Calibration of a Heterogeneous Network of Color and Depth Cameras

Idea for Canesta Vision Contest, Dec 2004

Sudipta N. Sinha { ssinha@cs.unc.edu}
Dept. of Computer Science, CB #3175, Sitterson Hall,
University of North Carolina at Chapel Hill,
Chapel Hill,NC 27599, Ph - (919)962-1885, (919)824-2647.

Abstract

The goal of this project is to investigate a practical method to recover the calibration of a heterogeneous network of conventional color and depth-sensing cameras. While a single depth camera provides depth measurements in a scene making innovative applications possible, we foresee even more powerful applications which will use a combination of measurements from multiple sensors. The heterogeneous camera network consisting of multiple color and depth sensors could be used in an integrated fashion as a single measuring device in 3D, which measures both depth and color in a coordinate system common to all the sensors. Applications of such camera networks could range from accurate 3D reconstruction of scenes, active 3D tracking for automated surveillance or recovering 3D measurements and locations from archived surveillance video footage for analysis of scenes and events.

1. Introduction

We refer to an arbitary configuration of cameras with overlapping viewing volumes, placed in a common environment as a *camera network*. Typical examples of camera networks are groups of indoor security cameras in buildings, many cameras recording stage events or sports action, or a set of web-cams or security cameras placed outdoors or in urban environments like parking lots. The real power of such camera or sensor networks can be reaped once the measurements of individual devices can be combined and represented in an integrated geometric reference frame. The calibration of all sensors in a camera network provides this vital information which will make all kinds of interesting application possible as cameras become more widely deployed in our day to day environment.

Conventional cameras which record intensity and color images typically rely on geometric feature detection and extraction as a key step in the calibration procedure. Such features which could be images of corners, straight lines, elliptical patches or simply small texture patch from images can only be captured in color encoded images. Depth camera neither records intensity nor color, hence such geometric features are simply not recorded by these cameras. As a result typical feature based calibration methods are ruled out for calibrating depth cameras. However a depth camera is good at measuring some special geometric information, something that a conventional camera cannot directly measure; these are depth discontinuities. Depth discontinuity edges constitute pixels where the depth

function undergoes a discontinuity due to occlusion in the scene. Depth discontinuity curves in images typically include silhouette edges and thus silhouettes of foreground objects can be reliably extracted from a depth camera especially if a depth image of the background scene is also available. In fact the foreground silhouette extraction problem is easier to solve using depth images rather than conventional color images.

We recently proposed an algorithm for recovering the full calibration of a camera network from processing silhouettes of dynamic objects from video sequences. The motivation behind this approach was a need for a practical calibration method which did not require physical access to the scene, or an offline phase of pre-calibration using calibration patterms or LEDs, which is used in most conventional calibration approaches. Moreover we wanted our approach to be robust to image quality and the photometric settings of different cameras, and their relative position and orientation. Our approach has these desirable properties. Since it is hard to extract high quality silhouettes from video sequences with complex backgrounds, we also designed our silhouette based calibration algorithm to be robust to noisy silhouettes.

It turns out that the approach that was developed for conventional cameras, holds great promise for depth cameras as well, since depth cameras can reliably provide silhouettes of foreground objects due to their inherent depth sensing capabilities. In fact a hybrid network of cameras could be conceptualized, where some of the cameras could be color cameras and others could be depth cameras. Since color and depth cameras sense complimentary information (intensity and color respectively), calibrating such groups of cameras is a difficult task. The idea we are proposing here is an extension to our calibration algorithm that will allow a depth camera to be seamlessly calibrated with respect to another depth camera or a color camera and vice-versa. Figure 1 shows a conceptual overview of a calibrated hybrid camera network which uses a combination of 4 cameras to sense depth and color from multiple viewpoints in an arbitrary scene.

2. Background

In [1], we presented a method to calibrate a camera network from synchronized video sequences of a dynamic object using only its silhouettes. This was done by robustly computing the epipolar geometry ie. the 2-view geometry from a pair of synchronized video sequences. This was done repeatedly for many view-pairs. The full metric calibration was then recovered once the epipolar geometry of a sufficient number of camera pairs were known.

