ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library
a435334

An Investigation on Motion Estimation using Gibbs Markov Random Fields

Stephanus Surijadarma Tandjung

School of Computer Engineering

A thesis submitted to the Nanyang Technological University
in fulfilment of the requirement for the degree of
Doctor of Philosophy

2006

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

An Investigation on Motion Estimation using Gibbs Markov Random Fields

Stephanus Surijadarma Tandjung

Abstract

Motion estimation algorithms are an essential part in many video processing applications. In this thesis, a "true" motion field, or a correct apparent motion, is pursued. It can be used for some advanced processing such as motion segmentation.

To obtain a "true" motion field, we focus on how to address the smoothness and discontinuities in the motion field. In detail, several approaches were proposed: (i) a discontinuity adaptive (DA) function was adopted to preserve the motion discontinuities; (ii) both implicit and explicit smoothness constraints are enforced for the motion field; (iii) proposals (i) and (ii) are simultaneously represented by Gibbs Markov Random Field (GMRF) and embedded into the Bayesian method.

We first deal with the motion discontinuity problem using a special DA function. We apply it to the dense motion field produced by optical flow equations using Horn and Schunck's approach. Experimental results show that motion discontinuities can be effectively preserved. Due to its effectiveness, we also apply this DA function to all the other approaches proposed in this thesis.

We then consider the smoothness of the motion fields. Basically there are two types of smoothness constraints: the implicit constraint and the explicit constraint. Implicit constraint is used for block motion models, such as block matching and Lucas and Kanade's block motion model. Explicit constraint, on the other hand, can be applied to most of motion estimation models. We are interested in enforcing both constraints on the motion fields. We use the Bayesian method to model our idea. The implicit smoothness is addressed in the likelihood probability and the explicit smoothness is addressed in the prior probability. Our analysis and experiments show that this could produce a better motion field and give more flexibility in adjusting the motion smoothness.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

We use the GMRF to represent the prior information. It is in fact the relationship of motion vectors of neighboring pixels. Hence, the smoothness and discontinuity problem can be addressed simultaneously to produce the "true" motion field.

In summaries, we improved Horn-Schunck's algorithm by preserving the discontinuities; we revised the conventional block matching algorithm into a smoothness adjustable block matching approach; we improved Lucas and Kanade's block motion model into a smoothness adjustable block motion model. All these algorithms are tested and verified by synthesized and real image sequences.

Key-words: motion estimation, Bayesian, Gibbs Markov random fields, gradient based motion estimation, matching based motion estimation, discontinuity.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Acknowledgments

I want to thank my supervisor Prof. Seah Hock Soon for giving me continuous support during the last part of my graduate study. His projects with the team have broadened my understanding about the problems in image sequences. And I want to thank Dr. Qian Kemao for guiding me in improving the quality of this thesis. I would also like to thank all members of AuReal and CACAni research groups: Lu Ji, Lili, Widya Andy, Wu Zhongke, Quah Chee Kwang, Xie Xue Xiang, Qiu Jie, Chen Quan, Xiao Xian, Zhang Chong, Dr. Tian Feng and Dr. Tony Chan and other members of CAMTECH.

I want to thank my previous supervisor Dr. Chong Man-Nang for introducing me to digital video technology: its problems and applications in industry. I also want to thank friends from our past research group under the same previous supervisor: Lim Keng Pang, Teddy Surya Gunawan, Dilip Krishnan, Showbik, and Pranay Kumar. I keep in mind the companies that they later joined or built: A*Star, GDC-Tech., Nirvana and DaVinci. I want to thank Miss Irene, who handles the laboratory nicely, so that we have a peace in mind for storing our program and data in the UNIX, Linux and Windows systems. We remember that our lab then was the Parallel Processing Laboratory under the School of Applied Science (SAS). Later we move to a new place and the laboratory was renamed PDCC. I would like to thank Evelyn, Irene and Ernest, who are like my sisters and brother in the laboratory and in Surabaya. I want to thank Wu Dajun and all my friends in SAS, who supported me in student related matters. I want to thank my friends who were once in the same laboratory in the Institute of Technology, Bandung (ITB): Ari Wahyudi and Wahendro Adi. We are all in the same school again at NTU. I want to thank Dr. Tati Mengko and lecturers at ITB, because of their inspiration during my master studies under them, I have a desire to get a further degree.

I wish to thank my friends at the Tokyo Institute of Technology (TIT): Dr. Trio Adiono and Prof. Kuneida and the team. Their VLSI laboratory and BeyondLSI Inc., a company under the incubation center in TIT, have given me much support and idea. Some of the contributions and future work in this thesis have been in-

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

fluenced by what we are doing during the implementation of a video phone system. One year with your laboratory has colored my thesis in a slightly different light from what I have written in my previous draft.

I want to thank my friends in the Indonesian Catholic Charismatic Group (KKIHS) and KKIS. With you, we can have a growing Christian life. I want to thank my parish's friends: Stephan's family for being a Christian family to me here. I want to thank my family and relatives in Surabaya and Jakarta, my friends in Singapore, Japan, and all over the world. Some of their name are Sony & Cory, Vinca, Lyna & Andy, Basuki. A want to thank both Dr. Siana Halim, a mathematician, who can identify the lack of precision in my mathematical writing, and Jani Rahardjo MBA at Petra Christian University, Surabaya. I want to thank all who I do not mention specifically. Please accept my apology those of you who have also suffered because I have spent so much time and finance in this thesis.

Stephanus Surijadarma Tandjung

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Contents

Abstract ii

1 Introduction 1

1.1 Thesis Contributions 2

1.1.1 Discontinuity Adaptive MRF for Optical Flow 2

1.1.2 Block Matching with Smoothness Prior Function 3

1.1.3 Block Motion Model for Optical Flow with Smoothness Prior Function 3

1.2 Thesis Organization 4

2 Review of Gibbs Markov Random Fields 6

2.1 Bayesian Framework 7

2.2 Labeling Problems 8

2.3 Markov Random Fields and Gibbs Distributions 9

2.3.1 Neighborhood System and Cliques 9

2.3.2 Markov Random Field 10

2.3.3 Gibbs Random Field 11

2.3.4 Example of a GMRF Model 12

2.4 Optimization Techniques 13

2.4.1 Gradient Descent 13

2.4.2 Iterated Conditional Modes 14

2.5 Applications using Gibbs Markov Random Field 14

2.5.1 Smoothing 15

2.5.1.1 Experimental Results 17

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2.5.2 Motion Detection 19
2.5.2.1 Experimental Results 20
2.6 Summary 23

3 Review of Motion Estimation Models 25
3.1 Two-Dimensional Motion Field Models 26
3.1.1 Region of Support 27
3.2 Bayesian Criteria 28
3.3 Literature Survey on Gradient Based Model 30
3.3.1 Motion Representation: Optical Flow 31
3.3.2 Estimation Criteria 32
3.3.2.1 Regularization 32
3.3.2.2 Block Motion Model 35
3.3.3 Search Strategies: Gradient-Based Optimization 35
3.3.3.1 Gauss-Seidel Method 36
3.3.3.2 Successive Over Relaxation 38
3.3.3.3 Hierarchical Optimization 39
3.4 Literature Survey on Matching Based Model 39
3.4.1 Motion Representation 39
3.4.2 Estimation Criteria 40
3.4.2.1 A Class of Criteria 40
3.4.3 Search Strategies 41
3.4.3.1 Matching Approaches 41
3.4.4 Supporting Techniques 44
3.4.4.1 Bilinear Warping 44
3.4.4.2 Warping Prediction 44
3.4.4.3 Block Segmentation 45
3.5 Literature Survey on Multiresolution Matching Based Model 46
3.5.1 Motion Representation 46
3.5.2 Estimation Criteria 48
3.5.3 Search Strategies 48
3.5.3.1 Subblock Multiresolution Motion Estimation 48

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3.5.3.2 Hierarchical Block Matching 49
3.5.4 Projection 49

4 Discontinuity Adaptive MRF Model for Optical Flow 50
4.1 Introduction 51
4.2 MAP-MRF Framework for Optical Flow 52
4.3 Discontinuity Adaptive MRF Model 54
4.4 Model Comparison 56
4.5 Experimental Results 57
4.5.1 Single Resolution Synthetic Sequence 57
4.5.2 Experiment on Preserving Discontinuities 59
4.5.3 Multiresolution Synthetic Sequence Experiment 60
4.5.4 Multiresolution Real Sequence Experiment 60
4.5.5 Noisy Sequence 62
4.5.6 Comparison with Robust Statistic 67
4.6 Conclusion 69

5 Block Matching with Smoothness Prior Function 71
5.1 MAP-MRF Framework 73
5.2 Problem and Solution 75
5.3 Realization 77
5.4 Algorithms 79
5.4.1 Conventional and Overlap Block Matching 81
5.4.2 Block Matching with Smoothness Prior Function 81
5.4.3 Block Matching with Discontinuity Adaptive Smoothness Prior Function 84
5.4.4 Block Matching with Smoothness Prior Function plus Reduced Block-Size 85
5.4.5 Multiresolution Block Matching with Smoothness Prior Function 86
5.5 Experimental Results 88
5.5.1 Strategies and Summaries 88
5.5.1.1 Conventional and Overlap Block Matching 88

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

5.5.1.2 Block Matching with Smoothness Prior Function . . 89
5.5.1.3 Block Matching with Discontinuity Adaptive Smoothness Prior Function 90
5.5.1.4 Block Matching with Smoothness Prior Function plus Reduced Block-Size 90
5.5.1.5 Multiresolution Block Matching with Smoothness Prior Function 90
5.5.2 Parameter Setting and the Influence of Explicit Smoothness . 91
5.5.3 Results 94
5.5.3.1 Peak Sequence 94
5.5.3.2 Yosemite Sequence 98
5.5.3.3 Suzie Sequence 103
5.5.3.4 Foreman Sequence 106
5.5.3.5 Traveling Salesman Sequence 107
5.5.3.6 Hamburg Taxi Sequence 112
5.6 Conclusion 116

6 Block Motion Model for Optical Flow with Smoothness Prior Function 122

6.1 MAP-MRF Framework 123
6.2 Realization 124
6.3 Block Motion Model using Smoothness Prior Function 125
6.4 Experimental Results 126
6.4.1 Strategies and Summaries 127
6.4.2 Parameter Setting and the Influence of Explicit Smoothness . 128
6.4.3 Results 131
6.4.3.1 Peak Sequence 131
6.4.3.2 Yosemite Sequence 131
6.4.3.3 Suzie Sequence 134
6.4.3.4 Foreman Sequence 137
6.4.3.5 Salesman Sequence 137
6.4.3.6 Hamburg Taxi Sequence 140

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

6.5 Conclusion 140

7 Conclusions **143**

7.1 Thesis Contributions 144

7.2 Topics for Future Research 145

7.2.1 Comparison of the Algorithms 145

7.2.2 Fast Block Matching using Smoothness Prior Constraints . . . 146

7.2.3 Multiscale Motion Estimation 146

7.2.4 Ratio Measure 146

Appendixes **147**

A Basic Iterative Methods **147**

A.1 Point Jacobi Method 147

A.2 Point Gauss-Seidel Method 148

A.3 Successive Overrelaxation Iterative Method 148

B Adaptive MRF **150**

B.1 Regularization and Discontinuities 150

B.2 Discontinuity Adaptive (DA) Model 152

C Measurement **154**

C.1 Peak Signal-to-Noise Ratio 154

C.2 Entropy 154

C.3 Average Angular Error 155

D Author's Publications **156**

Bibliography **159**

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

List of Figures

2.1 Successive Neighborhoods with Neighborhood Order Numbers on Two-Dimensional Lattice. 10
2.2 Cliques types. 10
2.3 1D Smoothing 17
2.4 2D Smoothing. From left to right and from top to bottom: The original data, the corrupted data, the smoothed result and the energy function during minimization process. 18
2.5 Sinusoidal Circle Images 20
2.6 Motion Detection: Detected Motion in Pink Labels for the Sinusoidal Circle Sequence. 21
2.7 Motion Detection: Detected Motion in Pink Labels for the Taxi Sequence. 22
2.8 Motion Detection: Detected Motion in Pink Labels for the Traveling Salesman Sequence. 22
2.9 Motion Detection: Detected Motion in Pink Labels for the Claire Sequence. 22

3.1 Effect of Affine Transform for Sinusoidal Circle Image. 27
3.2 Line Field Lattice with Neighborhood System and Cliques: (a)Line Lattice; (b) Neighborhood of a Line Site; (c) Neighborhood of a Pixel; (d) Pixel Site; (e) Line Sites; (f) Single-Element Line; (g) Four-Element Line Cliques; (h) Three-Element Clique. 29
3.3 Result of Robust Motion Estimation. 34

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3.4 Result of Block Motion Model Motion Estimation using Block-Size of 16×16. 36
3.5 Spatial Interpolation of Motion Vectors. 44

4.1 Interaction Function h_γ with Different γ (represented by the shape) and η^2 (the horizontal axis). 55
4.2 Laplacians's Weights 57
4.3 Ram Sequence and The Displacement Field. 58
4.4 Displacement Field for Synthetic Sequence with (a) $h = 1$, (b) $\gamma = 0.0001$. 58
4.5 Variation of Smoothness Results for Different γ 59
4.6 Preserving the Discontinuities in Suzie Sequence. 59
4.7 Results of Experiments for the Yosemite Sequence. 61
4.8 Results of Experiments for the Suzie Sequence. 63
4.9 Results of Experiments for the Foreman Sequence. 64
4.10 Results of Experiments for the Salesman Sequence. 65
4.11 Results of Experiments for the Taxi Sequence. 66
4.12 Results of Experiments using Robust Technique. 68

5.1 Motion Field Results. 74
5.2 Prior Energy to Solve Uniqueness Problem. 76
5.3 Influence of Implicit and Explicit Constraint for Many Types of Block Matching Algorithm. 76
5.4 Foreman Sequences with Different Weights. 77
5.5 Weight vs PSNR/Entropy 77
5.6 Comparison: Systems with and without Constraint. 79
5.7 Enlarging Block-Size. 79
5.8 Block Diagram of the System 80
5.9 Functional Block Diagram for Block Matching with Smoothness Prior Function Algorithm. 82
5.10 Iterative matching procedure, where $(1, -2)$ is found. 83
5.11 Initialization and Process of Smoothing the Motion Field 83

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

5.12 Progress of PSNR 84
5.13 Flowchart of Re-Sizing Block-Size Algorithm. 86
5.14 A Comparison between Single and Multiresolution Processes. 87
5.15 PSNR, Entropy and PSNR/Entropy vs Weight. 92
5.16 PSNR, Entropy and PSNR/Entropy vs Weight (cont.). 93
5.17 Results of Experiments for the Peak Sequence. 95
5.18 Results of Experiments for the Peak sequence. (cont.) 97
5.19 Results of Multiresolution Block Matching with Smoothness Prior Function for the Peak Sequence. 98
5.20 Results of Experiments for the Yosemite Sequence. 100
5.21 Results of Experiments for the Yosemite Sequence. (cont.) 101
5.22 Results of Multiresolution Block Matching with Smoothness Prior for the Yosemite Sequence. 102
5.23 Results of Experiments for the Suzie Sequence. 104
5.24 Results of Experiments for the Suzie Sequence. (cont.) 105
5.25 Results of Multiresolution Block Matching with Smoothness Prior for the Suzie Sequence. 106
5.26 Results of Experiments for the Foreman Sequence. 108
5.27 Results of Experiments for the Foreman Sequence. (cont.) 109
5.28 Results of Multiresolution Block Matching with Smoothness Prior for the Foreman Sequence. 110
5.29 Results of Experiments for the Salesman Sequence. 111
5.30 Results of Experiments for the Salesman Sequence. (cont.) 113
5.31 Results of Multiresolution Block Matching with Smoothness Prior for the Salesman Sequence. 114
5.32 Results of Experiments for the Hamburg Taxi Sequence. 115
5.33 Results of Experiments for the Hamburg Taxi Sequence. (cont.) . . . 117
5.34 Results of Multiresolution Block Matching with Smoothness Prior for the Hamburg Taxi Sequence. 118

6.1 Progress of Motion Field during Iteration of the Algorithm for the Suzie Sequence. 126

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

6.2 Progress of PSNR and Entropy vs Iteration. 127
6.3 PSNR, Entropy and PSNR/Entropy vs Weight. 129
6.4 PSNR, Entropy and PSNR/Entropy vs Weight (cont.). 130
6.5 PSNR, Entropy and PSNR/Entropy vs Parameter of Discontinuity. . 132
6.6 Results of Block Motion Model for Optical Flow Equation with the Peak Sequence. 133
6.7 Results of Block Motion Model for Optical Flow Equation with the Yosemite Sequence. 135
6.8 Results of Block Motion Model for Optical Flow Equation with the Suzie Sequence. 136
6.9 Result of Block Motion Model for Optical Flow Equation with the Foreman Sequence. 138
6.10 Results of Block Motion Model for Optical Flow Equation with the Salesman Sequence. 139
6.11 Results of Block Motion Model for Optical Flow Equation with the Hamburg Taxi Sequence. 141

A.1 Iteration Process for Three Methods 149

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

List of Tables

4.1 Result of Multiresolution Synthetic Sequence Experiment in PSNR and Entropy values. 60
4.2 Result of Multiresolution Real Sequence Experiment in PSNR and Entropy values. 62
4.3 Experimental Setup for MAP-MRF Labeling Model 62
4.4 Result of Multiresolution Noisy Real Sequence Experiment in PSNR 67
4.5 Result of Multiresolution Real Sequence Experiment using Robust Technique 68

5.1 PSNR, Entropy and Angular Error for the Peak Sequence. Notes: Conventional (Conv.), Discontinuity (Disc.), Reduced (Red.), Average Angular Error (AAE), Standard Deviation of the Angular Error (SD). 94
5.2 PSNR, Entropy and Average Angular Error for the Yosemite Sequence. (Notes: Conventional (Conv.), Discontinuity (Disc.), Average Angular Error (AAE), Standard Deviation (SD) 99
5.3 PSNR and Entropy for the Suzie Sequence 103
5.4 PSNR and Entropy for the Foreman Sequence 107
5.5 PSNR and Entropy for the Salesman Sequence 108
5.6 PSNR and Entropy for the Taxi Sequence. 112

6.1 PSNR and Entropy for the Peak Sequence. (Notes: Average Angular Error (AAE), Standard Deviation (SD)) 134
6.2 PSNR and Entropy for the Yosemite Sequence. (Notes: Average Angular Error (AAE), Standard Deviation (SD)) 134

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

6.3 PSNR and Entropy for Suzie Sequence using Standard Block Model. . 134

6.4 PSNR and Entropy for Foreman Sequence using Standard Block Model. 137

6.5 PSNR and Entropy for Salesman Sequence using Standard Block Model. 137

6.6 PSNR and Entropy for Hamburg Taxi Sequence using Standard Block Model. 140

7.1 Characteristics of Three Proposed Methods. (Chap.= Chapter, repr.= representation) 144

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

List of Algorithms

1 Iterative Conditional Modes 14
2 Gradient Based ME using DA function 67
3 Multiresolution Block Matching Algorithm with Smoothness Prior. . 87

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 1

Introduction

Motion estimation (ME) and compensation play a vital role in many digital video processing and computer vision algorithms. For image compression applications such as H.261, H.263, H.264 and MPEG coding algorithms, ME is used to achieve a better compression ratio by reducing temporal redundancy. In other video processing applications, motion estimation and compensation algorithms are used to restore and enhance digital images. In computer vision, ME is used to track and recognize objects under surveillance [1, 2].

An ME algorithm measures the displacement of pixels in image sequences. Often it is referred as apparent motion [3]. It is different from a 3D movement of the object and scene. It is also different from a 2D movement projected from a 3D scene in the image plane. Even so, correct apparent motion is difficult to be extracted using optical flow or block matching. For example a motion field which is obtained by a block matching algorithm that is based on the intensity alone does not reflect the real apparent motion. This is often observed in the textureless part of an image. It may be good enough for video compression, but it obviously needs to be improved if this motion field is used for some other purposes, such as motion segmentation and frame in-betweening.

In this thesis, we are interested in estimating the "true" motion, which reflects the correct apparent motion field. Basically, this requires that (i) motion for each object to be spatially smooth, (ii) motions for objects which have different direction must be preserved. To obtain a smooth motion for each object, smoothness constraints

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

could be forced onto the motion vectors. At the same time, this would violate the requirement (ii). Hence, motion discontinuities have to be considered and preserved.

In this thesis, bearing these two requirements in our mind, we make several proposals to improve motion estimation. Mixed constraints for smoothness and discontinuity adaptive function for discontinuities are some examples. To satisfy two different requirements simultaneously, we adopt the Bayesian method [4, 5] to model our motion estimation algorithms. Both requirements of smoothness and discontinuity can be treated simply as prior probability. In fact, both smoothness and discontinuity reflect the relationship of motion vectors of neighboring pixels, which could be elegantly realized by Gibbs Markov Random Fields (GMRF). The GMRF modeling is often used in the Bayesian method, forming a fairly general framework. This framework has been widely used in computer vision techniques [6], and is not new in motion estimation [7–9]. All the chapters in this thesis are described under the same framework.

1.1 Thesis Contributions

Bayesian estimation using GMRF is a generic modeling tool. It means that it can be used by any type of model representation, estimation criteria and optimization method. We can have a single model which can be implemented in gradient based or matching based motion estimation.

1.1.1 Discontinuity Adaptive MRF for Optical Flow

The generic model that involves Bayesian estimation and GMRF can be realized through gradient based method. This method produces a smoothed motion field due to smoothness constraint. But the motion field is oversmoothed, especially in the discontinuity area. We therefore use a discontinuity adaptive function to express motion discontinuities. This function is originally proposed by Li [10] for edge preserving problems such as image segmentation. As for the motion equation, we use the Horn-Schunk's optical flow approach. Our work results in an improved Horn-Schunk approach with discontinuities preserved. It can also be a comparable solution

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

to the robust technique [7]. Our work gives us an alternative to the gradient based motion estimators that are available publicly. This part is addressed in Chapter 4.

1.1.2 Block Matching with Smoothness Prior Function

The second contribution involves a block matching technique using GMRF model. We take the intensity difference in block matching as likelihood energy in the Bayesian model. This gives us a new point of view for the block matching. It also gives us the freedom to add stronger smoothness constraint into the motion field, which could be done by simply adding an explicit smoothness constraint as prior energy. This results in a new simultaneous way for addressing the smoothness constraints when both the implicit and explicit smoothness are addressed simultaneously. Is such a mixed constraint necessary? Does it give a better result rather than simply using implicit or explicit constraint? Our analysis and experiments give a positive indication.

We also suggest several implementations to make the proposed technique efficient. The first suggestion is to keep all the likelihood value in the computer memory to expedite the optimization process. Secondly, we only compute the likelihood values from the smallest block-size while the likelihood values of the bigger block-size can be simply synthesized from these computed values. Finally the algorithm is presented in a multiresolution framework and another framework similar to multiresolution (We call it a reduced block-size algorithm.). As it is well known, the concept of multiresolution provides a way for initialization and also speeds up the whole process. This part is addressed in Chapter 5.

1.1.3 Block Motion Model for Optical Flow with Smoothness Prior Function

Besides block matching, another approach for generating block motion vectors was proposed by Lucas and Kanade using a block of optical flow equations. This is a very interesting, simple and effective approach. The result could be obtained by simply solving two linear equations for each block. Implicit constraint is assumed.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Due to the similarity of this approach and the block matching approach, our idea of mixed constraints can also be applied to it. An interesting discovery is that we only need to slightly modify the two linear equations to realize our idea. Hence our algorithm does not only inherit the simplicity of the original algorithm; it also offers flexibility in adjusting the strength of explicit smoothness constraints. This elegant algorithm is also shown to be effective through various experiments. This part is addressed in Chapter 6.

1.2 Thesis Organization

The remaining chapters of this thesis are organized as the following.

Chapter 2 reviews Gibbs Markov Random Fields. The chapter begins by introducing the labeling problem, the neighborhood system, and the clique to facilitate the understanding of the Markov Random Field (MRF). Gibbs distribution, which is equivalent to MRF but emphasized global property, is explained. Optimization techniques are also introduced.

Chapter 3 reviews motion estimation models. They are divided into two groups: the gradient based and the matching based. Each of the groups is classified into motion representation, estimation criteria and search strategies. The last topic in this chapter is the multiresolution matching based model.

Chapter 4 presents the discontinuity adaptive MRF (DA-MRF) model for optical flow equation. The discussion begins with the MAP-MRF framework and the necessity of introducing DA-MRF into optical flow equation. An improved Horn-Schunck motion estimation algorithm is given. Finally, the experimental results are shown.

Chapter 5 presents block matching with smoothness prior function. The discussion begins with the MAP-MRF framework. The problem of the conventional block matching algorithm is highlighted and the solution is introduced. We propose a realization of the solution. Several algorithms are listed and the experimental results

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

are reported.

Chapter 6 presents block motion model for optical flow with smoothness prior function. The MAP-MRF framework is discussed. Realization and the algorithm are presented. This chapter implements the modified Lucas-Kanade motion estimation algorithm. Finally, the experimental results are shown.

Chapter 7 summarizes the results of this thesis, presents the major contributions and details a number of avenues for further research.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 2

Review of Gibbs Markov Random Fields

In this chapter we provide a basic knowledge needed to comprehend Gibbs Markov Random Fields (GMRF). GMRF is about the context dependent entities of image pixels or spatially correlated features, which can be used in the interpretation of visual information.

GMRF has been widely used to solve computer vision problems. If the problems are interpreted at a lower level representation, the applications of GMRF models include image/movie restoration and segmentation, surface reconstruction, edge detection, texture analysis, optical flow, and visual integration. If the problems are interpreted at a higher level representation, the applications of GMRF models include object matching and recognition.

GMRF theory shows us how to model prior probability of context dependent entities, such as an arrangement of object features or an interaction of motion vectors. The GMRF theory is often used in the field of estimation theory. The vision problems are formulated into objective functions so they become optimization problems.

We explain the Bayesian framework in Section 2.1. We present the labeling problem in computer vision in Section 2.2. We explain the definition of Gibbs and Markov random fields in Section 2.3. We explain the optimization algorithms in Section 2.4. We discuss the applications of GMRF in Section 2.5 and build several applications to show the usefulness of GMRF. Finally we summarize our discussion

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

in Section 2.6.

2.1 Bayesian Framework

According to E. T. Jaynes, the Bayesian framework provides a consistent way of reasoning about the world [11]. The world in computer vision is pixels, labels and human expectation. The pixels and human expectation are considered as the input to the reasoning system. The output of this system is represented by the labels.

A Bayesian framework has a prior distribution for expressing the human expectation and a likelihood distribution for interpretation of the pixels. The output is represented by the labels. The output is called posterior probability in *maximum a posteriori probability* (MAP) labeling and it is written as:

$$P(\mathbf{f}|\mathbf{d}) = \frac{P(\mathbf{d}|\mathbf{f})P(\mathbf{f})}{P(\mathbf{d})} \tag{2.1}$$

where $\mathbf{d}$ is data, $\mathbf{f}$ is labels, and $P(\mathbf{d})$ is the probability distribution of data $\mathbf{d}$. Consequently, the terms on the right-hand side are prior probability $P(\mathbf{f})$ and likelihood probability $P(\mathbf{d}|\mathbf{f})$.

Generally a good estimation of $\mathbf{f}$ from the given data $\mathbf{d}$ is that the posterior probability $P(\mathbf{f}|\mathbf{d})$ is maximized. It can be seen from Eq. 2.1 that this posterior probability is determined by two items: likelihood probability $P(\mathbf{d}|\mathbf{f})$ and prior probability $P(\mathbf{f})$. In this chapter we focus on reviewing the GMRF model, which can effectively construct the prior probability. This GMRF model has been widely adopted and shown to be an effective model.

For example, to solve motion estimation problem, let motion field f be an estimated realization of a random variable $\mathbf{F}$ in a posterior probability distribution, given image d_i and image d_{i-1}(the realization of a random variable $\mathbf{D}$). Using Bayes rule, the posterior probability distribution is:

$$P(f|d_i, d_{i-1}) = \frac{P(d_i|f, d_{i-1})P(f, d_{i-1})}{P(d_i, d_{i-1})}, \tag{2.2}$$

where P is a probability measure and the denominator is a normalizing constant.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

The numerator is the observation model and the prior model respectively. In MAP, the probability model becomes:

$$\begin{aligned} \widehat{f} &= \arg\max_f \{P(f|d_i, d_{i-1})\}, \\ \widehat{f} &= \arg\max_f \{P(d_i|f, d_{i-1})P(f, d_{i-1})\}. \end{aligned} \tag{2.3}$$

Given f and d_{i-1}, the motion-compensated prediction of d_i can be computed.

2.2 Labeling Problems

Before we define the GMRF, we first introduce the labeling problems in vision. Many applications in vision can be considered as labeling problems. For example, a motion detection algorithm for an image sequence is to assign a specific label to a site that is considered as moving.

The labeling problem is specified in a set of *sites* and a set of *labels*. Let S be the lattice with a set of m indices

$$S = \{1, \ldots, m\}. \tag{2.4}$$

Given a 2D image of size $w \times l$, its pixel location can be conveniently re-indexed by a single number k where it takes on value in $\{1, \ldots, m\}$. This notation of a single number site index will be used for images in this thesis unless otherwise treated.

A label is an event that may occur in a site. Let L be a set of *labels*. A continuous label set may correspond to the real number $\Re$ or an interval from low value, X_l, to high value, X_h:

$$L_c = [X_l, X_h] \in \Re. \tag{2.5}$$

For the discrete case, a label assumes a discrete value in a set of M labels

$$L_d = \{l_1, \ldots, l_M\}. \tag{2.6}$$

The *labeling problem* is to assign a label from the label set L to each of the sites

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

in S. The set

$$f = \{f(1), \ldots, f(m)\} \tag{2.7}$$

is called a labeling of the sites in S, where $f(s) \in L_d$ is a label for a site s.

Labeling is called a *configuration* in the terminology of random fields. If all the sites have the same label set L (L is the set of all possible labeling), then the configuration space is

$$\mathcal{F} = L \times L \ldots \times L = L^m \tag{2.8}$$

where m is the size of S.

2.3 Markov Random Fields and Gibbs Distributions

2.3.1 Neighborhood System and Cliques

Before we describe a vision problem using GMRF models, we need to define the neighborhood systems and cliques for sites. The explanations are in the followings. A neighborhood system, N_s, can be defined as the set of nearby sites within a radius of r (r-th order),

$$N_s = \{s' \in S | \text{dist}(pixel_{s'}, pixel_s) \leq r,\ s' \neq s\}, \tag{2.9}$$

where dist (A, B) denotes the Euclidian distance between A and B, and r takes an integer value. Neighborhood systems [12] satisfy the following conditions

- No site is its own neighbor; $s \notin N_s$
- Neighborhood membership is symmetric; $s \in N_{s'} \Leftrightarrow s' \in N_s$

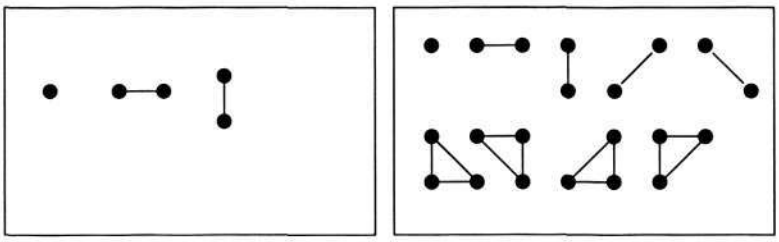
Figure 2.1 shows neighborhood systems, while the numbers indicate the outermost neighboring sites in the r-th order around a site s [13]. Often we only use the first or second order neighborhood system in motion estimation. For example, the following pair potential:

$$V(s, s') = (f(s), f(s')) = \lambda \|f(s) - f(s')\| \tag{2.10}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

9	8	7	6	7	8	9
8	5	4	3	4	5	8
7	4	2	1	2	4	7
6	3	1	s	1	3	6
7	4	2	1	2	4	7
8	5	4	3	4	5	8
9	8	7	6	7	8	9

Figure 2.1: Successive Neighborhoods with Neighborhood Order Numbers on Two-Dimensional Lattice.



(a) Cliques for First Order Neighborhood System

(b) Cliques for Second Order Neighborhood System

Figure 2.2: Cliques types.

is the first-order neighborhood system, where λ is the weighting function.

Besides the neighborhood system, a *clique* is used to define each relationship in the system. A *clique,* **c**, is a subset of sites in S that consists of either a single site $c = \{s\}$, or a pair of neighboring sites $c = \{s, s'\}$, or triple neighboring sites $c = \{s, s', s''\}$, and so on. While collections of the cliques are denoted by $C_1 = \{s \mid s \in S\}$, $C_2 = \{\{s, s'\} \mid s' \in N_s,\ s \in S\}$, $C_3 == \{\{s, s', s'''\} \mid s, s', s'' \in S$ are neighbors one to another $\}$.

We also define $C = \bigcup_{s=1}^{N} C_s$, as the collection of all cliques from the *r-th order neighborhood system.* Figure 2.2 shows the clique collections for the first and second order neighborhood systems.

2.3.2 Markov Random Field

Let $\mathcal{F} = \{F(1), \ldots, F(m)\}$ be a family of random variables defined on the set S. In the family F, each random variable $F(s)$ takes a value $f(s)$ in the label set L.

For simplicity, we write $F = f$, eliminating s, for the same event. This cor-

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

responds to a realization of a random variable. For a discrete label set L, the probability that random variable F takes the value f is denoted $P(f)$.

F is a *Markov random field* (MRF) on S with respect to a neighborhood system N if and only if the following two conditions are satisfied:

Positivity: $P(f) > 0, \forall f \in F$,

Markov property: $P\left(f\left(s\right)|f\left(S-s\right)\right) = P\left(f\left(s\right)|f\left(N\left(s\right)\right)\right), \forall s \in S$.

The first condition states that all possible realizations should have nonzero probability, while the second condition requires that the local conditional probability at a particular site s depends only on the values of the random field within the neighborhood $N\left(s\right)$ of that site.

So far the joint probability, $P(f)$, and conditional probabilities, $P(f\left(s\right)|f\left(N\left(s\right)\right))$, together specify an MRF. However, there is no obvious method to deduce the joint probability from the associated conditional probabilities. With Gibbs distribution, we can then mathematically track the joint probability of an MRF.

2.3.3 Gibbs Random Field

The same set of random variables F as in the above is said to be a *Gibbs random field* (GRF) on S with respect to N, if and only if its configurations obey a *Gibbs distribution.*

The Gibbs distribution is:

$$P(f) = \frac{1}{Z}\mathrm{e}^{\left(-\frac{1}{T}U(f)\right)} \tag{2.11}$$

where the U and Z are defined as follows:

$$U(f) = \sum_{c \epsilon C} V_c(f) \tag{2.12}$$

$$Z = \sum_{f \in \mathrm{F}} \mathrm{e}^{\left(-\frac{1}{T}U(f)\right)}. \tag{2.13}$$

where Z is a normalizing constant, T is the temperature and V_c may be any real function of variables that depends on the local configuration of the clique c. See

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Figure 2.2 for a configuration of cliques. Having $P(f)$ to measure the realization f, and having the temperature T to control the sharpness of the distribution, we can generate a "pattern".

We can summarize the random fields that the MRF is characterized by its local property whereas a GRF is characterized by its global property. The equivalence of these properties is established by the Hammersley-Clifford theorem, which states that *F is an MRF on S with respect to N if and only if F is a GRF on S with respect to N.* In other words, with respect to a specified neighborhood system a unique GRF exists for every MRF and a unique MRF exists for every GRF. This statement implies that we can define a random field based on the global probability measure or the local conditional probability measure. Thus the GRF and MRF models will be equivalent under a neighborhood system.

2.3.4 Example of a GMRF Model

The following introduces GMRF models for modeling image properties such as regions and textures. We look at their energy functions.

Contextual constraints on two labels are the lowest order constraints to convey contextual information. They are completely characterized by its clique functions and have the following energy functions:

$$U(f) = \sum_{s=1}^{m} \left\{ V_1(f(s)) + \sum_{r=1}^{c} V_{2,r}(f(s), f(s')) \right\} \tag{2.14}$$

where function V_1 is a potential function defined on cliques consisting of single pixel, function $V_{2,r}(f(s), f(s'))$ is a potential function defined on cliques consisting of pair of pixels, r is the pair number, and c is the total of available pairs and m is the area of a lattice.

In [14], Derin-Elliot model lets

$$V_1(f(s)) = \alpha_{f_s},\ V_2(f(s), f(s')) = \Theta I(f(s), f(s')), \tag{2.15}$$

where $I(a,b) = -1$ if $a = b$ and 1 if $a \neq b$, α and Θ are constants. We can generate a texture using this model.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Another model that is used in our thesis is smoothness prior. This is a generic contextual constraint. It assumes that many physical properties possess some coherence and generally do not change abruptly. Smoothness constraints are expressed as the prior probability. In the example section in this chapter the smoothness prior is used.

2.4 Optimization Techniques

In this section we briefly introduce some optimization techniques for maximizing an objective function. In this thesis, the objective function is the posterior energy. The optimization can be divided into two groups: the continuous and the discrete. For continuous optimization, it can be Gradient Descent or Graduated Non-Convexity For discrete optimization, it can be: Simulated Annealing, Iterated Conditional Modes (ICM), or Mean Field Annealing. Each group can also be categorized as global or local method where there is multiple local minima in the solution space, and it can be further categorized as constrained or unconstrained optimization. [6, 15].

2.4.1 Gradient Descent

When the energy function is differentiable, the simplest way to find $f^* = \arg\min_f U(f)$ is to perform "gradient descent". The update process is

$$f^{(t+1)} \leftarrow f^{(t)} - \mu \nabla U(f^{(t)}) \tag{2.16}$$

where $\mu > 0$ is a step size and $\nabla U(f)$ is the gradient of the energy function. The iteration stops if the algorithm converges to a point f for which $\nabla U(f) \approx 0$.

Currently, the minimization procedure in the proposed models is in the first order. If we can derive an efficient second order minimization, we can improve the current procedure.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2.4.2 Iterated Conditional Modes

Iterative Conditional Modes (ICM) is a "greedy" method, and is also a local minimization method. The label is usually discrete. Algorithm 1 shows the ICM optimization process.

Algorithm 1 Iterative Conditional Modes

1. Start with an initial configuration $f^0 \in L_d^m$.
2. For site s, choose the new label $f^{(t+1)}(s)$ that minimizes $U(f^{(t+1)}(s)) - U(f^{(t)}(s))$ locally.
3. Keep the new label in a separate variable until all the sites $s \in S$ have been visited once.
4. Update the labels for all the sites.
5. Do the iteration step number two to four until no further energy descent is possible.

2.5 Applications using Gibbs Markov Random Field

Many applications use GMRF for their prior model. The following is the list of applications: texture classification [16], texture modeling [17], segmentation [18], boundary and edge detection [19], image restoration [20], and optical flow estimation [21].

Two applications are given as examples. These examples seem irrelevant to motion estimation at a first glance. But with these examples, we will understand the whole procedure of modeling using GMRF. We have arranged the examples so that: (i) the examples give a complete modeling structure, (ii) the examples use 1D, 2D, continuous and discrete data to show the various applications. Many details are explained in these examples, which may not be elaborated in the following chapters if it is unnecessary, in order to highlight our novel proposals.

