3D RECONSTRUCTION

Shruti Sinha 3rd year EE 13115114

Krishna Murari 3rd year EE 13115068

Himanshu Shashank 3rd year EE 13115055

Subhrajit Mallick 3rd year EE 13115123

ABSTRACT

Recently the demand for 3D content for computer graphics, virtual reality medical imaging, computational science, Virtual Reality, media, and communication has taken a steep rise triggering a change in emphasis for the requirements. Already existing systems for constructing 3D models are built around specialized hardware which is not cost effective, and doesn't satisfy the requirement of new applications. This calls for the use of digital imaging facilities (like a camera). This work can also be done in software by extracting 3D from images sequences. The research of 3D reconstruction has always been a focus and difficulty. Using 3D reconstruction one can determine any object's 3D profile, as well as knowing the 3D coordinate of any point on the profile.

For instance, the lesion information of the patients can be presented in 3D on the computer, which offers a new and accurate approach in diagnosis and thus has vital clinical value. Binocular Stereo Vision method or Monocular cues methods can be used to obtain the 3-dimensional geometric information of an object from multiple images or one image to proceed for 3D construction.

Acknowledgement

This proposal describes about the research and development that was done and the people who helped us to accomplish the project. The project was carried out under **Electronics Section**, Indian Institute of Technology, Roorkee.

First of all, we would like to thank our staff advisor, **Mr. Kamal Singh Gotyan** for his guidance and support. His knowledge and ideas have given us a lot of inspiration. Secondly, we would like to thank our **Secretary Padmanabh Pande** for his ideas and suggestions, **Joint-Secretary Rahul Ratan Mirdha** for giving us information about the libraries involved in this project and assisting us in getting quad copter to get images of objects

Big thanks to all our friends and family for their gracious help and support. They have been our pillar of strength throughout this project.

Introduction

The work of 3D reconstruction has always been a focus and difficulty. Using 3D reconstruction one can determine any object's 3D profile, as well as knowing the 3D coordinates of any point on the profile. The 3D reconstruction of objects is a generally scientific problem and core technology of a wide variety of fields, such as Computer Aided Geometric Design, Computer Graphics, Computer Animation, Computer Vision, medical imaging, computational science, Virtual Reality, digital media, etc. For instance, the lesion information of the patients can be presented in 3D on the computer, which offers a new and accurate approach in diagnosis and thus has vital clinical value. Most work on 3-d reconstruction has focused on using methods such as stereovision or structure from motion, which require two (or more) images. Some methods can estimate 3-d models from a single image, but they make strong assumptions about the scene and work in specific set-tings only. Such methods are also called Monocular cues methods refer to use image (one, two or more) from one viewpoint (camera) to proceed 3D construction. It makes use of 2D characteristics (e.g. Silhouettes, shading and texture) to measure 3D shape. It takes a two-dimensional image and creates a three-dimensional "fly around" model, giving the viewers access to the scene's depth and a range of points of view. Binocular Stereo Vision obtains the 3-dimensional geometric information of an object from multiple images .The results are presented in form of depth maps. Images of an object acquired by two cameras simultaneously in different viewing angles, or by one single camera at different time in different viewing angles are used to restore its 3D geometric information and reconstruct its 3D profile and location.

Structure from motion (Sfm) has been successfully used for the reconstruction of increasingly large uncontrolled photo collections In a typical incremental SfM system two-view reconstructions are first estimated upon successful feature matching between two images, 3D models are then reconstructed by initializing from good two-view reconstructions, repeatedly adding matched images, triangulating feature matches, and bundle-adjusting the structure and motion.

We have a scene to construct and General Strategy is Triangulation which means matching a feature in at least 2 views to get 3D position. Instead of matching the entire image to each other match the features by using approximate nearest neighbour.

