Enhancing 3D reconstruction using Mobile Sensors Data

LaTeX

Krzysztof Wrbel

CollegeOrDepartment

University

A thesis submitted for the degree of

PhilosophiæDoctor (PhD), DPhil,..

year month

1. Reviewer: Name

2. Reviewer:

Day of the defense:

Signature from head of PhD committee:

Abstract

Put your abstract or summary here, if your university requires it.

To ...

Acknowledgements

I would like to acknowledge the thousands of individuals who have coded for the LaTeX project for free. It is due to their efforts that we can generate professionally typeset PDFs now.

Contents

List of Figures

List of Tables

GLOSSARY

Glossary

DAPI 4',6-diamidino-2-phenylindole; a fluorescent stain that binds strongly to DNA and serves to marks the nucleus in fluorescence microscopy

DEPC diethyl-pyro-carbonate; used to remove RNA-degrading enzymes (RNAases) from water and laboratory utensils

DMSO dimethyl sulfoxide; organic solvent, readily passes through skin, cryoprotectant in cell culture

EDTA Ethylene-diamine-tetraacetic acid; a chelating (two-pronged) molecule used to sequester most divalent (or trivalent) metal ions, such as calcium (Ca^{2+}) and magnesium (Mg^{2+}), copper (Cu^{2+}), or iron (Fe^{2+} / Fe^{3+})

1

Introduction

Mobile and wearable devices are becoming more and more popular. Modern smartphones despite having extremely good camera's also use advanced sensor's, like Accelerometers, Gyroscope, Magnetometer, Barometer etc.. There is also a big need and growing market of Augmented Reality (AR) and Virtual Reality(VR). That's why image analysis and recognition, as well as 3-D reconstruction techniques are really hot topic. Unfortunately algorithms that support these techniques are very time and memory consuming, that's why it's really hard to run them on mobile devices, which have many limitations in terms of CPU speed and RAM memory capacity.

1.1 Purpose of this thesis

Author of this document will present the reader with an overview of the idea of 3D reconstruction. This thesis also inculdes brief description of related research in this area. After short analysis of efficiency, accuracy and common problems of few chosen algorithms this thesis will propose their enhacment with data acquired with sensors, which can be found in smartphones. At the end author presents evaluation and discuss test results. TODO finish

1.2 Scope

The author researched, how Accelerometer, Gyroscope and Magnetometer can be used in order to improve Fundamental, Essential matrix and also relative Pose Estimation.

Unfortunately raw data sensors are really noisy and it's really hard to use them individually to enhance reconstruction. However there is a way to combine these data together in order to compensate error of each individual sensor. The term describing this process is called "Sensor Fusion". This data fusion allows to estimates in real time a relative or global(in term of earth magnetic field) rotation and translation of the device. TODO finish

1.3 Initial assumptions

The general process of 3-D reconstruction is quite broad, that's why the author of the thesis focus only on certain aspects of this topic. That's why author didn't write algorithms from the scratch, but built his algorithms on top of OpenCV library and "Relative Pose Esitmation" Open-source project setuped by .... In terms of sensor fusion, currently state of art approach is used by most of big Mobile Operating Systems(Android, iOS, Windows Phone). That's why author used Sensor Fusion API from API and only wrote what's needed it terms of getting rotation and translation of the smartphone camera, when acquiring images for his research. TODO finish

1.4 Thesis Outline

In Chapter 2 something something and so on
In Chapter 3
In Chapter 4
In Chapter 5
In Chapter 6
In Chapter 7
In Chapter 8

2

Fundamentals

In order to help user understand topics mentioned in this thesis, small theoretical background is needed.

2.1 3-D reconstruction in general

Today we have many devices, which are capable of 3D reconstruction. Most popular Kinect(**?**) is capable of real-time 3D cloud point generation, but that's very special case, because it uses 2 camera: RGB camera and IR depth-finding camera.