The first application is a smoothing algorithm. The second application, motion detection, uses discrete labels to separate moving and static objects.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2.5.1 Smoothing

Given a set of data, the objective is to smooth the data using GMRF. In Bayesian framework, the first task is to design the likelihood and prior probabilities of the smoothing process. The last task is to select the optimization strategy.

Given an image lattice S and noisy input image data $d = \{d(x, y) \in \Re\}$, the target is to determine the output of a smoothed image $f = \{f(x, y) \in L = \Re\}$. This smoothed images should satisfy the following conditions: (i) it should not be very different from the original image; this means the likelihood energy value $U(d|f)$ should be small, (ii) the output image is a smooth image, which means that intensity of neighboring pixels should not change abruptly; this is in fact the spatial relationship of neighbors, the prior probability $U(f)$.

We start with the maximum a *posteriori* (MAP) probability estimate:

$$f^* = \arg\max_f P(f|d) = \arg\min_f U(f|d). \tag{2.17}$$

The next step is to define the prior and likelihood probabilities. For prior probability, we use first order or 4-neighborhood system:

$$N(x, y) = \{(x-1,\, y)\,,\ (x+1,\, y)\,,\ (x,\, y-1)\,,\ (x,\, y+1)\}\,. \tag{2.18}$$

The prior probability is

$$p(f) = \frac{1}{Z}\mathrm{e}^{-U(f)} \tag{2.19}$$

where Z is the normalizing constant. The *prior energy function* is as following

$$U(f) = \sum_{(x,y)\in S} \sum_{(x',y')\in N(x,y)} g(f(x, y),\, f(x', y')), \tag{2.20}$$

where $g(\eta)$ is a quadratic function.

For likelihood probability, the model is

$$d(x, y) = f(x, y) + \epsilon \tag{2.21}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

with ϵ being the i.i.d or independent identically distributed Gaussian noise with mean zero and variance one, d being the data and f being the smoothed output.

$$p(d|f) = \frac{1}{Z}\mathrm{e}^{-U(d|f)} \tag{2.22}$$

is a Gaussian function with *likelihood energy function*

$$U(d|f) = \sum_{(x,y)\in S} \frac{[f(x,y) - d(x,y)]^2}{2\sigma^2} \tag{2.23}$$

After the likelihood and prior probabilities are defined, posterior distribution is

$$p(f|d) = \frac{p(f)p(d|f)}{p(d)} \propto p(f)p(d|f) = \frac{1}{Z}\mathrm{e}^{-U(f|d)} \tag{2.24}$$

with *posterior energy function*

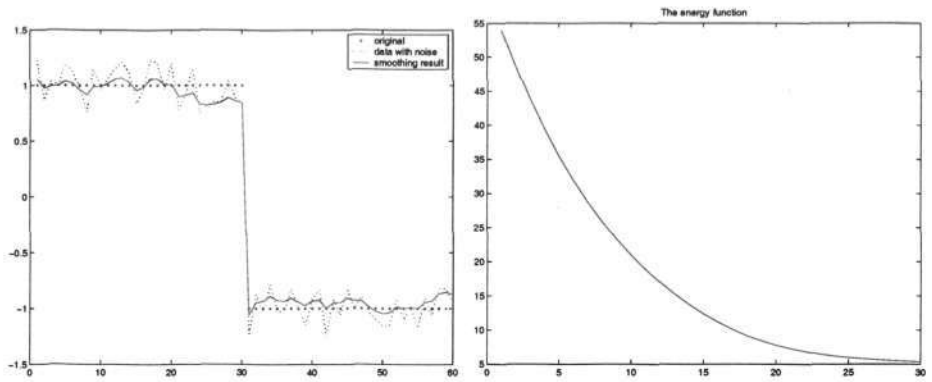
$$U(f|d) = \sum_{(x,y)\in S} \{(f(x,y) - d(x,y))^2/2\sigma^2 + \sum_{(x',y')\in N(x,y)} g(f(x,y), f(x',y')). \tag{2.25}$$

The rest is the task of the optimization algorithm to find an optimal configuration. A simple gradient descent method is used as is shown in the Equation 2.16. If the derivative $g'(\eta) = 2\eta h(\eta)$ and the gradient $\frac{\partial U(f|d)}{\partial f_i}$ have been defined, then the iteration is:

$$\begin{aligned} f^{(t+1)}(x,y) \leftarrow\ & f^{(t)}(x,y) - \\ & 2\mu(f^{(t)}(x,y) - d(x,y))/2\sigma^2 + \\ & \mu \textstyle\sum_{(x',y')\in N(x,y)} \left(f^{(t)}(x,y) - f(x',y')\right) h\left(f^{(t)}(x,y) - f(x',y')\right) \end{aligned} \tag{2.26}$$

Later, $h(\eta)$ is called an *interaction function* that the choices can be selected from [6]. If $g(\eta)$ is a quadratic function, then $h(\eta)$ is a constant. Followings are several experiments.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) The Original, the Corrupted Data and The Smoothed Result.

(b) Energy Function Minimization

Figure 2.3: 1D Smoothing

2.5.1.1 Experimental Results

1D Smoothing The data to be smoothed is corrupted by Gaussian noise. Our task is to smooth the noise and to preserve the edge. We use

$$h_\gamma = \frac{1}{[1 + \frac{\eta^2}{\gamma}]^2}, \tag{2.27}$$

an *adaptive interaction function* that will be used in Chapter 4 as a part of the prior energy function. Gradient descent minimization is used with the parameters: $\sigma = 1$, $\mu = 0.1$. The initialization for label f equals zero. The result is shown in Figure 2.3(a), where the original and the smoothed data are shown in the same graph. Figure 2.3(b) shows the reduced energy function during the minimization process.

2D Smoothing. Figure 2.4 shows the result of a 2D smoothing algorithm. We generate 2D well-shaped data with the added noise. Gradient descent minimization is used with the parameters: $\sigma = 0.75$, $\mu = 0.1$. The initialization for label f equals zero. The minimization algorithm reduces the energy function for 30 iterations.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

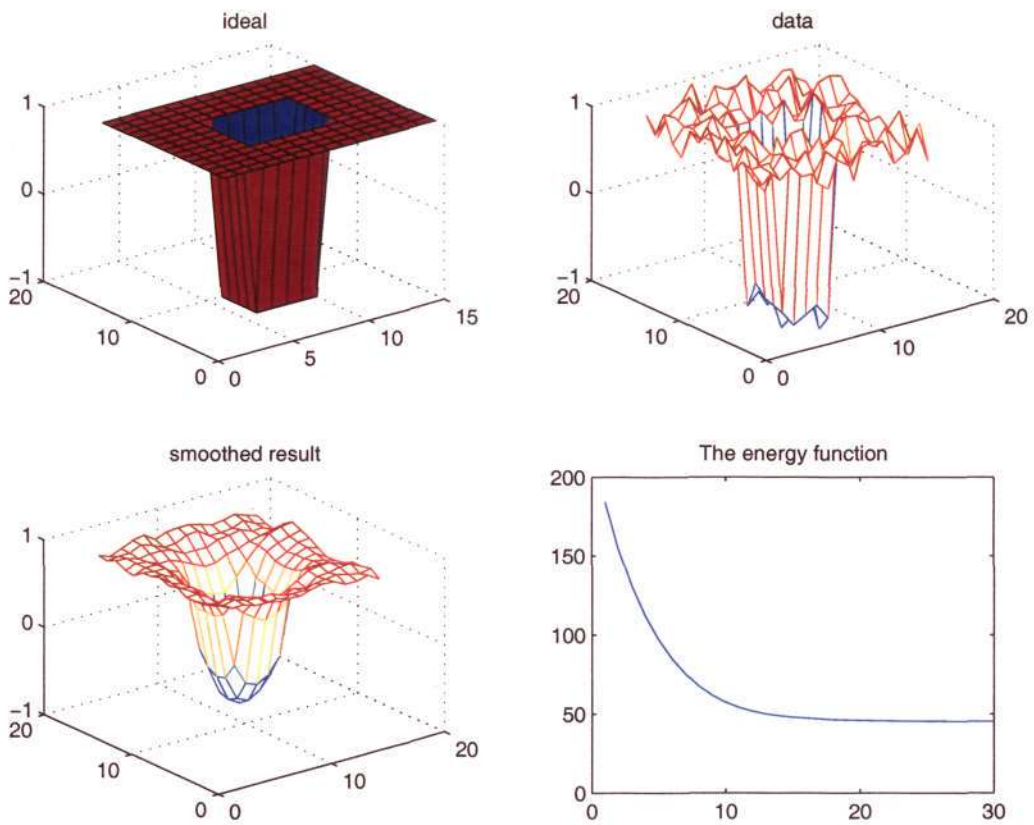


Figure 2.4: 2D Smoothing. From left to right and from top to bottom: The original data, the corrupted data, the smoothed result and the energy function during minimization process.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2.5.2 Motion Detection

Motion detection is a simple motion analysis problem. It detects the moving object using the pixel difference of image sequence. If the pixel difference is higher than a certain value, it may be a moving object on that site. To reduce wrong detections, the smoothness prior is used. The prior knowledge is the smoothness on moving labels.

With the smoothness purpose, the motion detection algorithm is designed using the GMRF model [22]. In this case, the modeling is a binary labeling problem, and the task is to label each pixel as a moving or static state. So, the label $f \in \{s, m\}$ has two states: s =static (set as logic 0) and m =moving (set as logic 1).

The Bayesian framework gives the posterior probability:

$$P(f|g) \propto P(g|f)P(f), \tag{2.28}$$

where the terms are likelihood probability, $P(g|f)$, and prior probability, $P(f)$, with $g = |d_i - d_{i-1}|$ that measures the intensity difference of pixel of image sequence at the same site.

Instead of probability, the criterion can be converted into energy functions:

$$U(f|g) \propto U(g|f) + U(f). \tag{2.29}$$

The $U(f)$ is the prior energy function and it is defined as the interaction between center pixel and its neighbors. This function can be defined as

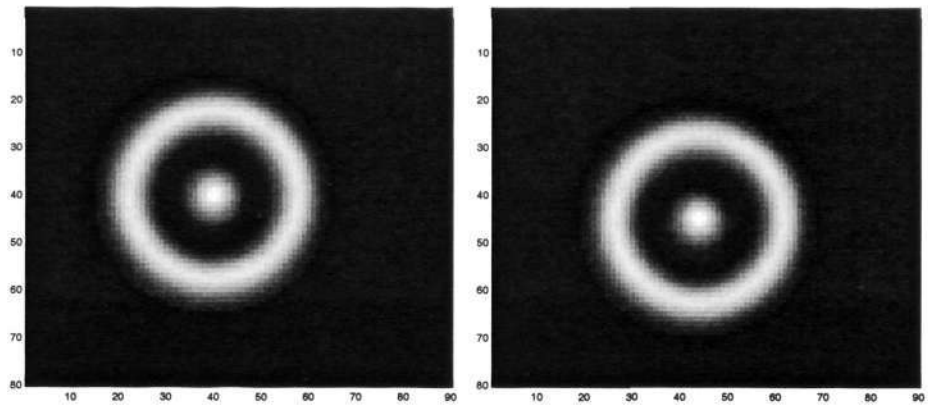
$$U(f) = \sum_{s' \in N} V(f(s), f(s')). \tag{2.30}$$

The potential $V(f(s), f(s'))$ will be $-\beta$ if the neighbor label is equal to the center, and $+\beta$ if they are not equal.

The likelihood energy is

$$U(g|f) = \frac{1}{2\sigma^2} \sum_{s \in S} \psi(g(s), f(s)) \tag{2.31}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) First Image (b) Second Image

Figure 2.5: Sinusoidal Circle Images

where $s \in S$ and

$$\psi(g(s), f(s')) = \begin{cases} -\alpha & if\ (g(s) > \mathbf{C}) \equiv f(s) \\ \alpha & if\ (g(s) > \mathbf{C}) \oplus f(s) \end{cases} \tag{2.32}$$

where $\mathbf{C}$ is a constant or can be empirically chosen from mean absolute difference of the frame. The $\psi(.)$ equals $-\alpha$ if the result of conditional logic for the intensity $g(s)$ equals $f(s)$. The $\psi(.)$ equals α if the result of conditional logic for the intensity $g(s)$ is exclusive-or to $f(s)$.

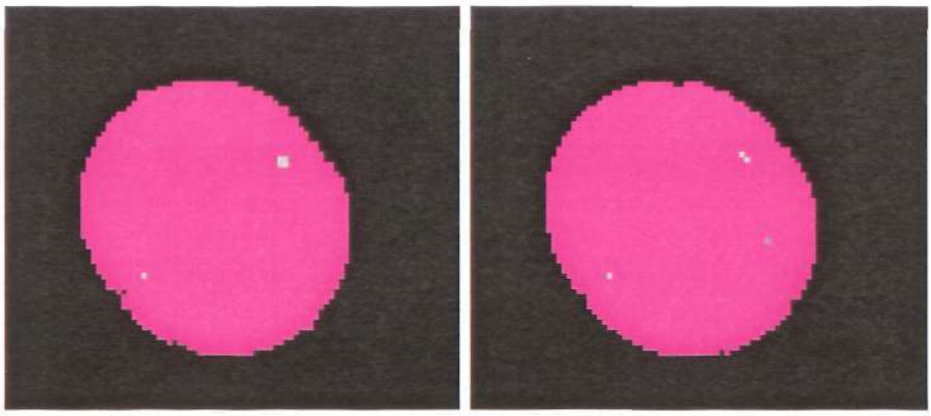
2.5.2.1 Experimental Results

After we have designed the energy functions and selected ICM to find the optimal solution, we start the experiments.

Synthetic Image Sequence. The sinusoidal circle images in Figure 2.5 are used to test the algorithm with parameters: α for the gain of likelihood probability and the order of the neighborhood system. The second image is a shifted image with displacement of $(3, 4)$ pixels.

The results of motion detection are shown as dark/pink labels in Figure 2.6(a) for experiment using $\alpha = 1$ and first order neighborhood system, and in Figure 2.6(b) for experiment using $\alpha = 1$ and second order neighborhood system. In some sites

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) First Order Neighborhood System

(b) Second Order Neighborhood System

Figure 2.6: Motion Detection: Detected Motion in Pink Labels for the Sinusoidal Circle Sequence.

from the data of this pair of images, pixels have the same/slightly different intensity value. In these sites the labels are set as a static label. The detection results around the edge are also different.

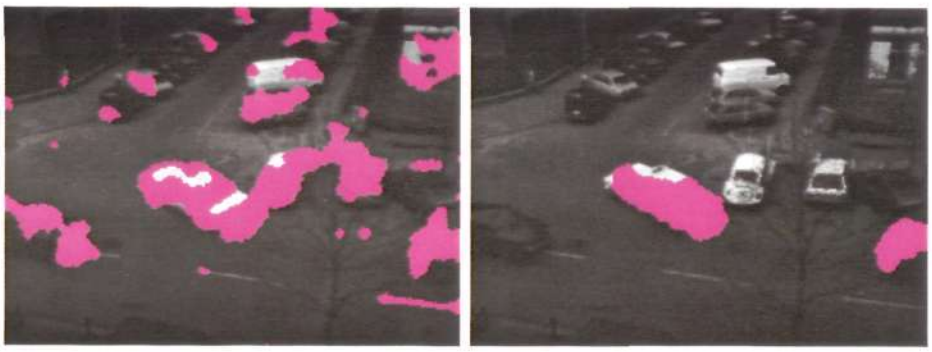
Real Image Sequences. Experiments with real images were conducted. The threshold value, C, is $3/255$. The value is empirically chosen to represent intensity-noise margin. Besides the single resolution framework, a multiresolution framework using three level resolutions is designed. All experiments must follow the condition that the iteration must be long enough or the energy function has reached near its minimum value.

Figure 2.7(a), with $\alpha = 1$ and the second order neighborhood system, shows the result of the single resolution motion detection and Figure 2.7(b) shows the result of the hierarchical motion detection. The results of other experiments with real images are shown in Figures 2.8 and 2.9.

If we look at the single resolution, the prior function, which relates a center to the neighbors, has a tendency to smooth the state of labels. So, we cannot find any single moving pixel, because the prior energy function tends to make them a group of pixels.

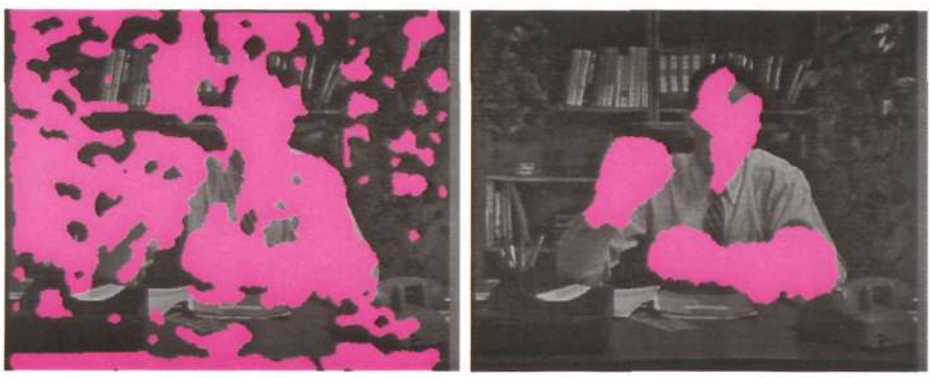
From the experiment, multiresolution motion detection gives better results. In

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) Single Resolution (b) Multiresolution Framework

Figure 2.7: Motion Detection: Detected Motion in Pink Labels for the Taxi Sequence.



(a) Single Resolution (b) Multiresolution Framework

Figure 2.8: Motion Detection: Detected Motion in Pink Labels for the Traveling Salesman Sequence.

(a) Single Resolution. (b) Multiresolution Framework

Figure 2.9: Motion Detection: Detected Motion in Pink Labels for the Claire Sequence.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

the Taxi sequence, Figure 2.7, the textureless white taxi can be detected as more solid moving labels in the multiresolution framework than the labels are found in the single resolution framework. But small moving objects cannot be detected. In the Traveling Salesman sequence, Figure 2.8, camera noise and moving shadows are eliminated in the multiresolution framework. In the Claire sequence, Figure 2.9, minor moving objects such as ears and hair are eliminated from the detection.

We can conclude that multiresolution framework can accommodate more neighbors and can be used to remove the less-significant moving object. However, when we need to detect the small moving objects then we can use simulated annealing (SA) to minimize the energy function. The SA enables the selection of labels that the energy value may climb out from valleys or local minima to later enter the valley which is considered as the global minimum.

2.6 Summary

This chapter introduced a way to transform a computer vision problem into an optimization problem. Some of them are considered as the labeling problem. The label itself has a neighborhood system which is defined by MRF. It is understood that they represent the local property while GRF has the global property.

Bayesian method includes likelihood and prior functions in the GMRF modeling. The likelihood function describes our expectation of the model, while the prior function is a representation of our prior knowledge. The prior function interprets the interactions between the center pixel and the neighbors.

The smoothing experiment, where the smoothness prior is used as our prior knowledge, illustrates one application of prior knowledge. In this experiment two excessively smoothed results are produced for 1D and 2D data. They show an existing problem in the smoothing algorithm using GMRF. Furthermore, the motion detection experiment has shown the differences between results from the single resolution and the multiresolution framework. Also, the experiment shows the ability of a multiresolution framework to extend the interaction of the pixels to produce labels that cannot be reached by the single resolution. This fact points to the use of mul-

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

tiresolution in further research. In the next chapter motion estimation algorithms will be reviewed.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 3

Review of Motion Estimation Models

This chapter reviews three topics: gradient based, matching based, and multiresolution matching based motion estimation. In addition with the brief review of Gibbs Markov Random Field in the previous chapter, this chapter completes the literature survey.

In general, research issues of optimization-based vision involve *problem representation*, the formulation for *an objective function* and *the optimization* objective. The first issue concern about representing the solution using label set for some pixel sites. For example, a block matching motion estimation approach has a label for a group of pixels and a gradient based approach has a label for a pixel. The second issue is an effort to map various object features and constraints into a real number. The third issue is the selection of optimization method that brings concerns such as local minima and its efficiency.

In each section of the motion estimation model, we classify a motion estimation model into three classes: *motion representation, estimation criteria* and *search strategies*. These classifications follow Konrad's style [3] that makes an understanding of the motion estimation model easier. The first class is the selection of a set of parameters to represent a motion. The second class considers the criteria to estimate a motion. The third class plans the search procedure.

This chapter is organized as following: in Section 3.1 we explain the two-dimensional motion field models. In Section 3.2 we explain Bayesian criteria. In Section 3.3 we discuss a gradient based model. In Section 3.4 we present a matching based model.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

In Section 3.5 we discuss a multiresolution matching based model.

3.1 Two-Dimensional Motion Field Models

Two essential models in motion estimation are a *motion model* and an *observation model*. The first model is to represent motion in an image sequence. The second model is a model relating motion parameters to image intensities.

Motion model decribes the motion of image points, i.e., the 2-D motion or apparent motion. Such a motion is a combination of projections of the motion objects in a 3-D scene and 3-D camera motion on a 2-D screen.

This motion depends on: (1) image formation model, e.g., perspective and orthographic projection, (2) motion model of 3-D object, e.g., rigid-body with 3-D translation and rotation, and 3-D affine motion and (3) surface model of 3-D object, e.g., planar and parabolic [3]. Two cases from these relationships have been used extensively in practice. For an orthographic projection and arbitrary 3-D surface undergoing 3-D translation, a 2-D vector describes the resulting 2-D instantaneous velocity. The vector at position s is :

$$f(s) = \begin{pmatrix} u(s) \\ v(s) \end{pmatrix}, \tag{3.1}$$

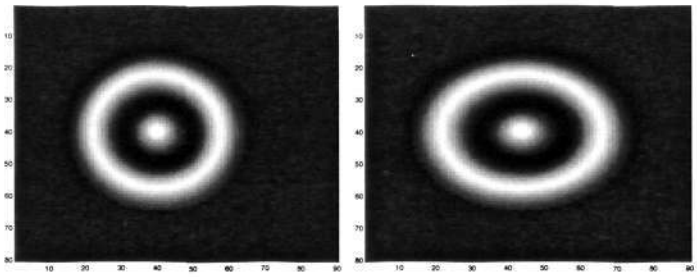
in the image plane where $[u(s), v(s)]^T$ is a vector of parameters depend on the camera geometry and 3-D translation parameters.

The motion can also be modeled by combining the projection model with the 3-D affine motion of a planar surface. This leads to the following 6-parameter affine model:

$$f(s) = \begin{pmatrix} u(s) \\ v(s) \end{pmatrix} + \begin{pmatrix} a_1(s) & a_2(s) \\ a_3(s) & a_4(s) \end{pmatrix} \begin{pmatrix} u(s) \\ v(s) \end{pmatrix}, \tag{3.2}$$

where $[u(s), v(s), a_1(s), a_2(s), a_3(s), a_4(s)]^T$ is a vector of parameters related to the camera as well as the 3-D surface and motion parameters. Increasing the number of parameters give more freedom for a pixel on site s to move to another location. Figure 3.1 shows the distortion that is possible for a circle image in Figure 3.1.a and

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) Original Image

(b) Image as a Result of Affine Transform

Figure 3.1: Effect of Affine Transform for Sinusoidal Circle Image.

the distorted circle in Figure 3.1.b.

If six parameters are not enough to describe a motion of image points, then more parameters can give more precision to the description of the motion [23].

3.1.1 Region of Support

The set of sites s to which a spatial motion model applies is called a region of support. The selection of a motion model and a region of support is one of the major factors determining the precision of the resulting motion parameter estimates.

The most constrained model is *global* motion representation. In this case, a single motion model applies to all image points. This model is suitable for the estimation of camera-induced motion as only a few parameters describe the motion of all image points [3]. Camera effects, zooming and panning use this model efficiently.

In contrast, the least constrained model only uses a *single* image point for a region of support. This model is called dense motion representation. Consequently, it can represent a very large dimension of the motion fields. This approach, however, raises computational complexity.

The commonly used region of support is a *rectangular* block of image points. The simplest case is where the blocks are non-overlapping and their union covers the whole image. This type of region has proven to be a very powerful model and is used today in all digital video compression standards, such as H.261, H.263, H.264, MPEG-1, MPEG-2 and MPEG-4.

A very active area of research today is about the arbitrary shaped region of

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

support [3]. Motion representation for each region consists of a set of motion parameters and the region boundary description. Compared with the models based on the rectangular blocks, this region of support, which is known as a region-based model, has the capacity to perform better image matching at the cost of a more complex representation.

3.2 Bayesian Criteria

Using the MAP discussed in Section 2.1, the observation model in this matching based approach is the model of displaced frame difference (zero-mean, Gaussian distribution) [24]:

$$P(d_i|f, d_{i-1}) = (2\pi\sigma^2)^{-\frac{m}{2}} \mathrm{e}^{\left\{-\sum_{s=1}^{m} \frac{[d_i(s)-d_{i-1}(s+f(s))]^2}{2\sigma^2}\right\}}, \tag{3.3}$$

where m denotes the total number of pixels in a frame and σ^2 denotes the variance of pixel's intensity. If the prior distribution is uniform, the posterior distribution will only depend on the observation model. For a gradient based motion model, the observation model relates to optical flow equation.

To combine prior knowledge into the estimate f, a prior distribution for the motion field must be selected. This distribution may favor the motion values that are near to some expected value $\overline{f}$ (adjacent sites):

$$P(f, d_{i-1}) = \frac{1}{2\pi\sigma_f^2} \mathrm{e}^{\left\{-\frac{\sum_{c\in C} V_c(f)}{2\sigma_f^2}\right\}}, \tag{3.4}$$

where $V_c = \left\|f - \overline{f}\right\|$, and $\|.\|$ denotes L2 norm. The scalar $\frac{1}{2\sigma_f^2}$ as a divisor of potential may be viewed as a *regularization constant* between observation and prior model. In the first-order Gibbs Markov random field (GMRF) as is shown in Figure 2.2, it has four of the following pair potentials. V_c becomes:

$$V_{(s,s')}(f(s), f(s')) = \lambda \left\|f(s) - f(s')\right\|, \tag{3.5}$$

where λ is a weighting constant.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

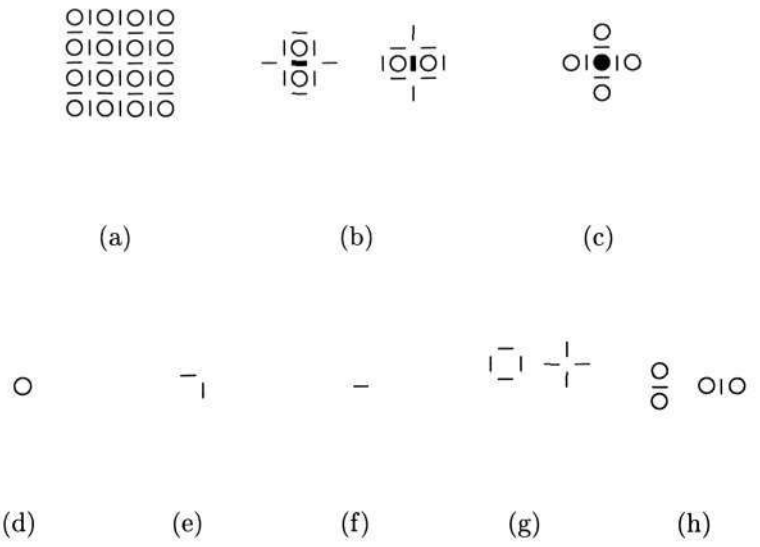


Figure 3.2: Line Field Lattice with Neighborhood System and Cliques: (a)Line Lattice; (b) Neighborhood of a Line Site; (c) Neighborhood of a Pixel; (d) Pixel Site; (e) Line Sites; (f) Single-Element Line; (g) Four-Element Line Cliques; (h) Three-Element Clique.

If the observation and prior models are combined, they create a well-defined MAP-GMRF framework. From this point, we can design the optimization strategy. The framework allows explicit modeling of discontinuities by suspending the smoothness in the prior model.

To handle discontinuities, the line and segmentation models can be added into the framework. The line model switches the line on and off according to the line neighborhood system and the cliques of the location under consideration [13, 25] as is shown in Figure 3.2. This is called a line process. The line process influences the prior model. The prior model is now defined by the following GMRF:

$$V_{(s,\,s',\,s\bowtie s')}(f(s),\,f(s'),\,l(s,\,s')) = \lambda \,\|f(s) - f(s')\|\,(1 - l(s,\,s')), \qquad (3.6)$$

where $s \bowtie s'$ means s and s' are neighbors, $l(s,\,s')$ is the line field between sites s and s', and f is the motion field.

Another way to consider discontinuities in the model is by the addition of a segmentation field [3, 25, 26]. A label field o can represent the segmentation, where

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

all pixels of the same region possess the same labels. The region of support for the smoothness constraint can be limited to pixels in the respective segment and the prior distribution of the label field can be modeled by the first or second-order GMRF. The potential is:

$$\begin{aligned} &V_{(s,\,s',\,r\bowtie r')}\left(f(s),\,f(s'),r(s),\,r(s')\right) = \\ &\lambda_f \left\|f(s)-f(s')\right\| \delta(o(s)-o(s')) + \lambda_o(1-\delta(o(s)-o(s'))) \end{aligned}, \tag{3.7}$$

where λ_f and λ_o are the weighting functions. The delta function δ in the first term suspends the smoothness constraint across region boundaries and the second term gives support to compact regions.

Beside the above purpose, the segmentation field itself can be the outcomes of some applications such as: MRF-based moving object detection [27], hierarchical MRF model for model-based multi-object tracking [28], motion segmentation which is based on dominant motion estimation and the detection of uncovered regions [29], and object segmentation in stereo image using cooperative line field [30].

For the next section, we follow a classification strategy that is introduced by Konrad, Dubois and Stiller [3,4,26,31,32]. Konrad is among the earliest contributor of MRF based motion estimation. The other contributors are Heitz, Bouthemy and Morris [5,33–35].

3.3 Literature Survey on Gradient Based Model

Gradient based motion estimation has been developed successfully in many areas of image processing and computer vision. Gradient based motion estimation requires a smoothness constraint to solve its ill-posed problem.

This section discusses the optical flow as the motion representation, regularization technique as the estimation criteria and several optimization techniques as the search strategies. The optimization techniques are Gauss-Seidel Method and Successive Over Relaxation. We end the discussion with some examples of applications.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3.3.1 Motion Representation: Optical Flow

Optical-flow/gradient-based is a paradigm to determine visual motion [6,24]. In this method, the recovered motion-flow is based on local spatial-temporal changes.

Optical flow uses intensity constancy, spatial and temporal coherence as its model. Let $d(x, y, t)$ denote the image intensity at the point (x, y) and time t. The intensity constancy gives the equation:

$$\frac{\partial d}{\partial x}\frac{dx}{dt} + \frac{\partial d}{\partial y}\frac{dy}{dt} + \frac{\partial d}{\partial t} = 0. \tag{3.8}$$

The optical flow vector at (x, y) is $f(x, y) = (\frac{dx}{dt}, \frac{dy}{dt})^T = (u, v)^T$. These vectors form an optical flow field for a given image. This single equation with two unknowns does not give a unique optical flow. An additional constraint must be imposed on the flow.

Equation 3.8 can be solved by minimizing:

$$V(d(x, y)|f(x, y)) = (ud_x + vd_y + d_t)^2 \tag{3.9}$$

where $\frac{\partial d}{\partial x} = d_x$ and $\frac{\partial d}{\partial y} = d_y$ with a consideration that the noise is in it. One of the early methods for solving this equation is from Horn and Schunck paper [36]. Besides their method, many others are available in the literature. Some of those methods are partial derivatives, Laplacians flow velocities, minimization technique with the smoothness constraint, and iterative method.

A study of eight optical flow algorithms designed by Anandan, Horn & Schunck, Lucas & Kanade, Nagel, Singh, and Uras et al, was conducted by Galvin et al [37]. Another thorough evaluation of optical flow techniques has also been presented in [38]. An implementation for matching two perspective views is presented by [39]. An optimal structure motion estimator is written by [40]. Feature-based motion estimation to match line drawings is presented by [41, 42].

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3.3.2 Estimation Criteria

3.3.2.1 Regularization

Smoothness is the constraint used to control the flow coherence in a neighborhood system. One way to impose this constraint is to minimize the squared magnitude of the gradient of the optical flow. By adding this constraint, the posterior potential function becomes:

$$V(f(x,y)|d(x,y)) = V(d(x,y)|f(x,y)) + \lambda V(f(x,y)), \tag{3.10}$$

where λ is a weighting factor and the constraint is:

$$V(f(x,y)) = \|\nabla u\|^2 + \|\nabla v\|^2 = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2. \tag{3.11}$$

The first right term in Equation 3.10 corresponds to the observation potential and the second right term corresponds to the prior potential of the MAP-MRF framework.

Integrating Equation 3.10 over all (x, y) gives the posterior energy function:

$$U(f|d) = \int\int [V(d(x,y)|f(x,y)) + \lambda V(f(x,y))]\mathrm{d}x\mathrm{d}y. \tag{3.12}$$

The posterior energy function gives a value to represent the state of the motion field. This concept will further be used to design the search strategy.

The Horn and Schunck [36, 43] laid the foundation of optical flow using their global smoothness approach. On-going research aims to modify the smoothness constraint to achieve some desired outcomes.

The drawbacks of the quadratic gradient-based model are:

1. Optical flow is not valid for large motions, because the calculated gradient values from a small number of pixels are used to lead the search strategy;

2. Optical flow is not valid for areas exhibiting very large luminance variations: occlusions, exposure, reflection patches, etc, because the gradient values are lost;

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3. For quadratic regularization, the results are over-smoothed, because the potential value will be higher when the gradient of optical flow is higher.

Besides the Horn-Schunck model, Heitz and Bouthemy [34] proposed an MRF interaction model which combines constraints from the gradient based and edge based paradigms. Nagel and Enkelmann [44] introduced *oriented smoothness* constraint. These approaches can be summarized as being based on adding the weight matrix that penalizes variations in the motion field depending on the spatial changes in gray-level content of the video. Later, Fogel [45] used the concepts of directional smoothness and adaptive weighting in a hierarchical formulation. The weight function in this formula is a function of the low pass filter parameter and the spatial resolution. The resolution parameter changes multiresolution representation. The detail can be seen in their respective papers.

Another alternative is to use a robust function such as Tukey, Huber, Andrews, and Hampel for the smoothness prior. The functions are statistical tools in which *outliers* are an issue. A robust procedure is an ability to be insensitive to the violation of assumptions or the robustness of a function to the *outliers*. For optical flow the robust method is proposed by Black in his Ph.D thesis [7–9]. Using $s = (x, y)$ and $f(s) = (u(s), v(s))^T$, Equation 3.9 becomes

$$V(d(s)|f(s)) = g(u(s)d_x(s) + v(s)d_y(s) + d_t(s)), \tag{3.13}$$

where $g(.)$ is a robust function that will be explained later. The posterior energy function is

$$\begin{aligned} U(u, v|d) \quad = \quad & \textstyle\sum_{s \in S} [\lambda g(u(s)d_x(s) + v(s)d_y(s) + d_t(s), \sigma_1) + \\ & \textstyle\sum_{s' \in N(s)} [g(u(s) - u(s'), \sigma_2) + g(v(s) - v(s'), \sigma_2)]\Big] \end{aligned} \tag{3.14}$$

where s is an element of sites S. In Equation 3.14, we need to choose appropriate values for σ that determine whether the data and smoothness error are considered outliers.

To see the need of a robust function, let us analyze the influence of a variable x to the value of its quadratic $g(x) = x^2$, and the first derivative $h(x) = 2x$. $h(.)$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

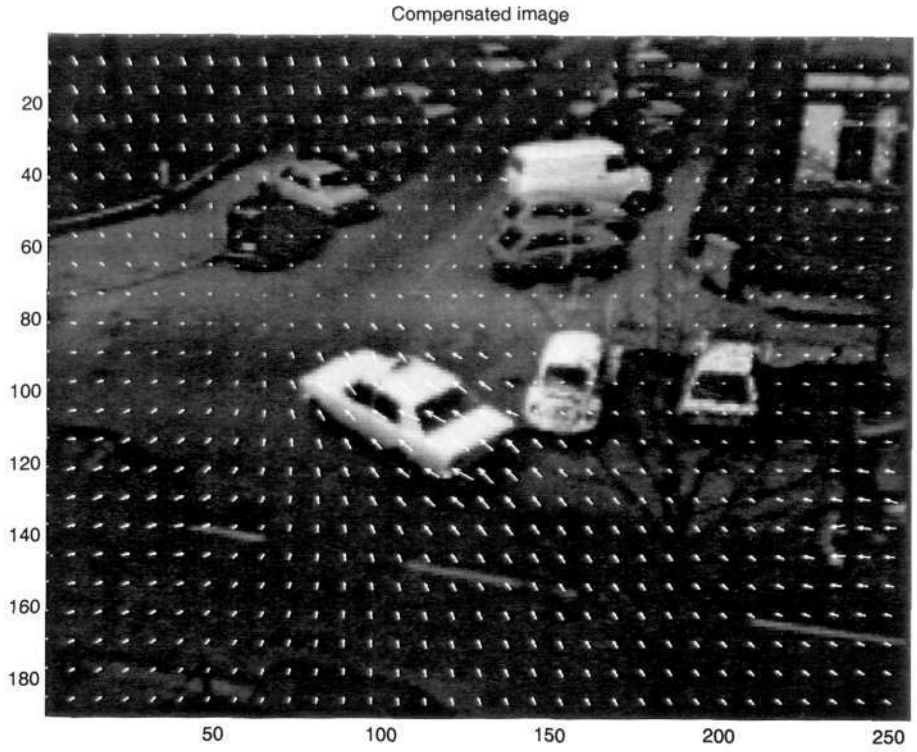


Figure 3.3: Result of Robust Motion Estimation.

is called *influence function* or *interaction function.* The influence function of data points in this case is linearly increasing and without bound.

To increase the robustness of a function, the g-function must be more forgiving about outlying measurement, where the influence of outliers tends to zero. One of the robust estimator for the g-function is Lorentzian estimator:

$$g(x,\sigma) = \log\left(1 + \frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right). \tag{3.15}$$

Its influence function or the derivative w.r.t x is

$$h(x,\sigma) = \frac{2x}{2\sigma^2 + x^2}. \tag{3.16}$$

Most of the regularization technique which implement optical flow equation will produce the dense motion field as is shown in Figure 3.3. This figure is generated by the Black's algorithm [7]. The search strategy is a successive over relaxation method that will be explained later.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3.3.2.2 Block Motion Model

The block motion model [46] is an example for the implicit smoothness. It assumes that the motion vector remains unchanged over a particular block of pixels, denoted by B. The likelihood energy function is:

$$U(d_i|d_{i-1}, u, v) = \sum_{s \in B} (d_x(s)u + d_y(s)v + d_t(s))^2 \tag{3.17}$$

where $d_x(s)$ is the partial derivative of intensity w.r.t x, $d_y(s)$ is the partial derivative of intensity w.r.t y, and $d_t(s)$ is the partial derivative of intensity w.r.t t.

