



THE UNIVERSITY
of ADELAIDE

**As-Projective-As-Possible Image Stitching with
Moving DLT**

by

Julio César Hernández Zaragoza

A thesis submitted for the degree of Doctor of Philosophy

in the

Faculty of Engineering, Computer and Mathematical Sciences
School of Computer Science

June 2014

Declaration of Authorship

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree. I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968. The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works. I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

In carrying out the research that underlies this thesis the following papers were published or are currently under review:

1. Julio Zaragoza, Tat-Jun Chin, Michael Brown and David Suter, "As-Projective-As-Possible Image Stitching with Moving DLT", in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Portland, Oregon, USA, June, 2013.
2. Julio Zaragoza, Tat-Jun Chin, Quoc-Huy Tran, Michael Brown and David Suter, "As-Projective-As-Possible Image Stitching with Moving DLT", in *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, November, 2013.
3. Quoc-Huy Tran, Tat-Jun Chin, Julio Zaragoza, Michael Brown and David Suter, "Outlier Rejection in Deformable Registration with Moving Least Squares", in *Transactions on Image Processing (TIP)*, manuscript submitted for review.

Julio César Hernández Zaragoza Signed: _____ Date: _____

“Anybody who has been seriously engaged in scientific work of any kind realises that over the entrance to the gates of the temple of science are written the words: ‘Ye must have faith.’”

Max Planck

THE UNIVERSITY OF ADELAIDE

Abstract

Faculty of Engineering, Computer and Mathematical Sciences

School of Computer Science

Doctor of Philosophy

by Julio César Hernández Zaragoza

The last ten years have witnessed important advances in image stitching algorithms. Such advances have allowed the development of several commercial tools that are based on or incorporate image stitching. Amongst these tools there are well known image editing suites like Adobe Photoshop, Microsoft's Image Composite Editor which is part of the web-based photo organization tool Photosynth, "dedicated" stitching software like Autostitch and its commercial counterparts AutoPano and AutoPano Giga, the image stitching functionality of the iOS from Apple, as well as the built-in stitching functionality of several off-the-shelf digital cameras.

The widespread availability of stitching tools often leads to the impression that image stitching is a solved problem. The reality is: many of these tools often fail to produce convincing results when given *non ideal data*, i.e., images that deviate from fairly *restrictive assumptions* of image stitching; the main two being that the photos correspond to views that differ purely by rotation, or that the imaged scene is effectively planar. Such assumptions underpin the usage of 2D projective transforms or homographies to align the photos. In the hands of the casual user, these conditions are often violated, yielding misalignment artifacts or "ghosting" in the results. Accordingly, many existing image stitching tools depend critically on post-processing routines to conceal ghosting.

This thesis proposes a novel estimation technique called Moving Direct Linear Transformation (*Moving DLT*) that is able to "tweak" or fine-tune the projective warp to accommodate the deviations of the input data from the idealised conditions. This produces "*as-projective-as-possible*" image alignments that significantly reduce ghosting without compromising the geometric realism of perspective image stitching. The Moving DLT technique lessens the dependency on potentially expensive post-processing algorithms.

In addition, this thesis also describes how Moving DLT can be performed in a “bundled” manner to simultaneously align multiple images in order to generate “long” panoramas while reducing the error propagation of the incremental stitching techniques. It is important to note that such a bundle adjustment formulation, which we call *Bundled Moving DLT*, is the first of its kind. There is no other bundle adjustment formulation that is able to simultaneously refine multiple non-rigid warps for image stitching.

The experimental results show that Moving DLT (and Bundled Moving DLT) can produce much better results than current state-of-the-art image stitching software and other recent methods for image stitching.

Acknowledgements

There is a lot of people that I would like to thank for accompanying me on the amazing adventure that this PhD has represented. First of all I would like to thank my principal supervisor: Dr. Tat-Jun Chin (TJ). It is no exaggeration to say that this research could not have been possible without the excellent guidance and support from TJ. Thank you so much for the countless hours of discussions, thank you so much for the ideas, corrections, suggestions, modifications and changes, but above all, thank you so much TJ for showing me how *great research* is done. I would not change such experience for the world.

I would also like to thank my co-supervisor Prof. David Suter for offering this amazing and life-changing opportunity, for taking me under his wing, for providing novel insights and giving fantastic advice on some of the initial and unpolished ideas that were presented over the course of three years.

