



Efficient Recursive Factorization Methods for Determining Structure from Motion

Yanhua Li, B.Comp.Sci.

Department of Computer Science

The University of Adelaide

Cooperative Research Center for

Sensor Signal and Information Processing

A thesis submitted for the degree of

Doctor of Philosophy

in the Department of Computer Science

The University of Adelaide

July 2000

Abstract

This thesis addresses the *structure from motion* (SFM) problem in computer vision. We study recursive algorithms for efficient shape and motion recovery at each frame of a sequence of images. In real-time applications where image data is extensive, efficiency of an SFM method becomes very important for estimating shape and motion at each frame. The proposed recursive method in this thesis improves the efficiency, both in computational cost and in storage, of a class of innovative SFM methods — the *factorization methods* (FMs).

Our work in this thesis may be viewed as an extension of the original [71], the sequential [55] and the paraperspective [60] factorization methods. A critical aspect of these factorization approaches is the estimation of the *shape space*, and their computational complexity is dominated by their shape space computing algorithms. If P object feature points are tracked through a sequence of F frames, the shape space updating complexity at each frame is $O(P)$ in the recursive least-squares (RLS) method proposed in this thesis, while that in the sequential FM is $O(P^2)$. In contrast, the batch-mode original and paraperspective FMs, which are not intended to be used frame by frame, compute the shape space at a cost of $O(FP^2)$ after all F frames are tracked. Further, the RLS shape space updating algorithm is an adaptive data driven algorithm. Hence it does not require storage of a large measurement or covariance matrix, while other FMs usually do.

The proposed recursive method uses the general affine camera model, while the original and the sequential FMs assumed an orthographic model. Like the paraperspective FM, we give Euclidean shape and motion recovery from the estimated shape space under one of the three specific affine camera models — orthography, weak perspective and paraperspective, in order to apply the recursive method to a wider camera motion

range.

We also extend the recursive method to accommodate the situation in which some feature points are occluded or leave the field of view during the sequence. The extended recursive method still retains the low computational complexity of $O(P)$ and is simpler than the occlusion solutions proposed in other FMs.

Experiments with real and synthetic image sequences confirm the recursive method's low computational complexity and good performance and indicate that it is well suited to real-time applications.

Contents

1	Introduction	1
1.1	Structure from Motion	2
1.2	Factorization Methods and our Motivation	4
1.3	Outline of the Thesis	9
1.4	Contributions	11
2	Review of Factorization Methods	13
2.1	The Original Method	13
2.1.1	The Bilinear Formulation and the Rank Theorem	14
2.1.2	Shape Space Estimation via SVD Factorization	17
2.1.3	Euclidean Shape and Motion Recovery	19
2.2	The Sequential Method	20
2.2.1	Formulation and Overview	20

2.2.2	Sequential Shape Space Computation	22
2.2.3	Sequential Metric Transformation	23
2.3	The Paraperspective Method	24
2.3.1	Paraperspective Bilinear Formulation	24
2.3.2	Paraperspective Metric Constraints	27
2.3.3	Paraperspective Motion Recovery	28
3	MSE Formulation for Shape Space	30
3.1	Affine Shape Space under GAP Model	30
3.1.1	Camera Models	31
3.1.2	The Bilinear Formulation under GAP	34
3.2	An MSE Formulation for Shape Space	36
3.2.1	Preliminaries	36
3.2.2	An MSE Cost Function	38
3.3	Properties and Comparisons	40
4	Recursive Computation of Shape Space	42
4.1	Recursive Least-Squares Technique	42
4.1.1	Matrix Inversion Lemma	43
4.1.2	Standard RLS Algorithm	43

4.2	RLS Shape Space Updating	47
4.2.1	Recursive Shape Space Updating	48
4.2.2	Affine Shape and Motion Recovery	53
4.2.3	Comparison	53
4.3	Performance Analysis	55
4.3.1	Synthetic Data Generation	55
4.3.2	Accuracy of the Shape Space Computation	56
4.3.3	Computational Time	57
5	Euclidean Reconstruction of Shape and Motion	62
5.1	Under Orthography	63
5.1.1	Synthetic Test	67
5.2	Under Weak Perspective	68
5.3	Under Paraperspective	72
5.3.1	Synthetic Test	73
5.4	Real Image Experiments	75
6	The Extended RLS Algorithm for Occlusion	86
6.1	The Extended RLS Algorithm	86
6.2	Performance of the Occlusion RLS Algorithm	90

6.3 Comparison with other Occlusion Solutions 90

6.3.1 The Original Method 90

6.3.2 The Paraperspective Method 91

7 Conclusion 96