American Journal of Engineering Research (AJER)
e-ISSN : 2320-0847 p-ISSN : 2320-0936
Volume-03, Issue-01, pp-140-148
www.ajer.org

Research Paper Open Access

3D Scene Reconstruction from Video Camera for Virtual 3D City Modeling

Surendra Pal Singh[1], Kamal Jain[1], V. Ravibabu Mandla[2]
[1]*Department of Civil Engineering, Indian Institute of Technology (IIT) -Roorkee, India*
[2]*School of Mechanical and Building Sciences, Vellore Institute of Technology (VIT)-University, India*

Abstract: The main purpose of this study is to explore the potential of normal digital video camera for virtual 3D City modeling. For any photogrammetric project work, image acquisition is the main issue. Cost, time, and success of any close range photogrammetric project is mainly dependent on image acquisition method. Video recording is an easy way to capture the large city area in less time. In the present study a simple method for 3D scene reconstruction by using digital video camera is developed for virtual 3D City modeling. The digital video camera used was Sony DSC HX7V camera for video recording. From this video data, image frames created and identified for suitable image frames for image based modeling. After processing some intermediate products were obtained and finally textured 3D model of area was created. Study area was Civil engineering department, IIT-Roorkee, India. Agisoft Photoscan software was used for this work. This paper covers the methodology, result, discussion, conclusion, advantages and limitations of the method.

Keywords: *- 3D scene, Computer vision techniques, Image based modeling, Virtual 3D City modeling,*

I. INTRODUCTION

The Virtual 3-D city model generation is a very hot research topic to engineering and non-engineering scientist. 3D city models are basically a computerized or digital model of a city contains the graphic representation of buildings and other objects in 2.5 or 3D. Demand of Virtual 3D City models is increasing day by day for various engineering and non-engineering fields. Now days, various methods are available to create Virtual 3D City model. Laser scanning and Photogrammetry are the main techniques. For 3D City modeling, Automatic and Semiautomatic; the two main techniques are used for data acquisition, [1]. For 3D City modeling, Image based techniques are more suitable than Laser based techniques due to cost and availability of data. For 3D City modeling, the main problem comes for image acquisition. To find the suitable position for capturing the image is a very important issue for Image based 3D city modeling. Due to this, there is a very high demand for suitable image acquisition system. Images are easily available to everybody at nominal cost. Handling of image based project is very cost effective and accuracy is also good.

For 3-D city modeling, Video recording is the main techniques for image acquisition. It has many advantages. Video is an easy obtainable and low cost data acquisition system, now a days; many researchers are showing interest in this field.

Some of the important previous works are summarized here:

Videogrammetry is a measurement technique which is mainly based on the principles of 'Photogrammetry, [2]. Videogrammetry refers to video images taken using camcorder or movie function on digital still camera. Video movie consists of sequences of images (or frames). If video speed is 25 fps (frame per second) and taken for 1 minute (i.e. 60 seconds), there are 25 frame per second or overall 1500 image.

Kawasaki et al., (1999), also worked for automatic modeling of a 3D city map from real-world video. They proposed an efficient method for making a 3D map from real-world video data. The proposed method was an automatic organization method by collating the real-world video data with map information using DP matching. They also made a system which can generate a 3D virtual map automatically in VRML format. [3]

Clip et al., (2008), designed a Mobile 3-D City Reconstruction system. It is an efficient flexible capture and reconstruction system for the automatic reconstruction of large scale urban scenes. This system is both backpack

and vehicle mounted allowing capture of interior or less accessible areas as well as large outdoor scenes. In this work, they propose an efficient system to capture the 3D-geometry of existing cities through computer vision techniques. This system can deliver 3D reconstructions of large urban scenes in near real time. This system is modular and man portable, it is able to record both from a backpack mounting for interior areas and from an automobile for exterior recording. GPS and INS was also used in this product. [4]

Figure1. 3D reconstruction from Video only with the back pack system (Source: [4])

Tsai et al., (2006), [5] developed a method for texture generation and mapping by using video sequences for 3D building models.