Besides TJ and David, I would also like to thank Dr. Qinfeng (Javen) Shi for all of the help, support and advice he offered me during my candidature, specially during the early days. But in particular, thanks a lot for all of the “non-research-related” comments and suggestions. Thanks a lot for having an open window when all of the other doors were closed.

During my time as a PhD student I performed some teaching activities as well. Such activities improved several aspects of my professional life like self confidence and communication skills, and performing these activities also made me realise how much I enjoy the “teaching experience”. I owe this great opportunity to Dr. Claudia Szabo who has given me this chance for the last 2.5 years. Oh and thanks a lot for the weekly chocs of course!

Also, I would like to thank my lab-mates, Sim, Xue, Huy, Trung, Guosheng and Alvaro for sharing the experience. Thank you guys for the useful discussions, help, tips and advice. Thanks a lot for the amazing and funny trips and meals, but most importantly, thanks a lot for the support and friendship.

Lastly, the most important person in the world that I would like to thank to, is my mom. Thank you so much for all of the love and support you have offered me from day one of my life. Thanks a lot for being my number one fan. Thanks a lot for always giving your best for me, for encouraging me, for helping me, for giving me advice, for always pushing me forward, but most of all, thank you so much for being my mom. *Te quiero hoy y siempre hasta el “delfinito” y de regreso y mil veces más.*

Contents

Declaration of Authorship	iii
Abstract	vii
Acknowledgements	ix
Contents	x
List of Figures	xv
List of Tables	xix
List of Algorithms	xxi
Abbreviations	xxii
1 Introduction	1
1.1 Background and Motivation	1
1.2 Contributions of the Thesis	7
1.3 Thesis Structure	9
1.3.1 Chapter 2	9
1.3.2 Chapter 3	9
1.3.3 Chapter 4	9
1.3.4 Chapter 5	10
1.3.5 Chapter 6	10
1.3.6 Chapter 7	10
2 Background	11
2.1 Introduction	11
2.2 The Homography Matrix	13
2.2.1 Plane Induced Homography	13
2.2.2 Rotation Induced Homography	16
2.3 Direct Linear Transformation (DLT)	18
2.4 Normalised Direct Linear Transformation	20
2.5 Image Stitching with Projective Transformations	22

2.6	Summary	25
3	Related Work	27
3.1	Introduction	27
3.2	Pairwise Stitching	28
3.3	Bundle Adjustment and Local Refinements	33
3.4	3D Reconstruction and Plane-Plus-Parallax	38
3.5	Direct Estimation of Flexible Warps	42
3.6	Arbitrary Camera Motions	47
3.7	Summary	48
4	As-Projective-As-Possible Warps	51
4.1	Introduction	51
4.2	Moving Direct Linear Transformation (Moving DLT)	52
4.3	Efficient Computation for Image Stitching	55
4.3.1	Partitioning into cells	55
4.3.2	Updating weighted SVDs	57
4.4	Comparing Moving DLT against (affine) Moving Least Squares for Image Stitching	58
4.5	Summary	62
5	Simultaneous Refinement of Multiple As-Projective-As-Possible Warps	63
5.1	Introduction	63
5.2	Selecting the Reference Frame	63
5.3	Bundled Moving Direct Linear Transformation (Bundled Moving DLT)	64
5.4	Comparing Bundled Moving DLT and Bundle Adjustment	67
5.5	Summary	69
6	Experiments and Results	73
6.1	Introduction	73
6.2	Comparisons with Flexible Warp Methods	74
6.2.1	Preprocessing and Parameter Settings	74
6.2.2	Qualitative Comparisons	75
6.2.3	Runtime Information	100
6.2.4	Quantitative Benchmarking	100
6.3	Comparisons with Bundle Adjustment	103
6.3.1	Stitching Full Panoramas without post-processing	103
6.3.2	Stitching Full Panoramas with post-processing	107
6.4	Summary	114
7	Conclusions	115
7.1	Future Work	116
7.1.1	Non-Rigid Image Registration	116
7.1.2	Non-Rigid Structure from Motion	118
7.1.3	Video Stabilisation	121
7.2	Drawbacks and Limitations	123

