3D Model Reconstruction from Multiple Views

Marcio Calixto Cabral

Abril 15, 2002

Work Plan for Topics in Computer Science MAC5701

Bibliography

[1] Aloimonos, John et al. , "A Multi-frame Approach to Visual Motion Perception" *Internacional Journal of Computer Vision, 6:3, 245-255*, Received June 26, 1989.

[2] J.-Y. Bouguet and P. Perona, "3D Photography on your Desk". *Proc. 6th Int. Conf. Computer Vision, Bombay, India, pages 43-50* January 1998.

[3] P. Debevec, C. Taylor and J. Malik, "Modeling and Rendering Architecture from Photographs: A Hybrid Geometry- and Image-Based Approach" *Siggraph* 1996

[4] Marc Pollefey, "Tutorial on 3D Modeling from Images" *In conjunction with ECCV2000* 26 June 2000, Dublin, Ireland

[5] Shashua, A "Projective Structure from Uncalibrated Images: Structure-from-Motion and Recognition" *"PAMI" 16, 8,778-790* August, 1994

[6] Shashua, A et al. "Trilinearity of Three Perspective Views and its Associated Tensors" *"IEEE" 920-925* 1995

[7] O.D. Faugeras. "Stratification of three-dimensional vision: projective, affine and metric representations" *Journal of the Optical Society of America, 12(3):465–484* 1995.

[8] Schweitzer, H et al. "Structure from Multiple 2D Affine Correspondences without Camera Calibration" *"IEEE" 258-263* 1996

Contents

Chapter 1

Introduction

The need for coherent and consistent representation of 3D models are required by a huge number of applications nowadays such as modeling, simulation, visualization and CAD/CAM. The visualization hardware has been evolving extremely fast which causes an urge for highly detailed and realistic models. Even though accurate and powerful modeling tools are widely available, the time it is required for creating such complex scenes is quite high, and thus tremendously increases the costs for this kind of activities.

Almost the majority of modeled objects are virtual representations of real objects in our day-to-day life. If one could automatically extract the 3D depth information contained in any particular scene that many are interested in just by taking a few pictures of it, a lot of expensive and time consuming modeling would be saved.

Extracting 3D depth information from multiple pictures in not an easy task. Camera calibration, image correspondence, 3D reconstruction are some of the issues that concern researchers.

Two techniques have been more widely used instead of others: Passive and Active acquisitions. The former one is more flexible and not as expensive as the later, but it is more computationally expensive. The later one is based on controlled lighting environments such as direct lighting and laser beams and, therefore, restricts the applicability while it simplifies the problem.

There is an important demand for flexibility in acquisition. So, calibration and restricted environments should not be used or restricted to a minimum. Many new applications require robust low cost acquisition systems that stimulate the use of consumer photo/video cameras.

Chapter 2

Historical Overview

Extracting depth information from pictures has been investigated for many years now. One the first works related to 3D depth extraction for model reconstruction was first introduced by John Aloimonos [1] in early 1989. He proposed a method for extracting particular features from a sequence of images and estimates its motion.

As this paper suggests, at first researchers in this area were more focused on using 3D depth information from image sequences for robot vision. Visual inspection is required for deciding wether a robot is able to reach a world object or not and consequently point containment problems arise. Therefore, visual quality is not the main concern here. However, nowadays the emphasis is changing. More highly accurate 3D model representations of real world objects are required in computer graphics, virtual reality and communications. While visual quality was left in background back then, now it is a major concern.

Another technique is based on shadow information extracted from pictures that is used by Bouget and Perona [2]. Proesmans et al. proposed another approach based on using pre-defined grid projection on the scene that is to be recovered. These two approaches are considered state-of-the-art techniques using active reconstruction.

Using the passive reconstruction approach, the goal is to increase the automation for reconstruction techniques. The end result would be a fully reconstructed scene or object by simply moving a camera around it.

One of the systems using these methods are *Facade* proposed by Debevec et al. [3]. This approach uses an approximate model as a start.

Another process recently developed was the one introduced by Marc Pollefey [4]. Using off-the-shelf cameras, his method consists of taking pictures around the object to be modeled. No calibration is required, which makes this system one of the most read to use and portable so far.

Chapter 3

Work Plan

This is the work plan for the class *MAC5701 - Tópicos em Ciência da Computação*. It is based primarily in scientific articles, thesis, tutorials and papers published in the past few years over the subject. Most of the articles and papers were selected and indicated by Professor Carlos Hitoshi Morimoto and Professor Antonio Elias Fabris. Some of them were selected by me and approved by the mentioned professors. These references are the knowledge base for my Master Thesis to be written.

During this semester a few papers are intended to be read. These are the papers:

1. A Multi-frame Approach to Visual Motion Perception by Aloimonos [1]

2. Projective Structure from Uncalibrated Images: Structure from Motion and Recognition [5]

3. Tutorial on 3D Modeling from Images [4]

4. Projective Structure from Uncalibrated Images: Structure-from-Motion and Recognition [5]

5. Trilinearity of Three Perspective Views and its Associated Tensors [6]

6. Stratification of three-dimensional vision: projective, affine and metric representations[7]

7. Structure from Multiple 2D Affine Correspondences without Camera Calibration[8]

$Week$	$3 - April$	$4 - April$	$1 - May$	$2 - May$	$3 - May$
$Paper$	$Ref1$	$Ref2$	$Ref3$	$Ref4$	$Ref5$

$Week$	$4 - May$	$1 - Jun$	$2 - Jun$	$3 - Jun$
$Paper$	$Ref6$	$Ref7$	$Ref8$	$Report$

These articles, papers and tutorials basically cover the main subjects involved in my Master Thesis. The paper by Aloimonos is considered to be one of the first on this subject. It introduces an algorithm for estimating motion from a sequence of images.

A series of papers by Amnon Shashua is also included. Shashua has been developing in the past few years state-of-the-art algorithms for recognizing structure from motion with uncalibrated systems, which in my opinion, is the natural way do it.

Olivier Faugeras has also been presenting several important papers for last few years. This particular one I am reading discuss the ways for representing the geometry of three-dimensional space when viewed by a system of cameras.

The tutorial from Marc Pollefeys is a great beginning for 3D Model acquisition with uncalibrated cameras. In the first few chapters, he introduces the mathematical notation to be used throughout the tutorial and after that each step for acquiring the 3D Model is explained.

When all the readings are finished, my knowledge will have improved quite a lot, providing a more solid foundation for the *Thesis* and further developments accompanied by it. For the duration of this course, these papers are going to be read and a brief description of each one is going to be written. These summaries will be the final *Report of Activities* to be delivered by June 25.