Gael et al., (2007), [6] explained a system for computing geo-referenced positions and orientations for non calibrated videos images of buildings. This method is based on the fusion of multimodal datasets, namely GPS measures, video sequences and rough 3D models of buildings. This is a method for registration of GPS, GIS, and Video data for urban scene modeling.

Pollefeys et al., (2000), [7] gave a method for 3-D model generation using video image sequence.

In 2008, M. Pollefeys and his team created a detailed real time urban reconstruction from Video. They used video data and GPS/GNS data. In this method, there were two main processing components. One was for video data input and another was computing component. After video data input, the data reading or data preparation is a processing component. In computing component, 2-D tracker (GPU) and 3-D tracker/Geo-location are the main track. Geo-located camera was used in this process. By using sparse scene analysis and multi-view stereo, depth map was generated which is very useful to create 3-D model of an area. After this triangular mesh texture map was generated which give a photorealistic textured 3-D model of that area. [8]

Fulton and Fraser, (2009), explained a method for automatic reconstruction of building by using a hand held video camera. In this method, a video recording was done for the building of interest. Video sequence were transferred into computer and saved as individual JPEG frames. Blurred frames were removed and non-blurred key frames were selected. Registered of these non-blurred key frames was done using phase correlation method, after this feature extraction was done. [9]

Zhang et al., (2009), gave a concept for consistent depth maps recovery from a video sequence. Video image sequence frames were used and depth maps from these frames were created. In this method, they used the Structure From Motion (SFM) to recovered the camera parameters, Disparity Initialization, Bundle optimization, and Space-Time fusion techniques was used to create depth maps. These depth maps are useful to create virtual 3-D model of an area or object. [10]

Tian et al., (2010), gave a concept of knowledge-based building reconstruction from terrestrial video sequence. They gave an automatic method for the reconstruction of building models from video image sequences. Building structure knowledge is used as a key factor. [11]

Hengel et al., (2007), developed a method and system, (named as Video Trace). VideoTrace is a system for interactively generating realistic 3D models of objects from video. The combination of automated and manual reconstruction allows VideoTrace to model parts of the scene not visible, and to succeed in cases where purely automated approaches would fail. In this system initially a frame from the input video sequence is taken and a partial tracing of the model takes place then the final model is overlaid on the video, and the result of rendering the final model is brought back into the original sequence. [12]

Singh et al., (2013), developed a multi-camera setup and method for camera calibration from video image frames. From video data, image frames were created for close range photogrammetric work. [13]
In India, Prof. Bharat Lohani and his team, (2012) from IIT-Kanpur, developed an Indigenous technique for Laser based mobile mapping system for 3D modeling. It creates a basic, simple and good 3D model of an area. [14]
Singh et al. (2013), [15] explains about techniques and applications of virtual 3D city modeling. 3D city model is also useful for e-Governance. [16]. Image based modeling is also suitable for building modeling for Virtual 3D City model generation. [17], [18], [19], [20].

Thus, it can be concluded that till now, there is no cost effective and easy to use system available for 3D City modeling. And there is a need for a method, which can be helpful for 3D City modeling by using video data. The main purpose of this work is to explore the potential of normal digital video camera for virtual 3D City modeling. In the present work, it is tried to develop a method for 3D scene reconstruction for 3D City modeling by using video data. For this work, the Agisoft Photoscan software was used for 3D scene reconstruction.

The main contribution of this research paper is to explore the potential of normal digital video camera for Virtual 3D scene reconstruction mainly for virtual 3D city modeling. This method is very fast and processing of image frames is automatic. So it is very easy to use for any kind of image based 3D modeling.

II. METHODOLOGY

Flow diagram of overall methodology can be seen in Figure. 2.