A	Image Sets Used in the Experiments	125
A.1	Images Used in Pairwise Stitching	125
A.2	Images Used For Stitching Full Panoramas	128
B	The Image Formation Process	131
B.1	The Camera Matrix	132
B.1.1	External Parameters	132
B.1.2	Camera Intrinsic	133
B.2	Relating Points in Two Views through the Camera Matrix	134
	Bibliography	137

List of Figures

1.1	An example of a high resolution panorama generated with Autostitch.	2
1.2	One of the panoramas generated by Google’s street views.	3
1.3	Illustrating the stitching problem when the images to align do not follow the prescribed conditions of the projective model.	4
1.4	A magnified (and cropped) view of Fig. 1.2 that shows some of the artifact errors caused by failed post-processing routines in image stitching.	6
2.1	A point in a 3D plane being projected into two image planes.	14
2.2	A homography matrix relating points from a planar view.	16
2.3	A 3D point projected into two views undergoing a pure (3D) rotation motion.	18
2.4	Demonstrating image stitching with the Direct Linear Transformation method. The input images correspond to views that only differ by rotation.	23
2.5	1D analogy of image stitching with projective warps.	24
2.6	Demonstrating image stitching with the Direct Linear Transformation method. The input images correspond to views that differ by rotation <i>and</i> translation.	26
3.1	Stitching long panoramas through incremental (pairwise) approaches.	29
3.2	An example of a panorama for video compression.	31
3.3	The gap closing technique for image stitching.	32
3.4	An example of a panorama generated by means of Brown and Lowe’s Bundle Adjustment method [12].	34
3.5	Example of the panorama recognition process of Brown and Lowe [12].	34
3.6	The bundle adjustment process of Marzotto [58] for panorama creation.	36
3.7	A “joiner” example from [94].	38
3.8	An example of a multi-view panorama.	39
3.9	The 3D reconstruction approach for image stitching of [99].	40
3.10	A video summary created with the plane+parallax approach of [36].	41
3.11	An overview of the Smoothly Varying Affine Image Stitching method from [51].	43
3.12	The hierarchical non-rigid image alignment process of [52].	44
3.13	An example of the content preserving warps for video stabilisation.	45
3.14	The dual homography warps for image stitching from [23].	46
3.15	Two panorama examples generated with the adaptive manifold mosaicking technique of Peleg <i>et al.</i> [65]. This example appears in [65].	47
3.16	The imaging process of the pushbroom camera modeled by the multi-perspective projection process of [62, 64, 66].	48
3.17	An anaglyph image constructed from a stereo panorama generated with the panoramic stereo imaging method of [66].	48
4.1	1D analogy of image stitching with as-projective-as-possible warps.	54

4.2	Demonstrating image stitching with the proposed Moving DLT method.	56
4.3	Histogram of number of weights $\neq \gamma$ for the cells in Fig. 2.6(b).	57
4.4	An affine-as-possible warp obtained by means of Moving Least Squares.	59
4.5	Image stitching results obtained with as-affine-as-possible warps generated with MLS.	60
4.6	Image stitching results obtained with APAP warps generated with Moving DLT.	61
4.7	Feature points extracted from the <i>temple</i> image pair.	61
5.1	Selecting the reference frame for Bundled Moving DLT.	64
5.2	Stitching results obtained with the proposed Bundled Moving DLT approach for panorama creation.	66
5.3	(Two page figure) Comparing the “raw” image alignment results of different techniques for panorama creation.	70
6.1	(Two page figure) Qualitative comparisons on the <i>railtracks</i> image pair.	78
6.2	(Two page figure) Qualitative comparisons on the <i>temple</i> image pair.	80
6.3	(Two page figure) Qualitative comparisons on the <i>bikes</i> image pair.	82
6.4	(Two page figure) Qualitative comparisons on the <i>construction site</i> image pair.	84
6.5	(Two page figure) Qualitative comparisons on the <i>train</i> image pair.	86
6.6	(Two page figure) Qualitative comparisons on the <i>garden</i> image pair.	88
6.7	(Two page figure) Qualitative comparisons on the <i>carpark</i> image pair.	90
6.8	(Two page figure) Qualitative comparisons on the <i>apartments</i> image pair.	92
6.9	(Two page figure) Qualitative comparisons on the <i>chess/girl</i> image pair.	94
6.10	(Two page figure) Qualitative comparisons on the <i>rooftops</i> image pair.	96
6.11	(Two page figure) Qualitative comparisons on the <i>couch</i> image pair.	98
6.12	Point cloud and average RMSE on the training set and the testing set as a function of inter-camera translational distance.	102
6.13	Panorama results without post-processing on the <i>construction site</i> image set.	104
6.14	Panorama results without post-processing on the <i>garden</i> image set.	105
6.15	Panorama results without post-processing on the <i>train</i> image set.	106
6.16	(Two page figure) Panorama results with post-processing on the <i>construction site</i> image set.	108
6.17	(Two page figure) Panorama results with post-processing on the <i>garden</i> image set.	110
6.18	(Two page figure) Panorama results with post-processing on the <i>train</i> image set.	112
7.1	Typical input data for Non-Rigid Image Registration tasks.	117
7.2	Examples of Non-Rigid Image Registration results obtained with Moving DLT	117
7.3	Non-Rigid Structure from Motion results obtained with Moving DLT	120
7.4	An example of the dense 3D reconstruction results of [25]	121
7.5	Comparing single and bundled camera paths for video stabilisations.	122
7.6	The image stabilisation process of Bundled Camera Paths.	123
A.1	<i>railtracks</i> image pair. Size of images: 2000×1500 pixels. Number of inliers after RANSAC is: 2753.	125
A.2	<i>temple</i> image pair [23]. Size of images: 730×487 pixels. Number of inliers after RANSAC is: 415.	126
A.3	<i>carpark</i> image pair [23]. Size of images: 653×490 pixels. Number of inliers after RANSAC is: 359.	126