2.1.1 Feature extraction and correspondence matching

2.1.2 Fundamental & Essential Matrices

2.1.3 Triangulation

2.1.4 Common problems

2.2 Structure from Motion

2.2.1 Relative Pose Estimation

2.2.2 Homography estimation

2.2.3 Projective Factorization

2.2.4 Bundle Adjustment

2.2.5 Common problems

2.3 Mobile Sensors overview

2.3.1 Accelerometer

2.3.2 Gyroscope

2.3.3 Magnetometer

2.3.4 Sensor Fusion

3

Related Work

There are many approaches to the problems, from raw one by one pixel analysis to high level abstraction of objects, light and shadows estimation[some references to each].However this thesis does not focus on high-level abstraction reconstruction, but focuses on refining relative poses estimation steps. In order to estimates especially first two relative positions of cameras essential matrix must be decomposed. Since basic epipolar geometry equation many scientist introduced a way to solve this linear problems, where the main differences were in terms of accuracy and speed. One of the first was 8-point algorithm[reference], which can be used to compute a fundamental matrix. This is done without any prior knowledge of the scene, as well as camera parameters. Still to find relative position knowledge of internal camera parameters is needed in order to calculate and decompose Essential Matrix. Later on 5-point algorithm approaches were introduced[references], which needed prior knowledge of internal camera parameters. David Nister in his paper shows that 5-point outperforms almost all similar algorithms in terms of accuracy and speed. Only 8-point algorithm can be competitive, when it comes to forward image sequences. One of the state-of-art real-time and roboust approaches is iterative 5-pt Algorithm created by Vincent Lui and Tom Drummond[].

Most of algorithms are very sensitive to presence of outliers. One of the most common approach is to use of RANSAC modelling[refining estimates], where iteratively subset of data is chosen to find a solution and then other points are checked, if they also satisfy equation with calculated solution.

Research group from Technishe Univeristt Berlin made an comparison and evaluation of methods, which were published at that point. It turned out that estimation of camera rotation is much more stable than translation. Also there are a lot of ambiguities in terms of choosing the correct solution of epipolar geometry equation.

In certain situation where external camera parameters as rotation and translation can be measured more accurate algorithms were proposed. In 2011 D. Scaramuzza from Zurich proposed a 1-point algorithm[reference], which shows how to describe and use model of camera mounted on a car to enhance 3D reconstruction. Introduced in 2013 4- point algorithm, which uses information of rotation angle in certain axis from additional sensor as show in paper[reference] can outperform even some versions 5-point algorithm. Lately scientists are creating more complex models to estimates relative stereometry. For instance group from Zurich proposed a way to enhance reconstruction with additional 6DOF sensor [Robust Real-Time Visual Odometry with a Single Camera and an IMU].

There are also different approaches like [Line-Based Relative Pose Estimation], where it's shown how to estimates the relative pose from 3 lines with two of the lines parallel and orthogonal to the third. Very accurate estimations also can be achieved when there is no camera rotation[Epipole Estimation under Pure Camera Translation*]. All these references shows that enhanced models help to achieve often faster more accurate solution.

Accuracy and speed are very important, when it comes to create systems capable of augmenting our reality. One of first successful systems for such situations were proposed by research group[PTAM]. They showed how two simultaneously working threads can be used to both create model of environment and use this knowledge to apply graphical effects to objects presented on stream camera video. Also some of these concepts where applied already even to robotic vision. Authors showed efficiency of proposed system for robot walking in cluttered indoor workspaces[MonoSLAM: Real-Time Single Camera SLAM].

The most important things, which can be concluded are rotation estimation is more stable than translation estimation and the more well described model of setup.

4

Concept

As indicated in similar research it's very attractive to use additional data to enhance reconstruction and reduce ambiguity in finding correct solution of 3D reconstruction. It also helps to achive faster, more stable and robust algorithms. This thesis will show how prior knowledge of rotation or translation can be used to faster process 3D reconstruction of series of images. However there are many algorithms, which relay on accuracy of additional rotation and translation data. In reality especially, when it comes to hand-held smartphones, collected data are very noisy and unstable. That's why this thesis also proposes enhancements of most popular algorytmic approaches, when noisy data are used.

