register the substance on every picture (Feature and Descriptor calculation, for case SIFT) [12],

2. Discovering the matches between two photos (locate the closest descriptor in the other picture of the sets),

3. Checking the geometry of the similar matches.

When the picture coordinating is done between the conceivable pictures the geometric chart is assembled.Fig.4.

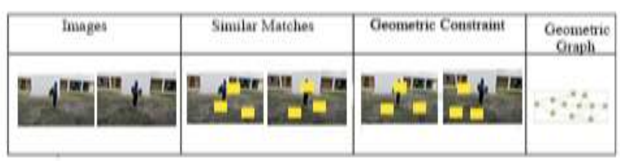

Fig 4: Three steps of Image Matching and final geometric graph

The position of a camera can likewise be figured from correspondences between 3D focuses and comparing projections in the picture plane [15]. This 3D-2D correspondence issue is known as Resection Fig.5. It comprises in assessing Pi (revolution, interpretation) with imperative geometry. It finds the Pi arrangement that minimizes the re-projection slips between the beams going through camera focus to 3D focuses and the 2D picture plane directions [21]. Once the cameras are connected with the Essential network then the 3D focuses are constructed.

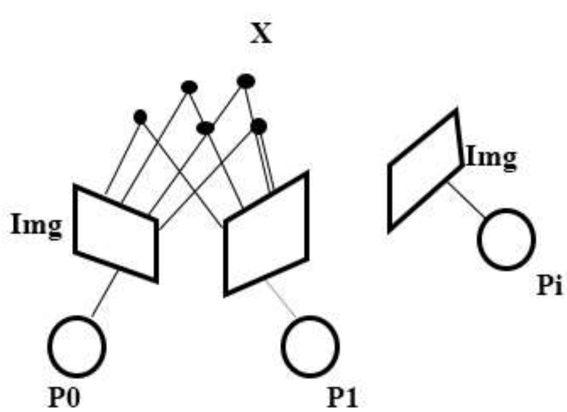

Fig.5. 2D- 3D Resection

Fig.6. Image Acquisition

E Multiple View Stereovision

Multiple View Stereovision (MVS) is the system for mapping picture pixel to 3D focuses to point cloud. The dense representation can be a dense point cloud or a dense mesh. Then it determines a 3D position for each corresponding pixel of the image series. MVS utilizes numerous pictures to diminish lapses and aides in achieving an exact model in co connection with pictures gave. One of the techniques utilized is the Patch methodology called PMVS (Patch Multi View Stereo) [20]. It is taking into account a seed developing method, which discovers relating patches in the middle of pictures and grows the district by an iterative extension and separating strides to uproot terrible correspondences. This methodology finds extra correspondences that were rejected or not found at the picture coordinating procedure. The vacant 3D territories relate to ineffectively textured picture.

V.RESULTS AND EXPERIMENTAL ANALYSIS

The experiment was carried out by placing the camera as said above and series of images where taken and was inputted for making the 3D model. The images retrieved are put into a folder and the program is run for making the 3D model from the 2D images. Fig.6.

The user has the option to select the image quality this is done because of the low computing power of the user's personal computer. The feature extraction is done and which is saved in a dataset and which will be used for the 3D dense reconstruction of the image.

The image set is loaded to convert the 2D images for image matching to find the similar matches among the series of images by removing the resection among images due to distortion due to movement of camera. Image of the tool box used is given below. Fig.7.

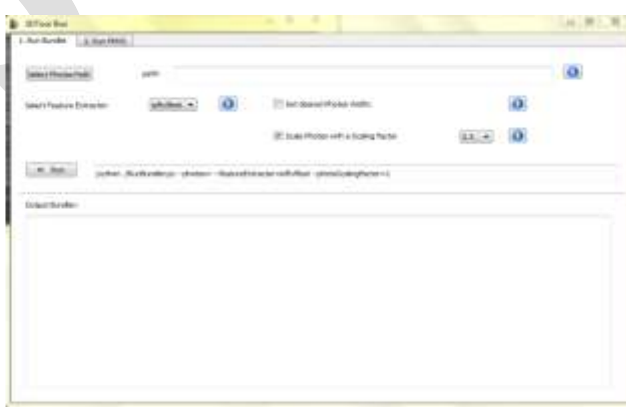

Fig.7. Toolbox

After the image extraction the PMVS (Patch Multi-view stereo) is done with the SIFT algorithm for reconstruction from the image feature stored. The user has the option to select the number of cluster from the images and the model is created in a ply format by a dense cloud formation which can be viewed in a Mesh lab. The missing pixels will be present that will be in a formation of hole. The working is show in the below image. Fig.8.

The output produced is better and not of high precision but in a satisfactory level, the computation time takes time based on the user's personal computer based on the ram. The output below in Fig.9 was done in a System having Windows 7 as Os having a 4GB Ram and 1GB Graphics Card, the computation time took around 1 hour for the given output and 5 minutes for smaller objects. E.g. Box. Fig.10.