3. Applications

We have presented an idea for a camera calibration approach that works with both conventional cameras as well as depth cameras. This holds great promise in the development and practical deployment of heterogeneous camera networks which consist of both color cameras as well as depth sensors. The advantages of recovering joint color and depth information from a scene are plenty. While the depth cameras produce explicit 3D depth measurements from a particular viewpoint, other color cameras can augment these depth measurements with color once the measurements are collected in a common 3D coordinate system.

We plan to use the calibration recoverd using the proposed approach to build a prototype system that captures color and depth samples independently from the scene but reconstructs a colored 3D point cloud from color and depth images captured by the system. Such reconstructed point clouds could be used in 3D-video applications or immersive viewpoint-free video based rendering applications.

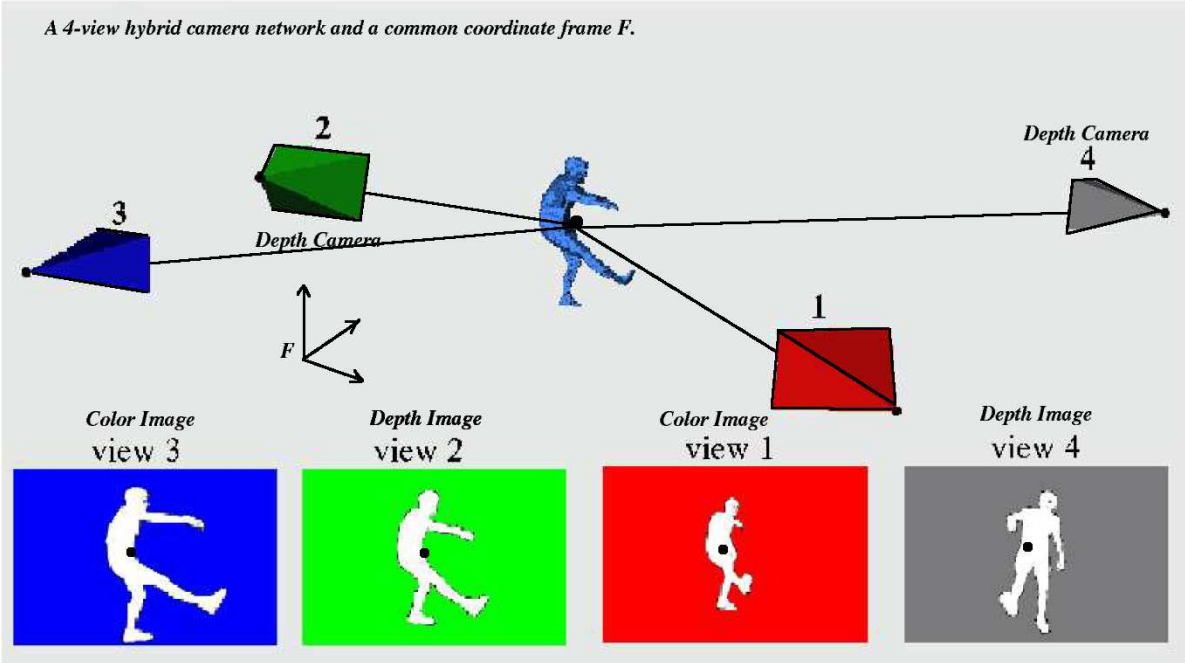

Figure 1: A conceptual heterogeneous camera network, where some cameras provide color images while other 3D depth sensors, measure depth in the scene. The four black pixels record independent color and depth from different viewpoints with different uncertainties. Geometric calibration will allow us to combine these measurements in a common coordinate frame F. This could be used in solving various hard computer vision problems: (1) Reconstruction of Colored 3D Points. (2) Resolving complex visibility and occlusion events. (3) Improving the accuracy of depth estimates from a depth camera by depth and color images from different viewpoints.

Multiple depth cameras could work together to recover more complete shape models of objects rather than a single depth camera which would suffer from occlusions and a lack of sufficient visibility. Multiple depth images of the object from different viewpoints would fill in holes created due to occlusions.

We feel the combined use of multiple color and depth cameras as a single 3D measuring device holds great promise for the future with applications in surveillance and 3D-digitizing events from video footage. We feel our calibration approach has great potential for making such applications possible and for making camera networks practical to use and easy to deploy.

References

[1] S.N. Sinha, M. Pollefeys and L. McMillan, "Camera Network Calibration from Dynamic Silhouettes", *In IEEE Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004, vol 1, pp 195-202, June 2004.*