The difference between the regularization technique and this model is in the handling of the 'ill-posed' problem. The result for this model is a sparse motion field. For example, in the Hamburg Taxi Sequence, the image is divided into 9×11 blocks. Each block is solved by computing the partials of the error U with respect to u and v and setting them equal to zero. We have

$$\begin{array}{rcl} \sum_{s \in B} \left(d_x(s)\widehat{u} + d_y(s)\widehat{v} + d_t(s)\right) d_x(s) & = & 0 \\ \sum_{s \in B} \left(d_x(s)\widehat{u} + d_y(s)\widehat{v} + d_t(s)\right) d_y(s) & = & 0 \end{array}, \tag{3.18}$$

where $\widehat{u}$ is the estimate of u and $\widehat{v}$ is the estimate of v. Solving the above equations simultaneously, it gives us:

$$\begin{bmatrix} \widehat{u} \\ \widehat{v} \end{bmatrix} = \begin{bmatrix} \sum_{s \in B} d_x(s)d_x(s) & \sum_{s \in B} d_x(s)d_y(s) \\ \sum_{s \in B} d_x(s)d_y(s) & \sum_{s \in B} d_y(s)d_y(s) \end{bmatrix}^{-1} \begin{bmatrix} -\sum_{s \in B} d_x(s)d_t(s) \\ -\sum_{s \in B} d_y(s)d_t(s) \end{bmatrix}. \tag{3.19}$$

The example result of the motion field is shown in Figure 3.4. This approach uses an inverse function. Later we will show that the motion field could be further improved using our proposed method in Chapter 6.

3.3.3 Search Strategies: Gradient-Based Optimization

At this point, we have the optical flow as a model representation and the estimation criteria. We now need a search strategy to solve the estimation problem.

Two separate approaches in using the optical flow exist. They are the explicit

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

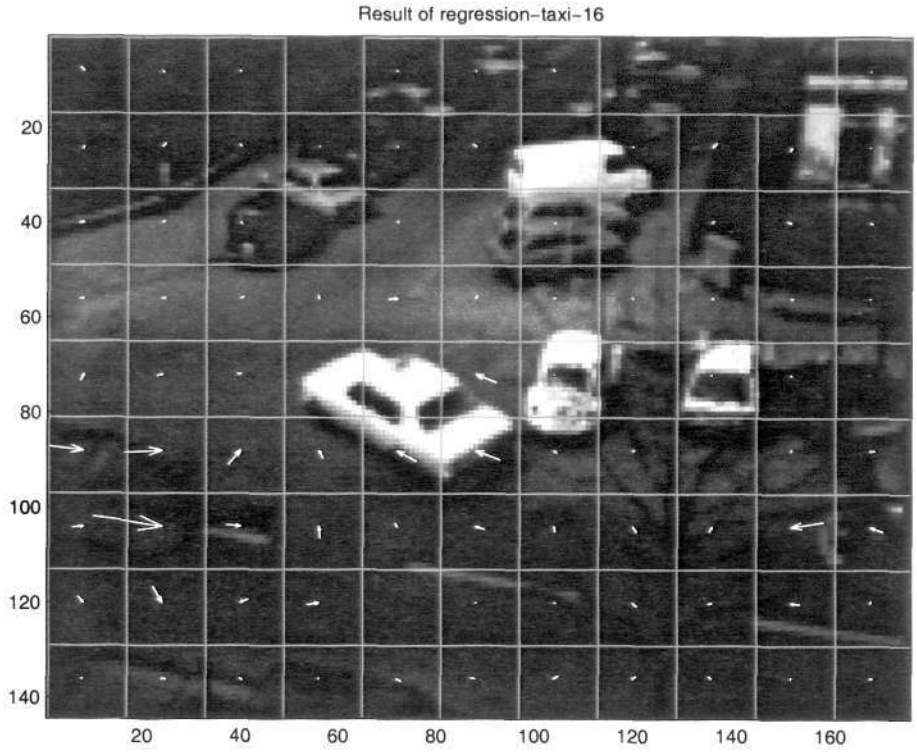


Figure 3.4: Result of Block Motion Model Motion Estimation using Block-Size of 16×16.

and implicit smoothness. Models such as those of Lucas and Kanade's approach [46] and its variant exhibit implicit smoothness. This approach is known as parametric motion fields. Usually, the *least-square technique* is used. On the other hand the explicit smoothness approach enforces the motion field based on the surrounding field.

For this gradient based optimization, we describe two techniques: Gauss-Seidel method and Successive Over Relaxation method.

3.3.3.1 Gauss-Seidel Method

Horn uses iterative Gauss-Seidel method [36,43]. A new set of velocities, $(u^{(t+1)}, v^{(t+1)})$, is calculated from the estimated derivatives and the average of the previous velocities, $(\overline{u}^{(t)}, \overline{v}^{(t)})$. The equations are:

$$u^{(t+1)} = \overline{u} - d_x[d_x\overline{u}^{(t)} + d_y\overline{u}^{(t)} + d_t]/(\alpha^2 + d_x^2 + d_y^2), \tag{3.20}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$v^{(t+1)} = \overline{v} - d_y[d_x\overline{v}^{(t)} + d_y\overline{v}^{(t)} + d_t]/(\alpha^2 + d_x^2 + d_y^2), \tag{3.21}$$

where α is a weighting constant, t is the iteration number and

$$\begin{aligned} d_x \approx \frac{1}{4}\{&d_{i,j+1,k} - d_{i,j,k} + d_{i+1,j+1,k} - d_{i+1,j,k} \\ &+ d_{i,j+1,k+1} - d_{i,j,k+1} + d_{i+1,j+1,k+1} - d_{i+1,j,k+1}\}, \end{aligned} \tag{3.22}$$

$$\begin{aligned} d_y \approx \frac{1}{4}\{&d_{i+1,j,k} - d_{i,j,k} + d_{i+1,j+1,k} - d_{i,j+1,k} \\ &+ d_{i+1,j,k+1} - d_{i,j,k+1} + d_{i+1,j+1,k+1} - d_{i,j+1,k+1}\}, \end{aligned} \tag{3.23}$$

$$\begin{aligned} d_t \approx \frac{1}{4}\{&d_{i+1,j+1,k+1} - d_{i,j,k} + d_{i+1,j,k+1} - d_{i+1,j,k} \\ &+ d_{i,j+1,k+1} - d_{i,j+1,k} + d_{i+1,j+1,k+1} - d_{i+1,j+1,k}\}, \end{aligned} \tag{3.24}$$

where i, j are the 2-D pixel index, k is the frame index, and the local averages $\overline{u}_{i,j,k}$ and $\overline{v}_{i,j,k}$:

$$\begin{aligned} \overline{u}_{i,j,k} = &\frac{1}{6}\{u_{i-1,j,k} + u_{i,j+1,k} + u_{i+1,j,k} + u_{i,j-1,k}\} \\ &+ \frac{1}{12}\{u_{i-1,j-1,k} + u_{i-1,j+1,k} + u_{i+1,j+1,k} + u_{i+1,j-1,k}\}, \end{aligned} \tag{3.25}$$

$$\begin{aligned} \overline{v}_{i,j,k} = &\frac{1}{6}\{v_{i-1,j,k} + v_{i,j+1,k} + v_{i+1,j,k} + v_{i,j-1,k}\} \\ &+ \frac{1}{12}\{v_{i-1,j-1,k} + v_{i-1,j+1,k} + v_{i+1,j+1,k} + v_{i+1,j-1,k}\}. \end{aligned} \tag{3.26}$$

are the Laplacian of flow velocities.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3.3.3.2 Successive Over Relaxation

Black [7] uses the successive over relaxation. The technique can be traced back to Jacobi and Gauss-Seidel method. The explanation on Jacobi is available in Appendix A.

Written in a complete form, the partial derivatives of the robust flow from Equation 3.14 are:

$$\frac{\partial E}{\partial u(s)} = \sum_{s \in S} \left[\lambda d_x(s) h\left(d_x(s)u(s) + d_y(s)v(s) + d_t(s),\, \sigma_1\right) + \sum_{s' \in N(s)} h\left(u(s) - u(s'),\, \sigma_2\right) \right], \tag{3.27}$$

$$\frac{\partial E}{\partial v(s)} = \sum_{s \in S} \left[\lambda d_y(s) h\left(d_x(s)u(s) + d_y(s)v(s) + d_t(s),\, \sigma_1\right) + \sum_{s' \in N(s)} h\left(v(s) - v(s'),\, \sigma_2\right) \right], \tag{3.28}$$

where $h(x, \sigma)$ is the Lorentzian influence function.

The iterative update equations for minimizing E at step $(t+1)$ are

$$u^{(t+1)}(s) = u^{(t)}(s) - \omega \frac{1}{T(u(s))} \frac{\partial E}{\partial u(s)}, \tag{3.29}$$

$$v^{(t+1)}(s) = v^{(t)}(s) - \omega \frac{1}{T(v(s))} \frac{\partial E}{\partial v(s)}, \tag{3.30}$$

where ω is an over-relaxation parameter. It is used to over-correct the estimate of $f^{(t+1)}$ or $\begin{bmatrix} u^{(t+1)}(s) & v^{(t+1)}(s) \end{bmatrix}^T$.

The terms, $T(u(s))$ and $T(v(s))$, are the upper bounds on the second partial derivatives of E:

$$T(u) \geq \frac{\partial^2 E}{\partial u^2(s)}, \qquad \forall s \in S, \tag{3.31}$$

$$T(v) \geq \frac{\partial^2 E}{\partial v^2(s)}, \qquad \forall s \in S. \tag{3.32}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

The second derivative is maximized when both the data and the smoothness errors are zero, which imply

$$T(u) = \frac{\lambda d_x^2}{\sigma_1^2} + \frac{4}{\sigma_2^2}, \tag{3.33}$$

$$T(v) = \frac{\lambda d_y^2}{\sigma_1^2} + \frac{4}{\sigma_2^2}. \tag{3.34}$$

3.3.3.3 Hierarchical Optimization

We have several purposes in applying the multiresolution scheme. Some of them are to accelerate the computation and to attain long motion vector length. The scheme influences the model, and in turn influences the result.

Gradient-based motion estimation assumes that the image motion is small; less than a pixel. To increase the length of estimated motion, a pyramidal representation of images is used. The reduction of the size of image is implemented as convolution and sub-sampling. The general effect desired is to reduce the high frequency components of the image at the coarser scales.

A simple multiresolution gradient-based scheme performs the following:

- Project motion field with interpolation from coarse to finer resolution,
- Warp the image using motion field,
- Compute new estimate of the motion field from this warped image,
- Update the motion field with the computed value above.

This algorithm uses pyramid image and pyramid flow fields.

3.4 Literature Survey on Matching Based Model

3.4.1 Motion Representation

Since motion is estimated based on the variation of intensities, we need a model to relate them. The observation models assume an existing relationship between

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

motion parameters and image intensity. The usual assumption is that the image intensity remains constant along a motion trajectory. Assuming a spatial sampling of the images, this condition can be expressed as follows:

$$d_i(s) = d_{i-1}(s + f(s)) + \epsilon, \tag{3.35}$$

where $d_i(s)$ is the image intensity of sequence number i on site s, f_s is a motion label of the image on site s, and ϵ denotes noise.

3.4.2 Estimation Criteria

Several estimation criteria which are being used in motion estimation are displaced frame difference (DFD) criteria, the frequency domain, regularization and Bayesian criteria [23]. The DFD and the Bayesian criteria are discussed in the following sections.

3.4.2.1 A Class of Criteria

A class of criteria arising from the constant-intensity assumption is called displaced frame difference —DFD, or displaced block difference — DBD, or displaced region difference — DRD, or displaced pixel difference — DPD, according to the region of support. These criteria try to minimize the error:

$$\varepsilon_i = \| d_i(s) - d_{i-1}(s + f(s)) \|_p, \ \forall s \epsilon \mathbf{R}, \tag{3.36}$$

where d is the intensity of image, s is the pixel site, i is the sequence number and $f(s)$ is the motion of a region, $\mathbf{R}$; this can be a frame, a block, a region or a pixel. p is the norm number.

To measure the magnitude of the prediction error, ε, the norm can be chosen from the family of norms (the absolute or the quadratic). We can also use the families of robust estimators: the L1 norm, the Huber min-max estimator, the Andrew's Sine, the Tukey's Biweight and the Lorentzian [7].

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3.4.3 Search Strategies

3.4.3.1 Matching Approaches

One of the search strategies is *full search block matching* [3, 47]. This strategy compares the motion compensated prediction error from various motion candidates f within the search region of the motion model. The best match results after the process of comparison is considered as the optimal estimate of f^*. This technique works for a block of pixels as a region of support.

If the region of support is only one pixel, as we have mentioned it before, it is commonly called a dense motion field. A technique called relaxation can be used to solve the dense motion field problem. This technique updates the motion vector of each site by some site-visiting order, e.g., the line scanning. A well-known deterministic relaxation technique is an iterated conditioned mode (ICM). For GMRF modeling, the conditional distribution of a single variable, $f^k(s)$, is completely specified by its neighborhood and it will be updated sequentially until the convergence is achieved. The distribution is:

$$f^k(s) = \arg\max_{f^k(s)}\{P(f^k(s)|f^{k-1}(s)\}. \tag{3.37}$$

Both stochastic (e.g. metropolis) and deterministic relaxation (e.g. iterated conditioned mode) can be used to find the optimal motion vector from its state space [23]. On the other hand, mean field theory is able to produce a result approximately the same as the stochastic model with a faster convergence time [48].

Another search strategy that has the ability to lower the computational burden is hierarchical optimization. In this strategy, the motion parameters are estimated at the lowest resolution first; at this stage the computation load is low, then the result is propagated to a higher resolution [31].

In addition to the reduced computational load, the smoothness constraints are evident at the same time when the motion parameters propagate into a higher resolution.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Sub-Optimal Search Scheme. The main idea is to give a prior knowledge to the search area. Jain and Jain [47] proposes a two-dimensional logarithmic search scheme and Ghanbari [49] extended this scheme. This search scheme relies on the assumption of convexity in estimation criteria. Another alternative scheme by de Hann *et. al.* [50] selects candidate vectors based on the available vectors from the spatial-temporal neighboring blocks. This can reduce the computational demand.

Fast Algorithms. Researchers define a fast algorithm according to their own definitions. Depending on the purpose of the implementation, a fast algorithm embeds a new constraint to reduce the computation task. Some researchers put emphasis on the communication rate [51–53]. It means that the motion estimation process has another limitation in the search strategy. In this case this process is a function of the rate of communication.

Other researchers modify the search strategy to reduce a search area. Those are such as the three-step search and its variants. Another good alternative, without reducing the search area, is Successive Elimination Algorithm (SEA). This is one interesting approach that harnesses advantage from mathematical inequality. Some details of various fast algorithms are emphasized below.

Three-step Search. This search strategy is implemented with the assumption that the search area is smooth enough. Further computation reduction approaches such as the diamond search, the cross-search, and the block-based gradient descent search algorithm (BGGDS) [54] work under the same assumption.

This assumption reduces the computation time by jumping to a planned-group of candidates, e.g. when the length of search array is maximum ± 7: the first group of candidates can be ± 4, then the second group of candidates can be ± 2, and the last group of candidates are ± 1.

Successive Elimination Algorithm. A clear understanding of how the block matching algorithm is computed provides a good ground in trying to reduce the computation task. As it was said, the reduction involves the block based structure, all algorithms that use the DBD can be slightly modified to increase the speed

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

of computation. The interesting approach is the use of a mathematical inequality in [55] and [56]. The mathematical inequality is:

$$|\|X\| - \|Y(i,j)\|| \leq \|X - Y(i,j)\|, \tag{3.38}$$

where $\|X\| = \sum_k |x_k|$ is the vector representing the block in the current frame, while the vectors $Y(i,j)$ refer to the possible candidate blocks in the previous frame. The left term in Successive Elimination Algorithm (SEA) is pre-computed efficiently using procedure described in [55]. Assume that in the middle of an exhaustive search algorithm (ESA), the motion vector (m,n) is the best matching vector found so far and that vector (i,j) is the next vector to be considered. If $|\|X\| - \|Y(i,j)\|| > \|X - Y(m,n)\|$, a test that can be performed using previously computed quantities, then from Equation 3.38, we see that (i,j) is not a viable candidate for the optimal motion vector and that, therefore, the norm does not need to be computed. If this test fails, then the norm $\|X - Y(i,j)\|$ does need to be computed. To further reduce the computation or increase the speed of estimation, a fast calculation of the norm sum is used. The calculation involves two steps: (1) keeping the sum of each column and save them for all row strips, (2) calculating block by summing the saved sum of each column.

The performance evaluation, as mentioned in [55], shows that the average number of search point per block is 144, whereas the exhaustive search algorithm needs 1089 and the three-step search needs 33 search points per block for a particular block of the Salesman sequence.

A further speed-up is achieved by using partial distortion elimination (PDE) which tests the partial sum against the current minimum after each row instead of each computation. In the paper [56], the speed-up from exhaustive search (ESA) for SEA+PDE is 29.40 instead of 8.07 for SEA. PDE procedure can be added to block matching algorithm, whether we use fast calculation or the normal way to calculate the norm sum.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

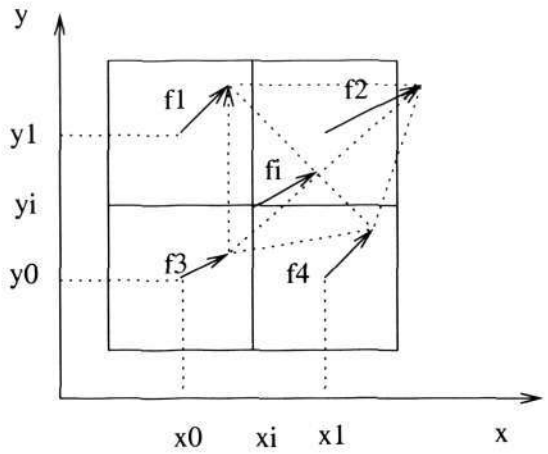


Figure 3.5: Spatial Interpolation of Motion Vectors.

3.4.4 Supporting Techniques

We list some techniques to improve the quality of the estimated motion field: bilinear warping, warping prediction and block-segmentation. The bilinear warping is useful for the dense motion field. The warping prediction is a trick to reduce blocky effect of block matching. Block-segmentation is another fast segmentation that divides the block into two regions, each with one of two given distinct motion vectors.

3.4.4.1 Bilinear Warping

Bilinear warping has two types of directions, backward and forward. It is cheaper to compute backward motion vector than the other way. For the backward motion, the tail of vectors lie somewhere between four sites. A bilinear calculation is able to generate the intensity value that would be warped to the compensated image.

3.4.4.2 Warping Prediction

Warping prediction is implemented to produce the dense motion field in [57]. Suppose that the motion vector exists in the middle of the block, the bilinear geometric transformer transforms the motion field into a smooth motion field over the image. The result is the smoothed-compensated image. Figure 3.5 shows four motion vectors and an interpolated motion vector.

The equation for interpolated motion vector f_i as a percentage of distance from the desired motion vector and f_1 is:

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$f_i = f_1(a(1-b)) + f_2(ab) + f_3((1-a)(1-b)) + f_4(b(1-a)), \tag{3.39}$$

where $a = \frac{x_i - x_0}{x_1 - x_0}$ and $b = \frac{y_i - y_0}{y_1 - y_0}$.

The concept of the block matching algorithm is that an object moves in a whole block. The matching algorithm which implements a fixed block size suffers conceptually when a motion discontinuity exists inside the block. Making the block smaller is one of the partial solutions. Another possibility is to use a concept from the field of robust statistic that has been mentioned in Section 3.3.2.1 to remove outliers from their influence to estimation criteria.

3.4.4.3 Block Segmentation

Orchard [58] assumes the existence of two distinct motions f_1 and f_2 for each block as the best two among the candidate motion vectors. This assumption is able to divide the block into two regions with their respective motion vector. The way it divides the block is implemented using a prior function. This function relates its value to the length L of a motion boundary.

Elias [25] modifies the Orchard's approach with a deterministic approach. The conventional block matching provides the initial data, the motion field f. Each pixel is assigned a motion vector and its corresponding estimation error, E. The estimation error, which is calculated by an overlapping block, is the lowest. With this arrangement, the field of estimation error and the directional filter can be used to identify the existence of discontinuity in the region. As it happened when a rise to a ridge in the estimation error field, E, exists, the orientation must be toward the discontinuity.

Line fields are used to do segmentation for the blocks. For vertical segmentation, each row must have one active edge, so that each row r is divided into two segments at position l_r. The likelihood function is the displaced frame difference with its respective chosen: f_1 on the left, the active edge, and f_2 on the right. The prior function calculates the length of active edges. On the whole the energy function consists of the likelihood function and the prior function.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3.5 Literature Survey on Multiresolution Matching Based Model

Multiresolution represents the varieties of spatial density of grid sites. It also has many choices of smoothing kernels for pyramid generation.

3.5.1 Motion Representation

The multiresolution framework finds the motion field via a reduced computational task. The hierarchical representation is a common framework for multiresolution in which images are arranged in pyramidal structure.

When motion models are arranged in hierarchical representation [3,59], they use variable-size of the region of support. In such a representation, motion estimation can be modeled at multiple levels of detail, making it a possibility to extract the coarse characteristic first and subsequently adding finer details. The main reason for hierarchical representation is to speed up the estimation process. Moreover the hierarchical representation reduces violation of the underlying assumption such as a locally linear intensity variation. Also, the hierarchical representation eliminates the spatial high frequency aliasing component undergoing large motion [60].

In a hierarchical algorithm the basic operation is data pre-filtering. Usually, during each stage of filtering, data are sub-sampled to provide a more compact representation at a resolution level. The Gaussian and Laplacian pyramids of images have previously been used in hierarchical motion estimation [25, 31, 32, 61]. Once estimation is obtained at some level, it must be transformed to the next higher resolution level for subsequent improvement.

Hierarchical representation, i.e. a pyramid based image, can be constructed by a Laplacian operator [62]. The main advantage of the pyramid representation is a rapidly decreasing image size. However, the algorithmic process in this representation becomes a main disadvantage that makes theoretical analysis complicated [63]. A Laplacian pyramid is considered as a bandpass filter. Another filter used to generate a pyramid is the smoothing/low-pass filter.

The combination of an affine model and a multiresolution processing is presented

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

in [64]. In [64], S. Kruger commented that a small analysis window is unlikely to contain sufficient intensity variation for unique motion to be determined, whereas a large window will contain image variation incompatible with the chosen model. This is also known as the aperture problem [65]. Robust motion estimation can be achieved if the motion analysis is performed using large analysis windows. The affine transformation is the best choice of the higher-order motion model because of its representational compactness and reasonable model accuracy. This motion model can also be estimated in a computationally efficient manner.

Computation efficiency (speed), accuracy and robustness are the issues that make multiresolution representation attractive. P. Anandan, J. R. Bergern and K. J. Hanna in [60] suggest that it is efficient and necessary to reduce the high frequency aliasing component undergoing large motion. This can be done by using multiresolution estimation techniques.

Odobez and P. Bouthemy [66] use multiresolution least mean squares scheme to incrementally estimate the motion model through coarse-to-fine refinement, Tziritas [61] implements detection and localization of moving object in multiresolution representation.

Multi-grid uses a quad-tree decomposition technique where the block-size is decomposed according to the local requirements. The repeated procedure will split the region until a certain smaller block-size for a certain region that is suitable to be split. Chan [67] proposed the original scheme of this approach. The initialization is given by conventional block-matching with a uniform size. For each block, the estimation error is compared to a certain threshold. The ones which are not below the threshold value are split and the procedure tries to find another motion vector. The weaknesses of this scheme are the unreliability of smaller block-size, and the fixed threshold is not suitable for a wide-range of images.

Dufaux and Kunt [68] modify the procedure for splitting block, instead of doing a new search, by using prior knowledge of four neighboring blocks to calculate a desired motion vector. The children blocks have a choice of choosing motion vectors from the neighboring blocks.

In block based area, Burt and Adelson's pyramid representation of the im-

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

age [62] initiates pyramid motion estimation. A further improvement referred as 'neighborhood-blocks motion vector estimation using pyramidal data structure' was proposed by Zan et. al [69]. The conventional pyramid motion estimation has the equation that may be expressed as

$$f^p(s) = 2.f^{p+1}(s/2) + \Delta f^p(s), \tag{3.40}$$

where $f^k(s)$ is the motion vector on a site s and on a resolution level k, and Δf is the refined length of motion vector.

The equation says that, at the finer level of the pyramid, the motion vectors from the coarse level are multiplied by a factor of two, and a further refinement process will be executed. Instead of a single motion vector, their proposal involves the eight neighbors of a motion vector from the coarse image to be included in finer level and a further refinement for all of them are compared together through its matching algorithm. The computational complexity of their proposal is a maximum $9\times$ more than the conventional pyramid motion estimation because it involves the 8 neighbors. With this proposal, the result shows reduced false motion vectors.

3.5.2 Estimation Criteria

The estimation criteria of multiresolution are basically the same as in the single resolution. As the motion representation is already in multiresolution, the estimation criteria are automatically implemented for each resolution.

3.5.3 Search Strategies

3.5.3.1 Subblock Multiresolution Motion Estimation

To increase the computation speed, we can use sub-block multiresolution motion estimation to enhance the block matching performance [52]. The image is arranged as a pyramid image. The search algorithm uses 16×16 macroblocks, and each macro-block contains four 8×8 subblocks. Each motion vector must record the sum absolute difference (SAD) to be used later.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

The algorithm starts from a coarse resolution in a three-resolution pyramid with subblocks of 8×8 pixels and macroblocks of 16×16 pixels and candidate motion vector length with a maximum of ± 4. The SAD of each macroblock is calculated at the same time using the summation of the four subblocks. In every resolution five motion vectors are available for each macroblock. The result is propagated to the finer resolution for further minimization based on the five candidates with an addition of ± 1 to the length of each candidate. The maximum length is $(((4 \times 2) + 1) \times 2) + 1$.

3.5.3.2 Hierarchical Block Matching

To improve vector reliability, Bierling [70] proposed a hierarchical block matching scheme. Every scale uses a specific block-size, such as a smaller supporting region for the lower resolution. For flat moving regions, using this scheme, we can recover a moving field based on the results achieved from a coarser scale. A static region may be initiated as a moving region at the finer scale. A suggested solution is to use motion detection to detect the static regions.

3.5.4 Projection

To transfer the motion field, the simplest equation is "projection using duplication":

$$f^p(s) = 2.f^{(p+1)}(s/2). \tag{3.41}$$

The equation is similar to Equation 3.40 without a further improvement.

A better scheme uses "projection using interpolation". For this scheme, we have four pixels from one parent after the projection using duplication. According to the pixel position, four-diagonal-data averaging is executed for the odd-odd pixel position; Two-horizontal-data averaging is executed for the even-odd pixel position; Two-vertical-data averaging is executed for the odd-even pixel position; The unchanged value is implemented for the even-even pixel position.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 4

Discontinuity Adaptive MRF Model for Optical Flow

It is a well-known fact that the optical flow equation is ill-posed. To make the equation solveable, a prior constraint is needed. The smoothness prior is usually added to make it well-posed. This strategy is also equivalent to the regularization method, which is a special case of the MRF model. An often used smoothness prior is the quadratic function. Though it makes the equation solvable, it often produces over-smoothed labels. For example, even if discontinuities are in the motion field, the result will be smoothed due to the smoothness constraint. Hence the line field representing these discontinuities is a useful switch to cancel the smoothness constraint. As an alternative approach, the usage of *an adaptive interaction function* (AIF) is proposed here. AIF is a modified version of the quadratic function. Using AIF, the line field is not necessary. This approach is shown to be effective [6] and we have adopted it into the optical flow problem using the MAP-MRF framework. It will be shown that our approach produces a sharp motion field at the discontinuity area and we can set the strength of discontinuity while keeping its stability.

Section 4.1 gives an introduction. Section 4.2 shows the optical flow equation with smoothness constraint. Section 4.3 explains the use of the discontinuity adaptive model to modify the smoothness constraint. Section 4.5 shows the experimental results. Section 4.6 gives the conclusion.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

4.1 Introduction

Optical flow estimation is an ill-posed problem. A constraint is needed to solve the problem [36, 43]. The addition of a smoothness constraint can solve the problem. However if the constraint is a quadratic function, the estimation result is over-smoothed. To tackle the appearance of this new problem, several solutions are available to date: motion estimator using line process model [31], motion estimator using truncated quadratic function [71], and motion estimator using robust formulation [7–9, 72]. Another solution we proposed is a motion estimator using *a discontinuity adaptive* smoothing model. This smoothing model is created by Li [6].

A line field as a discontinuity label has been introduced in Chapter 3. The line field is considered as a discrete MRF labeling. The presence of both smoothing and discontinuity functions cause instability during the optimization process and it is difficult to choose a weight for each function. Moreover, this is a mix between the continuous and discrete labeling.

The simplest discontinuity function embedded into a smoothing function is a truncated quadratic function. This function needs a threshold value. The value will decide whether the smoothing function truncates the interaction between data and its neighbors. A heuristic approach is usually used to set the threshold value.

In another solution, robust statistics have played an important role to preserve the discontinuities. The robust method handles the *outliers.* Robust technique loosely means that a system built with this technique is insensitive to the nature of data, parameters or environment. For robust technique, the estimator's examples are Lorentzian, Geman-McClure, Cauchy and Leclerc estimators. Regarding the studies about the unification of line processes, outlier rejection and robust statistics can be found in [73].

A *discontinuity adaptive* smoothing model is another solution, which is based on an analysis of the Euler equation associated with the energy minimization in MRF and regularization models [6, 74, 75]. A comparison study between robust statistics and *discontinuity adaptive* smoothing has been carried out by Li. It shows that the two methods are interchangeable [10]. Nevertheless, we need to establish a solution of the optical flow using MAP-MRF framework. Hence we choose Li's *discontinuity*

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

adaptive MRF model, which is, to the best of our knowledge, the first time being used. We will also show that it produces a sharper motion field at the discontinuity area than if we use the quadratic function.

4.2 MAP-MRF Framework for Optical Flow

In a MAP-MRF framework, a posterior probability as is written in the Equation 2.1 states that the task is to estimate the motion field f given a pair of images d_i, d_{i-1}. The likelihood probability corresponds to the optical flow constraint and the prior probability corresponds to the smoothness constraint.

Instead of using the probabilities, we can describe the MAP-MRF labeling as energy functions. Let f be the u or v from the continuous label set. The posterior energy function of optical flow with smoothing constraint is:

$$\begin{aligned} U(u,v|d_i,d_{i-1}) &= \textstyle\sum_{s\in S}\left(d_x(s)u(s)+d_y(s)v(s)+d_t(s)\right)^2/2\sigma^2+ \\ &\quad \textstyle\sum_{s\in S}\sum_{s'\in N(s)}\left(g\left(u(s)-u(s')\right)+g\left((v(s)-v(s')\right)\right) \end{aligned} \tag{4.1}$$

where $g(\eta)$ is the potential function, s' is a site from the second order neighborhood system (our model under discussion). The first term is the likelihood potential due to independent Gaussian noise in the image intensities and the second and third terms are the prior potentials due to the prior distribution of the MRF.

The main difference between our proposed algorithm and the algorithm represented by Equation 3.12 is in the prior potentials. Our prior potential function is an adaptive function instead of a fixed function such as the Horn-Schunck algorithm. The function adapts its characteristic according to the amplitude difference of vector neighbors. In this way, an improved motion field is achieved.

There is also a difference between our proposed algorithm and the robust estimation, Equation 3.14. The major difference is in the likelihood function. The robust estimation approach uses one more function to calculate the robust likelihood energy value. We avoided another function because it introduces complexity in the selection of parameters. We are satisfied with the use of a standard deviation parameter, σ, to control the noise influence. Hence, our algorithm has a reduced

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

complexity and an easier setting.

After we have the energy function, our task is to minimize it to find the optimal u^* and v^*. The least square technique requires us to calculate the energy function's partial derivatives with respect to each of the parameters. A minimum position is necessarily a stationary point where the gradient $\nabla U(u, v)$ must be a zero vector. Therefore, the following system of equations must be satisfied:

$$\begin{aligned} \frac{\partial U(u(s),v(s))}{\partial u(s)} &= 2d_x(s)\left(d_x(s)u(s) + d_y(s)v(s) + d_t(s)\right)/2\sigma^2 + \\ &\quad \textstyle\sum_{s'\in N(s)} g'(u(s) - u(s')), \\ &= 0, \qquad \forall s \in S, \end{aligned} \tag{4.2}$$

$$\begin{aligned} \frac{\partial U(u(s),v(s))}{\partial v(s)} &= 2d_y(s)\left(d_x(s)u(s) + d_y(s)v(s) + d_t(s)\right)/2\sigma^2 + \\ &\quad \textstyle\sum_{s'\in N(s)} g'(v(s) - v(s')), \\ &= 0, \qquad \forall s \in S, \end{aligned} \tag{4.3}$$

where g' is the derivative of g.

If the prior model is quadratic, $g(\eta) = \eta^2$, and we define local averages:

$$\overline{u(s)} = \sum_{s'\in N(s)} w(s')u(s'), \tag{4.4}$$

where w is the weight constant ($\overline{v(s)}$ is defined in the same way), we have:

$$\begin{aligned} \tfrac{1}{2}\frac{\partial U(u(s),v(s))}{\partial u(s)} &= d_x(s)\left(d_x(s)u(s) + d_y(s)v(s) + d_t(s)\right)/2\sigma^2 \\ &+ \left(u(s) - \overline{u(s)}\right), \qquad \forall s \in S, \end{aligned} \tag{4.5}$$

$$\begin{aligned} \tfrac{1}{2}\frac{\partial U(u(s),v(s))}{\partial v(s)} &= d_y(s)\left(d_x(s)u(s) + d_y(s)v(s) + d_t(s)\right)/2\sigma^2 \\ &+ \left(v(s) - \overline{v(s)}\right), \qquad \forall s \in S. \end{aligned} \tag{4.6}$$

Solving these equations using an iterative scheme from Gauss-Seidel method yields

$$u^{(t+1)}(s) \leftarrow u^{(t)}(s) - \left\{ d_x(s)\frac{d_x(s)\overline{u(s)}^{(t)} + d_y(s)\overline{v(s)}^{(t)} + d_t(s)}{1 + \left(d_x^2(s) + d_y^2(s)\right)/2\sigma^2} \right\}, \tag{4.7}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$v^{(t+1)}(s) \leftarrow v^{(t)}(s) - \left\{ d_y(s) \frac{d_x(s)\overline{u(s)}^{(t)} + d_y(s)\overline{v(s)}^{(t)} + d_t(s)}{1 + \left(d_x^2(s) + d_y^2(s)\right)/2\sigma^2} \right\}. \tag{4.8}$$

The initial values, $u^{(0)}(s)$ and $u^{(0)}(s)$, are zero.

Horn-Shunck has used Laplacians to calculate the prior term, $g'(.)$ of Equation 4.2 and 4.3. In the MRF model, it is classified as the second-order neighborhood system. In the next section we will discuss a technique from the MRF model to manipulate this prior term and to preserve the discontinuities.

4.3 Discontinuity Adaptive MRF Model

Li [6] proposed a simple and effective discontinuity adaptive scheme for a MRF model. He derived an important *necessary condition* for an MRF prior potential functions to be able to deal with discontinuities. Based on this *necessary condition*, he defined the Discontinuity Adaptive (DA) smoothness model in terms of the Euler equation constrained by a class of *adaptive interaction functions* (AIF).

If the derivative of g is expressed as $g'(\eta) = 2\eta h(\eta)$, then the *necessary condition* for any regularization model to be adaptive to discontinuities is

$$\lim_{\eta \to \infty} |g'(\eta)| = \lim_{\eta \to \infty} |2\eta h(\eta)| = C, \tag{4.9}$$

where $C \in [0, \infty)$ is a constant. If C is 0, then it prohibits smoothing at discontinuities, but $C > 0$ allows limited smoothing. We have to make an interaction $h(\eta)$ small for a large $|\eta|$ and an interaction approaches 0 as $|\eta|$ goes to ∞. Further detail is provided in Appendix B.

Li introduces four possible choices of AIFs with the corresponding *adaptive potential functions* (APFs) and the bands. We select one from his book [6]:

$$h_\gamma = \frac{1}{[1 + \frac{\eta^2}{\gamma}]^2}, \tag{4.10}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

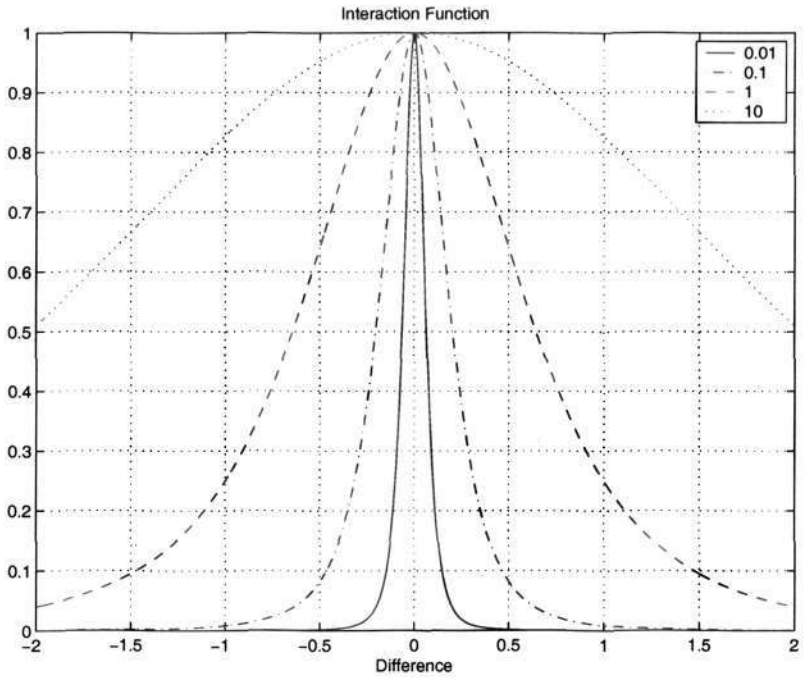


Figure 4.1: Interaction Function h_γ with Different γ (represented by the shape) and η^2 (the horizontal axis).

the potential function for this AIF is:

$$g_\gamma = -\frac{\gamma}{1+\frac{\eta^2}{\gamma}}, \tag{4.11}$$

and the band of convexity is

$$B_\gamma = (-\sqrt{\frac{\gamma}{3}}, \sqrt{\frac{\gamma}{3}}). \tag{4.12}$$

Figure 4.1 shows the interaction function in comparison with several γ. In practice, it is not always necessary to know the explicit definition of g_γ, except for analyzing the convexity of $U(f)$.

For the quadratic smoothness prior, $g(\eta) = \eta^2$, the interaction function $h(\eta)$ is always one, which means that it is not adaptable to discontinuities. We hence use the *adaptive interaction* function instead of the *constant interaction* function for our motion estimation algorithm because it preserves discontinuities. Based on this theoretical analysis, a better motion field can be achieved.

If the *necessary condition* is embedded into the energy function of the Equation

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

4.2 and 4.3, we have

$$\frac{1}{2}\frac{\partial U\left(u(s),v(s)\right)}{\partial u(s)} = d_x(s)\left(\left(d_x(s)u(s)+d_y(s)v(s)+d_t(s)\right)/2\sigma^2\right. + \sum_{s'\in N(s)}(u(s)-u(s'))h(u(s)-u(s')), \qquad \forall s\in S \quad (4.13)$$

$$\frac{1}{2}\frac{\partial U\left(u(s),v(s)\right)}{\partial v(s)} = d_y(s)\left(d_x(s)u(s)+d_y(s)v(s)+d_t(s)\right)/2\sigma^2 + \sum_{s'\in N(s)}(v(s)-v(s'))h(v(s)-v(s')), \qquad \forall s\in S \quad (4.14)$$

We can embed the AIF into Equation 4.4. So, the weight constants are now the AIF:

$$\overline{u(s)} = \sum_{s'\in N(s)} h(u(s)-u(s'))u(s'). \quad (4.15)$$

Using the same iterative scheme (Equation 4.7), we have the optical flow algorithm using the discontinuity adaptive model.