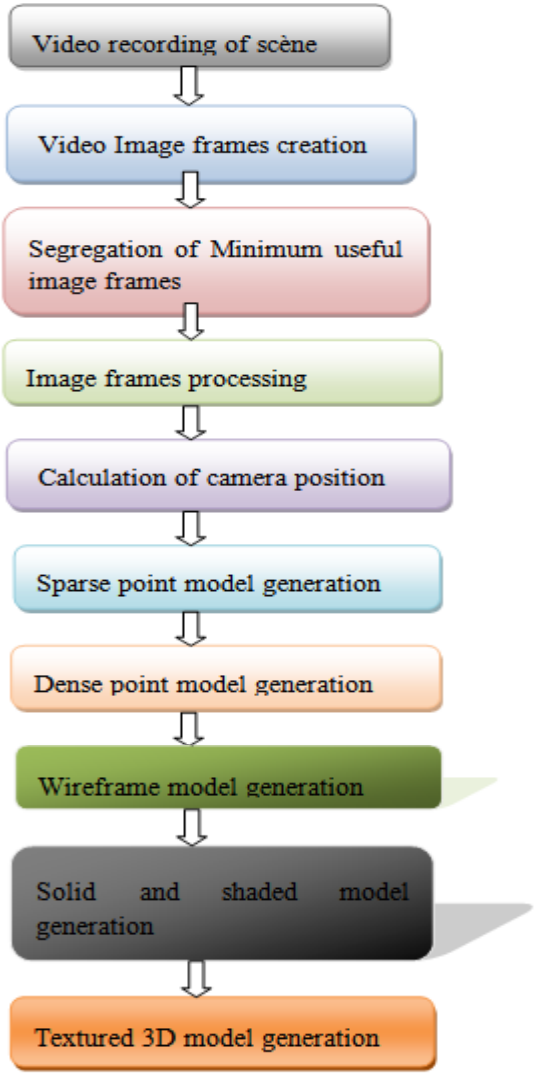

Figure 2. Flow diagram of methodology for 3D scene reconstruction from video camera.

To create the 3D scene reconstruction from video camera, following steps are followed:

1- Video recording of scène
2- Video frame creation
3- Segregation of Minimum useful image frames
4- Image frames processing
5- Calculation of camera position by SfM
6- Sparse point model generation
7- Dense point model generation
8- Wireframe model generation
9- Solid and Shaded model generation
10- Textured model creation

2.1. VIDEO RECORDING OF SCENE

Video recording of Department of Civil Engineering was done using multi camera set up. This multi camera set up is developed by Singh et al. (2013) [13]. Video recording should be taken with slow moving speed. The direction of camera should be parallel for façade modeling. In this research work, the Sony DSC HX 7V, camera was used.

Figure 3. Sony DSC HX7V Digital Camera

The Sony DSC HX7V digital camera is 16.2 mega pixel resolution for image and has Exmor R CMOS Image sensor. It has 10x optical zoom. This digital camera can record full HD video at 1920×1080 resolution. It has capacity to create 50 frames per second (FPS).

2.2. VIDEO FRAME CREATION

After video recording of a scène, the video frames were created. All video camera has the feature "frames per second" (FPS) or "Frame Rate". It is the characteristic feature of any video camera. It defines the frequency (or rate) at which the camera device produces unique consecutive images called frames.

In this work, free software "Free Video to JPG Converter" was used. This software also has the capacity to control the frame rate. It means, if Video recording time is for 5 minutes, (5×60= 300 seconds) and camera has 50 FPS. So in five minutes video (300×50=15000), 15000 images fames are created. With this software, one can change the frame rate. According to user requirements, one can choose every 10, 30, 50,100, 500, 1000 frames, or in every 1,2,5,10,20 seconds. One can also choose total number of frames as 10, 20, 50,100,200,500, or one can also extract every frame from video.

Figure 4, is showing some video image frames created by this software.

Figure 4. Some video image frames

2.3. SEGREGATION OF MINIMUM USEFUL IMAGE FRAMES

After extraction of video image frames, the minimum useful image frames for 3D image modeling was selected. One can use all created image frames, but computer system will heavy loaded during processing and create problem during image frames processing. Consecutive images or frames have more overlapping area. Therefore images frames after some interval were selected; depending on overlapping of area in image frames.

2.4. IMAGE FRAMES PROCESSING

For 3D model generation, these image frames were used for 3D point creation. After processing of these video image frames, 3D points and model was obtained. Image alignment is a main processing for this work.