A.4	<i>apartment</i> image pair [23]. Size of images: 1632×1224 pixels. Number of inliers after RANSAC is: 2634.	126
A.5	<i>chess/girl</i> image pair [51]. Size of images: 1824×1368 pixels. Number of inliers after RANSAC is: 1381.	127
A.6	<i>couch</i> image pair [51]. Size of images: 1824×1368 pixels. Number of inliers after RANSAC is: 1329.	127
A.7	<i>rooftops</i> image from [51]. Size of images: 320×240 pixels. Number of inliers after RANSAC is: 161.	127
A.8	<i>bikes</i> image pair. Size of images: 1632×1224 pixels. Number of inliers after RANSAC is: 2561.	128
A.9	<i>construction site</i> dataset. Size of images: 2000×1329 pixels. Number of inliers between the pair of images used for the pairwise experiments in Section 6.2: 5068. Number of inliers between the images used for the panorama experiments in Section 6.3: 12616.	128
A.10	<i>garden</i> dataset. Size of images: 2000×1329 pixels. Number of inliers between the pair of images used for the pairwise experiments in Section 6.2: 4567. Number of inliers between the images used for the panorama experiments in Section 6.3: 10693.	129
A.11	<i>train</i> dataset. Size of images: 2456×1632 pixels. Number of inliers between the pair of images used for the pairwise experiments in Section 6.2: 4231. Number of inliers between the images used for the panorama experiments in Section 6.3: 13380.	129
B.1	The image formation process through the pinhole camera model.	131
B.2	Example of the simplified camera intrinsic parameters.	133
B.3	Relationship between the projections or pictures of a 3D point in two different images.	135

List of Tables

2.1	Hierarchy of 2D linear transformations in projective space.	12
6.1	Average RMSE (in pixels) and % outliers over 20 repetitions for 5 methods on 11 image pairs. See Figs. 6.1 to 6.11 to view the qualitative stitching results. . .	101

List of Algorithms

1	Direct Linear Transform algorithm for the estimation of a 2D projective warp (homography).	21
2	Normalised Direct Linear Transform algorithm for the estimation of a 2D projective warp (homography).	22
3	Moving Direct Linear Transform algorithm for the estimation of a 2D as-projective-as-possible warp.	54
4	Simultaneous refinement of multiple as-projective-as-possible warps for panorama creation.	67

Abbreviations

2D	Two-Dimensional
3D	Three-Dimensional
APAP	As-Projective-As-Possible
Bundled Moving DLT	Bundled Moving Direct Linear Transformation
CPW	Content Preserving Warps
DHW	Dual-Homography warps
DLT	Direct Linear Transformation
MRF	Markov Random Field
Moving DLT	Moving Direct Linear Transformation
MLS	Moving Least Squares
SLAM	Simultaneous Localisation and Mapping
SVA	Smoothly Varying Affine
SVD	Singular Value Decomposition

To my mom, for a lifetime of love and support.