4.4 Model Comparison

The robust method for optical flow that has been discussed in Section 3.3.2.1 is a comparable method with our method under discussion. In Figure 3.3 we have shown the estimation result as well. In both methods, we have the same problem to be solved, but the strategy is different. Robust method looks at the *outliers,* while our method looks at the *discontinuity* of data.

In the MAP-MRF framework, the main differences of our method with the robust method are in the likelihood term. This term, the optical flow constraint, uses quadratic functions, but the robust method uses the robust function. Our argument is that we are handling discontinuity of the motion field. So we are only concerned to find the minimum of the potential value of optical flow constraint. As a result we should only consider parameters related to discontinuity function.

In the robust method that we use as a reference, it needs a parameter for the likelihood function, but our approach does not require it. This offers an advantage

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

1/12	1/6	1/12
1/6	−1	1/6
1/12	1/6	1/12

Figure 4.2: Laplacians's Weights

that a fewer number of parameters makes it an easier way for us to set the conditions for desired results.

4.5 Experimental Results

To test the *adaptive* interaction function, we compare the Laplacians weights (Figure 4.2) of Horn-Schunck model for its $\overline{u}$ and $\overline{v}$ of Equation 4.4 with the model using

$$h(\eta) = \frac{1}{[1 + \frac{|\eta|}{\gamma}]} \tag{4.16}$$

for its $\overline{u}$ and $\overline{v}$ of Equation 4.15. We also use Horn-Schunck's approach to estimate the derivatives of image brightness.

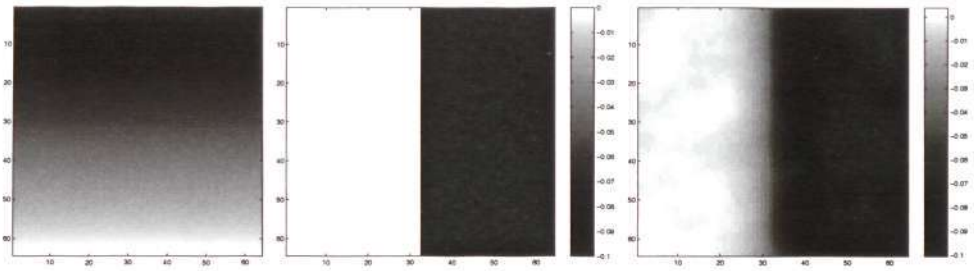
The test involves synthetic image sequences and is followed by experiments using several multiresolution real image sequences. For the experiment using multiresolution real image sequences, the total resolution levels depend on the image size. Our experiment uses four to seven levels. For multiresolution sequence, we process the images using a 5×5 gaussian kernel and a standard deviation of 0.5.

We test the algorithm using noisy image and we also compare our proposed method with the Black's robust technique.

4.5.1 Single Resolution Synthetic Sequence

Figure 4.3 shows the Vertical Ram sequence using a synthetic displacement field. The image has two parts: non-moving part on the left and moving part on the right. The moving parts are displaced 0.1 pixel vertically using interpolation technique. Using this sequence, the motion estimator under test is able to calculate the motion

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

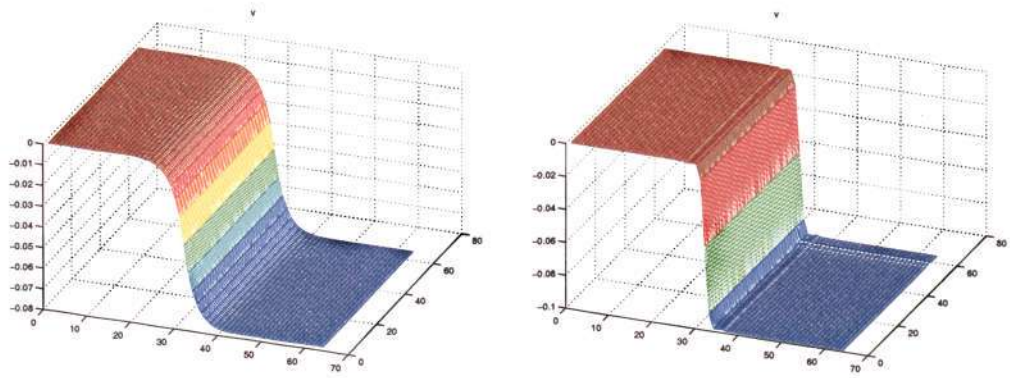


(a) The Ram Image

(b) The original displacement field

(c) The estimated displacement field

Figure 4.3: Ram Sequence and The Displacement Field.



(a) Constant Interaction Function

(b) Adaptive Interaction Function

Figure 4.4: Displacement Field for Synthetic Sequence with (a) $h = 1$, (b) $\gamma = 0.0001$.

field accurately. Therefore, we can analyze the discontinuity aspect at the sites where both parts merge.

The results of setting the *constant interaction* function and *adaptive interaction* function are shown in Figure 4.4. The constant interaction shapes the motion field as smooth as possible without discontinuities. In contrast, the *adaptive interaction* will give a varied constraint to the smoothness according to the differences between data and their neighbors. Therefore, it can preserve the discontinuities. Furthermore, setting $\gamma = 0.0001$ causes a crisp discontinuity. Figure 4.5 shows the differences when we modified γ.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

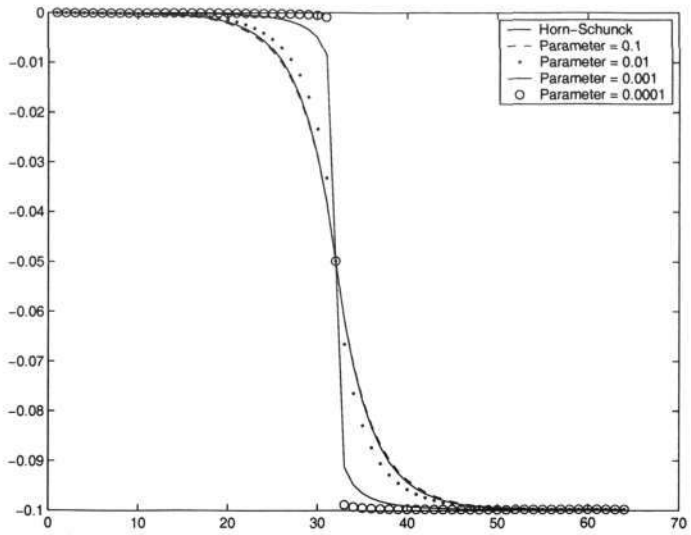
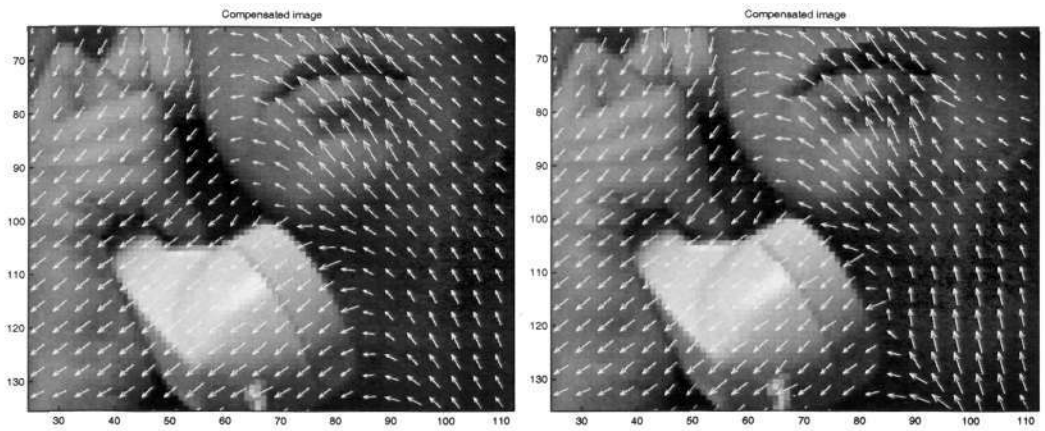


Figure 4.5: Variation of Smoothness Results for Different γ



(a) Constant IF (b) Adaptive IF

Figure 4.6: Preserving the Discontinuities in Suzie Sequence.

4.5.2 Experiment on Preserving Discontinuities

Suzie sequence in Figure 4.6 is presented to show how the AIF preserves the discontinuity by comparing the result of the AIF with the constant IF (CIF). We use $\gamma = 0.001$ for a crisp discontinuity. We can identify the effect of preserving on the motion field of the moving phone and Suzie's neck. At these merging sites a crisp motion field is not found in the CIF's result, but is found in the AIF's result.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

NAME	Yosemite Sequence			
	dB	Entropy	Average Angular Error	Standard Deviation
Constant	31.97	2.24	13.76°	10.43°
Adaptive	33.19	2.41	8.42°	6.71°

Table 4.1: Result of Multiresolution Synthetic Sequence Experiment in PSNR and Entropy values.

4.5.3 Multiresolution Synthetic Sequence Experiment

An experiment using yosemite sequence is reported here. The PSNR, entropy, average angular error and its standard deviation [38] are used to report the results. Table 4.1 shows the measurement results. The explanation of average angular error is given in Appendix C.3.

We use a yosemite image and its ground truth motion field to produce the second image. Using the adaptive model, we can get the improved PSNR value but with a reduced entropy value. This experiment also shows the improved *average angular error* and its *standard deviation* after we use adaptive model. This shows that our proposed model is working as expected. Figure 4.7 shows the result.

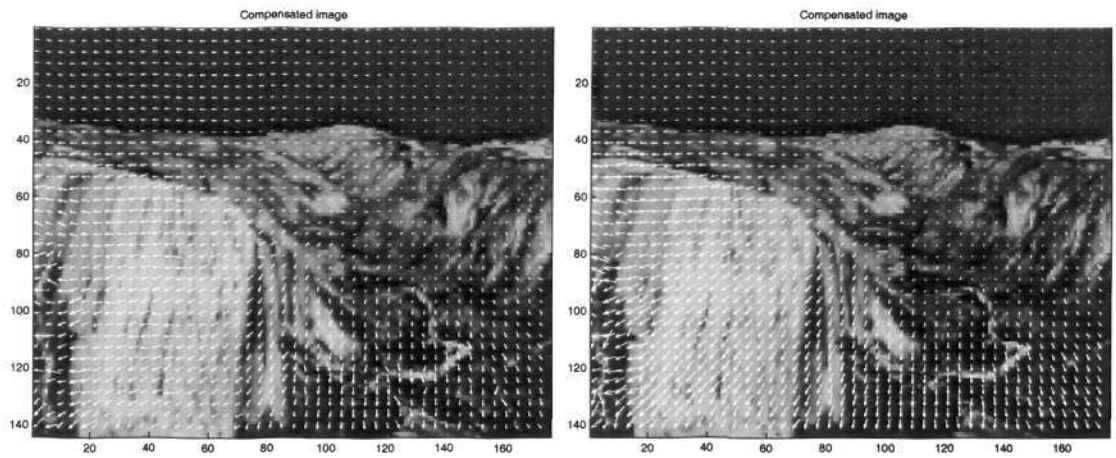
4.5.4 Multiresolution Real Sequence Experiment

Experiments are set up to verify the correctness of the Equation 4.13 and 4.14, and to find out how much the embedded *adaptive interaction* function improves the motion field. The PSNR and Entropy values are reported.

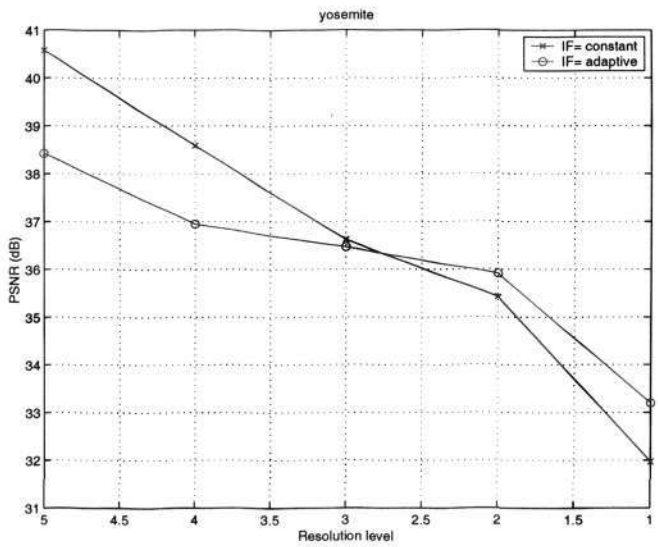
We have developed an algorithm called A GRADIENT-BASED MOTION ESTIMATION (ME) USING DISCONTINUITY ADAPTIVE (DA) FUNCTION written in Algorithm 2 to handle the images and to execute the motion estimator. This algorithm uses Equations 4.7, 4.8, 4.15 and 4.16.

We use four pairs of image sequences: foreman_148 / foreman_149, salesman_009 / salesman_010, suzie_000 / suzie_004, and taxi_000 / taxi_001. All sequences are selected so that we have a mixture of image pairs with large and small movements. In this case, the Foreman sequence has the longest motion vectors among others, so the sequence has more different data that produces the lowest

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) Horn-Schunck Motion Estimation (b) DA Horn-Schunck Motion Estimation



(c) PSNR vs Resolution Level

Figure 4.7: Results of Experiments for the Yosemite Sequence.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

NAME	CONSTANT		ADAPTIVE	
	dB	Entropy	dB	Entropy
Suzie	36.93	2.24	37.79	2.41
Foreman	31.76	3.86	33.10	3.34
Salesman	35.63	0.68	36.84	0.72
Taxi	36.86	1.03	37.31	1.34

Table 4.2: Result of Multiresolution Real Sequence Experiment in PSNR and Entropy values.

Variables	Value
Standard Deviation,σ	0.01
AIF's Parameter, γ	0.01
Adjustment Difference Margin	0.1
Maximum Number of Iteration	60

Table 4.3: Experimental Setup for MAP-MRF Labeling Model

PSNR as shown in Table 4.2. The Salesman sequence has several object movements with a textured background and the Suzie sequence has many movements with a large and textured region. These are reasons why we chose the Salesman and Suzie sequences in the previous experiments. The Taxi sequence has many movements on texture-less background.

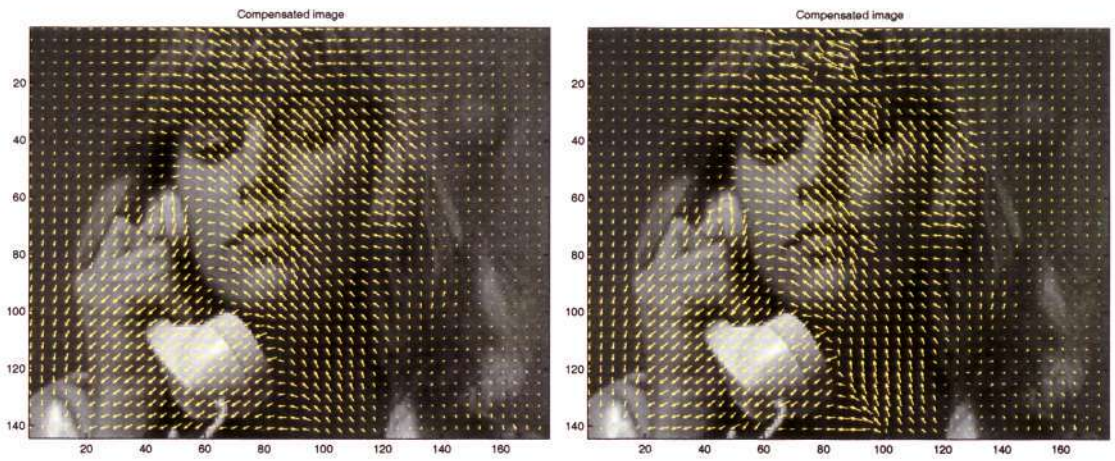
The parameter settings are shown in Table 4.3. We set the same parameters for all the sequences. We present a result of Horn-Schunck motion estimation, a result of our algorithm and a diagram of PSNR versus level of the image in Figures 4.8, 4.9, 4.10, and 4.11. Each diagram has two lines of the PSNR value of the compensated images. The line represents a *constant* interaction function experiment and the dot represents an *adaptive* interaction function experiment.

From all the experiments, the behavior of our proposed model is stable. The performance depends on the setting of the AIF's parameter. If we set γ to a high value, then a smooth motion field is achieved. And a low value of γ will give more discontinuities.

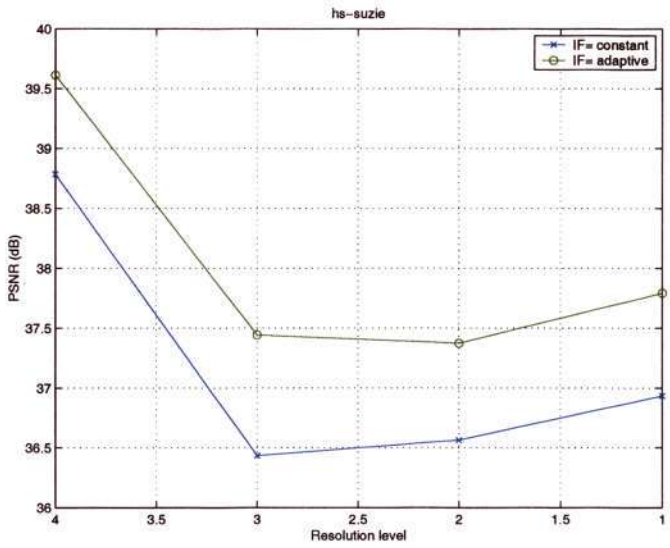
4.5.5 Noisy Sequence

In the previous section, the results of all images are always as expected when we adjust the parameter values. To further check the model stability, we add normal

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



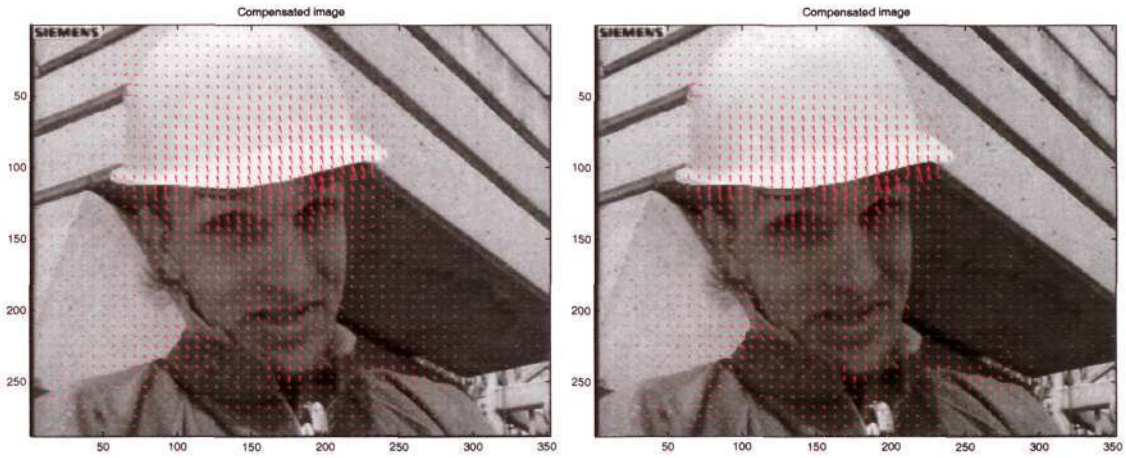
(a) Horn-Schunck Motion Estimation (b) DA Horn-Schunck Motion Estimation



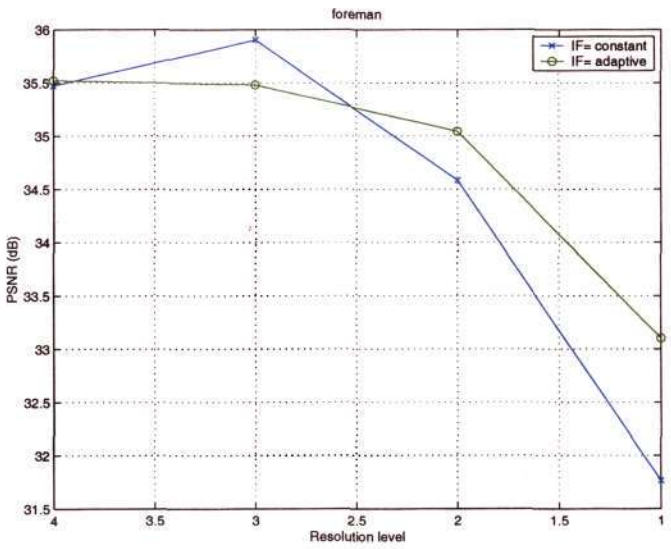
(c) PSNR vs Resolution Level

Figure 4.8: Results of Experiments for the Suzie Sequence.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



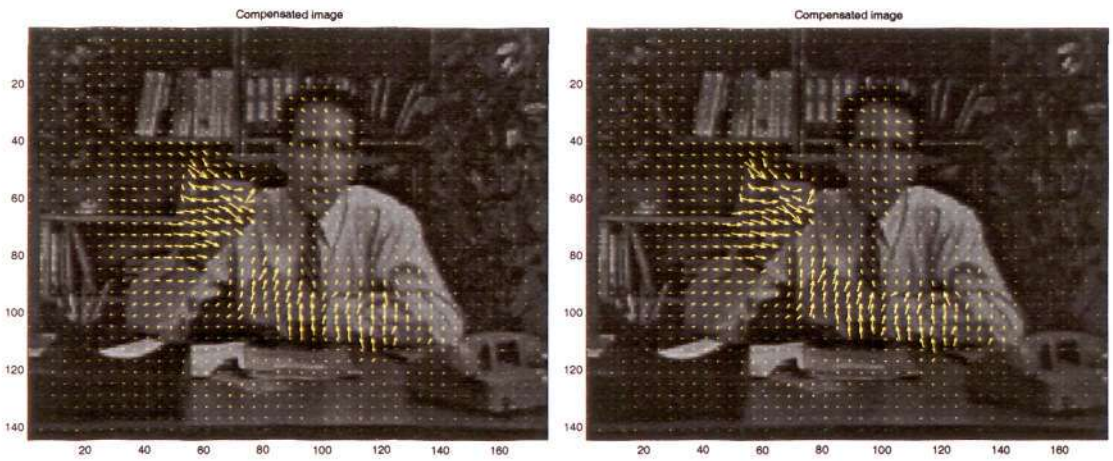
(a) Horn-Shunck Motion Estimation (b) DA Horn-Schunck Motion Estimation



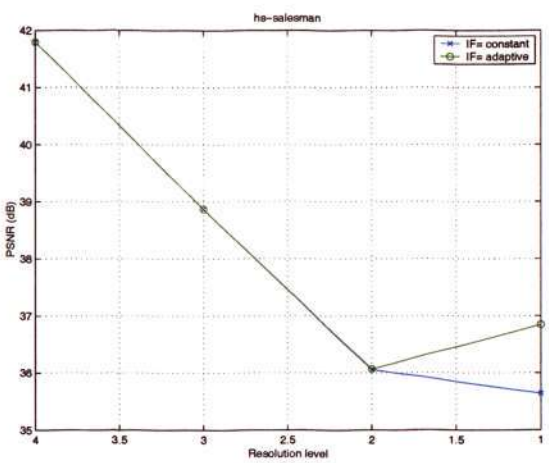
(c) PSNR vs Resolution Level

Figure 4.9: Results of Experiments for the Foreman Sequence.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



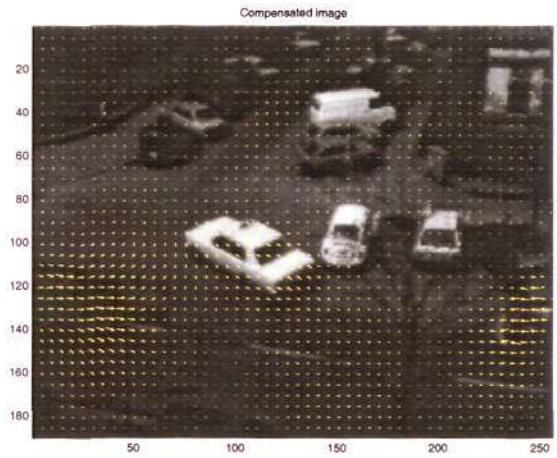
(a) Horn Schunck Motion Estimation (b) DA Horn-Schunck Motion Estimation



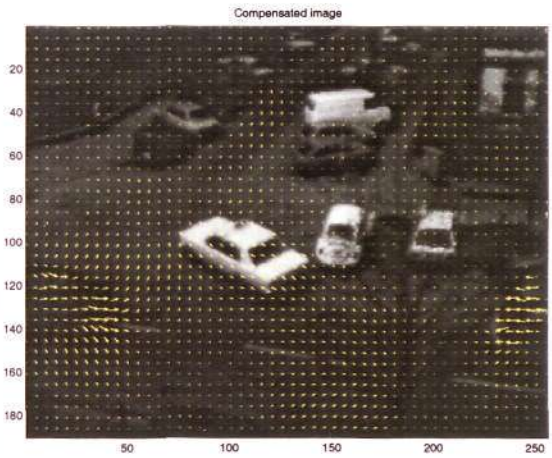
(c) PSNR vs Resolution Level

Figure 4.10: Results of Experiments for the Salesman Sequence.

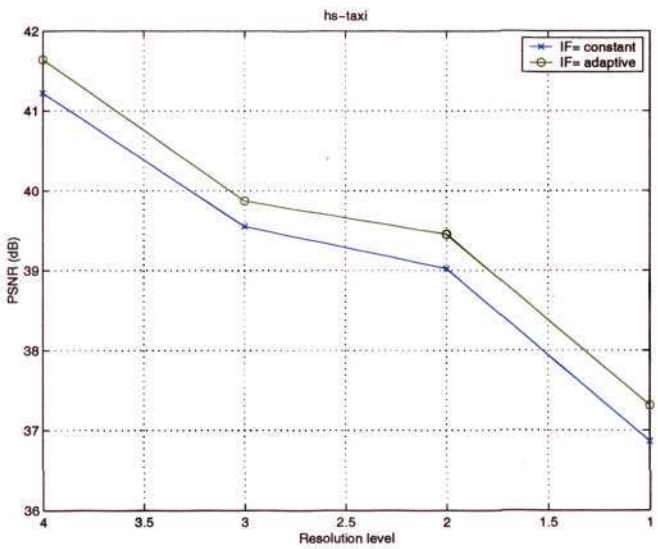
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) Horn-Schunck Motion Estimation



(b) DA Horn-Schunck Motion Estimation



(c) PSNR vs Resolution Level

Figure 4.11: Results of Experiments for the Taxi Sequence.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Algorithm 2 Gradient Based ME using DA function

1. Init:
   Generate multiresolution sequence,
   Set parameter of AIF, γ, if needed,
   Set Standard Deviation, σ,
   Set the Adjustment Difference Margin. The difference between the previous and the current adjustment values,
   Set Maximum Number of Iteration.
2. Calculate:
   Spatial and temporal gradient,
   Interaction values and local averages.
3. Update motion field using iterative scheme.
4. Repeat step no. 2-3 until the adjustment difference margin or maximum number of iteration requirements are achieved.
5. Propagate the motion field down to the higher resolution using bilinear transform using 3×3 blocksize.
6. Repeat step no. 2-5 until the highest resolution is reached.

Name	Constant	Adaptive
	dB	dB
Suzie	32.56	33.18
Foreman	28.64	31.20
Salesman	32.34	32.54
Taxi	32.54	32.95

Table 4.4: Result of Multiresolution Noisy Real Sequence Experiment in PSNR

random noise with $\sigma = 0.03$ into the images. Table 4.4 lists the PSNR result of the experiments. It shows the reduced PSNR values with the similar motion field as the one without noise. Our proposed method also increases the PSNR values.

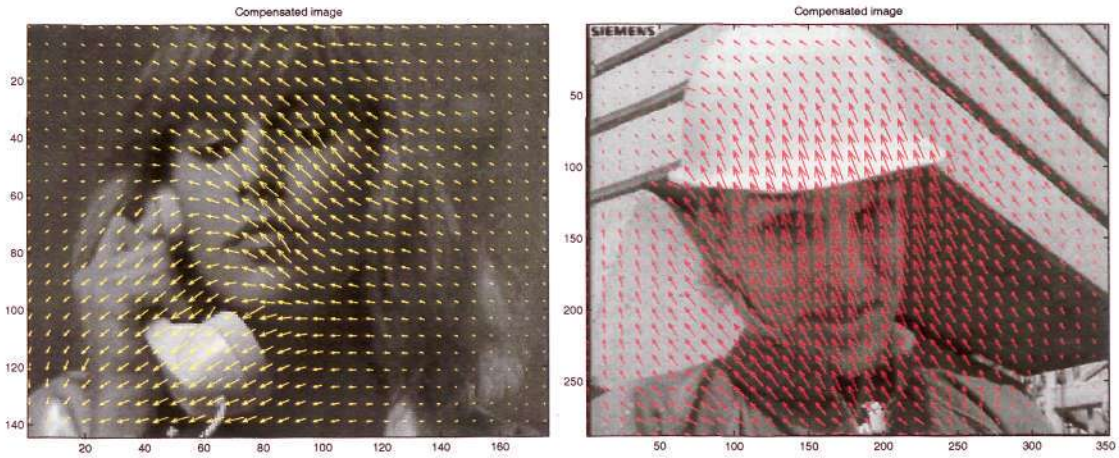
4.5.6 Comparison with Robust Statistic

The robust technique that is explained briefly in 3.3.2.1 is used to compare with our approach. Li's has studied the DA-MRF versus the robust technique for the range image segmentation. Here we make a comparison for motion estimation problem. For our experiments, the parameters are selected so we can generate the result similar to our DA-MRF model.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

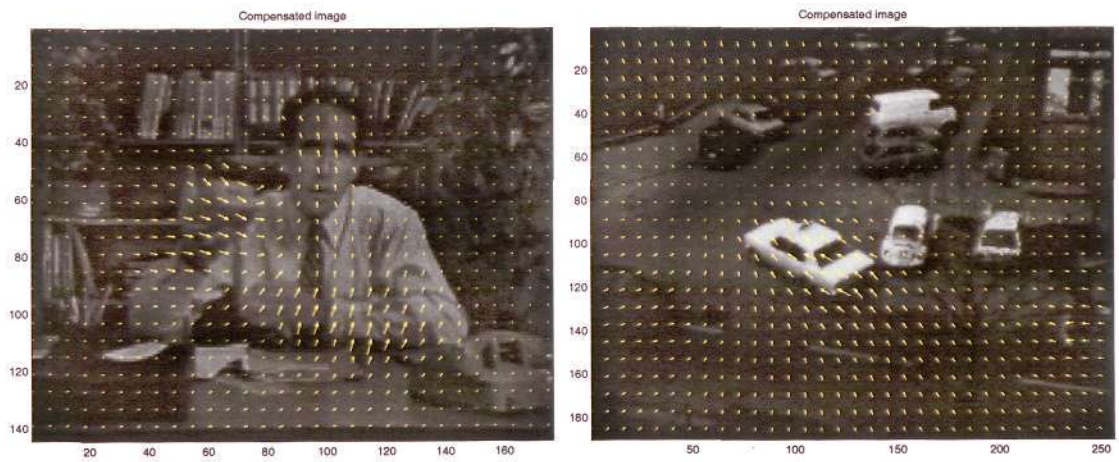
Name	PSNR	Entropy
	dB	bits
Suzie	35.26	1.71
Foreman	28.80	3.05
Salesman	34.77	0.32
Taxi	36.10	0.39

Table 4.5: Result of Multiresolution Real Sequence Experiment using Robust Technique



(a) The Suzie Sequence.

(b) The Foreman Sequence.



(c) The Salesman Sequence.

(d) The Taxi Sequence.

Figure 4.12: Results of Experiments using Robust Technique.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Table 4.5 shows the measurement result. Figure 4.12 shows the motion field. The result of Foreman sequence shows the smoothed motion field with less PSNR value than our proposed model. Actually, results with lower PNSR and lower entropy value are expected. The result of Salesman sequence gives less PSNR value than our model. The result of Suzie and Taxi sequences give less PSNR values and lower entropy. We conclude that the robust technique also produces a comparable result to our model.

We have explained that our model uses DA-MRF only in the smoothness term. It implies that the parameters involved are less than the robust technique. As we know that robust technique uses the robust functions for both the likelihood and prior terms.

4.6 Conclusion

This chapter introduces a novel optical flow model embedded with an *adaptive interaction* function (AIF) of Li [6] using the MAP-MRF continuous labeling framework. The AIF is a non-linear function and makes the smoothness prior adaptable to a motion vector and its surrounding.

This proposed model uses the optical flow model that is proposed by Horn-Schunck. Our proposed model calculates the local averages differently. The calculation has Li's discontinuity adaptive (DA) function. Since Li's DA-MRF model is employed here, the proposed model is derived using the MAP-MRF framework. So our model is a new optical flow model that preserves discontinuities and is described using the GMRF framework. From the literature survey, our proposed model is similar to that from robust statistics.

To deal with the gradient based motion estimator, we need to develop a multiresolution framework. This framework calls motion estimation algorithm to estimate every level of the multiresolution image. The result of the coarse level is the initial value for the finer level.

From this experiment we identify that our model is adjustable under the variation of parameter values: standard deviation, σ and AIF's parameter, γ. The expected

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

results from the given parameters are correct.

We also add noise in the sequences. We notice that the algorithm produces motion field similar to those without noise.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 5

Block Matching with Smoothness Prior Function

In the previous chapter, we have discussed gradient based motion estimation and at the same time we have proposed the implementation of Discontinuity-Adaptive MRF to adjust the result of estimation. In this chapter, we apply the MRF model, the iterative minimization strategy and the Discontinuity-Adaptive MRF (DA-MRF) from the gradient based motion estimation to matching based motion estimation.

We have a hypothesis that block matching based motion estimation also needs the smoothness prior knowledge in order to improve its performance to generate a "true" motion field. The meaning of the "true" motion field has been explained in the third paragraph of Chapter 1. This hypothesis has motivated our research and the correctness of the hypothesis is shown in this chapter.

As we know that block matching algorithm is a simple approach for motion estimation. It is the reason that it has been used for many practical purposes. It serves to initialize the different kind of motion estimation algorithm [24, 76]. It is also commonly used in several compression technologies such as H.263 for video phone and MPEG [77] for DVD.

We found that the block matching algorithm just produces a motion field with the lowest error and in many cases it does not represent the correct motion of the object. Though it can be used by some compression algorithms, it cannot be used for other purposes such as moving object segmentation, old movie restoration, or

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

in-between technique where a "true" motion field is required. Because of these we often design post-processing algorithms to adjust the motion field obtained from the block matching algorithm.

In the context of implementing constraint into an algorithm, block matching implicitly assumes a smooth motion field. However it is not always enough to guarantee a good motion field, especially in the textureless areas. We hence introduce the explicit smoothness prior function to the motion field, which will be shown to be effective.

To merge the smoothness prior function, several algorithms are designed in this chapter. So, our contributions are to design the block matching algorithm using

- smoothness prior function,
- discontinuity adaptive smoothness prior function,
- smoothness prior function plus reduced block-size algorithm, and
- smoothness prior function in multiresolution.

They are described in a single MAP-MRF model with likelihood and prior functions. The design emphasizes on the use of smoothness prior functions and the role of initialization value.

The experimental results verify the correctness of the hypothesis. If the result has shown an improvement that human perception can agree, the algorithm is acceptable. However, evaluating using human perception is said to be subjective. For example, the human perception suggests that the motion field must be available in the middle of large moving textureless area. Often, in those areas, the motion field cannot be generated by the conventional block matching.

The PSNR value and the entropy value are also measured. The PSNR value may rise or fall affected by the algorithms and the image sequence. The entropy value is reduced for a smoother motion field.

Section 5.1 describes the energy functions using the MAP-MRF framework. Section 5.2 describes the problem and its solution. Section 5.3 shows the realization of the idea. Section 5.4 describes the proposed algorithms. Section 5.5 shows the experimental results. Lastly, Section 5.6 gives a conclusion.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

5.1 MAP-MRF Framework

We use the MAP-MRF framework to model the proposed approach. The framework has Equation 2.1 as the posterior probability. The likelihood probability can be a measure on the correlation between two frames d_i and d_{i-1} and a motion field f. But researchers have added a certain test according to their assumption, for example: Heitz and Bouthemy [33] have involved an edge detector and a validity test of the motion vectors in the likelihood probability. Dubois and Konrad [32] have defined a prior probability using an occlusion field in the motion discontinuity model. Lim and Chong [76] have described the relationship between the occlusion and motion fields.

Instead of using the probability, we can describe them as energy functions. We minimize this energy function to get an estimate of the motion field or the label, $\widehat{f}$. It is written as

$$\{\widehat{f}\} = \arg\min_{f} \{U_d\left(d_i|f,\, d_{i-1}\right) + \lambda U_f\left(f|d_{i-1}\right)\}, \tag{5.1}$$

where λ is a weight. In this equation we have the likelihood energy function, U_d, related to the block matching algorithm, and the prior energy function, U_f, related to the smoothness prior function.

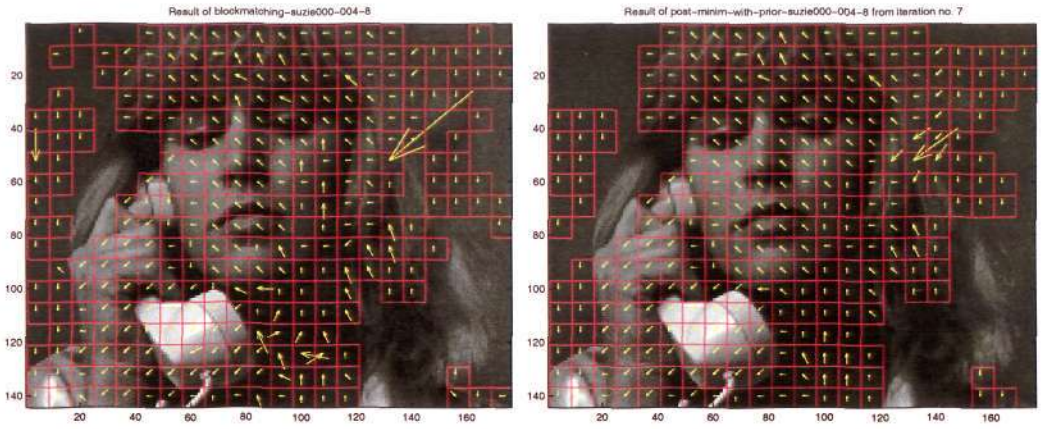
The conventional block matching algorithm can be considered as the maximum likelihood estimation in which the energy function has no prior knowledge. The energy function becomes

$$\{\widehat{f}\} = \arg\min_{f} U_d\left(d_i|f,\, d_{i-1}\right). \tag{5.2}$$

In detail, the U_d is the sum of difference value of blocks from image d_i and image d_{i-1} which is shifted by f. In this model, the state of the motion field only reaches the minimum energy value. Meanwhile, each motion vector is not correlated to other motion vectors. As a frame of an image, the result often does not show that the object's movements are the same with the human perception.

