Thesis or Article

JeffGWood@mavs.uta.edu

April 19, 2016

Contents

Chapter 1

Introduction

Through the development of applications such as augmented and virtual reality, object / scene reconstruction and visual effects, the process of generating images from an arbitrary vantage point can be found in a variety of applications. In this Thesis (or Article) I will discuss various methods for Image Creation from an arbitraryvantage point, which can be accomplished by two main methodologies of Geometric Construction and Image Synthesis. While both methods use stereo correspondance of multiple images, they differ in the way information is stored and used.

Geometric Construction (GC) contains information about the real-world spatial properties (Coordinates in space, Color), thus viewing results are non-constrained in vantage point. Image Synthesis (IS) relies on image properties (pixel displacement) and is thus viewing results are imited in the possible vantage points.

Symbols and Notation

Symbol	Description
$\mathbf{v}$	*Vectors* in *lowercase* bold
$\mathbf{M}$	*Matrices* in *uppercase* bold
$\mathbf{u}$	2-dimensional image coordinate
$\mathbf{x}$	3-dimensional coordinate
$\tilde{\mathbf{x}}$	3-dimensional *homogeneous* coordinate
${}^{A}\mathbf{x}$	3-dimensional coordinate expressed in reference frame A
${}^{C}_{B}\mathbf{M}$	Change from of reference frame B to reference frame C
${}^{D}\mathbb{S}$	Spacial reference frame D

Chapter 2

Backround

2.1 Stereo-vision

Stereo-vision refers to the extraction of 3-dimensional information from multiple viewpoints. The advantage of stereo-vision over *monocular-vision* involving only single images is that quantities such as size and position can be calculated based on the position of features common to both images. Though it can involve multiple cameras resulting in multiple viewpoints, the majority of research is limited to only 2-view principles as they can be extened 3-views or more.

NOTE: When discussing concepts and principles specific or limited to only 2 view points we will prefix terminology with *stereo-* and when discussing topics that apply to any number views we will use the prefix with *mutli-*.

Change of Reference

Stereo-vision often involves expressing 3d points from different frames of reference (traditionally referred to *left* and *right*) in a single reference frame. As such it is necessary to be able to express coordinates in a given reference frame in any other reference frame.

Coordinates given in ${}^{R}\tilde{\mathbf{x}}$ can be expressed in ${}^{L}\tilde{\mathbf{x}}$ by the geometric transformation:

$$ {}^{L}\tilde{\mathbf{x}} = \left[\begin{array}{c|c} \mathbf{R} & \mathbf{t} \\ \hline 0 & 1 \end{array}\right] \cdot {}^{R}\tilde{\mathbf{x}} \tag{2.1} $$

where ${}^{L}_{R}M$ is also the geometric transformation necessary to transform ${}^{L}\mathbb{S}$ into ${}^{R}\mathbb{S}$.

Fundamental Matrix

Intrinsic Calibration Matrix

Essential Matrix

Chapter 3

Process

The system in question contains 3 main components

1. Image Acquisition System
   - Webcam / Kinect set-up
   - If Webcam should also contain Image-Processing module for:
     - Feature Identification
     - Point-correspondance
     - Sub-Pixel interpolation
2. Point Cloud Processing
   - Should take inputs
   - Should produce point-clouds as one of the output
   - (Possible) Options for Surface Reconstruction include:
     - Calculation of surface Normal through PCA
     - Mesh construction through Delaunay trianglulation
     - Parametrization of Bezier surface through linear-least squares.