2.5. CALCULATION OF CAMERA POSITION

After Image Alignment, camera position obtained for each video image frame. Fundamental matrix (3×3 matrix, which relates the corresponding points in stereo images) used and the intrinsic parameter of camera were calculated. After that, a solution for camera position was obtained. [21], [22].

Figure 5. Building façade with video frame positions

2.6. SPARSE POINT CLOUD GENERATION

3D point cloud created by using Structure from Motion (SfM) techniques. Feature detection and feature matching is the key concept to produce sparse 3D point cloud from image sequence [23]. Figure 6, is showing the sparse point cloud model of an area from Civil Engineering department of IIT-Roorkee.

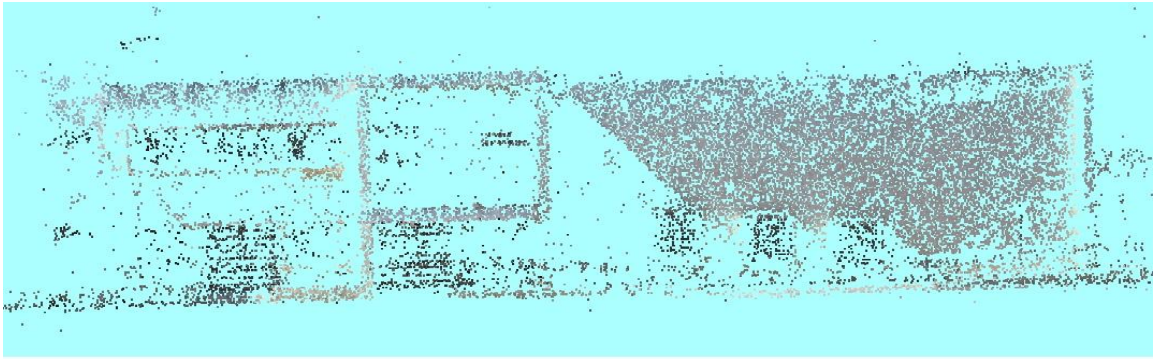

Figure 6. Sparse Point cloud model

2.7. DENSE POINT CLOUD GENERATION

The number of point cloud increased using more efficient methods of feature matching techniques and dense point cloud model of an area were created. It produces more details of area.
Figure 7, showing the dense point cloud model of area.

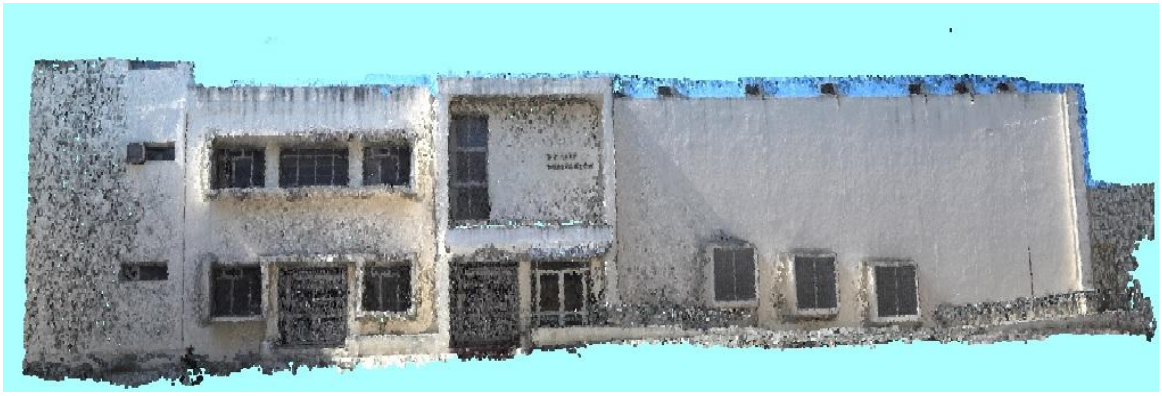

Figure 7. Dense Point cloud model

2.8. WIREFRAME MODEL GENERATION

Wireframe model also created with the help of these 3D points. Wireframe model is useful to produce exploded and perspective view more easily. This model is also useful for viewing the model from any desired point by changing line of sight. Figure-8, showing the wireframe model of an area.