For an illustration, Figure 5.1(a) shows the result of the conventional block matching algorithm, which is chaotic and does not reflect the "true" motion. In

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) Result of Block Matching Algorithm

(b) Result of Block Matching Algorithm using Smoothness Prior Function.

Figure 5.1: Motion Field Results.

contrast, Figure 5.1(b) shows the result of the block matching algorithm with the smoothness prior function. The improvement toward a "true" motion is obvious. A noticeable single long motion vector which is shown on Figure 5.1(a) has been adjusted. The details of adding the smoothness prior will be introduced in the following sections. This observation is the basic reason that we propose our algorithm with both implicit and explicit constraints. In this figure, more iterations will further smooth the motion field.

If we write the prior energy function in detail, then U_f becomes

$$U_f(u,v) = \sum_{s \in S} \sum_{s' \in N_s} \left(g\left(u(s) - u(s')\right) + g\left(v(s) - v(s')\right) \right) \tag{5.3}$$

where s and s' are block sites, N_s is neighbors of block s, and $(u(s), v(s))$ is a motion vector. We can select $g(\eta) = \eta^2$ for the prior function. During minimization, this prior energy functions will be in balance with the likelihood energy functions.

However, if the need arises to preserve the discontinuity, we can implement the potential function as is shown in Equation 4.11. We can use this function if the result is already smooth and we need to enhance the discontinuity. If the condition is not met, the current non-smoothed motion field will be wrongly preserved. Next, we investigate the problems in the conventional block matching and its solutions.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

5.2 Problem and Solution

During the minimization process of the block matching algorithm, sometimes it is difficult to select a *unique* motion vector using only the lowest likehood energy. This is due to a condition that data or the feature of the region given to the algorithm were insufficient. In other words, they do not have the uniqueness. For one block, even if the uniqueness can be achieved and the "true" motion vector can be found, a slight noise may give a wrong result.

Two improvements can be incorporated to increase the possibility for information *to be sufficient* or to enable motion field *to be unique.* To make the sufficient information and the unique result, we may enlarge the block-size. This approach is considered as adding a stronger implicit constraint. However in some condition, this still may not be sufficient. Moreover, this approach needs more computational tasks. Another solution is to consider the smoothness constraint of neighboring motion vectors. The constraint is realized as the prior energy function. With both likelihood and prior energy functions, we have the posterior energy that is more probable to be unique. We illustrate this unique condition in Figure 5.2. The left figure shows several labels that have the same energy value. With the prior energy function the energy value for those labels are modified. This condition is shown in the right figure. This approach is considered as adding an explicit constraint. The advantage of this approach is the interaction between neighboring vectors. From the left to the right figures, Figure 5.3 shows: a motion field when the block matching with a certain block-size is insufficient to produce a good estimation; a motion field when the large block-size in the overlap block matching is still insufficient; a motion field in which the prior energy function has tremendously changed the state of the motion field into an acceptable state.

After we have recognized that imposing the explicit constraint brings a better result, we need to decide the weight of the smoothness prior function. The weight is an important parameter that will affect the smoothness of the result as is shown in Figure 5.4. The best weights are in the area where the balance between the likelihood and prior function are maintained. Every pair of images has its own good weight. It is also for each specific resolution.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

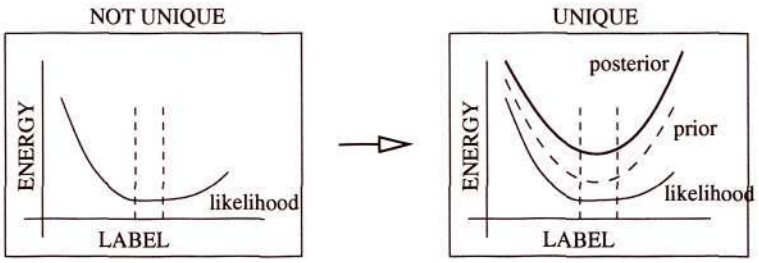
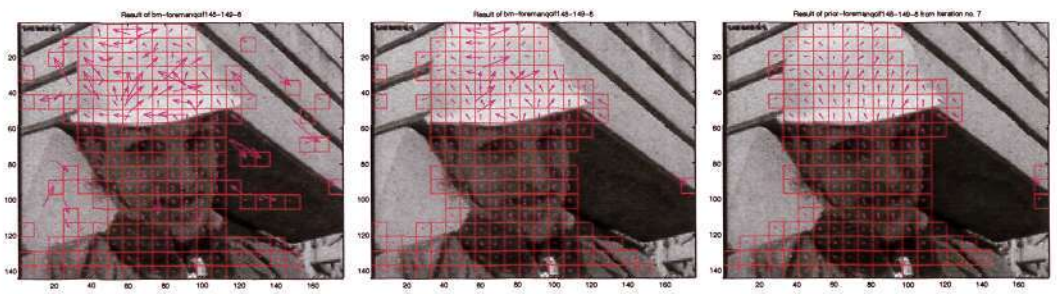


Figure 5.2: Prior Energy to Solve Uniqueness Problem.



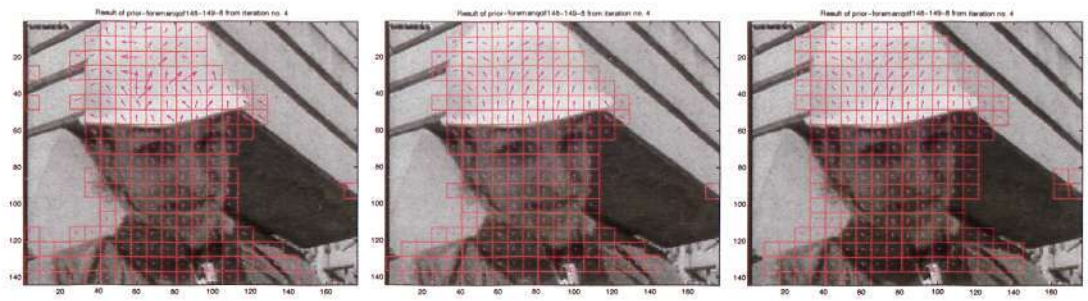
(a) Conventional Block Matching

(b) Overlap Block Matching

(c) Block Matching with Smoothness Prior Function

Figure 5.3: Influence of Implicit and Explicit Constraint for Many Types of Block Matching Algorithm.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) Weight = 0.01. (b) Weight = 1. (c) Weight = 100.

Figure 5.4: Foreman Sequences with Different Weights.

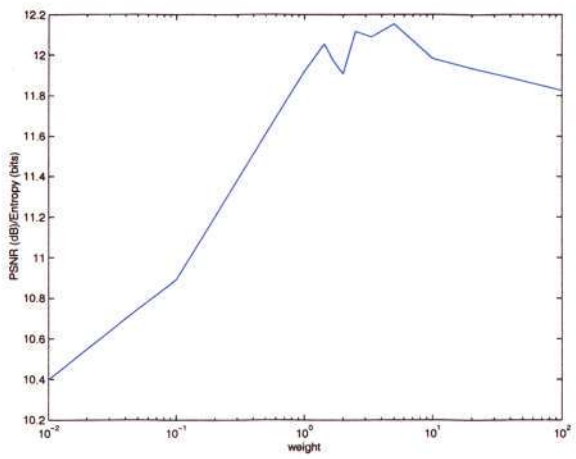


Figure 5.5: Weight vs PSNR/Entropy

Currently we use a diagram to roughly showing the area that we can select as a good weight. The diagram, *weight versus PSNR/Entropy,* is shown in Figure 5.5. The diagram identifies the peak that gives the highest value of PSNR and the lowest value of Entropy. For example, in that figure, it shows that the value between 1 to 10 is favorable for this image.

5.3 Realization

In the previous sections we have explained the problem in the conventional block matching algorithm. After we have redesigned the block matching model to solve the pointed problem, our current step is to design a complete system and verify our solution through various experiments. The complete system consists of an initialization procedure, a type of the minimization technique, and procedures to reduce

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

the computational complexity.

We initialize the motion field with the result of the full search block matching algorithm. We choose the full search block matching algorithm because the hardware solutions are already available in the market. Our algorithm can be an extension of it.

We adopt Iterated Conditioned Mode (ICM) to minimize the energy and search the motion vectors because ICM is fast in execution and easy to implement. However, as a local minimization technique, the result depends heavily on the initial values. The result will be very different if we initialize the algorithm with zero values. Moreover it might not develop a motion field during the iteration because of the imposed smoothness. This shows the importance of the initialization or deficiency of the ICM algorithm.

Besides initialization and minimization technique, we have to reduce the computation complexity by

- keeping the likelihood values in memory for use later and
- calculating the likelihood values using the smallest block-size.

Keeping all calculated likelihood energy values in memory, so that further computations can refer to it again, reduces the computational cost. Those values are available when we compute the full search block matching algorithm. When a block matching algorithm finishes the calculation of the likelihood energy and finds the best candidate vector, we usually keep only the best candidate vector and delete all other calculated likelihood energy values. In our algorithms, we need to keep those values. We illustrate this procedure in Figure 5.6. In this block diagram, we can calculate the likelihood energy, followed by OPTIMIZATION A, which simply takes the label with the lowest energy as the desired motion vector. The process is the same as the conventional block matching. In this process the smoothness constraint is not constrained and the chaos of the motion field is often obtained. If we add the smoothness constraints, the result will be very different and more desirable. This can be fulfilled by OPTIMIZATION B, where the posterior energy, instead of likelihood energy, is minimized. For this purpose, we need to store all the likelihood

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

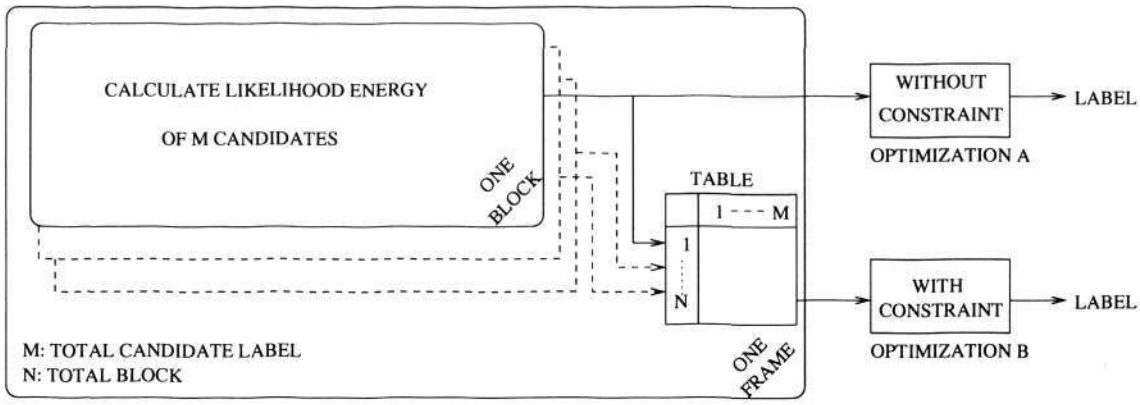


Figure 5.6: Comparison: Systems with and without Constraint.

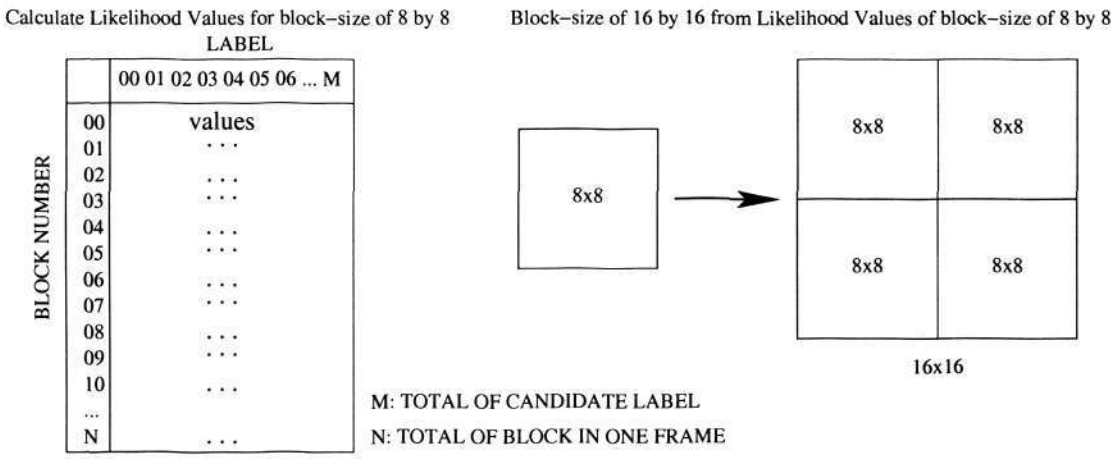


Figure 5.7: Enlarging Block-Size.

energy values in the memory before OPTIMIZATION B is carried out.

Calculating the likelihood values using the smallest block-size provides a basis to obtain the values for the larger block-size. For example the summations of the difference values of four blocks with 8×8 block-size can produce a summation of the difference value of a block with 16×16 block-size. The illustration is in Figure 5.7.

5.4 Algorithms

In this section we explain all the algorithms we proposed and used. We begin with the conventional and overlap block matching algorithm, then the proposed algorithms: block matching with smoothness prior function algorithm, block matching with discontinuity adaptive smoothness prior algorithm, block matching with

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

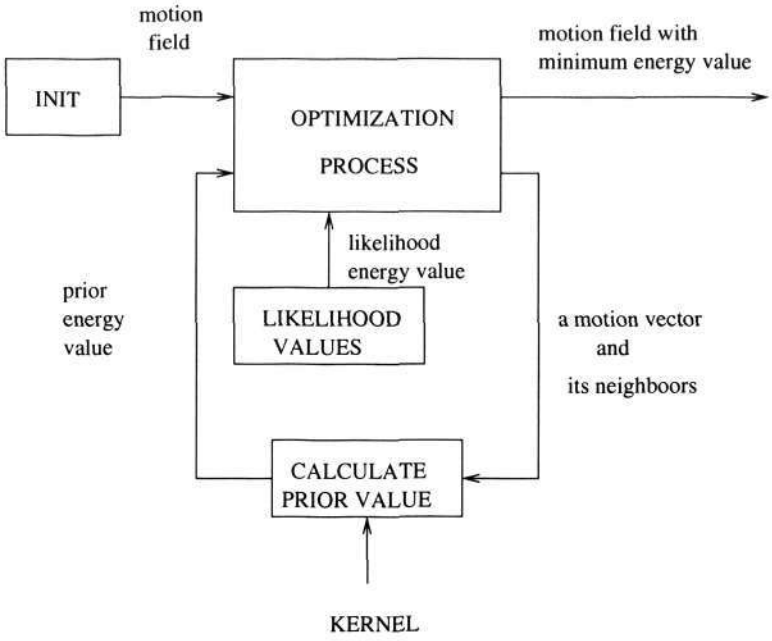


Figure 5.8: Block Diagram of the System

smoothness prior function algorithm plus reduced block-size, and multiresolution block matching with smoothness prior function algorithm. In the following sections, the explanations of each algorithm are arranged from the simple to the complicated one.

Actually, in this chapter we are implementing the generic MAP-MRF model that makes us easier to transfer the techniques found in gradient-based approaches to the matching-based. In our case, the smoothness prior function and the concept of discontinuity adaptive MRF, the multiresolution strategy can be seamlessly transferred to the block matching algorithm.

We illustrate the relationship of those algorithms in a block diagram shown in Figure 5.8. In the block diagram, the block INIT and LIKELIHOOD VALUES are the results of the conventional block matching algorithm or the overlap block matching algorithm. The block INIT can be chosen from a different resolution as in the multiresolution scheme. While the smoothness prior function needs the CALCULATE PRIOR VALUE block to calculate the function, the OPTIMIZATION PROCESS block iteratively updates the state of the motion field. If a discontinuity adaptive function is required, the KERNEL is given with a different configuration.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

5.4.1 Conventional and Overlap Block Matching

The conventional block matching algorithm uses full search strategy to estimate the motion field. This search strategy generates the likelihood values for all candidate labels and the motion field. The purposes of using this algorithm are:

- to calculate all likelihood values,
- to initialize the proposed algorithm using its estimation result.

Both data must be saved for further processing in the proposed algorithm.

Because conventional block matching has no preferences of the result, often the result is chaotic. Several conditions are given in the procedure to reduce the chaos. For example, if the likelihood values are equal for all the candidate vectors then we prefer the vector equals $(0, 0)$. Besides that, the candidate vectors computed first may have higher preference than those computed last.

One of the reasons for using the full search strategy is the need of likelihood values for the proposed algorithm. Full search strategy is used to calculate the values for all candidate labels. So, those values must be kept in a memory. The minimization process of the proposed algorithm refers to this memory everytime it needs to know the likelihood value. For QCIF image, there are $22 \times 18 = 396$ blocks, with 8×8 block-size and a search area of ± 16. It needs 3D-memory $33 \times 33 \times 396$ units to store the likelihood values.

The estimation result from the overlap block matching algorithm is used to compare our proposed approach. The overlap block matching algorithm calculates the likelihood values using a larger block-size of 16×16. In this way more area of pixels is under consideration. As a result, we need more computation time. This technique may increase the smoothness of the motion field.

5.4.2 Block Matching with Smoothness Prior Function

In the conventional block matching, the likelihood function is the only data source to find the best label. However, our proposed model, a block matching with the

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

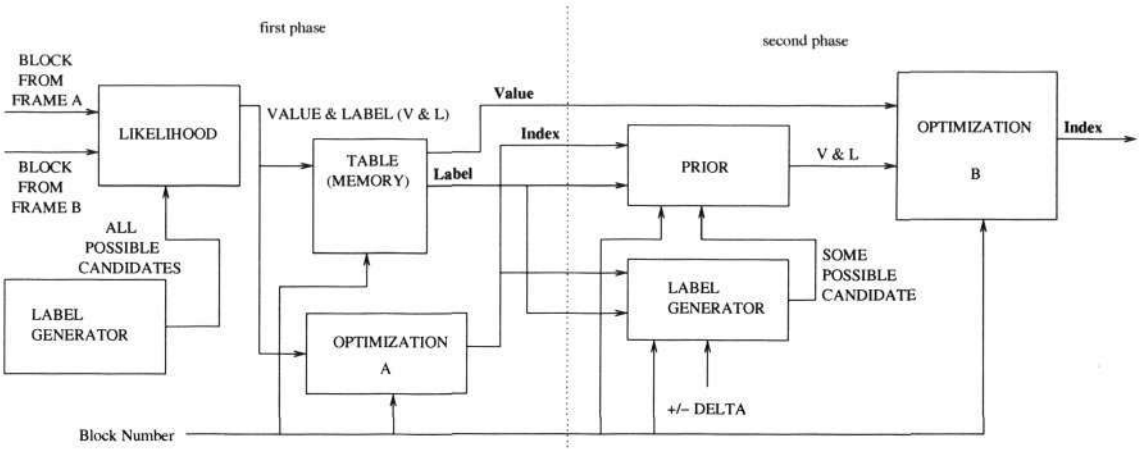


Figure 5.9: Functional Block Diagram for Block Matching with Smoothness Prior Function Algorithm.

smoothness prior function algorithm, calculates posterior values by adding the likelihood and prior energy values.

The block diagram in Figure 5.9 shows the functional blocks of this algorithm. A label generator generates all possible candidate labels from which the likelihood-block calculates the likelihood values. The table or memory keeps all the calculated value for all sites in a frame. The optimization-A block gives the first motion field. This motion field is considered as a product of optimization with the minimum likelihood energy. This first phase is the same as a conventional block matching, except all likelihood energy data are stored in the memory. In the second phase, the prior-block modifies labels from the first phase with small delta values and it also calculates the prior-value. The optimization-B block gives the final motion field. This is considered as a product of optimization with minimum posterior energy.

In the second phase, we introduce several candidate vectors with values around the current vector into the optimization algorithm. The calculated new prior energy values are based on the several candidate vectors. For example, Figure 5.10 shows candidate vectors from the center vector $(2, -1)$, where the prior energy values are calculated. From there the energy leads it to another vector $(1, -2)$. After we add them with the likelihood energy values, we get the posterior energy values. The minimization process finds a motion vector with the lower posterior energy value.

One of the differences between this algorithm and the conventional block matching is a step that we have to keep the likelihood energy values indexed by its labels

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

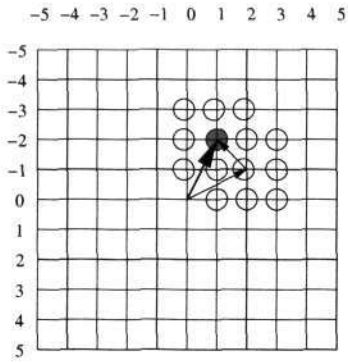


Figure 5.10: Iterative matching procedure, where $(1, -2)$ is found.

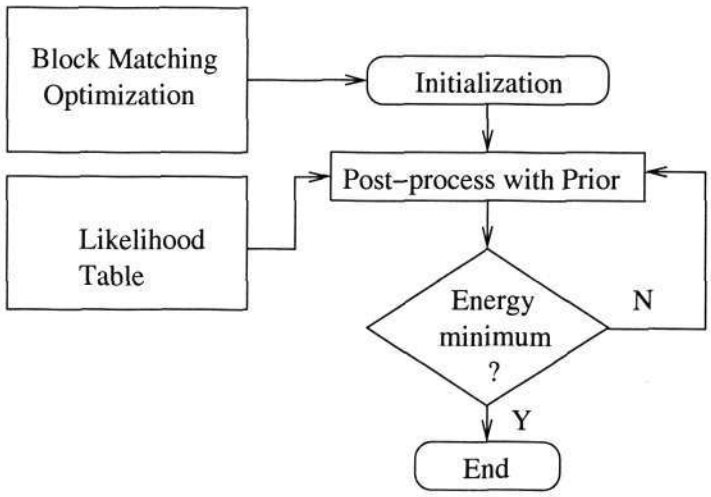


Figure 5.11: Initialization and Process of Smoothing the Motion Field

whenever a block matching is executed for a block. Without this step, the iteration procedure is a costly task. After all the blocks have been visited, we have a motion field. We then enter into the post-processing algorithm to update the motion field. This is illustrated in Figure 5.11. The post-processing algorithm minimizes the new posterior energy values and gives the final motion field.

We also have to consider the number of iterations. The number must give a possibility for the algorithm to reach the desired distance from the current motion vector value. For example, if a candidate vector is around $\pm delta$ value from the current vector value, the maximum distance from that value is multiplication of the number of iterations with a $\pm delta$.

Figure 5.12 illustrates typical progress of PSNR during a minimization process of the Suzie sequence. This figure is from the experiment using a 8×8 block-size. During the first few steps the PSNR drops then it rises and reaches the oscillation

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

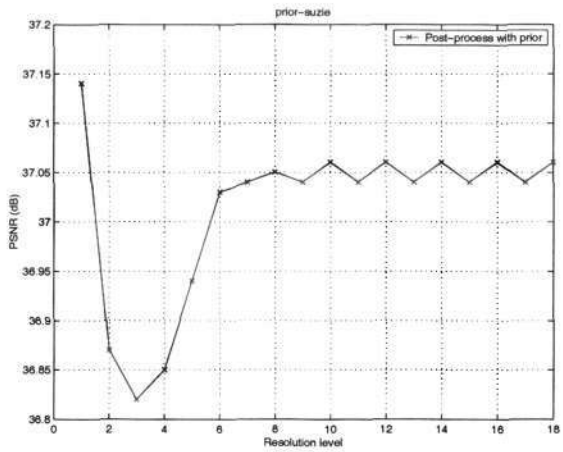


Figure 5.12: Progress of PSNR

state. The initialization from the block matching always gives high PSNR. If a motion vector that has a high match is wrong in the evaluation which is influenced by the smoothness prior function, the motion vector will change its state. During this change, the balance between the likelihood and the prior energy functions is in progress. If the balance has been reached, the minimization process enters the oscillation condition. The main reason for the oscillation is the integer type of the motion vector. We had observed that these behaviors vary although they have the same trend.

5.4.3 Block Matching with Discontinuity Adaptive Smoothness Prior Function

The algorithm is similar as the block matching with a smoothness prior function in Section 5.4.2. The difference is on the smoothness prior function. Instead of η^2 we use the Discontinuity Adaptive (DA) potential function. The DA potential function has been written in Equation 4.11. The purpose is the same as in the previous chapter, to preserve discontinuity.

Similar to the previous section, this algorithm relies on the initial values and if the initial values are already in chaos, the DA potential function will blindly preserve the error of the initial values. It mistakes some obviously wrong initial values as edges of motion field and preserve them. Hence one suggestion is to use this DA

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

potential function after a smoothed motion field has been developed. In this way we reduce the smoothness and this results in a better motion field for a certain area.

5.4.4 Block Matching with Smoothness Prior Function plus Reduced Block-Size

The experiments for the algorithm in the Section 5.4.2 have shown that most of the PSNR values are higher if we use a smaller block-size than one with a larger block-size. But a meaningless motion field may be present. For example, the condition that the vectors matched to the locations where they have the smallest energy values may not be correct. This condition is called local minima. To take advantage of this fact that block matching using a larger block-size often gives a reasonable motion field, the process of minimization using 8×8 block-size could be initialized by the result of minimization using 16×16 block-size.

Such a process with two types of block-sizes can be deemed as transferring smoothness of a motion field from the implicit to explicit constraint. , i.e., a block matching with the smoothness prior using 16×16 block-size, i.e., a block matching with the smoothness prior using 8×8 block-size. This is possible because of the smoothness term in the energy function.

Usually the motion field obtained by conventional block matching with 8×8 block-size is not good, as is shown in Figure 5.3(a). The improvement of the result is limited even when the block-size is extended to 16×16 in the overlap block matching, as is shown in Figure 5.3(b). However we can improve the result effectively if (i) the initial values are provided by conventional block matching with 16×16 block-size and (ii) the smoothness prior knowledge is forced in the motion field.

To run this algorithm, first of all, we need the calculated likelihood values from the conventional block matching algorithm with several types of block-sizes. Secondly, the algorithm is initialized with the result of block matching algorithms using the larger block-size. Thirdly, we continue with the minimization process. Lastly, we use a smaller block-size for another minimization process. The flowchart of this procedure with two block-sizes is shown in Figure 5.13. If it is needed, we can design an algorithm with more than two block-sizes.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

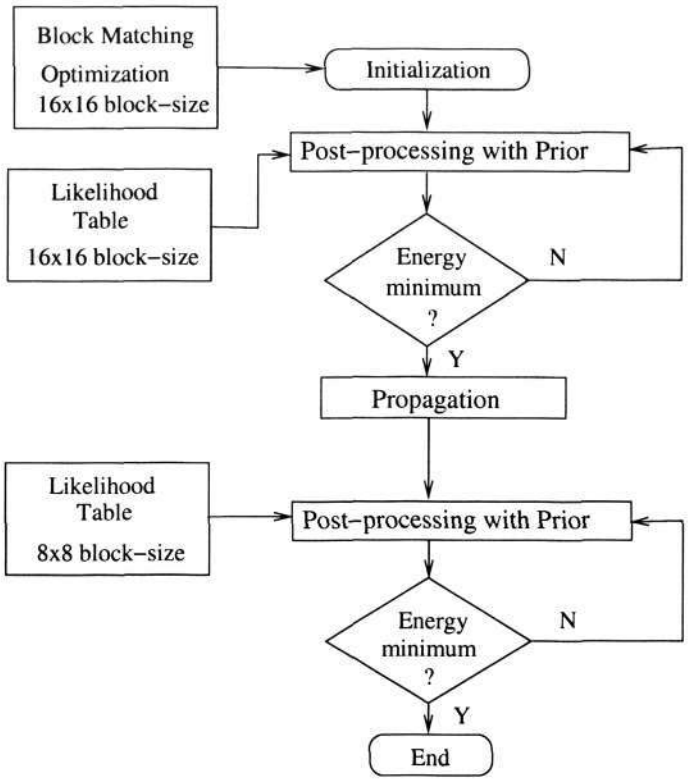


Figure 5.13: Flowchart of Re-Sizing Block-Size Algorithm.

5.4.5 Multiresolution Block Matching with Smoothness Prior Function

In the previous section, we discussed the block matching with a smoothness prior algorithm. That algorithm uses a single image size for all steps: the initialization and minimization steps. We notice that if we reduce the block-size from 16×16 to 8×8, then the total number of blocks is increased by four times, e.g. $11 \times 9 = 99$ to $22 \times 18 = 396$. If we do a full search for these blocks, we need more computation time. Because of this reason, we need to design a multiresolution scheme.

Our multiresolution scheme is shown in Algorithm 3. This scheme may put our motion field into the different state compared to the result of the other algorithms. As a result, the PSNR and entropy value are achieved differently. The multiresolution scheme calculates a full search algorithm for a smaller image size then it is followed by the minimization process on the same image size. The motion field is propagated to a higher resolution image before the minimization process refines it.

Figure 5.14 illustrates the differences between "Block Matching with Smoothness

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Algorithm 3 Multiresolution Block Matching Algorithm with Smoothness Prior.

1. Build multiresolution images.
2. Do full search block matching at the lowest image resolution.
3. Do minimization using a posterior energy function at the current image resolution.
4. Propagate the motion field into higher image resolution.
5. Do minimization using a posterior energy function at the current image resolution.

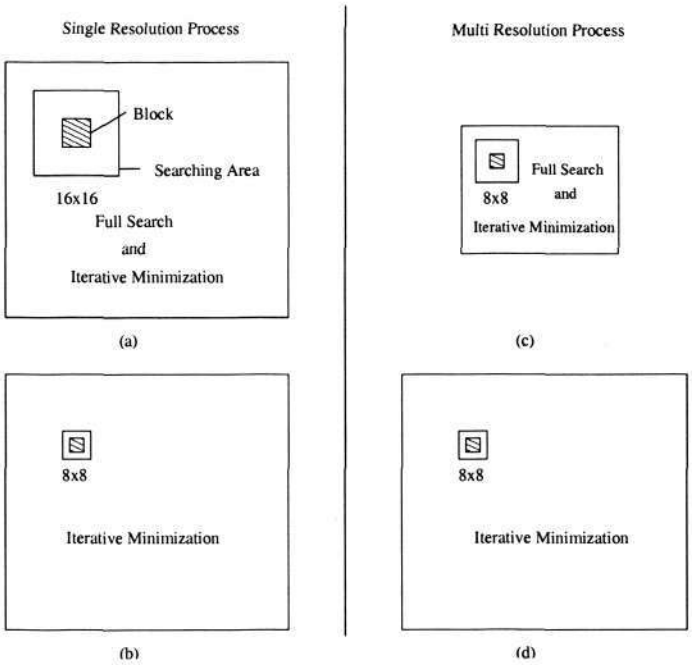


Figure 5.14: A Comparison between Single and Multiresolution Processes.

Prior plus Reduced Block-Size Algorithm," and "Multiresolution Block Matching with Smoothness Prior Function Algorithm". Figures 5.14 (a) and (b) show two steps of the first algorithm. The first algorithm uses all the pixels to find the motion field. The second algorithm uses less pixels to find the motion field. Figures 5.14 (c) and (d) show two steps of the second algorithm. Moreover, the second algorithm still uses an integer type motion field as the first algorithm. So the multiresolution part gives rougher estimation than the first algorithm. However, if the minimization process in the high resolution can recover the rough estimation, it would be good.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

5.5 Experimental Results

In the experiments, we use two synthetic sequences: the Peak and Yosemite sequences, and four real sequences: the Suzie, Foreman, Salesman and Hamburg Taxi sequences. All sequences comprise the QCIF 176×144 pixels images. For all experiments, the motion fields are of an integer type.

The strategies and summaries of the experiments and results are presented here.

5.5.1 Strategies and Summaries

Our experiments examine motion estimation algorithms from the conventional block matching algorithm to the more complicated ones. The list are:

- Conventional block matching using block-sizes of 16×16 and 8×8, respectively;
- Overlap block matching using block-size of 16×16 for the calculation while block-size of 8×8 is for the result;
- Block matching with smoothness prior energy function using block-sizes of 16×16 and 8×8, respectively;
- Block matching with discontinuity adaptive smoothness prior energy function using block-sizes of 16×16 and 8×8, respectively;
- Block matching with smoothness prior energy function plus reduced block-size;
- Multiresolution block matching with smoothness prior energy function.

There are a total of six experiments.

5.5.1.1 Conventional and Overlap Block Matching

Here we compare the result of block matching using block-sizes of 16×16 and 8×8.

We observe that:

- PSNR value increases if the block-size is reduced,
- Entropy value increases if the block-size is reduced,

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

- More incorrect motion vectors appear when we use a smaller block-size.

This shows that one of the problems is the lack of information for a smaller block-size. This has been mentioned before in the *problem and solution* section. However a smaller block-size means a finer motion field, which is often desired.

Because the lack of information is considered as the problem, the simplest approach tries to overcome it by overlap block matching. The result of overlap block matching algorithm shows:

- The motion fields are smoother than conventional block matching using 8×8 block-size,
- The PSNR and entropy values are roughly between the results of conventional block matching using block-sizes of 16×16 and 8×8.

In this experiment, we show all the compensated images and their motion fields for block-sizes of 16×16, 8×8 and the overlap scheme.

5.5.1.2 Block Matching with Smoothness Prior Function

We first use the conventional block matching algorithm for motion estimation, then it is followed by our block matching with the smoothness prior algorithm. We show the result as a compensated image with the motion field overlaid on top of it.

The setting of conventional block matching is listed here. The search strategy is the full search type. The search area is from a location shifted by $(-16, -16)$ to a location shifted by $(+16, +16)$. Two block-sizes, 16×16 and 8×8 have been chosen.

Block matching with smoothness prior algorithm uses $\{-2, -1, 0, 1, 2\}$ as the delta values. In each iteration, these values are to be added to the current value to create new candidate vectors.

The likelihood and prior energy values are normalized. The prior energy is calculated from the total square difference between the center and each of the four neighbor vectors. The weight for likelihood energy is 1.

We deduce that the block matching with smoothness prior function can produce smooth motion field without too much sacrifices on the PSNR values.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

5.5.1.3 Block Matching with Discontinuity Adaptive Smoothness Prior Function

All the experiments in the block matching with smoothness prior function use the quadratic function. This experiment shows a result from the block matching algorithm with a discontinuity potential function as its smoothness prior. We have set the parameter of discontinuity function $\gamma = 10$ and have used the same observation weight as those with the quadratic function. In this way, the results can be compared.

Some of the results show that the energy values have maintained the same PSNR value, while some of the motion vectors have been modified. The experiment shows the importance of the initial value. The discontinuity adaptive potential function keeps the discontinuity of the initial data. This suggests that the discontinuity adaptive smoothness prior energy function can be executed only after the smoothness of the motion field has been reached.

5.5.1.4 Block Matching with Smoothness Prior Function plus Reduced Block-Size

In this experiment, the results of the block matching algorithm in which block-sizes of 16×16 are used for initialization. We found that the result is smooth, although at the last stage the block-size is 8×8. In this case *initialization* takes a major role to shape the motion field, although we use the same energy functions and algorithm.

We conclude that the algorithm propagates the smoothness achieved using a larger block-size to a smaller block-size. The smoothness of the motion field that is propagated is what we desire. So we can still get smooth motion field using smaller block-size.

5.5.1.5 Multiresolution Block Matching with Smoothness Prior Function

We use two image resolutions, a quarter QCIF image and a QCIF image for these multiresolution experiments. The procedure involves several steps. First, an image sequence is re-sized to a smaller image. Second, full search motion estimation is

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

executed using 8×8 block-size with a search area of ± 8 on that image. Third, a minimization process is executed. The result is propagated to a larger image size using bilinear transforms and is followed by the minimization process at that level. The experiment results for each step are presented in the sections below.

5.5.2 Parameter Setting and the Influence of Explicit Smoothness

As a complete model, the weight, λ of Equation 5.1 must be set. The value can be centered at value 1 as the balance value. It works if the values from likelihood and prior functions have been normalized.

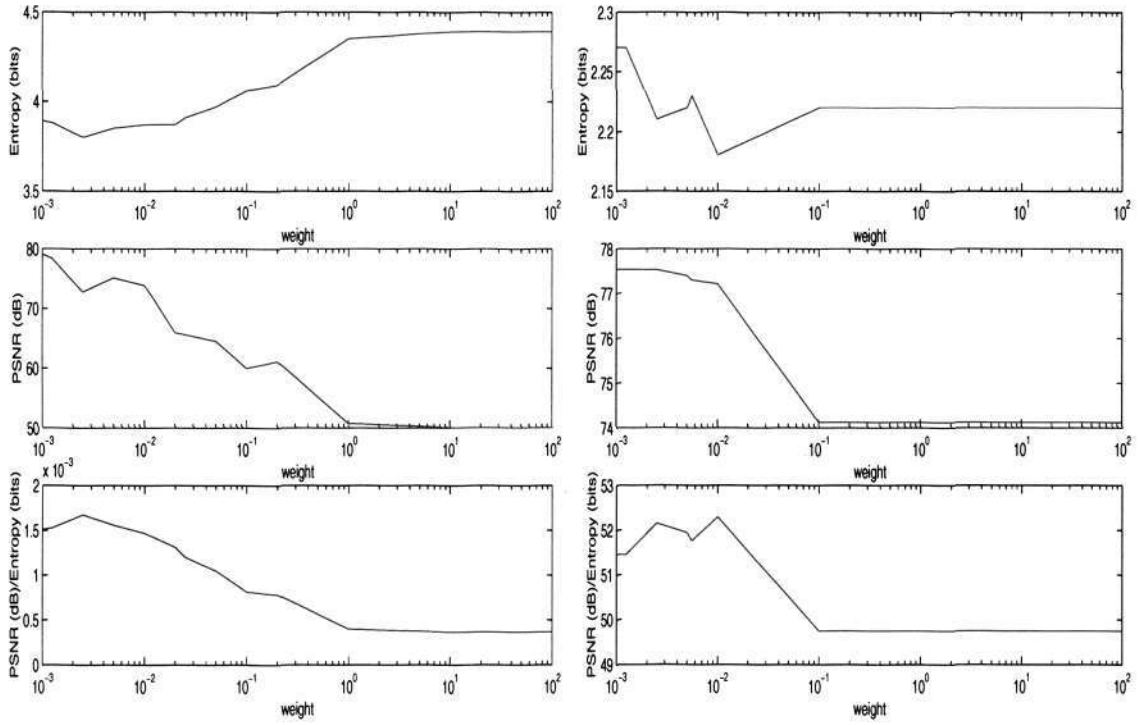
Figure 5.15 reports the weight vs PSNR, Entropy, and PSNR/Entropy from their respective image sequences. The weight value ranges from a small value of 0.001 to a large value of 1 or 100. If the weight value is 0.001, then it is similar to the result of conventional block matching in which the best PSNR is achieved. If the weight value is 1, then the result will be very smooth or reaches a critical area where any motion vector can give a similar energy value. The majority of figures indicate that the PSNR and the Entropy value decreases when the weight value is increased.

We can check the compromise between PSNR and Entropy with:

$$Ratio = \frac{PSNR}{Entropy^{\kappa}}, \tag{5.4}$$

where κ is to modify the strength of entropy influence. This modification is useful because for some sequences they cannot produce a ratio with the peak that we desire (The κ is set to $1, 0.1$ or 0.01 according to the shape of entropy). The peak from the ratio points to the weight that can give high PSNR and low entropy value. This ratio is an illustration to our hypothesis that we need both implicit and explicit smoothness to achieve a good estimation result. So, in every image sequence, we have a weight value that deliver the reasonable compromise of the two types of smoothness. From our synthetic sequences the weight values are near 0.01. From our real image sequence experiments, the weight value is roughly near 0.1 for the Suzie, Foreman and Salesman sequences and 0.01 for the Taxi sequence.