4.1 Requirements

Proposed methodology needs as the input series of images with additional information about position of the camera - euclidan rotation and optionally translation. Usage of smartphone is actually not necessary. Any camera with SensorFusioned accelerometer and gyroscope(magnotometer is optional and as discussed in 2.5??? has its up and downsides) capable of storing pictures and sensor data can be used. During algorithm runtime either both rotation and translation informations are used or just rotation, which as indicated in ....??? is less noisy than translation estimation. Internal camera parameters need to be calculated before reconstruction process is began. Additional sensor data can be unaccurate and noisy.

4.2 Enhancing fundamental equation

Taking standard fundamental geometry equation and relative camera based system (P = [I—0], P' = [R—t]) we can note that:

$$x'^T * K^{-T} * [T]_x * R * K^{-1} = 0 \tag{4.1}$$

It's also good to note that:

$$[T]_x = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} where T = [t_x, t_y, t_z] \tag{4.2}$$

As we were discussing both rotation and translation can be distorted with noise. This can be written as:

$$R = R_{init} * R_{error} \tag{4.3}$$

where R_{init} is rotation matrix from measured angles and R_{error} is rotation matrix of angles errors. Analogically

$$[T]_x = [T_{init}]_x * [T_{error}]_x \tag{4.4}$$

4.2.1 Rotation enhancements

4.2.2 Rotation & translation enhancements

4.3 Enhancing essential equation

4.3.1 Rotation enhancements

4.3.2 Rotation & translation enhancements

4.4 Pose estimation

4.4.1 Rotation enhancements

4.4.2 Rotation & translation enhancements

4.4.3 Alternative 3-point algorithm for translation finding

4.5 Reconstruction process strategy

Description of steps in whole reconstruction pipeline depending on which data are present and what accuracy and convergence speed we need. Approaches: 1) Known rotations and translations -¿ feature finding and outliers removal -¿ triangulation/pose estimation -¿ with or without BA (to reduce outliers further) 2) Noisy rotations and translations -¿ feature finding and outliers removal -¿ with or without BA -¿ super convergance 3) Noisy rotations and translations -¿ feature finding and outliers removal -¿ dR and dT estimation from modified essential decomposition -¿ with or without BA -¿ better accuracy and robustness 4) Noisy rotation -¿ feature finding and outliers removal -¿ translation from essential decomposition or pose estimation -¿ with or without BA -¿ better accuracy and robustness 5) Noisy rotation -¿ feature finding and outliers removal -¿ translation from essential decomposition or pose estimation -¿ with or without BA -¿ better accuracy and robustness 6) Known rotations -¿ feature finding and outliers removal -¿ Alternative 3-point algorithm for translation finding -¿ triangulation/pose estimation -¿ with or without BA (to reduce outliers further)

5

Implementation

5.1 Environment

5.2 Project Structure

5.2.1 Android application

5.2.2 OSX CMake base project

5.3 Important Implementation Aspects

5.3.1 Custom Sensor Data File format

5.4 Graphical User Interface

5.5 Documentation

6

Evaluation

6.1 Test Environment

6.1.1 Synthetic Data

6.1.2 Real world Data

6.2 Usability

6.3 Performence

6.3.1 Accuracy comparison

6.3.2 Time comparison

7

Conclusion

7.1 Summary

7.2 Dissemination

Who uses your component or who will use it? Industry projects, EU projects, open source...? Is it integrated into a larger environment? Did you publish any papers?

7.3 Problems Encountered

7.4 Future work

8

Materials & methods

References

Declaration

I herewith declare that I have produced this paper without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This paper has not previously been presented in identical or similar form to any other German or foreign examination board.

The thesis work was conducted from XXX to YYY under the supervision of PI at ZZZ.

CITY,