Augmented Reality System with Planar Homographies using Building Façade

Nan-Hung Cho (Jack)

Master of Engineering (Research)

University of Technology Sydney
Faculty of Engineering and Information Technology

School of Electrical and Data Engineering

August 2017

Principle Supervisor: A/Prof. Qiang Wu

Co-Supervisor: A/Prof. Jian Zhang

Acknowledgement

My two years in University of Technology Sydney have been an apprenticeship under my principle supervisor, Qiang Wu. I would like to express my sincere gratitude for his continual guidance, technical consultation and discussion during my research period. His guidance and insight have shaped this thesis, and I remain deeply grateful for everything that he had taught me.

I would like to thank my co-supervisor, Jian Zhang, for his help and support while Qiang was absent, and my other colleagues for their help on answering my questions regard to technical issues and related mathematical theory.

Finally, I dedicate this work to my beloved family, who offered unconditional support and their mentor have always been there for me. Thank you so much.

Table of Content

List of Figures and Tables

List of Acronyms

Acronym	Definition
AKAZE	Accelerated KAZE
AR	Augmented Reality
CMR	Computer Mediated Reality
FED	Fast Explicit Diffusion
FPS	Frame per second
HMD	Head-mounted display
kNN	K-nearest Neighbour
LDB	Local Difference Binary
M-LDB	Modified-Local Difference Binary
NNDR	Nearest Neighbour Distance Ratio
POI	Point-of-interest
RGB	Red, Green, Blue
SIFT	Scale-invariant Feature Transform
VR	Virtual Reality
KAZE	KAZE feature as described in [1]

Abstract

In recent years, there has been widespread adoption of AR technology with various applications in the urban environment. The essence of an AR application includes the estimation of camera and object placement, which is still a challenging research problem. The first issue is that the structure of the real-world environment is usually very cluttered and involves many unforeseen uncertainties which cannot be generalised as a model. Secondly, the reprojection of the real-world scene to the camera's viewpoint flattens 3D information down to a pixel level, which causes what is known as perspective projection.

Fortunately, it is possible to recover the pixel level of the camera pose via planar homography. Methods commonly used in typical AR applications to recover the homography are based on feature-point matching, but many of them still suffer from the problem of perspective and camera movement.

By taking characteristics of the urban environment, such as buildings which are widely seen on the street view, we propose a simple, yet efficient, approach to recover the camera pose from a single image. Our approach exploits the building façade of rectilinear structures portrayed by man-made structures (e.g. windows and brick structure). Using lines from these structures, our proposed algorithm can produce a rectified homography for each detected building façade, so we can then recover the fronto-parallel view of each building façade. Since these planar homographies are invertible, we can use them to recover the camera pose of the building façade for the object emplacement task. Finally, we implemented a tracker-based AR application that uses the idea of a fronto-parallel view that results in better matching effectiveness.