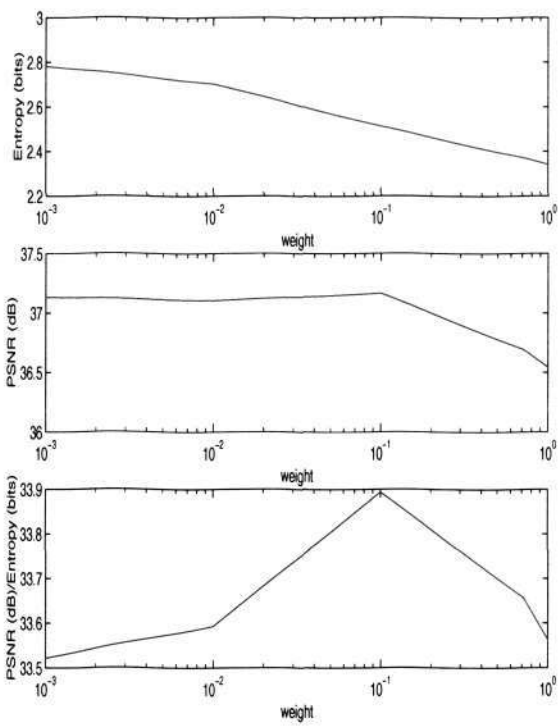
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



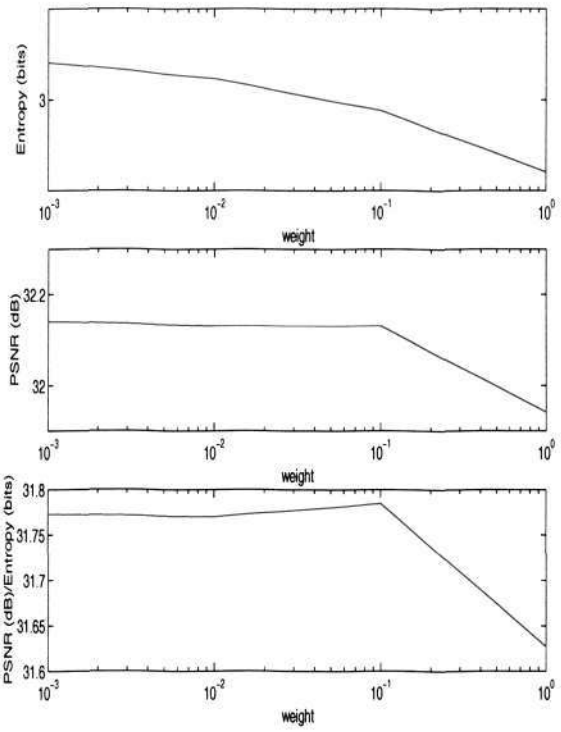
(a) Peak Sequence (b) Yosemite Sequence

Figure 5.15: PSNR, Entropy and PSNR/Entropy vs Weight.

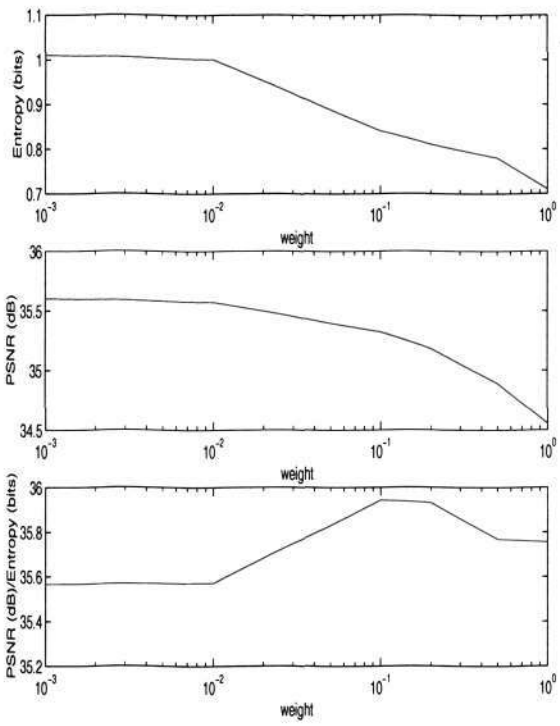
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



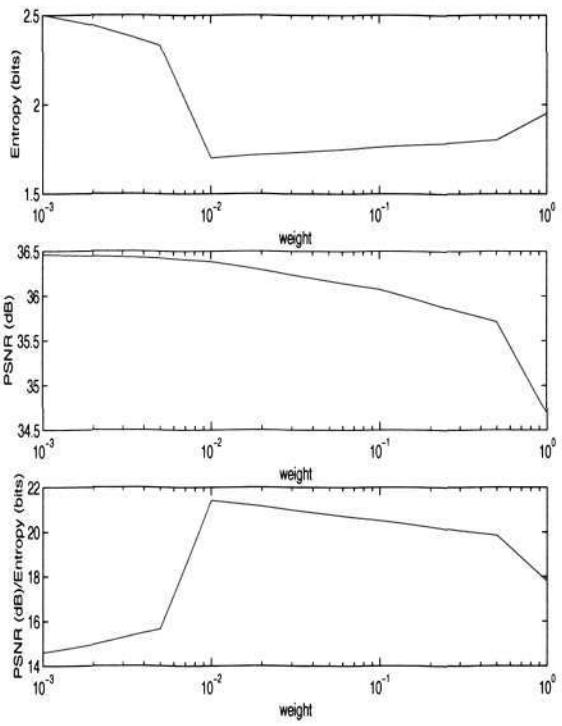
(a) Suzie Sequence



(b) Foreman Sequence



(c) Salesman Sequence



(d) Taxi Sequence

Figure 5.16: PSNR, Entropy and PSNR/Entropy vs Weight (cont.).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

	PSNR (dB)		Entropy (bits)		AAE/SD	
	16	8	16	8	16	8
Conv. BM	$\sim$	$\sim$	2.53	2.50	0.54°/2.33°	1.58°/8.39°
Overlap BM	-	$\sim$	-	2.46	-	1.19°/7.21°
Prior	86.15	80.17	2.52	2.65	0.54°/2.33°	1.58°/8.38°
Discontinuity	86.15	80.17	3.63	3.92	5.26°/6.90°	1.10°/6.54°
Reduced	-	92.17	-	2.63	-	0.54°/2.33°
Red. & Disc.	-	92.17	-	3.83	-	5.25°/7.94°
Multiresolution	-	48.74	-	4.01	-	15.42°/22.16°
$\sim$ is infinite, $-$ is not available						

Table 5.1: PSNR, Entropy and Angular Error for the Peak Sequence. Notes: Conventional (Conv.), Discontinuity (Disc.), Reduced (Red.), Average Angular Error (AAE), Standard Deviation of the Angular Error (SD).

At this stage, the ratio measure is only used to show that the explicit smoothness is helpful when a compromise between PSNR and Entropy is desired. However, quantitative analysis of the properties of the ratio measure will be more useful and will be studied further in the future research.

5.5.3 Results

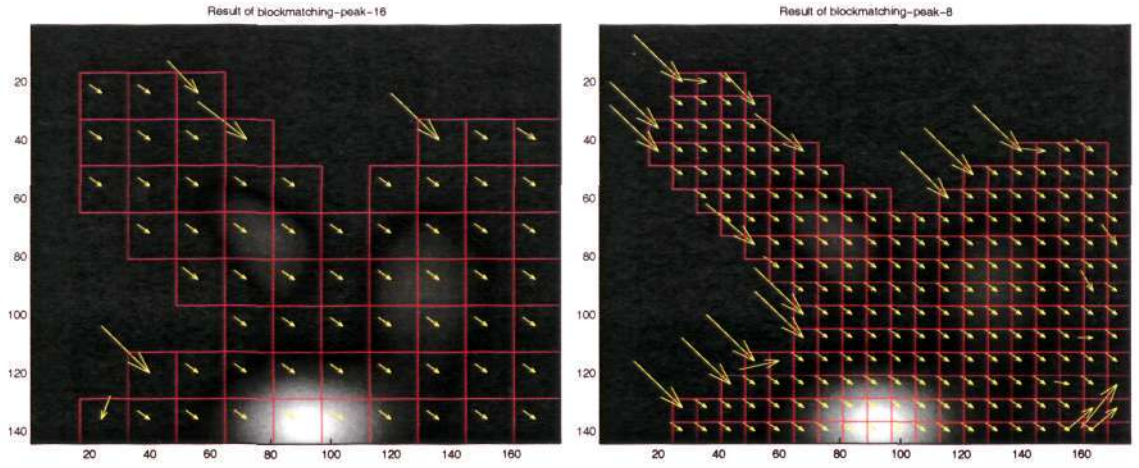
This section lists all the results and the implication of each result. The implications are divided into three groups: the conventional and overlap block matching, the block matching with smoothness prior function, and the multiresolution block matching with smoothness prior function.

5.5.3.1 Peak Sequence

Table 5.1 shows all the experimental results for the Peak Sequence. The PSNR values are high because it is a synthetic sequence with a floating point type.

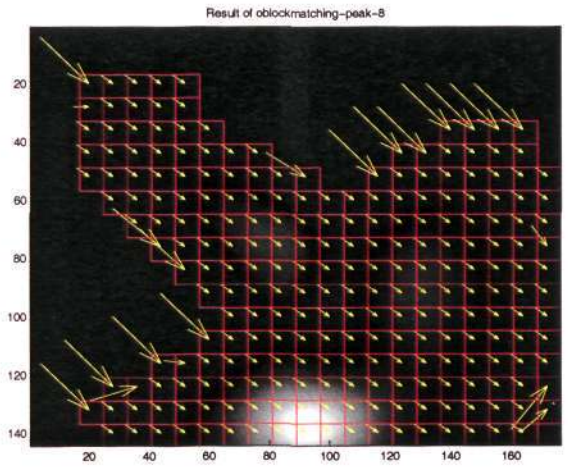
This sequence is generated by *peak.m,* a function from MATLAB. The second image is shifted by vectors $(3, 4)$ pixels from the first image. The value of the vectors become ground truth. This is a whole frame movement. Because the majority of pixel intensities are zero, many parts are considered static and they have no motion field.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) Block Matching for Block-size of 16 by 16

(b) Block Matching for Block-size of 8 by 8



(c) Overlap Block Matching

Figure 5.17: Results of Experiments for the Peak Sequence.

Conventional and Overlap Block Matching Figure 5.17 shows the motion field of the conventional block matching using different block-sizes and the overlap block matching. Several incorrect motion vectors are found, especially near the edge of the image frame in all of the results. The result of overlap block matching shows its inadequacy to improve the quality of the result.

The AAE for the 16×16 block-size is smaller than the AAE of 8×8 block-size. This condition is expected because large block-size increase implicit smoothness. For the experiment using 8×8 block-size, the overlap block matching has a smaller AAE than the conventional one.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Block Matching with Smoothness Prior Function Figure 5.18 shows the smoothing results from three different algorithms. The block matching algorithm using quadratic smoothness prior function gives a smoother result than the other algorithms. The block matching using discontinuity adaptive smoothness prior function preserves the discontinuity, so the result is different from the previous algorithm. It shows the effectiveness of the discontinuity function. For the block matching using smoothness prior function plus reduced block-size, the result shows the progress of adjusting the motion field on the seventh iteration. A longer iteration produces a similar result as the block matching using quadratic prior function.

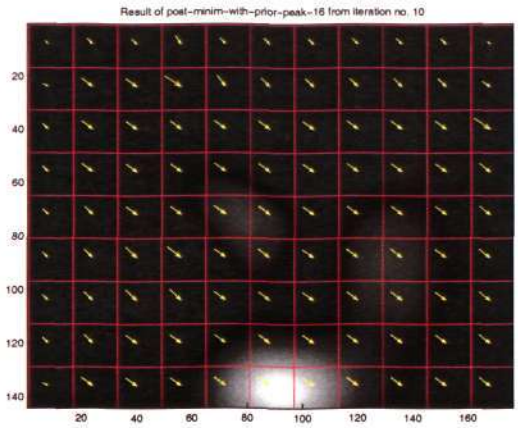
In this sequence, the minimization technique modifies the motion field so that it is smooth for the whole frame. The motion vectors have the same likelihood value because it is a whole frame movement. In this condition, the prior energy value takes a role as a driving force to adjust the motion field.

The progress of PSNR values during the iteration can be explained as following. First, a motion vector reduces its length if its neighbors have smaller motion vectors. In effect it reduces the PSNR value. But after the second iteration, the likelihood energy value becomes dominant and it modifies the motion vector for lower energy. It progresses until the end of the iteration. The Entropy value increases after using the prior function, because the vectors are varied by the function. The PSNR value is also decreased.

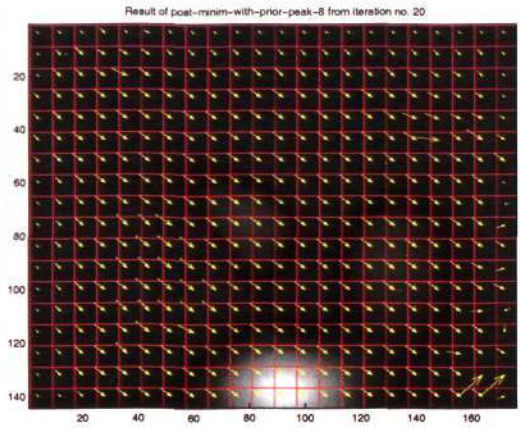
The AAE value of experiment using discontinuity is smaller than the one using square smoothness model. The discontinuity function may choose better vectors in this condition.

Multiresolution Block Matching with Smoothness Prior Function Figure 5.19 shows the result of conventional block matching, the result of the propagation between the resolution levels and the result after the minimization process on the low and high resolutions. Many wrong motion vectors are produced by the full search strategy. The minimization process slightly reduces the number of wrong motion vectors. The compensated image at high resolution image has an PSNR of 48.74 dB. The AAE value has the highest value than the other models.

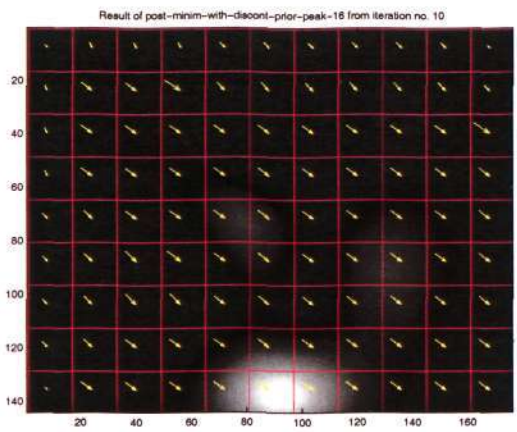
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



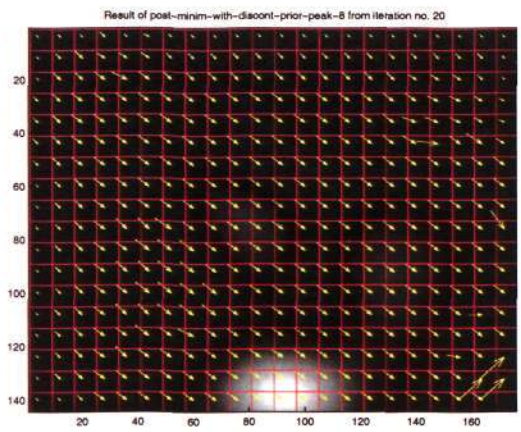
(a) Using Prior Function for Block-Size of 16 by 16



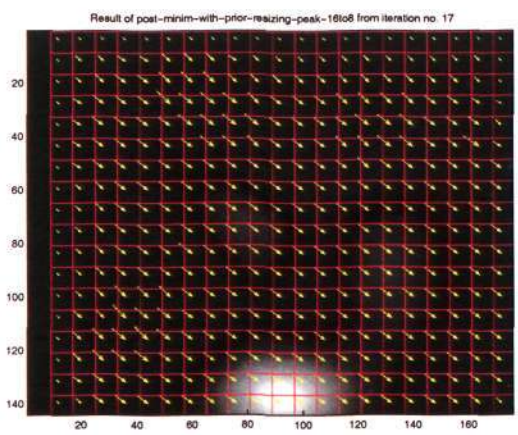
(b) Using Prior Function for Block-Size of 8 by 8



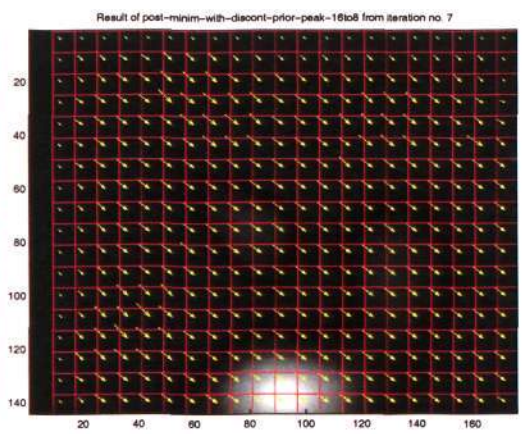
(c) Using Discontinuity Adaptive Smoothness Prior Function for Block-Size of 16 by 16



(d) Using Discontinuity Adaptive Smoothness Prior Function for Block-Size of 8 by 8



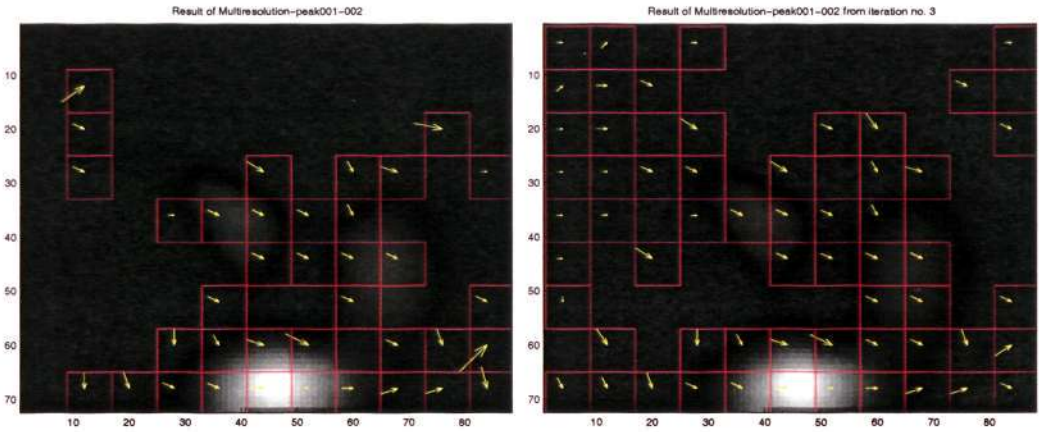
(e) Using Smoothness Prior Function plus Reduced Block-Size



(f) Using Discontinuity Adaptive Prior Function plus Reduced Block-Size

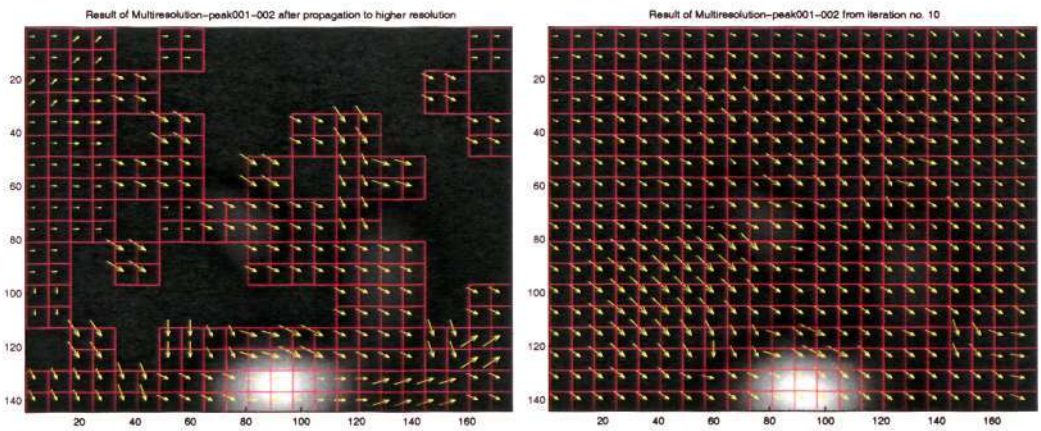
Figure 5.18: Results of Experiments for the Peak sequence. (cont.)

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) After Full Search in Low Resolution

(b) After Minimization Process in Low Resolution



(c) The Result of Propagation in High Resolution

(d) After Minimization Process in High Resolution

Figure 5.19: Results of Multiresolution Block Matching with Smoothness Prior Function for the Peak Sequence.

This type of matching is for a speed reduction. The results are usually lower in quality.

5.5.3.2 Yosemite Sequence

Table 5.2 shows all the experimental results for the Yosemite Sequence. The yosemite motion field is around one or two may not be good enough to be used here because the vectors are an integer type. The PSNR values are high because we do not process the images using filters, and the pixel values are a floating point type. The AAE values are high because we use integer type of vectors.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

	PSNR (dB)		Entropy (bits)		AAE/SD	
	16	8	16	8	16	8
Conv. BM	77.64	77.90	2.19	2.14	55.60°/47.28°	36.42°/32.41°
Overlap BM	-	77.85	-	2.10	-	37.42°/32.97°
Prior	76.24	77.93	1.84	2.27	55.62°/47.28°	36.42°/32.41°
Disc.	76.26	76.57	1.59	1.73	55.62°/47.28°	36.42°/32.41°
Reduced (Red.)	-	76.30	-	1.86	-	39.06°/26.81°
Red. & Disc.	-	76.30	-	1.85	-	39.06°/26.81°
Multiresolution	-	75.57	-	2.42	-	57.29°/36.23°

Table 5.2: PSNR, Entropy and Average Angular Error for the Yosemite Sequence. (Notes: Conventional (Conv.), Discontinuity (Disc.), Average Angular Error (AAE), Standard Deviation (SD)

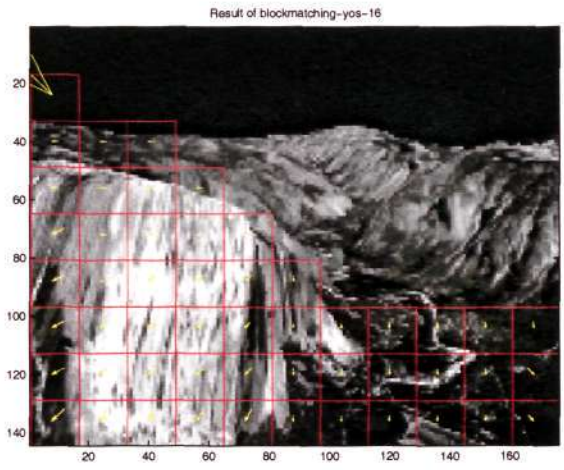
Conventional and Overlap Block Matching Figure 5.20 shows the motion field after we processed the yosemite sequence. We do not use the cloud as is suggested by Black. The matching function produces an odd result in the empty area shown by a long motion vector.

The AAE values of 8×8 block-size experiment is better than the 16×16 block-size. The overlap block matching experiment produces high AAE error. The reason is that the vectors of yosemite sequence are always different between neighbors.

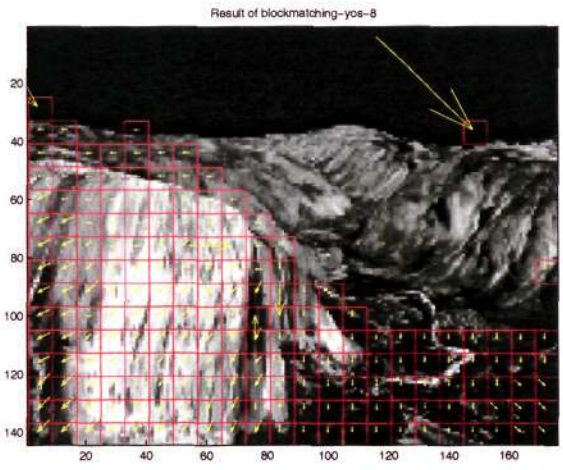
Block Matching with Smoothness Prior Function The prior and discontinuity model cannot produce a further improvement, because the pixels use floating point types and the candidate vectors are limited to integer values. The improvement we notice is the unavailable error motion vectors in the black background area using DA model experiment as is shown in Figure 5.21. It was the cloud area. The improvement of the PSNR value using prior function is minor. The AAE value has no improvement.

Multiresolution Block Matching with Smoothness Prior Function The problem in the multiresolution framework for the yosemite sequence is the small motion vectors. The multiresolution framework relies on the lower resolution for its initialization. Figure 5.22 shows the results. We can see the influence of the prior function in (b) and (d). The reasons for the small motion vectors that do not appear are that prior energy values are less than likelihood energy values.

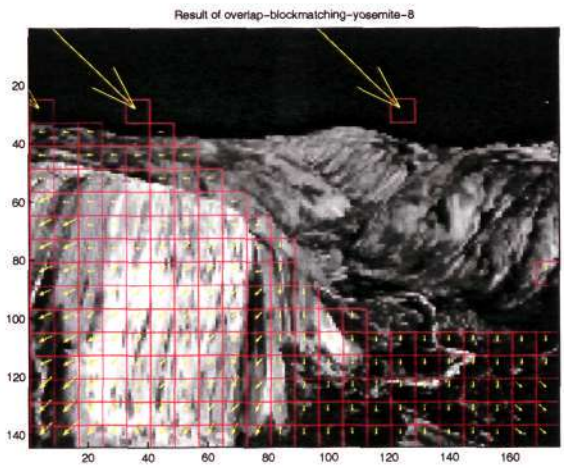
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) Block Matching for Block-Size of 16 by 16



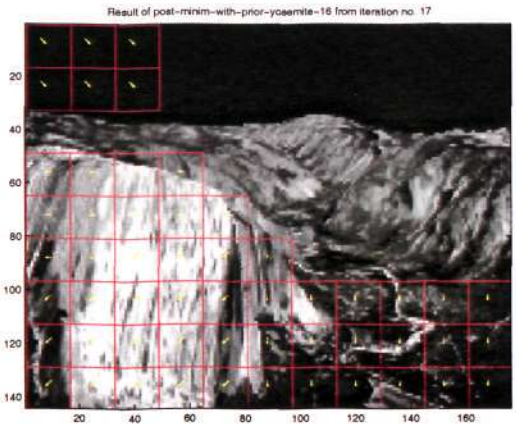
(b) Block Matching for Block-Size of 8 by 8



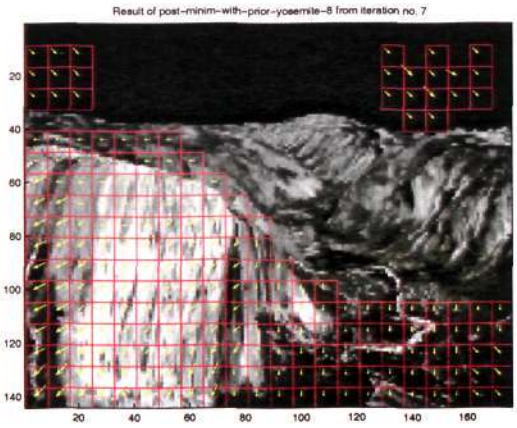
(c) Overlap Block Matching

Figure 5.20: Results of Experiments for the Yosemite Sequence.

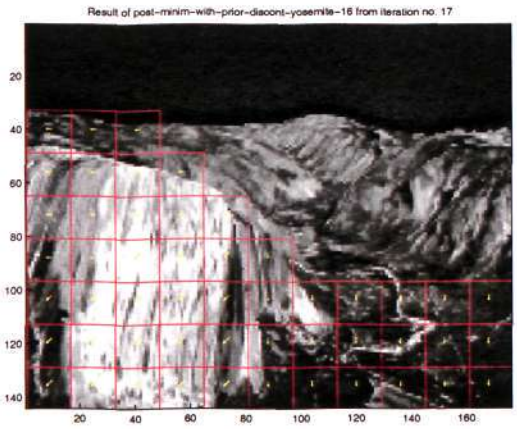
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



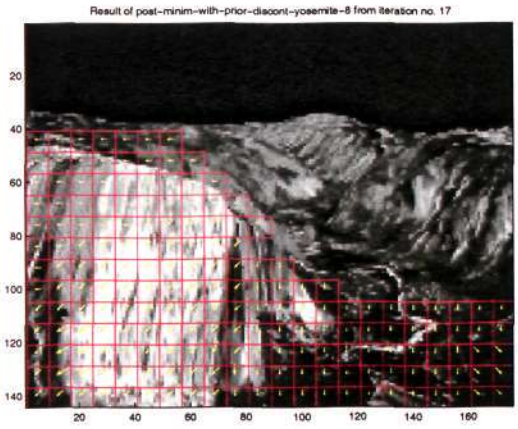
(a) Using Prior Function for Block-Size 16 by 16



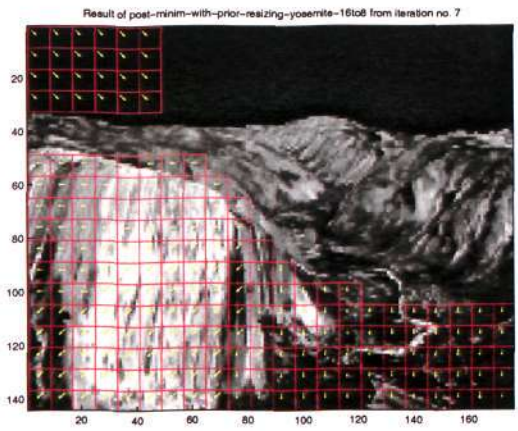
(b) Using Prior Function for Block-Size 8 by 8



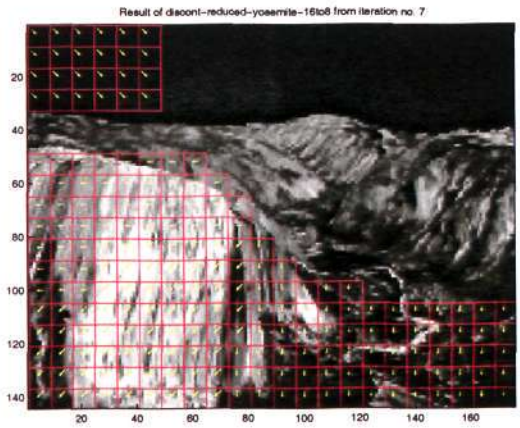
(c) Using Discontinuity Adaptive Smoothness Prior Function for Block-Size 16 by 16



(d) Using Discontinuity Adaptive Smoothness Prior Function for Block-Size 8 by 8



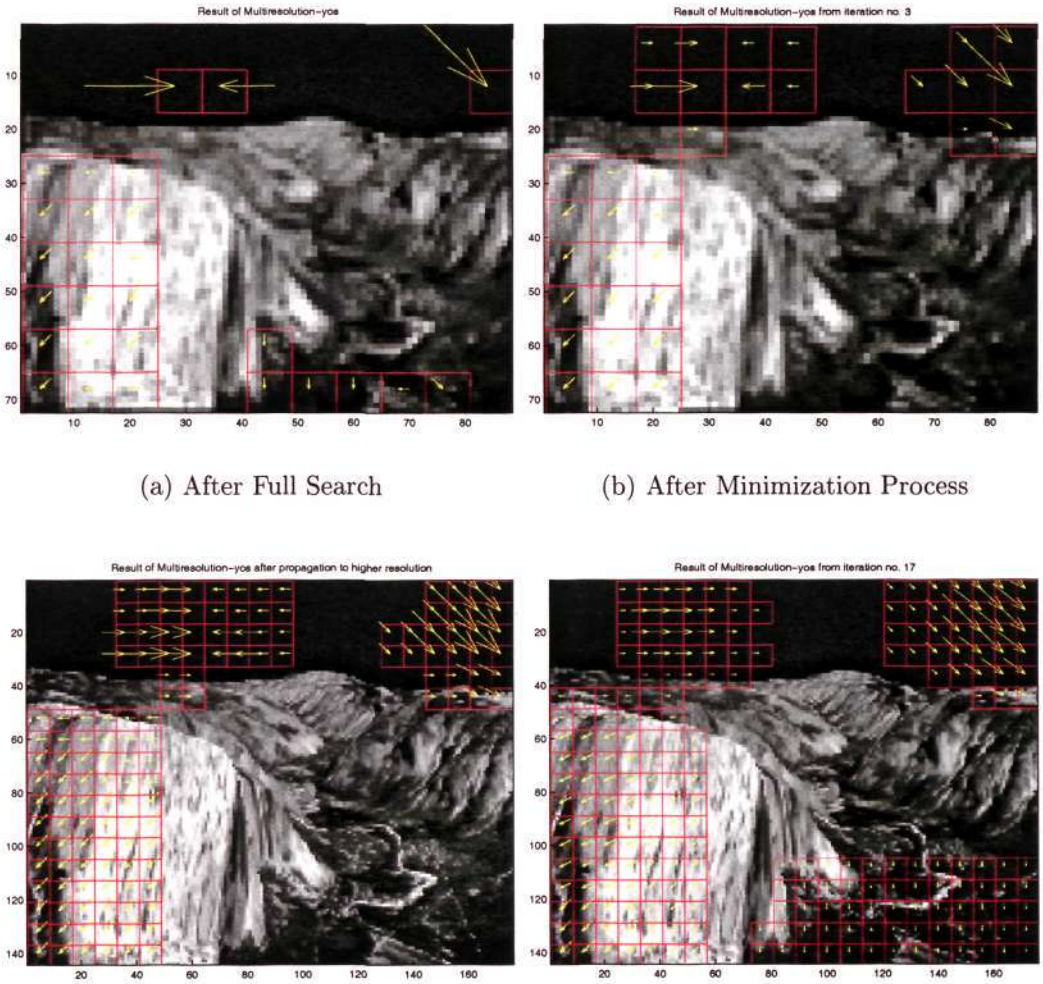
(e) Using Smoothness Prior Function plus Reduced Block-Size



(f) Using Discontinuity Adaptive Prior Function plus Reduced Block-Size

Figure 5.21: Results of Experiments for the Yosemite Sequence. (cont.)

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) After Full Search

(b) After Minimization Process

(c) The Result of Propagation

(d) After Minimization Process

Figure 5.22: Results of Multiresolution Block Matching with Smoothness Prior for the Yosemite Sequence.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

	PSNR (dB)		Entropy (bits)	
	16	8	16	8
Conventional BM	36.49	37.14	2.53	2.81
Overlap BM	-	37.09	-	2.54
Prior	36.52	37.12	2.53	2.82
Discontinuity	36.52	37.12	2.53	2.82
Reduced	-	36.94	-	2.31
Reduced & Discontinuity	-	36.82	-	2.31
Multiresolution	-	33.31	-	1.55

Table 5.3: PSNR and Entropy for the Suzie Sequence

5.5.3.3 Suzie Sequence

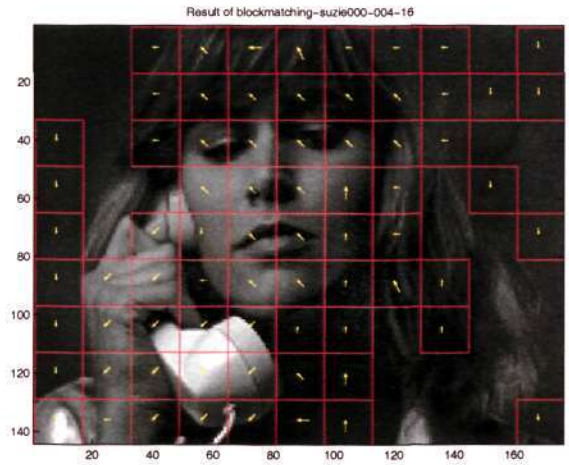
Table 5.3 shows all the experimental results for the Suzie Sequence.

Conventional and Overlap Block Matching Many movements with a difference of 90° and having good textures appear in the Suzie Sequence as shown in Figure 5.23. A smooth motion field is found in the result from the algorithm that uses the 16×16 block-size. Using the 8×8 block-size gives us more incorrect motion vectors. The result with overlap block shows a better motion field than the conventional one.

Block Matching with Smoothness Prior Function Figure 5.24 shows all the results of the algorithms with smoothness prior function for the Suzie sequence. Block matching with quadratic smoothness prior function increases smoothness of the motion field by adjusting its length and orientation. So it gives a better result than the conventional one. For the one with the DA prior function, several motion vectors on the cheek and chin are better than those obtained with the quadratic prior function. For block matching using the smoothness prior function plus reduced block-size, the motion field of the initial values influences the final result.

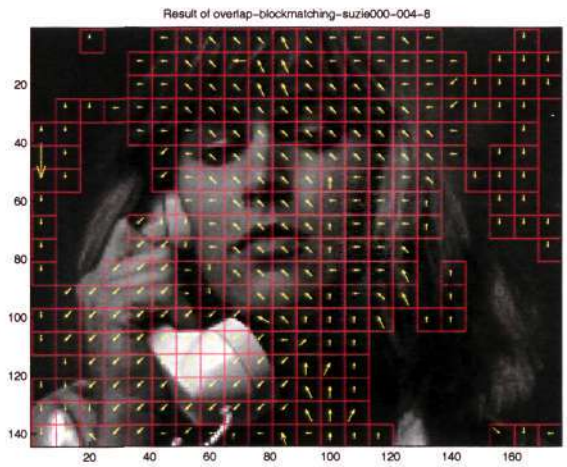
Multiresolution Block Matching with Smoothness Prior Function Figure 5.25 shows the result of conventional block matching, the result of propagation between resolution level and the result after minimization process in the low and high resolutions. The movement in this sequence is small. For some locations,

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) Block Matching for Block-Size of 16 by 16

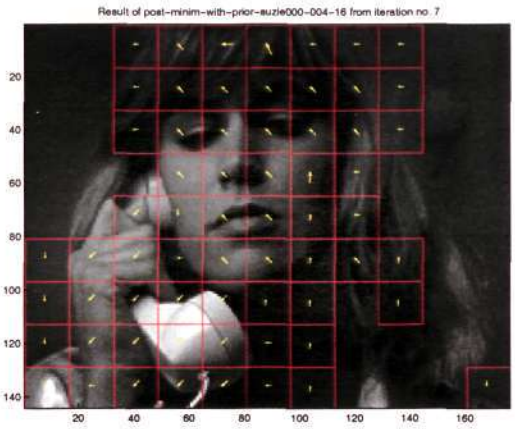
(b) Block Matching for Block-Size of 8 by 8



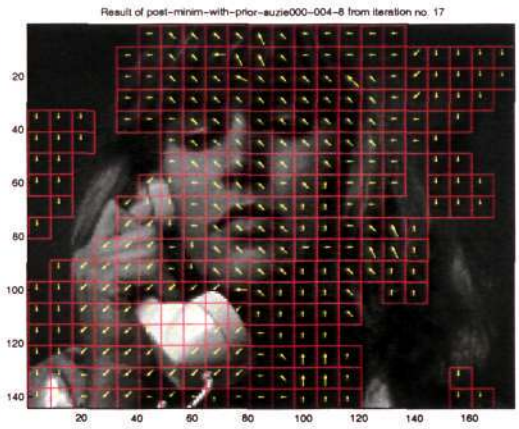
(c) Overlap Block Matching

Figure 5.23: Results of Experiments for the Suzie Sequence.