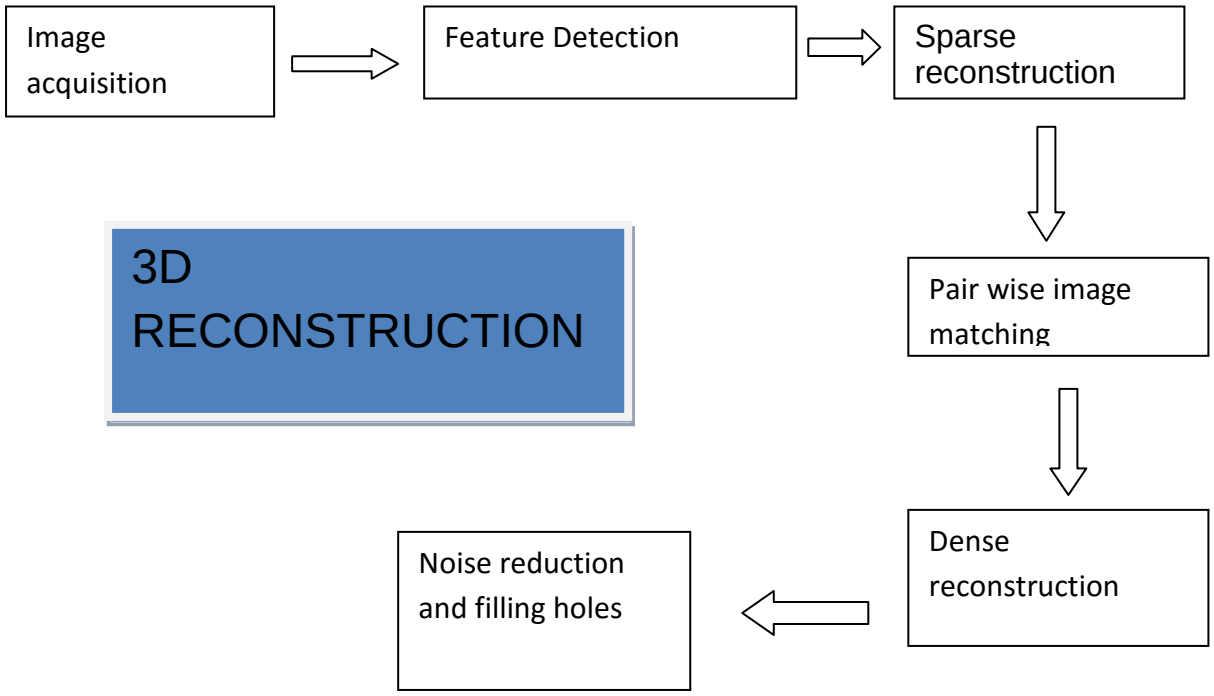


Image acquisition

2D digital image acquisition is the information source of 3D reconstruction. Commonly used 3D reconstruction is based on two or more images, also it may only employ one single image sometimes. There are various types of methods for image acquisition that depends on the occasions and purposes of the specific application. Not only the requirements of the application must be meet, but also the visual disparity, illumination, performance of camera and the feature of scenario should be considered. Usually image is taken from several viewpoints.

a) From 4 views up to several hundreds

b) 20~50 images on average

Features detection and pair wise matching:

The aim of feature extraction is to gain the characteristics of the images, through which the stereo correspondence processes. We have used SiftGPU feature detection functionality. SiftGPU is an implementation of SIFT for GPU. SiftGPU processes pixels parallel to build Gaussian pyramids and detect DoG Key points. Based on GPU list generation, SiftGPU then uses a GPU/CPU mixed method to efficiently build compact key point lists. Finally key points are processed parallel to get their orientations and descriptors.This can be done in commercially available softwares like VisualSfm.

Sparse reconstruction:

Get 3D points from extracted features and creates a 3D mesh of object.

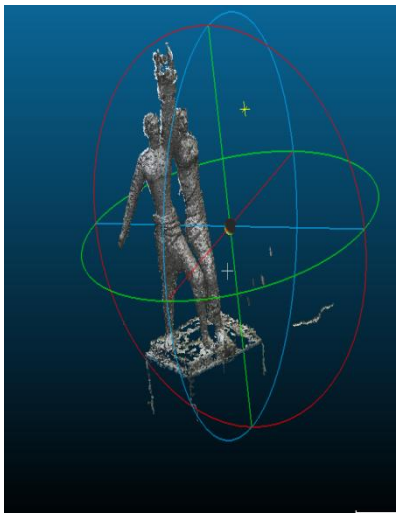
Dense reconstruction by using Yasutaka Furukawa's CMVS/PMVS

This software Clustering Views for Multi-view Stereo (CMVS) takes the output of structure-from-motion (SfM) software as input, then decomposes the input images into a set of image clusters of manageable size. An MVS software can be used to process each cluster independently and in parallel, where the union of reconstructions from all the clusters should not miss any details that can be otherwise obtained from the whole image set.

Noise cleaning and editing:

The output of dense reconstruction is WRL file which can be opened and viewed using any VRML viewer. Background noise is subtracted and finally we get 3D textured image of object.

Input Image



Output

Conclusions

Sfm approach reconstructs 3D model of Smooth objects easily If we have multiple images of object from different views.3D reconstruction is a hard problem however, solutions do exists but they need to be adapted to specific environment. I t requires to gather the good points from many views otherwise holes appear .With high resolution data this method works well (except texture less areas) sufficiently in many cases.

Limitations

More information is needed with Classical assumption that Objects are “smooth.” Optimization problem .One needs to find the “best” smooth consistent object. Difficulties are hand held camera, outdoor environment .Challenges for the Future are Shinny materials: metal, porcelain. We need many images from different view angles of object.

Future Scopes

The possibility for improvement in 3D reconstruction is given below;

1) 3D model of shiny objects and texture less objects:
   Current method poses problem with shiny and texture less objects .However multiple images from different angles and dense reconstruction method can solve this problem up to some extent.
2) Reduction in number of images required :
   Above approach uses many images to extract features; however 3D model can be reconstructed from using advanced algorithms from just single image or few images.

3) 3D depth reconstruction can be used for autonomous navigation of aerial vehicles.

References:

1) http://ccwu.me/vsfm/doc.html

2) http://www.cs.cornell.edu/~asaxena/learningdepth/ijcv_monocular3dreconstruction.pdf

3) http://www.cvfoundation.org/openaccess/content_iccv_2015/papers/Zheng_Structure_From_Motion_ICCV_2015_paper.pdf

4) http://ptak.felk.cvut.cz/sfmservice/websfm.pl?menu=cmpmvs