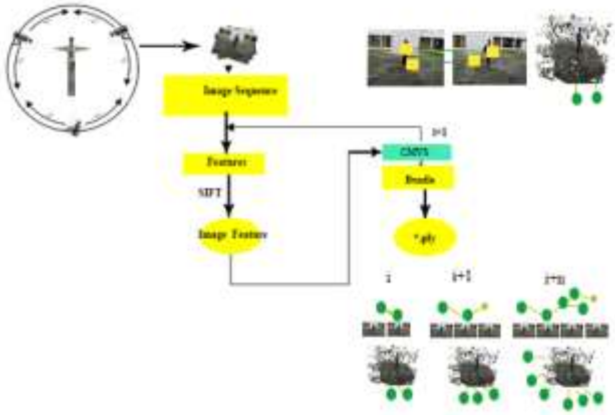

Fig.8. Workflow

Fig.9 3D dense model of a Person

Fig.10. 3D dense model of Box

V. PERFORMANCE METRICS

The performance of the system implemented is shown is graph. In which the first graph Fig.11. Shows the performance of the system with time vs. images.

Fig.11. Time Vs Images

The second graph Fig.12. Shows the comparison between other existing systems and the newly proposed system.

Fig.12. Existing system Vs Proposed System

The Third and final graph Fig.13. Shows the time with respected to the graphics processor used in making different model or scenes.

Fig.13 Time Vs Graphics Processor

V. CONCLUSION

This is an easy to use application to perform 3D computerized duplicates of the picture arrangement. A low-cost, which opens a direct access to a solution for Image –Based Modeling to anyone who has a digital camera. A SIFT algorithm is the only drawback since it uses a feature descriptor. The SFM is a fast and simple with an easy data processing which is able to satisfy the needs of a project. The 3D reconstruction from images has become a lot easier, in which the user can use it will full efficiency. The only drawback is the time required for processing but that depends on the user having a higher computing power, at a mostly a higher Ram around 16GB with a 4GB graphics card can make a far more difference in computing.

REFERENCES:

[1] Richard A. Vincent the Sanborn Map Company Michael Ecker Light Detection and Ranging (LiDAR) Technology Evaluation

[2] Daniel Huber and Burcu Akinci Using Laser Scanners for Modeling and Analysis in Architecture, Engineering, and Construction

[3] Pierre Moulon Alessandro Bezzi A free solution for Three-Dimensional Documentation

[4] David Tingdahl and Luc Van Gool, "A Public System for Image Based 3D Model Generation", Computer Vision/Computer Graphics Collaboration Techniques 5th International Conference, MIRAGE 2011

[5] Erickson, M., Bauer, J., and Hayes, W., "The Accuracy of Photo-Based Three-Dimensional Scanning for Collision Reconstruction Using 123D Catch," SAE technical Paper 2013-01-0784, 2013, doi:10.4271/2013-01-0784

[6] Constructing 3D human model from front and side images Yueh-Ling Lin, Mao-Jiun J. Wang

[7] S. Agarwal - N. Snavely - I. Simon - S. M. Seitz - R. Szeliski, Building Rome in a day, in ICCV 2009, 72-79.

[8] H. Bay – A. Ess - T. Tuytelaars – L. Van Gool, SURF: Speeded Up Robust Features, in CVIU 2008, 346-359

[9] A. Bezzi – L. Bezzi – D. Francisci – R. Gietl, L'utilizzo di voxel in campo archeologico, in Geomatic Workbooks,6, 2006.

[10] J. -M. Frahm – P. Georgel - D. Gallup – T. Johnson – R. Raguram – C. Wu – Y.-H. Jen – E. Dunn - B. Clipp - S. Lazebnik, Building Rome on a Cloudless Day, in ECCV 2010, 368-381.

[11] Y. Furukawa – B. Curless – S. M. Seitz – R. Szeliski. Towards Internet-scale multi-view stereo, in CVPR 2010, 1434-1441.

[12] D. G. Lowe, Distinctive image features from scale-invariant keypoints, in IJCV 2004, 91-110.

[13] D. Nister, An Efficient Solution to the Five-Point Relative Pose Problem, in IEEE Trans. Pattern Anal. Mach. Intell. 2004, 756-777.

[14] N. Snavely – S. M. Seitz – R. Szeliski. Modeling the World from Internet Photo Collections, in IJCV 2008, 189- 210.

[15] Bae, S.-H., Balakrishnan, R., And Singh, K. 2008. ILoveSketch:As-natural-as-possible sketching system for creating 3D curve models. In Proceedings of the ACM symposium on User.interface software and technology (UIST), 151–160.

[16] AUTODESK, 2009. Maya. http://www.autodesk.com/maya.

[17] A Rayleigh reconstruction/interpolation algorithm for 3D ultrasound Jo~ao M. Sanches , Jorge S. Marques

[18] Reconstruction of 3D human body pose from stereo image sequences based on top-down learning

[19] A neural network for recovering 3D shape from erroneous and few depth maps of shaded images Mohamad Ivan Fanany , Itsuo Kumazawa

[20] A Flexible And Automatic 3d Reconstruction Method

Shunyi Zhenga , Zongqian Zhana, Zuxun Zhanga

[21] Xiao, J., Baker, S., Matthews, I., Kanade, T.: Real-Time Combined 2D+3D Active Appearance Models. In: CVPR, IEEE. (2004) 535-542

[22]Jiang, H.:3D Human Pose Reconstruction Using Millions of ICPR,IEEE. (2010) 1674-1677