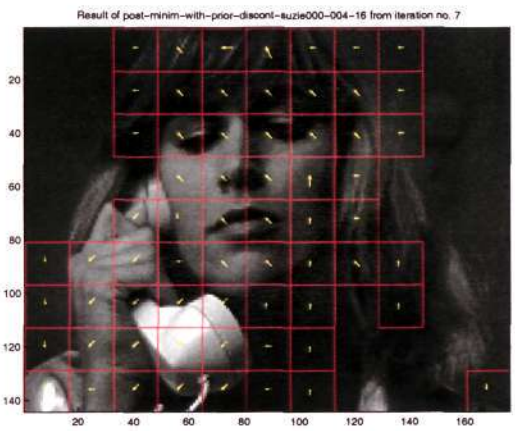
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



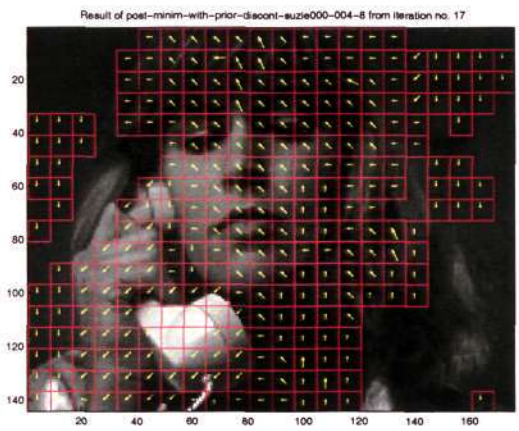
(a) Using Prior Function for Block-Size 16 by 16



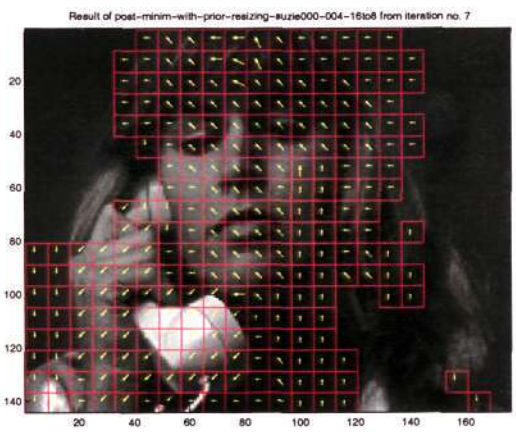
(b) Using Prior Function for Block-Size 8 by 8



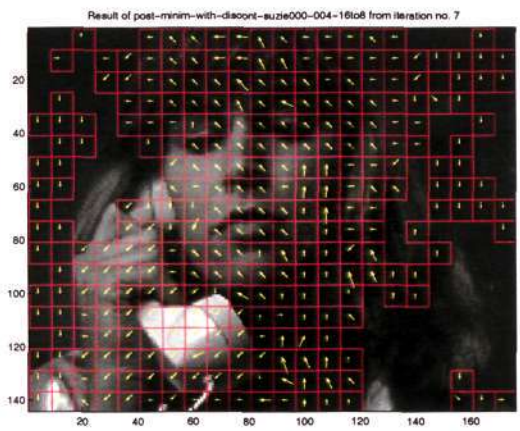
(c) Using Discontinuity Adaptive Smoothness Prior Function for Block-Size 16 by 16



(d) Using Discontinuity Adaptive Smoothness Prior Function for Block-Size 8 by 8



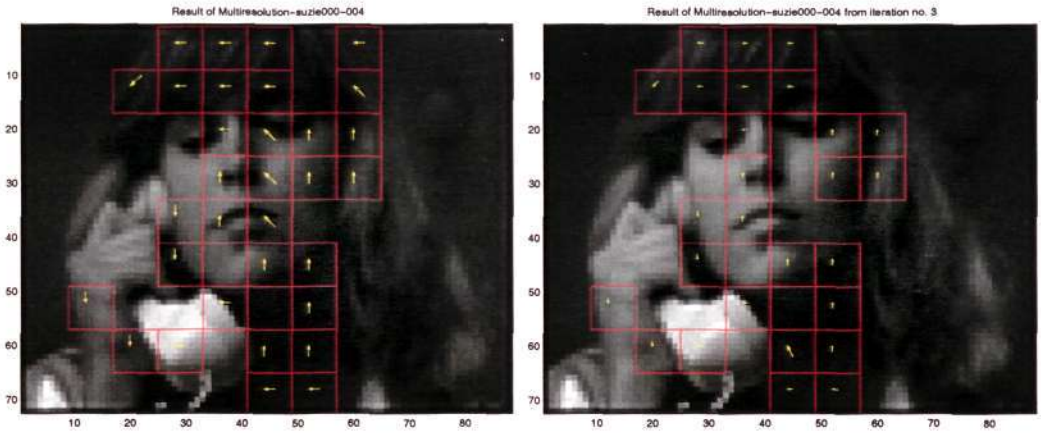
(e) Using Smoothness Prior Function plus Reduced Block-Size



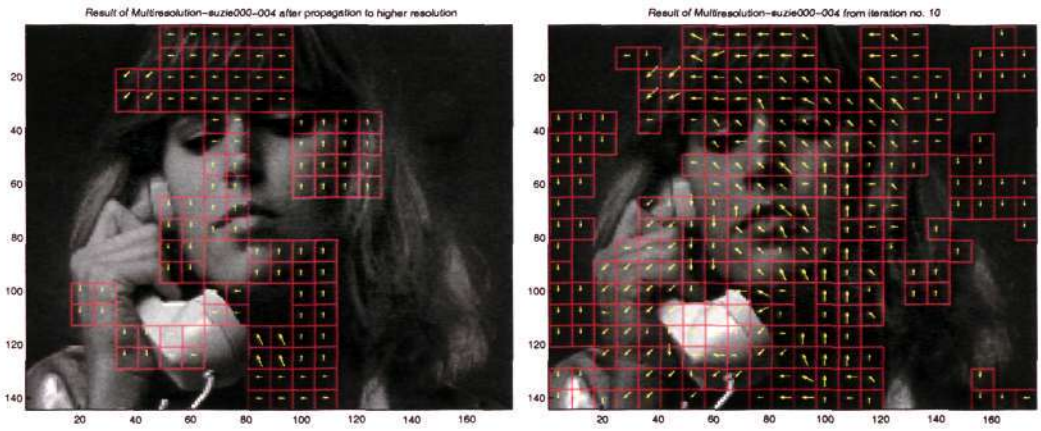
(f) Using Discontinuity Adaptive Prior Function plus Reduced Block-Size

Figure 5.24: Results of Experiments for the Suzie Sequence. (cont.)

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) After Full Search (b) After Minimization Process



(c) The Result of Propagation (d) After Minimization Process

Figure 5.25: Results of Multiresolution Block Matching with Smoothness Prior for the Suzie Sequence.

the minimization process will remove wrong motion vectors. The slightly decreased PSNR value is expected. The compensated image on the higher resolution image has PSNR of 33.31 dB. This is the lowest PSNR, because the algorithm only gives small adjustment to the current motion vector. It stops at the local minima.

5.5.3.4 Foreman Sequence

Table 5.4 shows all the experimental results for the Foreman Sequence.

Conventional and Overlap Block Matching Figure 5.26 shows the experiment with the Foreman Sequence. This sequence has long motion vectors. The helmet

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

	PSNR (dB)		Entropy (bits)	
	16	8	16	8
Conventional BM	31.49	32.98	2.85	3.93
Overlap BM	-	32.14	-	3.17
Prior	31.38	32.07	2.84	3.17
Discontinuity	31.38	32.03	2.85	3.17
Reduced	-	31.70	-	2.82
Reduced & Discontinuity	-	31.72	-	3.01
Multiresolution	-	30.66	-	2.25

Table 5.4: PSNR and Entropy for the Foreman Sequence

has less texture. This area is prone to produce incorrect motion vectors. The rim of the helmet is where the different lengths of the motion vectors exist. The overlap block matching has removed some of the wrong motion vectors.

Block Matching with Smoothness Prior Function Figure 5.27 shows the result for Foreman sequence. For this sequence, the algorithm decreases the PSNR values because of the influences of the smoothness prior. So, the algorithm is working as expected. Many improvements can be observed at the white helmet area which is textureless. We can see several improvements of the motion field on the face, the background and the helmet area. The one with the discontinuity adaptive potential function looks better.

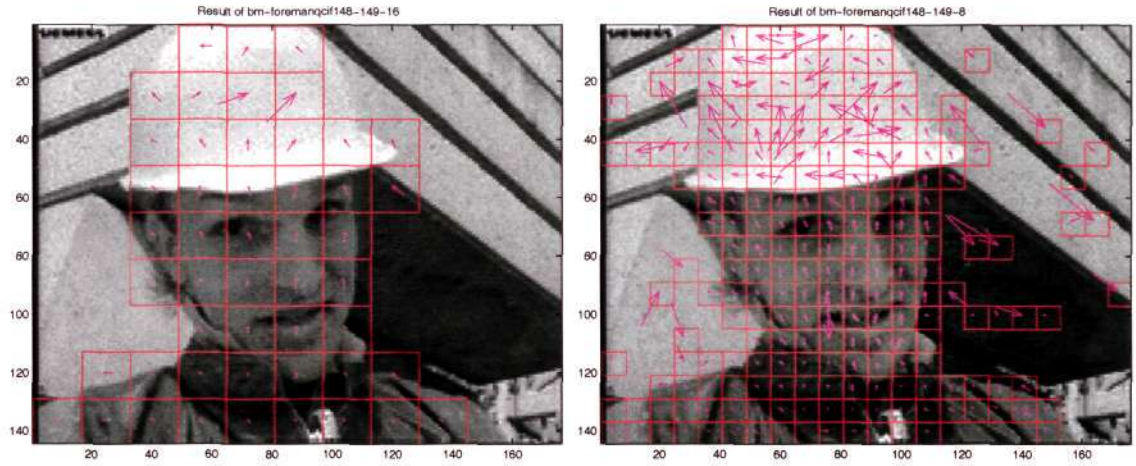
Multiresolution Block Matching with Smoothness Prior Function Figure 5.28 shows the result of conventional block matching, the result of the propagation between resolution levels and the result after minimization process on the low and high resolutions. The compensated image at higher resolution image has PSNR of 30.66 dB.

5.5.3.5 Traveling Salesman Sequence

Table 5.5 shows all the experimental results for the Salesman Sequence.

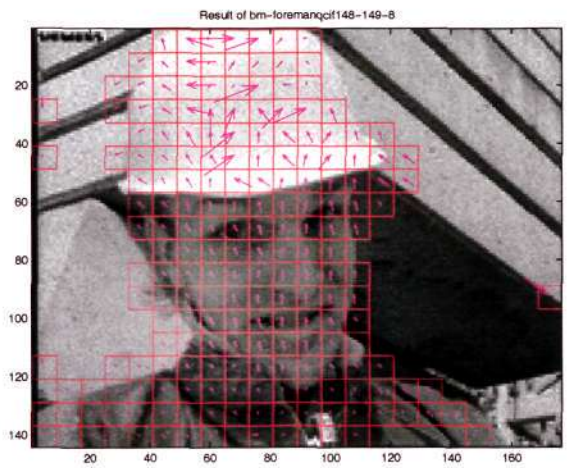
Conventional and Overlap Block Matching This sequence, shown in Figure 5.29, has several small moving objects and static background. This sequence is very interesting, because it has textureless area; it has a small box; and it has a blurred

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) Block Matching for Block-size of 16 by 16

(b) Block Matching for Block-size of 8 by 8



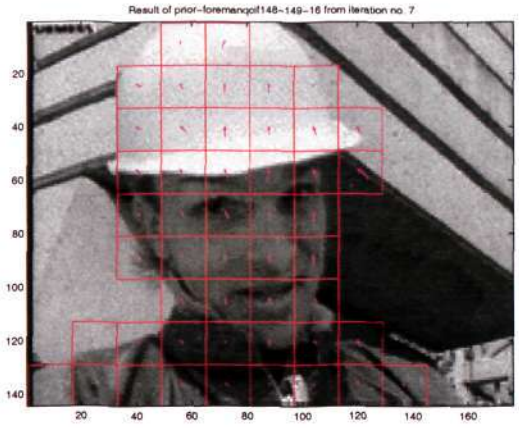
(c) Overlap Block Matching

Figure 5.26: Results of Experiments for the Foreman Sequence.

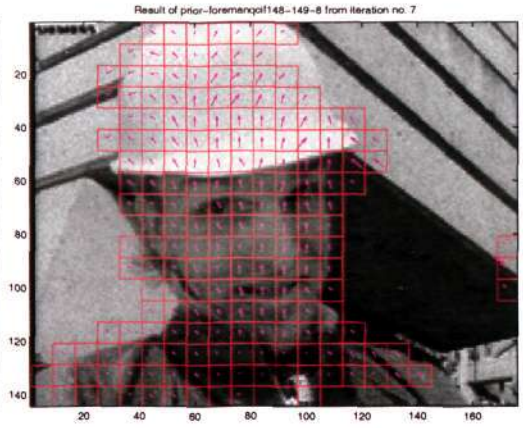
	PSNR (dB)		Entropy (bits)	
	16	8	16	8
Conventional BM	33.46	35.60	0.65	1.01
Overlap BM	-	34.56	-	0.81
Prior	33.46	34.95	0.65	0.84
Discontinuity	33.48	35.08	0.62	0.81
Reduced	-	36.21	-	0.65
Reduced & Discontinuity	-	36.04	-	0.65
Multiresolution	-	34.83	-	0.45

Table 5.5: PSNR and Entropy for the Salesman Sequence

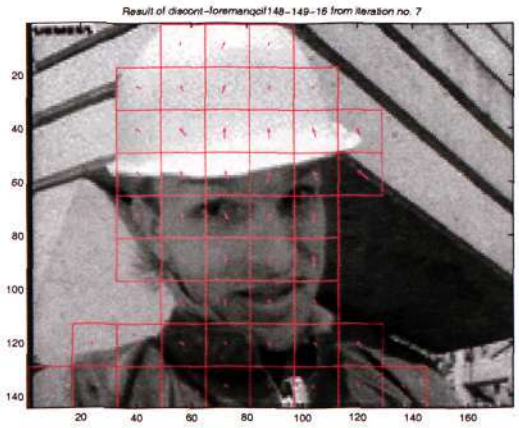
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



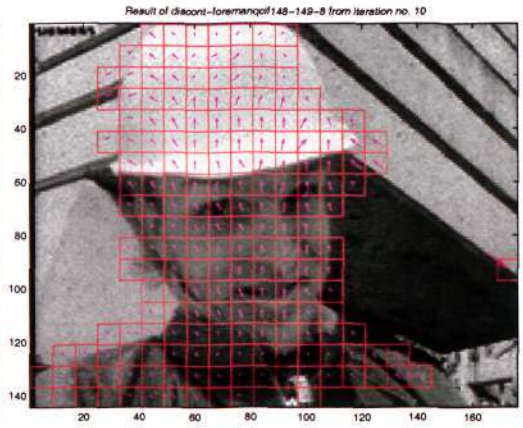
(a) Using Prior Function for Block-Size of 16 by 16



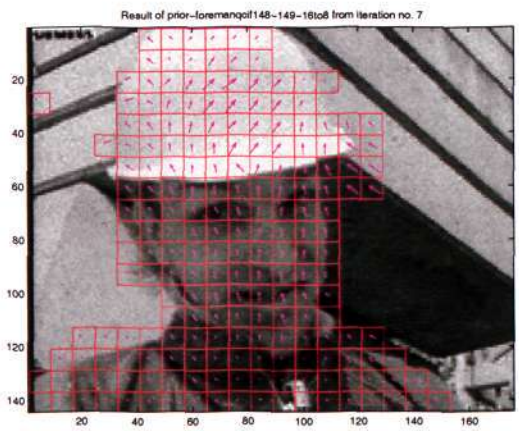
(b) Using Prior Function for Block-Size of 8 by 8



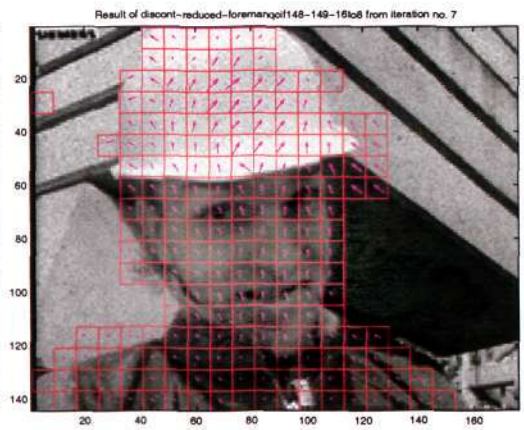
(c) Using Discontinuity Adaptive Smoothness Prior Function for Block-Size of 16 by 16



(d) Using Discontinuity Adaptive Smoothness Prior Function for Block-Size of 8 by 8



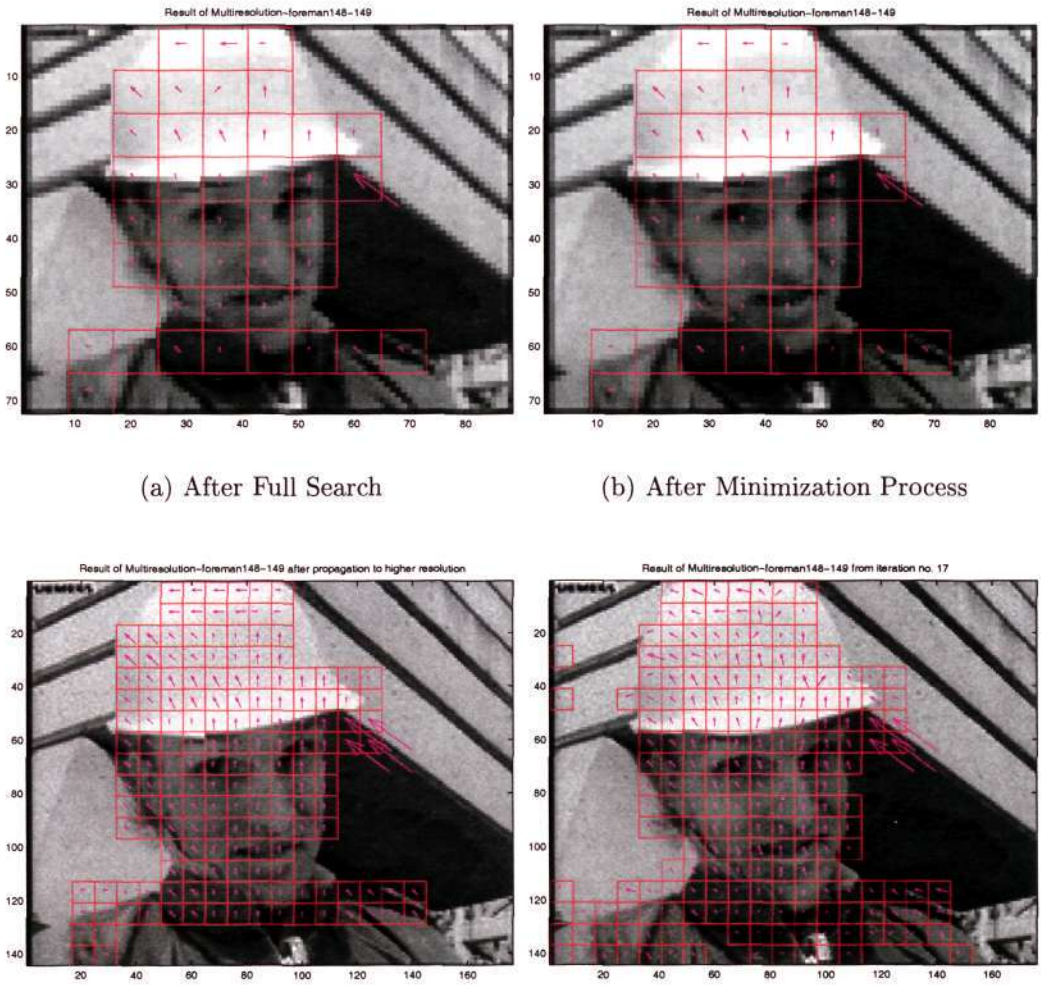
(e) Using Prior Function plus Reduced Block-Size



(f) Using Discontinuity Adaptive Prior Function plus Reduced Block-Size

Figure 5.27: Results of Experiments for the Foreman Sequence. (cont.)

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) After Full Search

(b) After Minimization Process

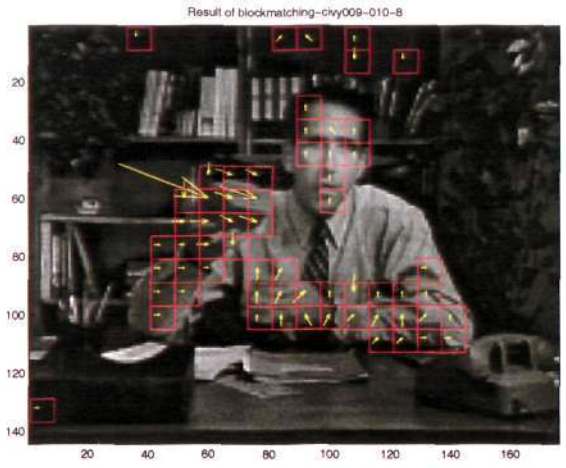
(c) The Result of Propagation

(d) After Minimization Process

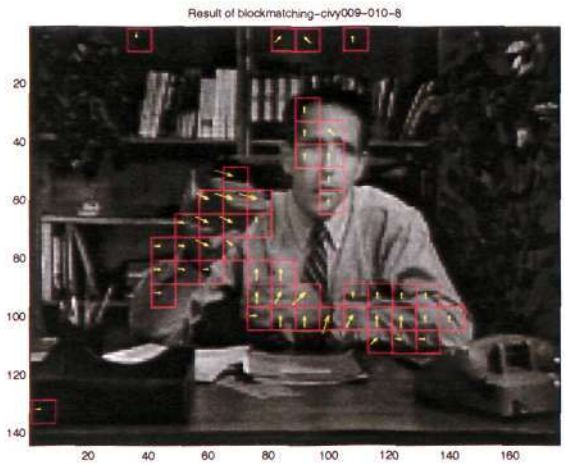
Figure 5.28: Results of Multiresolution Block Matching with Smoothness Prior for the Foreman Sequence.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

(a) Block Matching for Block-Size of 16 by 16



(b) Block Matching for Block-Size of 8 by 8



(c) Overlap Block Matching

Figure 5.29: Results of Experiments for the Salesman Sequence.

moving hand. The shadow on the table can trigger the generation of motion vector. The result of the overlap block matching experiment near the right face of the man shows a problem with the overlap block matching technique compared to the result of the conventional block matching.

Block Matching with Smoothness Prior Function Figure 5.30 shows all the results of the algorithms with smoothness prior function for the Salesman sequence. Block matching using the quadratic smoothness prior function adjusts the motion field nicely, but matching with discontinuity adaptive produced a slightly higher

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

	PSNR (dB)		Entropy (bits)	
	16	8	16	8
Conventional BM	34.75	36.46	0.89	2.51
Overlap BM	-	36.38	-	1.29
Prior	34.92	35.93	0.65	1.78
Discontinuity	34.92	36.38	0.65	1.70
Reduced	-	35.90	-	0.65
Reduced & Discontinuity	-	35.88	-	0.65
Multiresolution	-	36.03	-	0.95

Table 5.6: PSNR and Entropy for the Taxi Sequence.

PSNR value. Block matching using the smoothness prior function plus reduced block-size gives the highest PSNR value.

Multiresolution Block Matching with Smoothness Prior Function Figure 5.31 shows the result of conventional block matching, the result of the propagation between resolution levels and the result after minimization process on the low and high resolutions. As the moving object is smaller than the block-size, it is easily removed by the minimization process. The compensated image at higher resolution has PSNR of 34.87 dB.

5.5.3.6 Hamburg Taxi Sequence

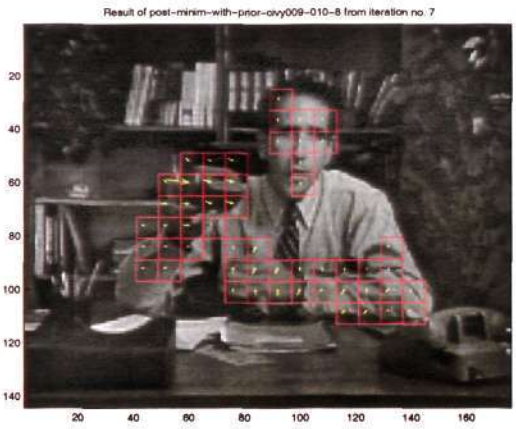
Table 5.6 shows all the experimental results for the Taxi Sequence.

Conventional and Overlap Block Matching In Figure 5.32, the Hamburg Taxi Sequence has many small-object movements, some with less than a pixel length. The taxi moves diagonally. They pose difficulties to our integer type algorithm. The dark color portions of the moving cars on the left and right sides give another problem. Some noise movement is also there. The overlap block matching result reduces wrong motion vectors and reduces the PSNR value.

Block Matching with Smoothness Prior Function The movement in this taxi sequence is small. Figure 5.33 shows two small group of vectors on the left and the right of the image. In one step of the iteration the PSNR is improved. It remains at the same value because of the small movement. This algorithm is shown to be

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

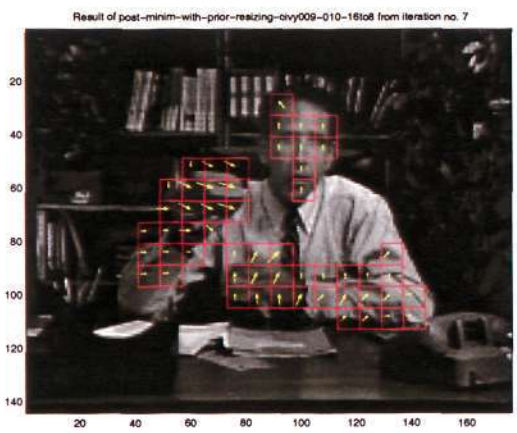
(a) Using Prior Function for Block-Size of 16 by 16



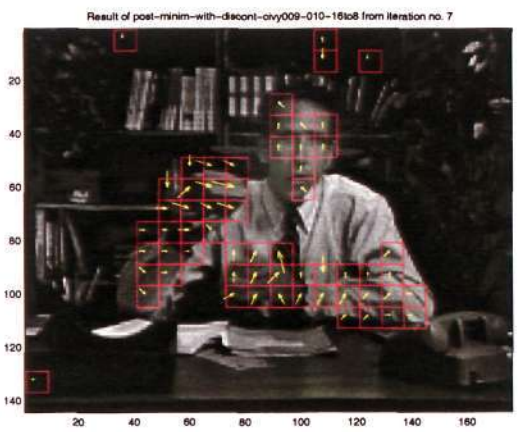
(b) Using Prior Function for Block-Size of 8 by 8

(c) Using Discontinuity Adaptive Smoothness Prior Function for Block-Size of 16 by 16

(d) Using Discontinuity Adaptive Smoothness Prior Function for Block-Size of 8 by 8



(e) Using Smoothness Prior Function plus Reduced Block-Size



(f) Using Discontinuity Adaptive Prior plus Reduced Block-Size

Figure 5.30: Results of Experiments for the Salesman Sequence. (cont.)

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

(a) After Full Search

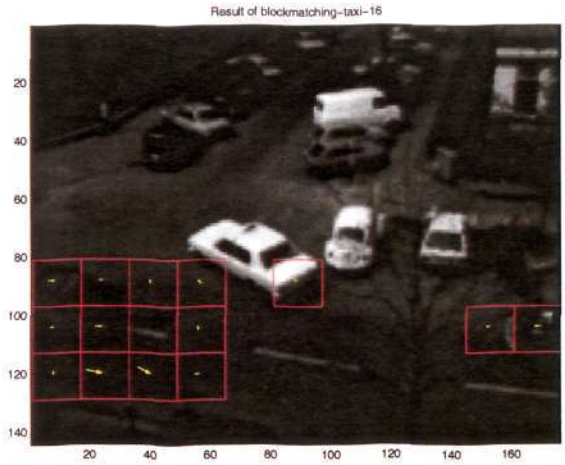
(b) After Minimization Process

(c) The Result of Propagation

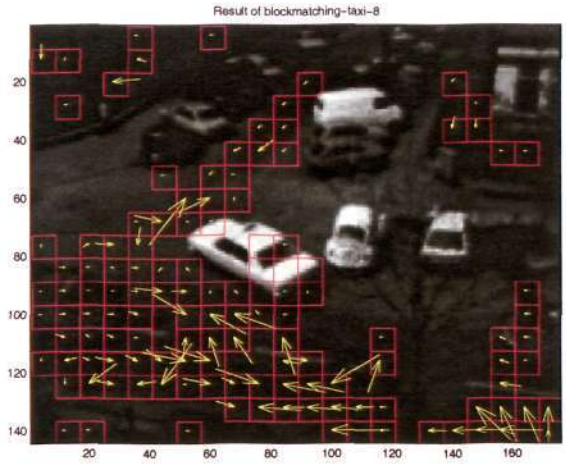
(d) After Minimization Process

Figure 5.31: Results of Multiresolution Block Matching with Smoothness Prior for the Salesman Sequence.

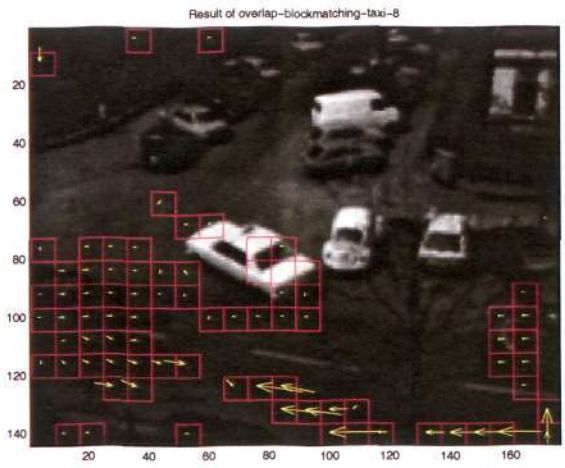
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) Block-size of 16 by 16



(b) Block-size of 8 by 8



(c) Overlap Block Matching

Figure 5.32: Results of Experiments for the Hamburg Taxi Sequence.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

effective, compared to the result of conventional block matching. It removes many false motion vectors. The PSNR progress has the same characteristic as the other sequences. The algorithm reduces the PSNR value from the beginning for several iterations before it increases the PSNR value.

From this experiment we also conclude that by using larger block-size, the block matching algorithm produces a smoother motion field.

Multiresolution Block Matching with Smoothness Prior Function Figure 5.34 shows the result of conventional block matching, the result of the propagation between resolution levels and the result after minimization process on the low and high resolutions. The PSNR value of the compensated image is 35.82 dB for the higher resolution. This algorithm gives another state of the motion field with the lowest PSNR value.

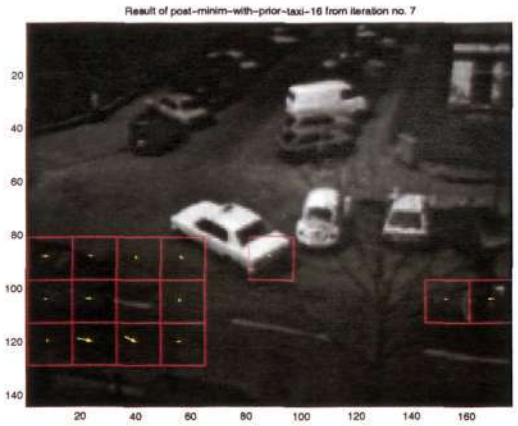
5.6 Conclusion

The idea of bringing the MRF model from the gradient into the matching based technique has been proven to be a good approach to improve the quality of the motion estimation of the conventional block matching algorithm. Improving the quality of the motion field means promising higher level processes or applications. The proposed algorithm adjusts the motion field so that the correct motion field may be achieved. For the desired quality of the motion estimation, the PSNR may decrease, but at the same time it should reduce the Entropy value. These are simple metric tools to help our objective judgment of the motion field.

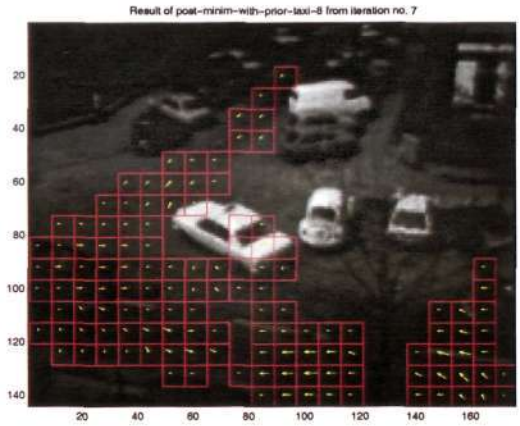
One of driving forces to introduce this proposal is the availability of hardware for the conventional block matching algorithm. This might motivate the hardware designer to modify their design for higher level applications. Other driving force is the need to have a better initialization for the dense motion motion estimation algorithm.

We hold the hypothesis that the generic motion estimation using GMRF model will provide a well defined model for block matching based motion estimation.

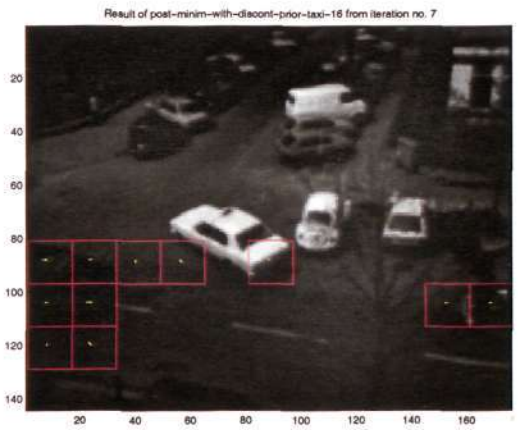
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



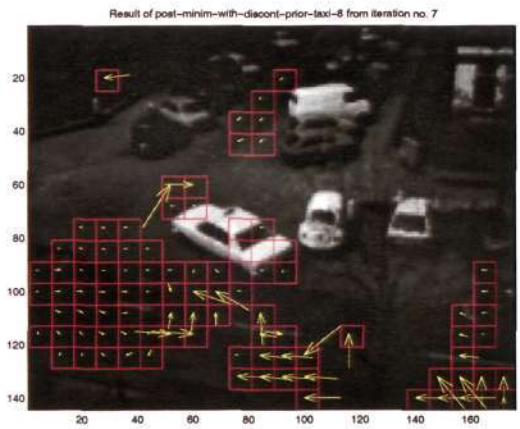
(a) Using Prior Function for Block-Size of 16 by 16



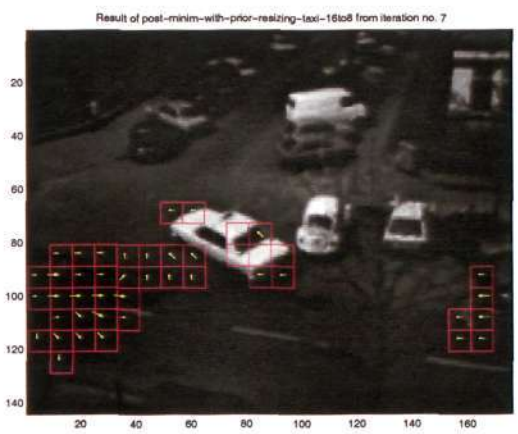
(b) Using Prior Function for Block-Size of 8 by 8



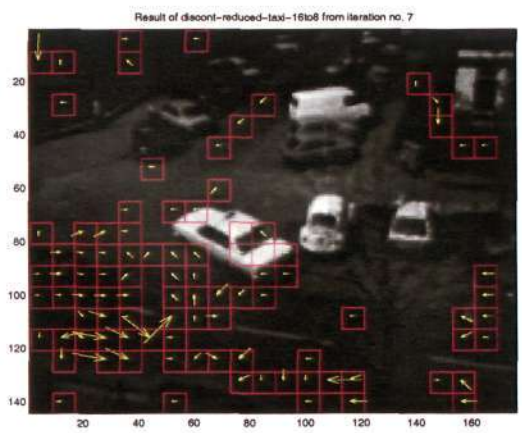
(c) Using Discontinuity Adaptive Smoothness Prior Function for Block-Size of 16 by 16



(d) Using Discontinuity Adaptive Smoothness Prior Function for Block-Size of 8 by 8



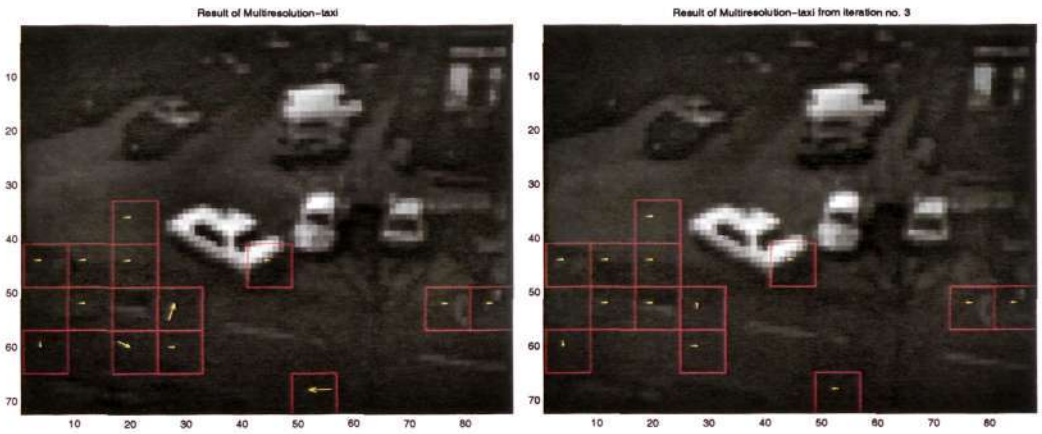
(e) Using Prior Function plus Reduced Block-Size



(f) Using Discontinuity Adaptive Prior Function plus Reduced Block-Size

Figure 5.33: Results of Experiments for the Hamburg Taxi Sequence. (cont.)

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) After Full Search

(b) After Minimization Process

(c) The Result of Propagation

(d) After Minimization Process

Figure 5.34: Results of Multiresolution Block Matching with Smoothness Prior for the Hamburg Taxi Sequence.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Throughout the design and the experiments conducted, we are convinced that this is true.

The contributions of the work in this chapter are as following

- We introduced smoothness prior function to block matching algorithm. It was also shown experimentally that this smoothness prior function gives better results. To the best of our knowledge, this is the first investigation on mixed constraints in matching based motion estimation.

- A measurement about usefulness of mixing the constraints using weight versus the PSNR/Entropy diagram is introduced. In all experiments using real image sequences, an optimal weight between 0.1 and 10 is evident.

- We developed an algorithm to implement the idea of mixed constraints. Many factors are also considered in the algorithm to improve the source for data initialization such as the DA for the discontinuity adaptive function, the reduced block-size and the multiresolution.

- Detailed comparison is made on synthetic image sequences as well as real image sequences. In all the comparisons, we found our approach gives better results than conventional block matching by the PSNR value, entropy value and visual inspection.

- The local minimization using ICM plus good initial data is the desireable approach for this model.

- The technique using the smoothness prior function plus reduced block-size has a better performance than the multiresolution algorithm where the candidates for lower resolution matching are integer pixel motion vectors. The integer value causes many unmatched motion vectors in the lower resolution level. It can be improved if the half-pixel motion vector is used in the same resolution level. This error is propagated to the higher resolution levels. Because the algorithm uses local minimization technique that relies on the initial value, the results differ significantly.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

- The algorithms need extra memory to store the likelihood value but it reduces repeated computations for the minimization step.

