

UNIVERSITY OF KENTUCKY

MASTER THESIS

**Universal Real-Time XYZ rectified
Reconstruction**

Author:
Sen Li

Supervisor:
Daniel L. Lau

May 30, 2016

Chapter 1

Introduction

3D reconstruction aims to reproduce the 3D profile of real objects as accurate as possible, which require accurate X/Y/Z coordinate values in three dimensional space for every single point of a profile. Since the Kinect brought low-cost depth cameras into consumer market, with PrimeSense 3D sensing technology as the core depth determination principle for its first generation, great interest has been invigorated into RGB-D sensors. In the mean time, optical and perspective distortion become a problem that stops from getting a good view. On most wide angle prime lenses and many zoom lenses with relatively short focal lengths, especially cheap low quality lenses, barrel distortion would typically be present.

In this research, a novel method is proposed for accurate real-time rectification and 3D reconstruction of universal RGB-D cameras.

1.1 3D Reconstruction

Three dimensional (3D) profile measurement technologies have been developed by various means, as summarized by Curless and Seitz [1], among which the non-contact optical methods are widely applied into reality as consumer RGB-D camera. Within the non-contact optical category, as well as 3D reconstruction using multiple images, there are two levels of distinctions[2], as shown in the 3D profile acquisition taxonomy diagram is given in Figure 1.1.

Figure 1.1

The first: active methods and passive methods. Their classifications are decided by the control of light sources. Active methods need special light sources control as part of the strategy to get 3D information, while on the other hand, passive techniques could work with whichever reasonable available ambient light. With a special known illumination offering more information to simplify some of the steps for 3D information acquiring process, active methods tend to be computationally less demanding. Both of the famous consumer PrimeSense and Kinect V2 3D cameras, which are calibrated by the new proposed approach, are using active methods.

The second: single-vantage points methods and multi-vantage points methods. The second distinction is determined by the number of vantage points. With a single vantage system, reconstruction is done based on single view point. In the case that there

are multiple viewing or illumination components, all of them would be positioned very close to each other so that they could ideally coincide. For multi-vantage points methods, several viewpoints, with possible controlled illumination source positions, are involved. As contrast with the single-vantage points methods, the multi-vantage systems need the different components to be positioned far enough from each other.

Among those non-contact optical methods, structured-light and time-of-flight methods are of the most practical importance. As will be discussed shortly, the PrimeSense technology and SeikowaveLCG camera use Structured light methods, and the Kinect V2 camera uses Time-of Flight.

Structured Light

Structured light (SL) based techniques are famous for its fast speed. It is composed of one camera and one light pattern projector[3]. The projector projects a series of special known patterns onto a target, and the camera captures the corresponding images, which contain special information corresponded to the patterns from the projector. A decoding algorithm would be used to extract world coordinate information of the target object from the captured images, by analyzing the relationship among the camera, the projector and the target object using triangulation.

Being after accuracy, the most important issue for structured light method comes to the question of, how to design the projected patterns. In other words, how to design the coding algorithm and its corresponding decoding strategy will decide the final quality of the reconstructed 3D profile. Various classified SL pattern strategies have been proposed, and are still being studied.

PrimeSense SL speckle pattern

The PrimeSense 3D camera uses an infrared projector to project an infrared speckle pattern onto a target , as shown in figure 1.2,

Figure 1.2

and an infrared camera to capture images of the target. By comparing part by part to reference patterns, that were captured previously at known depths and stored in the device, the per-pixel depth could be looked up based on the reference pattern that the projected pattern matches best.

After the per-pixel depth data determined from the infrared sensor, the next step would be to correlate to a calibrated RGB data, which will generate a popular unified representation of target's profile: point cloud, a collection of points with XYZ 3D coordinates and RGB color data. What's more, the surface normals of the target's profile are also stored in every single point of the point cloud data.

Seikowave SL Phase Measuring Profilometry pattern

SeikowaveLCG 3D camera consists of a Charge-Coupled Device (CCD) camera and a Digital Micro-mirror Device (DMD) projector. 2D image pattern strategies are always preferred for a fast scan if a is involved. [4][5] And the multi-shot pattern Phase Measuring Profilometry (PMP) strategy was used for its properties of robust and accuracy. With PMP information encoded in the structured light pattern projected onto the target, the CCD camera could capture a series of images that contains PMP informatio. Triangulation analyzing could be used to extract the 3D world coordinates for each points of the target profile, by a determination of the relationships among CCD camera, DMD projector, and the target object. A system configuration of PMP application is given in figure 1.3.

Figure 1.3

PMP method uses either vertical or horizontal sinusoid patterns, which could be described as:

$$(Projector) \text{In}(x, y) = A + B * \cos(2\pi * f * y - 2\pi * n / N) \quad (1.1)$$

where (xp, yp) denotes the coordinates of every single pixel in the projector, In denotes the intensity of the corresponding pixels, Ap and Bp are constants, f is the frequency of sine wave. The subscript n represents the index of phase shift, while capital N is the total number of phase shift.

Time of Flight

1.2 Universal RGB-D camera calibration

2. principles to get D for Z.(structured light illumination; Time of Flight; comparison)

3. From Z to X/Y

1.3 Contributions of this thesis

1.4 Summation

Chapter 2

Real-Time 3D Rectification

What is real-time: solving inverse problem and arctangent

2.1 From Depth to Z-coordinate

2.2 From Z to X/Y

2.3 3D Reconstructions