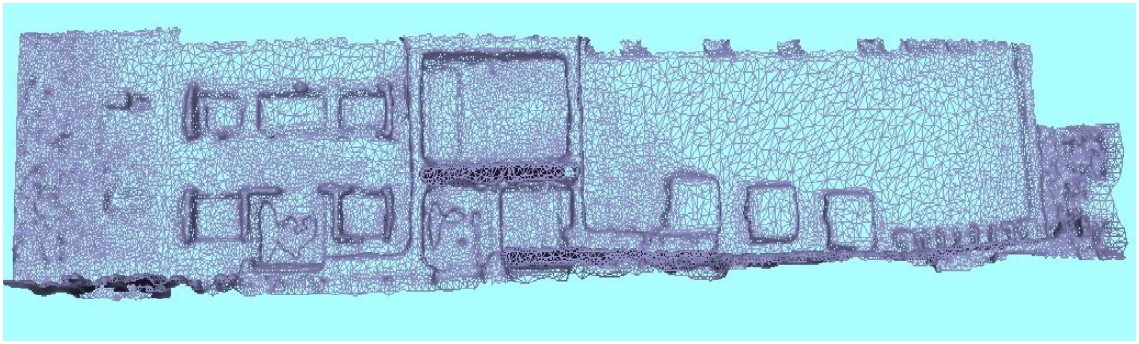

Figure 8. Wireframe model

2.9. SOLID AND SHADED MODEL GENERATION

Wireframe model does not represent an actual solid and has no surface and volume information. Thus, solid model and shaded surface model created. Solids model is a complete and unambiguous representation of a precisely enclosed and filled volume. Shaded model gives near about actual model of an area.
Figure 9, and Figure 10 is showing the Solid model and Shaded model.

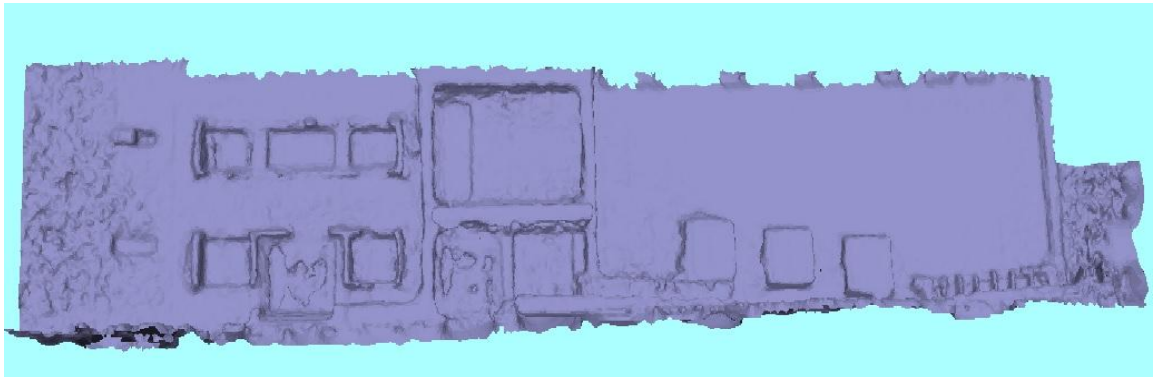

Figure 9. Solid model generation

Figure 10. Shaded model generation

2.10. TEXTURED MODEL GENERATION

At finally, generic texture method applied, which allows parameterization of texture atlas for arbitrary geometry and make final textured 3D model of an area. It gives photo-realistic representation of a scene.

Figure 11. Textured 3D model of O.P.Jain Auditorium building of Civil Engineering Department, IIT-Roorkee, India.

III. RESULT AND DISCUSSION

By using this proposed method, 3D scene obtained of O.P.Jain Auditorium side of Department of Civil Engineering, Indian Institute of Technology, Roorkee, India. Our main goal was to show the feasibility of 3D scene reconstruction from hand held digital video camera. This proposed method will be very useful for image based virtual 3D city modeling by using normal digital video camera. Normal digital video camera is very cheap source of image acquisition and easily to handle by user. In less time, more data can be obtained. So project cost will reduce. In this way, it can conclude that this method will be helpful for 3D city reconstruction.
In Indian city, this technique will be very useful to make virtual model of any city. This technique is very cost effective to create virtual 3D City model, compare to any other Geomatics techniques such as Aerial Photogrammetry, Satellite Photogrammetry or Laser scanning based techniques.

