

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

ABSTRACT	1
1. INTRODUCTION	2
1.1. OVERVIEW.....	2
1.2. BACKGROUND INFORMATION	3
<i>Proprioceptive sensors</i>	7
<i>Computer vision technique</i>	7
<i>Hybrid approaches</i>	8
1.3. PROBLEM STATEMENT.....	9
1.4. PUBLICATION LIST	11
1.5. STRUCTURE OF THE THESIS	11
2. MATHEMATICS PREREQUISITE	12
2.1. GEOMETRIC TRANSFORMATION AND HOMOGRAPHY.....	12
2.2. PROJECTIVE SPACE AND HOMOGENOUS COORDINATES	13
2.3. CAMERA PROJECTION MODEL.....	15
<i>Intrinsic Matrix</i>	16
<i>Extrinsic Matrix</i>	17
<i>Complete camera model</i>	17
3. RELATED WORKS	20
3.1. PROPRIOCEPTIVE SENSOR-BASED AR	20
3.2. COMPUTER VISION-BASED AR	21
<i>Marker-based AR</i>	21
<i>Markless-based AR</i>	23
3.3. PLANAR RECTIFICATION.....	25
3.4. LITERATURE REVIEW SUMMARY.....	28
4. PLANAR FAÇADE RECTIFICATION	29
4.1. LINE SEGMENT DETECTION	31
4.2. GAP FILLING AND LINE SEGMENT EXTENSION	32
<i>Gap-filling algorithm</i>	32
<i>Line segment extension algorithm</i>	36
4.3. ORTHOGONAL PAIR MATRIX GENERATION	39
4.4. DOMINANT PLANE RECTIFICATION	40
<i>Recovery of a façade view via RANSAC</i>	40
<i>Pure rotation recovery</i>	43
4.5. EXPERIMENT	44

	<i>Experimental setup</i>	45
	<i>Experimental data</i>	45
	<i>Performance measurement</i>	46
	<i>Experimental Results and Discussion</i>	48
5.	FAÇADE-BASED AR	54
5.1.	OVERVIEW.....	54
5.2.	AKAZE.....	55
5.3.	TRACKER IMPLEMENTATION.....	57
	<i>Feature detector threshold adjustment</i>	57
	<i>kNN + NNDR</i>	58
	<i>Symmetry Test</i>	59
	<i>Homography estimation + outlier removal</i>	59
5.4.	FRONTO-PARALLEL VIEW VIA INTER-FRAME HOMOGRAPHY	59
5.5.	TRACKER RESULT VERIFICATION AND REFINEMENT	60
	<i>Problems of RANSAC in homography estimation</i>	60
	<i>Analysis of rigid body transformation via a homography</i>	61
	<i>Kalman Filter</i>	62
5.6.	EXPERIMENTAL RESULTS	65
	<i>Descriptiveness evaluation</i>	65
	<i>Tracker effectiveness</i>	67
5.7.	DISCUSSION AND CONCLUSION	67
6.	CONCLUSION	69
6.1.	FUTURE WORKS	70
7.	REFERENCE	71

List of Figures and Tables

Fig. 1 A man is fitting hot-desks using a tablet device	3
Fig. 2 Taxonomy of CMR application	4
Fig. 3 A typical environment setup for a VR game	5
Fig. 4 – A house is rendered onto a marker	8
Fig. 5 A virtual creature “Yoshi” is rendered on the office desk [9]	9
Fig. 6 – A rectangle in projective space	13
Fig. 7 – A sequence of images shooting at the same building under different camera poses.....	14
Fig. 8 – An example of lens effects. In this thesis, rectilinear lens is assumed.	15
Fig. 9 – Some examples of AR in an urban environment	23
Fig. 10 – Perspective rectification of a building façade.....	26
Fig. 11 – An overview of the planar rectification algorithm.....	30
Fig. 12 – Illustration of the line segment detection algorithm.	31
Fig. 13 – Illustration of gap filling and the line segment extension algorithm.	32
Fig. 14 – Result of line segment and gap-filling algorithm. The red lines are the gap-filling result from the original line segment result, indicated by the green colour.	33
Fig. 15 – An illustration of collinearity testing: (a) is the selected line segment pair to be tested, a line segment is defined by its two endpoints; (b) ~ (j) represent different cases formed by the endpoints of the line segment pair.	34
Fig. 16 – Gap distance testing between line segments.....	34
Fig. 17 – An example of line segments which satisfy the gap distance threshold: (a) will be merged together; while (b) will not be merged because the total length of line segments is smaller than its gap distance.....	35
Fig. 18 – Sample result of line segment extension. Red lines are the result of line extension of the original ones shown in red colour.....	36
Fig. 19 – Point P_x lies on line segment $P_{1,2}$, can be represented using parametric equations.	37
Fig. 20 – Intersection of line segments by solving its linear system of parametric equations.	37
Fig. 21 – An illustration of a line segment, $L_{1,2}$ extended to its nearest-neighbour line segments.....	39

Fig. 22 – An illustration of a sample orthogonal pair matrix. Each element in the matrix indicates if a pair of line segments by its index is orthogonal. Notice that the diagonal elements are always 0 since comparing to the line segment is invalid.	40
Fig. 23 –Dominant planar façade classification via RANSAC.....	43
Fig. 24 – Rectification sample images (a) ~ (d) consist of only one face, and (e) ~ (f) are pairs of images with two faces.....	46
Fig. 25 – Sample experimental result of (a) line gap-filling and (b) line segment extension.	48
Fig. 26 – Line segment and gap-filling on a grid image.....	49
Fig. 28 – Sample rectified view result on various buildings’ façades.	50
Fig. 29 – Augmentation results: user manually choose the diagonal points, indicated as red and yellow markers in the original image, so then we can construct a rectangle region by back-projection from the fronto-parallel view. Once the corners are found, a 2D image can be replaced on the region of augmentation.	51
Fig. 30 – Augmentation result.....	52
Fig. 31 - (a) an example of non-planar facade; (b) and (c) are the result of the fronto-parallel view. While it successfully recovers the buildings at the back, it fails to recover the non-planar building facade at the very front.	53
Fig. 32 – (a) indoor tunnel example as an extreme case for planar rectification. While the proposed algorithm managed to recover one side of the wall shown in (b), it does not recover the other side of the wall.	53
Fig. 33 – Framework of our façade-based AR system.....	54
Fig. 34 – Local Binary Difference test illustration.	56
Fig. 35 – Tracker framework	57
Fig. 36 – Descriptiveness evaluation: 1 st row is the image pair in normal view while 2 nd row is in fronto-parallel view;.....	66
Fig. 37 – Descriptiveness evaluation: The vertical axis represents the squared Euclidean distance while the horizontal axis is the threshold value that controls the number of features. The blue line represents the normal view while the orange line represents the fronto-parallel view.	66
Fig. 38 – Experimental results on correspondences matching.....	67

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.