IV. CONCLUSION

Advantages and Limitations of this method can be summarized as given below:
Advantages: The main key advantages for this method:

- Automatic processing,
- Photo-realistic 3D scene,
- Time and Cost of project will reduce.
- Image frames created by video; can be used for any kind of Image based 3D modeling software.

Limitations: The main limitations are given below:

- For Video recording of a City, weather conditions and light condition should be favorable.
- Crowd conditions should be avoided.
- Video camera quality should be high for good quality of 3D model.
- Higher resolution camera gives good photo-realistic 3D model.
- All video image frames are not useful.
- Processing time depends on number of video image frames.
- Speed of video recording is also play important role for image frames quality.
- If the building façade is large then camera is kept at large distance from the building.

The possible application of this research work is also very important. Video recording of any city will be very easy and cost of project will reduce. Anyone can use multi camera set up for 360^0 recording of area. Camera configuration can be change according to software requirements. Integration of GPS/INS is also possible in this research work. For future work, video camera based mobile mapping vehicle is also possible for fast 3D City modeling.

Virtual 3D city modeling has many important applications. 3D city model is useful to identify the encroachment in municipality, Virtual tourism, Historical fort and building conservation etc. So the demand of 3D city modeling is increasing day by day. Normal digital video camera has the good potential to create virtual 3D city model. Finally, it concluded that, this proposed method will be very useful to create 3D city model by using video camera in less cost and less time. Texture quality of 3D model is also very good. This method is fully automatic and very easy to use.

V. ACKNOWLEDGEMENT

This work was supported by the MHRD grant (also called as Institute Fellowship; duration From Jan.-2009 to Dec.-2012) given by Indian Institute of Technology, Roorkee, Uttarakhand, India. Also thanks to my friends and staff of Civil Engineering Department of IIT-Roorkee for support us during this work.

REFERENCES

[1] Forstner, W., “3-D City Models: Automatic and Semi-automatic Acquisition methods”, Photogrammetric week, Wichmann, Karlsruhe. 1999.
[2] Gruen, Armin. "Fundamentals of videogrammetry—a review”, Human movement science, 1997, Page-155-187.
[3] Kawasaki Hiroshi, Yatabe Tomoyuki, Ikeuchi Katsushi, Sakauchi Masao, “Automatic Modeling of a 3D City Map from Real-World Video”, 1999, Institute of Industrial Science, University of Tokyo, Japan.
[4] Clipp Brian, Raguram Rahul, Frahm Jan-Michael, Welch Gregory, Pollefeys Marc, “A Mobile 3D City Reconstruction System”, 2008, UNC Chapel Hill, ETH, Zurich.

[5] Tsai Fuan, Chen Cheng-Hsuan, Liu Jin-Kim and Hsiao Kuo-Hsing, "Texture Generation and Mapping Using Video Sequences for 3D Building Models", Innovations in 3D Geo Information Systems, 2006, 429-438.
[6] Gael Sourimant, Luce Morin,Kadi Bouatouch, "GPS, GIS and Video fusion for urban modeling". CGI, May 2007.
[7] Pollefeys M., Koch R., Vergauwen M., Van Gool L., "Automated reconstruction of 3D scenes from sequences of images", ISPRS Journal of Photogrammetry & Remote Sensing, 55, 2000, 251–267.
[8] Pollefeys M., Nistér D., Frahm J.M., Akbarzadeh A., Mordohai P., Clipp B., Engels C., Gallup D., Kim S.J., Merrell P., Salmi C., Sinha S., Talton B., Wang L., Yang Q., Stewénius H., Yang R., Welch G., Towles H., "Detailed Real-Time Urban 3D Reconstruction from Video", International Journal of Computer Vision, 78: 2008, 143–167.
[9] Fulton J.R. and Fraser C.S., "Automated Reconstruction of Buildings using a Hand Held Video Camera", S. Jones, K. Reinke (eds.), Innovations in Remote Sensing and Photogrammetry, Lecture Notes in Geoinformation and Cartography, Springer-Verlag Berlin Heidelberg, 2009, 393-404.
[10] Zhang Guofeng, Jia Jiaya, Wong Tien-Tsin, and Bao Hujun, "Consistent Depth Maps Recovery from a Video Sequence", IEEE Transactions on pattern analysis and machine intelligence, Vol. 31, No. 6, June, 2009, 974-988.
[11] Tian Yixiang, Gerke Markus, Vosselman George, and Zhu Qing, "Knowledge-based building reconstruction from terrestrial video sequence", ISPRS Journal of Photogrammetry and Remote Sensing, 65, 2010, 395-408.
[12] Hengel Anton van den, Dick Anthony, Thormahlen Thorsten, Ward Ben, Philip H, and Torr S., "VideoTrace: Rapid interactive scene modelling from video", ACM Transactions on Graphics, Vol. 26, No. 3, 2007, 86.
[13] Singh, Surendra Pal, Mandla, V.Ravibabu, Jain, Kamal, "Design and Calibration of Multi Camera setup for 3D City modeling,", International Journal of Engineering Research & Technology (IJERT), Vol. 2, Issue 5, May 2013.
[14] Salil Goel , Manohar Yadav, Anjani K. Singh and Bharat Lohani, "An Indigenous development of land based mobile mapping system", India Geospatial Forum, Hyderabad, 22-24, Jan, 2013.
[15] Singh, Surendra Pal, Mandla, V.Ravibabu, Jain, Kamal, "Virtual 3D City Modeling : Techniques and Applications", International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-2/W2, ISPRS 8th 3DGeoInfo Conference & WG II/2 Workshop, 27 – 29 November, 2013, Istanbul, Turkey.
[16] Singh, Surendra Pal, Mandla, V.Ravibabu, Jain, Kamal, "Virtual 3-D City Model for G-Governance", Regional Conference on Geomatics for Good-Governance (G-Governance-2011) (13-14 Sep, 2011), Indian Society of Geomatics and Kashmir University, Srinagar, (J&K), India.
[17] Shashi M. and Jain Kamal, "Use of Photogrammetry in 3D Modeling and visualization of buildings", Asian Research Publishing Network (ARPN) - Journal of Engineering and Applied Sciences, Vol. 2, No. 2, April 2007, ISSN 1819-6608.
[18] Singh, Surendra Pal, Mandla, V.Ravibabu, Jain, Kamal, "3-D Building Modeling from Close Range Photogrammetry", Geomatrix'12 , An International Conference on Geo-spatial technologies and Applications, 26-29, February 2012, CSRE, Indian Institute of Technology Bombay (IITB).
[19] Singh, Surendra Pal, Mandla, V.Ravibabu, Jain, Kamal, "Image Based 3D Modeling of Campus (Department of Civil Engineering, IIT Roorkee, Uttarakhand, India) by Using SketchUp", International Journal of Advanced Civil Engineering and Architecture Research, Volume 2, Issue 1, pp. 65-71, Article ID Tech-168. September, 2013.
[20] Singh, Surendra Pal, Mandla, V.Ravibabu, Jain, Kamal, "Virtual 3D Campus Modeling by using Close Range Photogrammetry", American Journal of Civil Engineering and Architecture, Vol.1, No.6, 2013, Page 200-205, doi: 10.12691/ajcea-1-6-9.
[21] Richard Hartley, "Estimation of relative camera positions for uncalibrated cameras", Proceedings of European Conference on Computer Vision, 1992.
[22] Richard Hartley and Andrew Zisserman, "Multiple View Geometry in computer vision", Cambridge University Press, 2003, ISBN 0-521-54051-8.
[23] Noah Snavely, Steven M. Seitz, Richard Szeliski, "Photo tourism: Exploring photo collections in 3D," ACM Transactions on Graphics (SIGGRAPH Proceedings), 25(3), 2006, 835-846.