- The algorithms using smoothness prior function calculate the likelihood values using the smallest block-size. As a result, the likelihood values of larger block-size can be obtained from the available likelihood values.

The contributions are realized as the proposed algorithms after having a decision that the conventional and overlap block matching techniques are inadequate algorithms to produce a good motion field. The main algorithm is the block matching with a smoothness prior function. This algorithm is shown to improve the result of motion estimation. With a suitable weight, λ, this algorithm can give a high PSNR value with a low entropy value. This hypothesis is verified by the experiments using different weight λ and we found a peak in the PSNR/Entropy ratio diagram. The location of the peak shows the preferred weight.

The behavior of the algorithm is further modified by the discontinuity adaptive smoothness prior function. This algorithm preserves the discontinuity among the neighborhood motion field. In our experimental results, we show that this preserving behavior prevents disorder of the initial values of the motion field. We suggest that this discontinuity adaptive smoothness prior function be activated after a smooth motion field has been developed.

The block matching using the smoothness prior function plus reduced block-size is tested. This is an experiment with the initial values. The initialization is from a large block size that makes the result smoother than block matching using only smoothness prior function. The results resemble the motion field from the large block-size with additional details.

The last algorithm deals with a multiresolution strategy. This results in motion fields with most of them having lower PSNR value and lower entropy value than other algorithms. This result is triggered by the already smooth motion field from the lower resolution, so the motion field at the higher resolution has reached a certain smoothness, but the likelihood value is lower. We conclude that the result has reached another local minima.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

From all the algorithms we have presented, the key algorithm is the block matching with the smoothness prior function in which a simple quadratic equation is used. The other algorithms are variants of this key algorithm.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 6

Block Motion Model for Optical Flow with Smoothness Prior Function

Modifying and adding smoothness prior function into the established motion estimation algorithm have been the main topic in this thesis. The combination between implicit and explicit smoothness has been explored in the previous chapter with the topic: the matching based algorithm for a block motion field. In this chapter, the gradient based block motion model [24, 46] is analyzed and modified with the same objective as in the previous chapter.

We expect an estimation result that would give us a "true" motion field by using this modified model. The meaning of the "true" motion field has been explained in the introduction.

After the experiments, we conclude that the smoothness prior function can fulfil our expectation. The high value of PSNR that is desired combines with the low Entropy produces a unique motion field. Furthermore, the discontinuity function is also implemented to give us an improved result.

We explain the MAP-MRF framework in Section 6.1 and the realization in Section 6.2. Implementation of the model is given in Section 6.3. Section 6.4 reports all results. Finally, Section 6.5 gives a conclusion.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

6.1 MAP-MRF Framework

The same MAP-MRF framework as in the other chapter is used. The posterior probability shown in the Equation 2.2 will have the best value if the motion field f can warp the image d_{i-1} to image d_i correctly.

Given an image sequence, each image is divided into blocks. For each block, instead of the probability function, the energy function is used to model the motion estimation problem. So, the problem is written as

$$U(d_i, d_{i-1}|f) = U(f|d_i, d_{i-1}) + E(f). \tag{6.1}$$

As the desired estimation result is the "true" motion field, this modification gives a prior constraint to the motion field. The common constraint is smoothness.

The Lucas-Kanade block motion model [46] uses implicit constraint for their equation. We will mention their equation as the likelihood energy function. So, in this case, the likelihood energy function is interpreted as a function of a block of optical flow equations:

$$U(u, v|d_i, d_{i-1}) = \sum_{s \in B} (d_x(s)u + d_y(s)v + d_t(s))^2, \tag{6.2}$$

where (u, v) is the elaboration of f and d_x, d_y, and d_t are image derivatives at each pixel in the block B, and s is the location of pixel. The image derivatives are calculated according to Horn-Schunck. From the result of our research presented in the previous chapter, the implicit constraint is not enough to produce the desired motion field.

The explicit constraint to control the behaviors between blocks in a frame is added into the equation. This is called the prior energy function. So, the prior energy is a function of the difference between the center block and its neighboring blocks:

$$U(f) = \left\| \overline{f} - f \right\|, \tag{6.3}$$

where $\overline{f}$ is the local average of neighboring blocks. Both energy functions are merged into one posterior energy function.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

6.2 Realization

Combining the likelihood and prior functions, Equation 6.1 becomes

$$U(u,v|d_i,d_{i-1}) = \sum_{s\in B}\left(d_x(s)u + d_y(s)v + d_t(s)\right)^2 + \lambda(u-\overline{u})^2 + \lambda(v-\overline{v})^2, \quad (6.4)$$

where λ is the weight and B is the block. The equation can be solved in several ways. However, we use the partial derivative technique to solve this equation [24].

Computing the partials of the energy function respectively to u and v, and setting them equal to zero, we have

$$\begin{aligned} \textstyle\sum_{s\in B}(d_x(s)u + d_y(s)v + d_t(s))d_x(s) + \lambda(u-\overline{u}) = 0, \\ \textstyle\sum_{s\in B}(d_x(s)u + d_y(s)v + d_t(s))d_y(s) + \lambda(v-\overline{v}) = 0. \end{aligned} \quad (6.5)$$

Solving the equations simultaneously, we have

$$\begin{bmatrix} \widehat{u} \\ \widehat{v} \end{bmatrix} = \begin{bmatrix} \sum_{s\in B} d_x(s)d_x(s) + \lambda & \sum_{s\in B} d_x(s)d_y(s) \\ \sum_{s\in B} d_x(s)d_y(s) & \sum_{s\in B} d_y(s)d_y(s) + \lambda \end{bmatrix}^{-1} \begin{bmatrix} \lambda\overline{u} - \sum_{s\in B} d_x(s)d_t(s) \\ \lambda\overline{v} - \sum_{s\in B} d_y(s)d_t(s) \end{bmatrix} \quad (6.6)$$

where $\overline{u}$ and $\overline{v}$ are the averages of neighboring motion vectors or the values which are acquired by Gaussian kernel. The initial values of motion vectors are zero. This equation is an improvement of the original equation as is shown in Equation 3.19. There are many ways to realize $\overline{u}$ and $\overline{v}$. This will be explained later. Besides that, the weight of the smoothness prior, λ, influences the smoothness of the result. If $\lambda = 0$, the Lucas-Kanade, Equation 3.19, is obtained. If λ is given, the smoothness is considered.

Instead of calculating $\overline{u}$ and $\overline{v}$, an adaptive interaction function from the Discontinuity Adaptive (DA) MRF that has been mentioned in the previous chapters can be used. The purpose of this modification will be to preserve the discontinuity of the motion field. We can have an adaptive interaction function of Equation 4.10 where η is the difference of motion vectors between a block of pixels and its block-neighbor,

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

and γ is the parameter of discontinuity function. In this way the interaction between a block of pixels and its neighbors is not a constant ratio such as averaging or using a kernel. If we set the parameter $\gamma = 0$ for the adaptive interaction function, the discontinuity function becomes an averaging function.

As we have explained in the comments of our solutions, we highlighted the difference between our equation and the original equation, Equation 3.19. In words, the main difference is the consideration of neighboring motion vectors, $\overline{u}$ and $\overline{v}$. The consideration is always carried out during the iterative calculation to obtain the optimal $\begin{bmatrix} \widehat{u} \\ \widehat{v} \end{bmatrix}$. Furthermore, the calculation of $\overline{u}$ and $\overline{v}$ is adaptive to their values.

6.3 Block Motion Model using Smoothness Prior Function

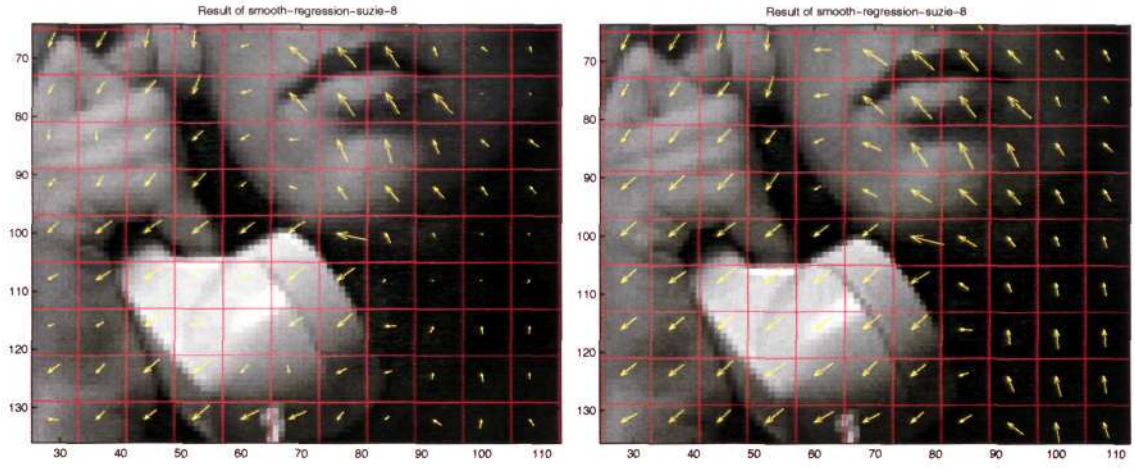
Calculating Equation 6.6 is a straightforward task. However we need to iterate the equation, because the $\overline{u}$ and $\overline{v}$ are unknown in the beginning. For every iteration a new $\overline{u}$ and $\overline{v}$ are computed. After a few iterations, the motion field is adjusted according to the smoothness prior function. The first iteration produces motion field as shown in Figure 6.1(a) and the last iteration produces motion field as shown in Figure 6.1(b).

To calculate the $\overline{u}$ and $\overline{v}$, we can select the second order neighborhood system. The kernel can be:

$$\begin{bmatrix} 1/16 & 1/8 & 1/16 \\ 1/8 & 1/4 & 1/8 \\ 1/16 & 1/8 & 1/16 \end{bmatrix}. \tag{6.7}$$

If we involve the DA-MRF, then each constant in the kernel can be replaced by $h_\gamma(.)$. This function can be interpreted as in Equation 4.16. The value varies according to the difference of the center motion vector from the neighboring blocks. For our purpose, we select the first order neighborhood system for the model with DA-MRF.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) The Motion Field after the First Iteration

(b) The Motion Field after Tenth Iteration.

Figure 6.1: Progress of Motion Field during Iteration of the Algorithm for the Suzie Sequence.

After several iterations the values of PSNR and Entropy go into a stable state as shown in Figure 6.2. We notice that seven iterations are enough to get the stable motion field. For the dynamics of the values of PSNR and Entropy, they depend on the weight of the prior function and the type of image sequence.

The computation of d_x, d_y and d_t follows the finite differences of Horn and Schunck [36] in Section 3.3.3.1, so that all the results can be compared with our previous chapter that also uses the gradient based model.

6.4 Experimental Results

The same series of real image sequences from the previous chapters and synthetic image sequences are used for our experiments. We have three types of experiments: *Standard Block Motion Model*, *Block Motion Model using Smoothness Prior Function* and *Block Motion Model Using Discontinuity Adaptive Function*. For the first experiment, two motion estimation algorithms with different block-size are executed. The result will be used as a comparison tool for our proposed model. The second and the third experiments are for the 8×8 block-size.

Besides those experiments, the experiments on the weight and the parameter of

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

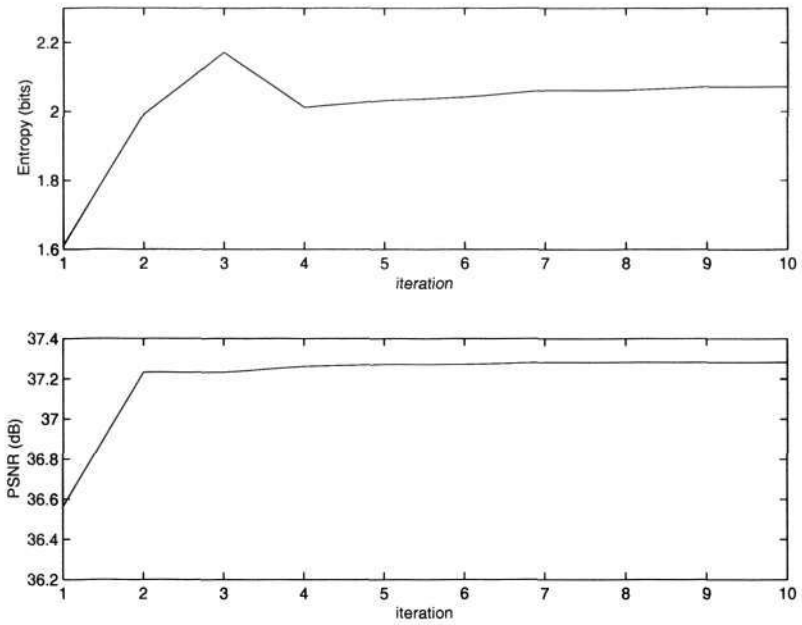


Figure 6.2: Progress of PSNR and Entropy vs Iteration.

the discontinuity function are conducted. The optimal values of the weights and the discontinuity parameters are used for the second and third experiments.

6.4.1 Strategies and Summaries

The first part of experiments, the standard block motion model, uses 16×16 block-size, and the second uses 8×8 block-size. In addition to these block-sizes, the overlap model is implemented to show its inadequacy. Those results are used for comparisons. We can comment again as in the previous chapter that the figures show that estimation using smaller block-size produces many wrong motion vectors.

In summary, the motion field from the block motion model using smoothness prior in the second experiment illustrated the effect of the smoothness prior. It showed the desirable results. The motion field from the block motion model using discontinuity adaptive function in the last experiment is different from the motion field of the second experiment at the area where the action is active to preserve the discontinuity.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

6.4.2 Parameter Setting and the Influence of Explicit Smoothness

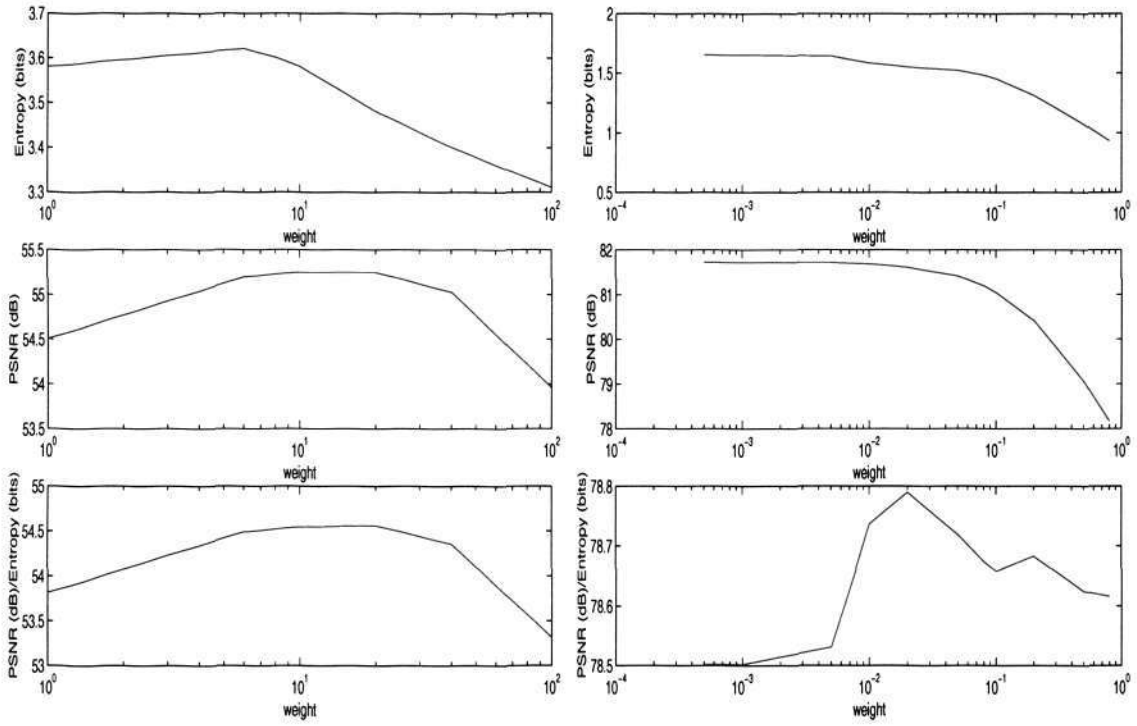
The algorithm using smoothness prior function needs a weight, λ, to balance the likelihood and prior energy functions. If the weight is zero, then the model is without the smoothness prior function. If we increase the weight, the PSNR value drops and the Entropy value also drops. The low value of Entropy is desirable but the low value of PSNR is undesirable.

To get the desired highest PSNR and the lowest Entropy, the compromise between PSNR and Entropy value is represented by Equation 5.4 where κ is to modify the strength of influence from the entropy. Usually κ is one, however in certain cases, such as the Suzie and Foreman sequences, we need to set the κ to reduce or to increase the influence of Entropy value.

The experimental results with varied weight are shown in Figure 6.3. The Peak Sequence is a whole frame movement, the *ratio* plotting suggests 10 as the weight value. The Yosemite Sequence has many small movement vectors that have slightly different values than its neighbors, the ratio plotting suggests 0.01 as the weight value. It has a tiny peak. The Suzie Sequence is a sequence with many smooth movements all around the frame. The *ratio* plotting suggests a weight value in the hundreds range. The *ratio* plotting for the Foreman Sequence has a peak value if the weight is a thousand. The *ratio* plotting for the Salesman Sequence with small moving objects suggests a weight in the hundreds range. The *ratio* plotting for the Taxi Sequence also suggests a weight in the hundreds range. The existence of the peak from the ratio diagram shows a rough estimation of optimal weight to keep the PSNR high and entropy low. Besides that, it shows that implicit smoothness can change the motion vector without sacrificing the PSNR value too much.

For the third experiment, the *discontinuity parameter,* γ, also plays an important role to control the smoothness of the motion field. Figure 6.5 displays the experiment with the parameter from 10^{-3} to 10^{3}. In the Suzie sequence, the ratio suggests a parameter of ten. In the Foreman sequence, one is the best ratio. In the Salesman sequence, ten is the best ratio. In the Taxi sequence, one is the best ratio. From

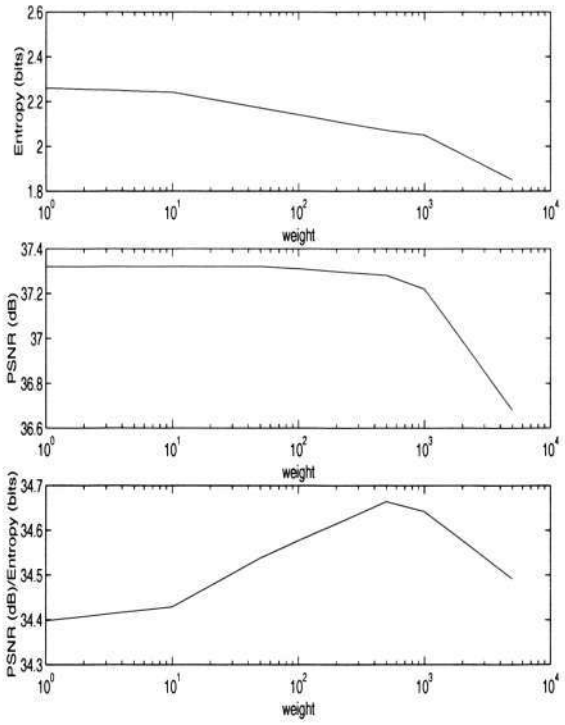
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



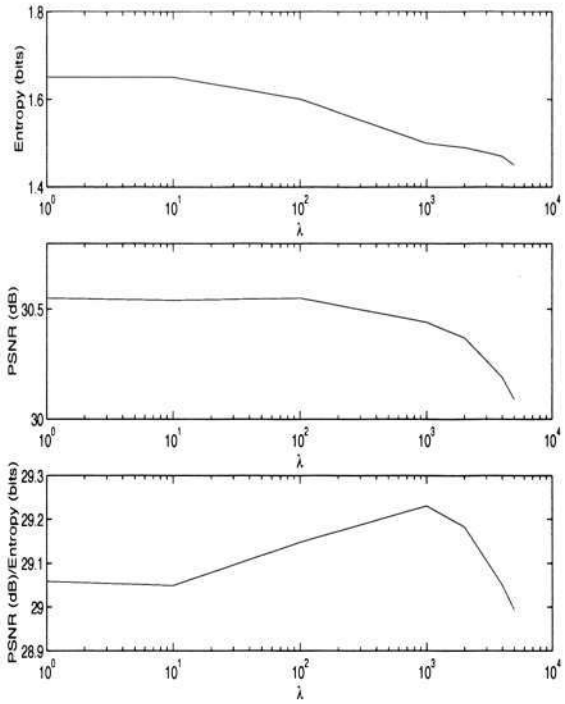
(a) Peak Sequence (b) Yosemite Sequence

Figure 6.3: PSNR, Entropy and PSNR/Entropy vs Weight.

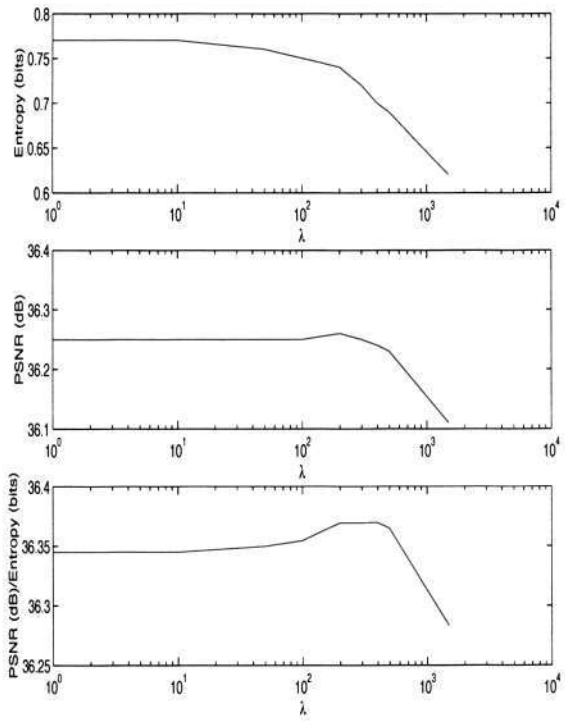
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



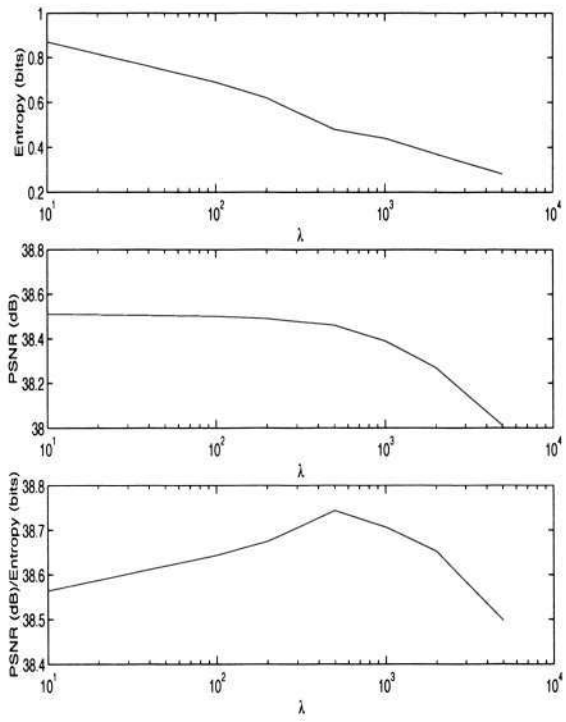
(a) Suzie Sequence



(b) Foreman Sequence



(c) Salesman Sequence



(d) Taxi Sequence

Figure 6.4: PSNR, Entropy and PSNR/Entropy vs Weight (cont.).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

the experiments we can roughly suggest one is the best choice.

6.4.3 Results

Motion fields from the results of different algorithms are shown here. The PSNR and Entropy values can be checked in Section 6.4.2. The results here are chosen from the weight and the interaction function parameters that give the highest value of the PSNR/Entropy.

The result of Lucas-Kanade algorithm that uses implicit smoothness has no control on the value of the neighborhood vectors. However, our approach using smoothness prior function has a control on the neighborhood values. So, the result will be smooth to the surrounding values.

6.4.3.1 Peak Sequence

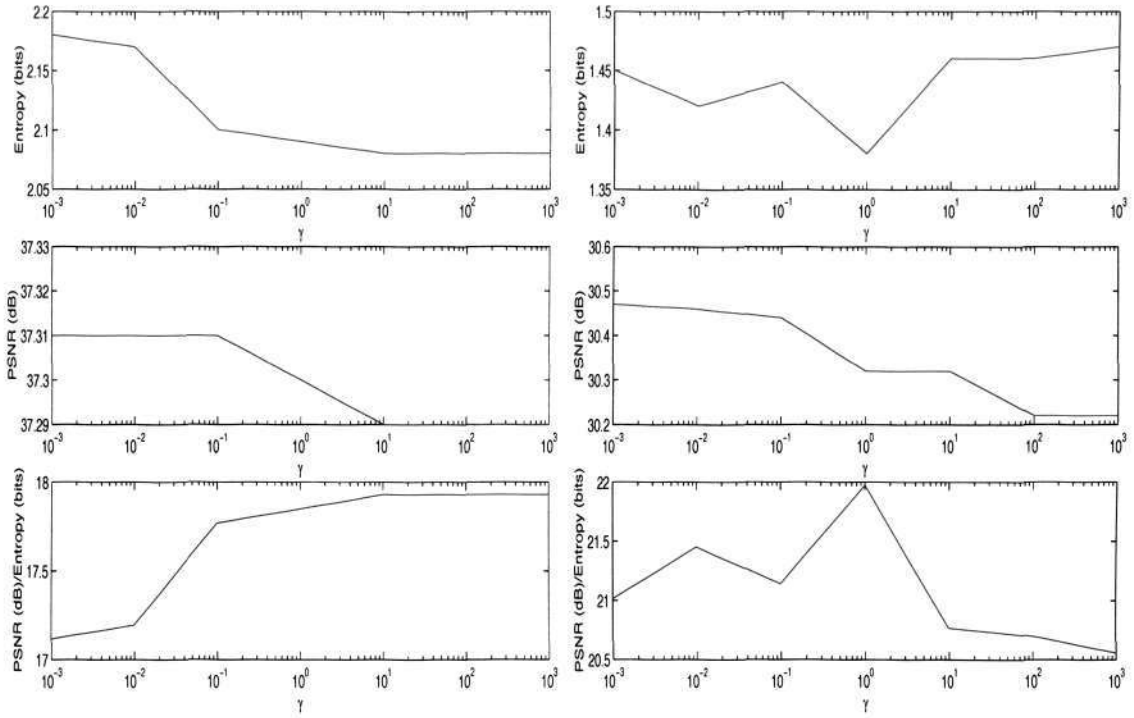
The peak sequence (Figure 6.6) is a synthetic image sequence. Some blocks have zero motion vectors. In the figures given, these blocks are not shown. In the experiment using a 16×16 block-size, the correct motion vectors are in the textured area. But at the edges of the peaks, the motion estimator gives a small motion vector as there is not much error to drive the estimator to the correct motion vector. With smaller block-size, some blocks cannot get the correct motion vectors. The overlap block also cannot produce a better motion field. These results show the need of the smoothness prior function.

The result of estimation using smoothness or discontinuity adaptive prior function shows the influence of the function on the motion fields at the black background. If the iteration is long enough, the whole frame would have the motion vector. The PSNR value will rise because the edge will have a better motion vector. However, the PSNR value for this sequence is already very high. Table 6.1 shows the PSNR, entropy and the average angular error and its standard deviation.

6.4.3.2 Yosemite Sequence

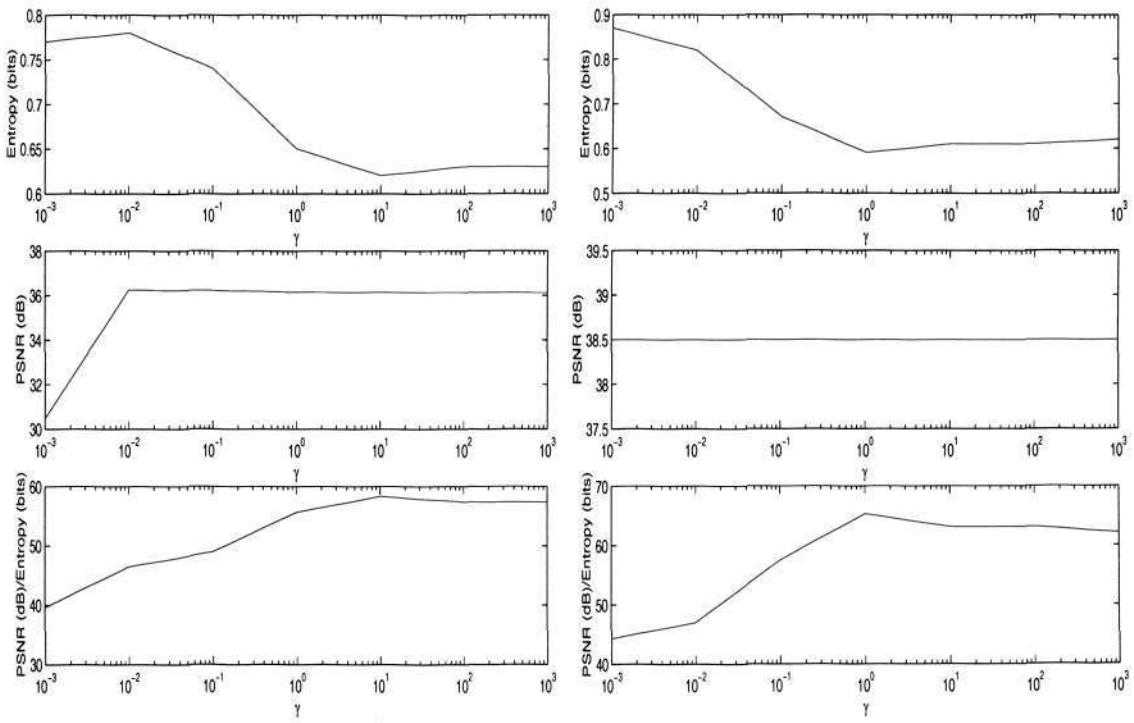
Figure 6.7 and Table 6.2 show the experimental results using Yosemite sequence. The angular error is reduced by using prior model. This shows the advantage of

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



(a) Suzie Sequence

(b) Foreman Sequence

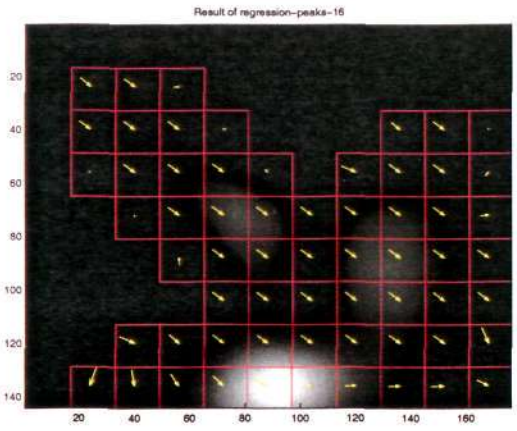


(c) Salesman Sequence

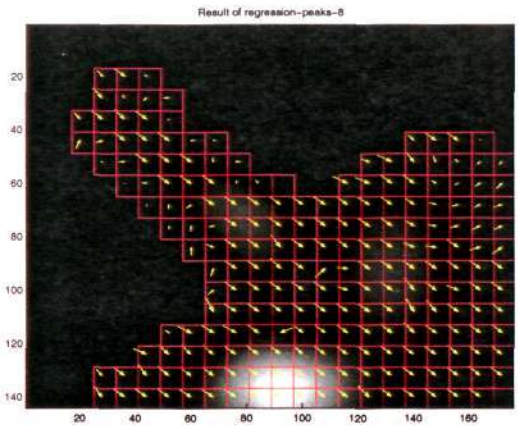
(d) Taxi Sequence

Figure 6.5: PSNR, Entropy and PSNR/Entropy vs Parameter of Discontinuity.

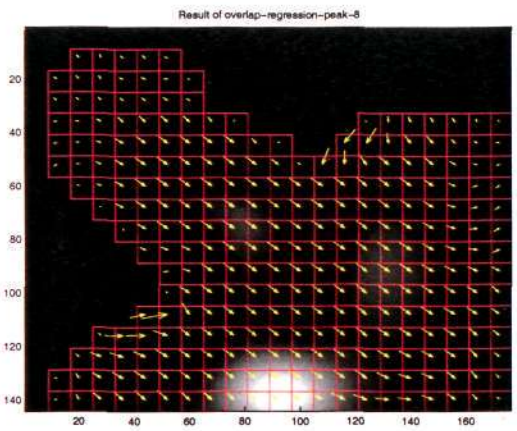
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



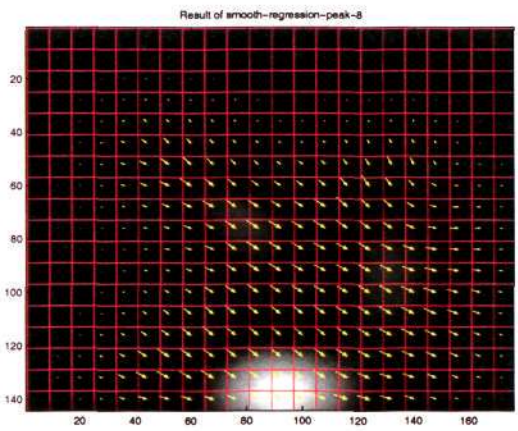
(a) Block Motion for Block-Size of 16 by 16



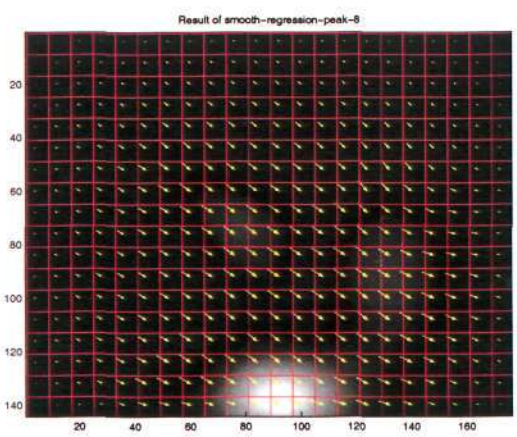
(b) Block Motion for Block-Size of 8 by 8



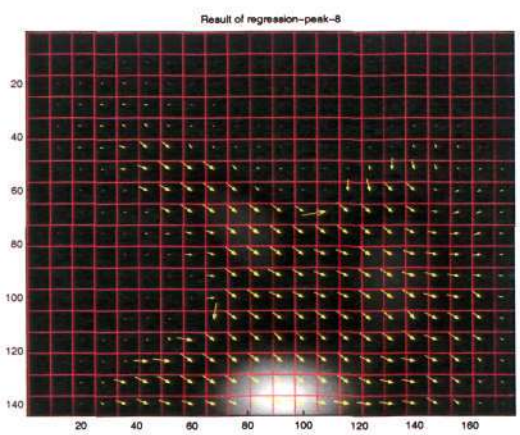
(c) Overlap Block Motion



(d) Using Smoothness Prior Function with 10 iterations.



(e) Using Smoothness Prior function with 1000 iterations.



(f) Using Discontinuity Adaptive Prior Function

Figure 6.6: Results of Block Motion Model for Optical Flow Equation with the Peak Sequence.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

SEQUENCE	PSNR	ENTROPY	AAE/SD
	dB	bits	
Standard 16	54.57	3.51	12.50°/23.88°
Standard 8	48.40	3.70	16.32°/27.81°
Overlap	55.92	3.68	23.94°/24.10°
Prior_10	54.50	3.79	34.76°/27.77°
Discontinuity	52.35	3.30	52.16°/26.35°

Table 6.1: PSNR and Entropy for the Peak Sequence. (Notes: Average Angular Error (AAE), Standard Deviation (SD))

SEQUENCE	PSNR	ENTROPY	AAE/SD
	dB	bits	
Standard 16	81.25	1.42	53.15°/42.39°
Standard 8	80.95	1.62	35.59°/30.92°
Overlap	81.53	1.51	36.29°/30.74°
Prior_10	78.16	0.93	33.92°/27.73°
Discontinuity	78.88	0.93	34.09°/28.31°
Prior_1000	79.00	1.25	37.26°/28.41°

Table 6.2: PSNR and Entropy for the Yosemite Sequence. (Notes: Average Angular Error (AAE), Standard Deviation (SD))

using smoothness prior function.

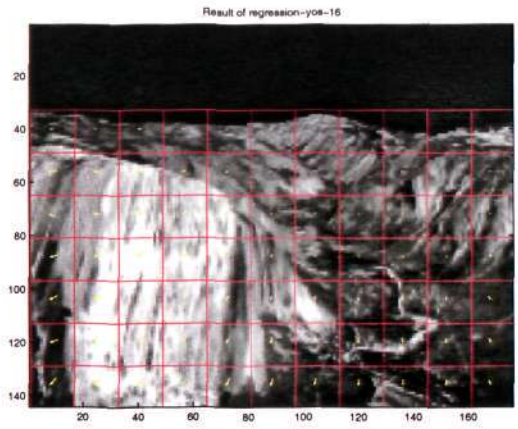
6.4.3.3 Suzie Sequence

This sequence (Figure 6.8) has texture, small motion and discontinuity. We notice several wrong motion vectors on the flat cheek. In overlap block motion model, the motion vector cannot be improved at that location. Using the smoothness prior function, we can expect an improvement at that location. Using the discontinuity adaptive prior function, we get a better motion field near the phone-neck area, where vector discontinuity value is high.

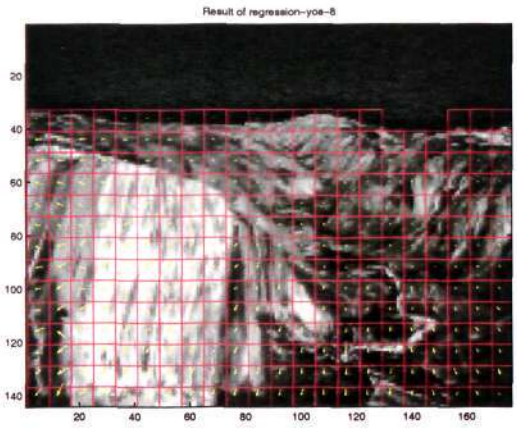
SEQUENCE	PSNR	ENTROPY
	dB	(bits)
Standard 16	36.84	2.12
Standard 8	37.32	2.26
Overlap	35.79	1.85
Prior	36.68	1.85
Discontinuity	34.42	1.34

Table 6.3: PSNR and Entropy for Suzie Sequence using Standard Block Model.

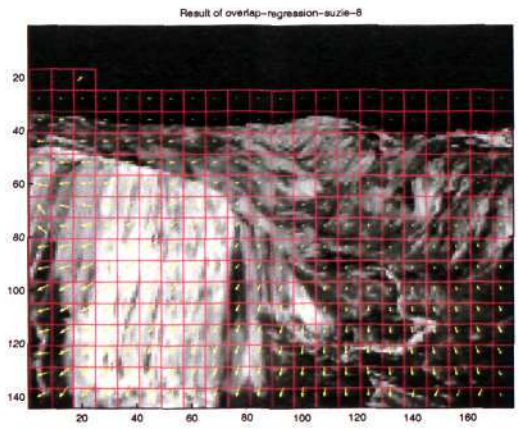
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



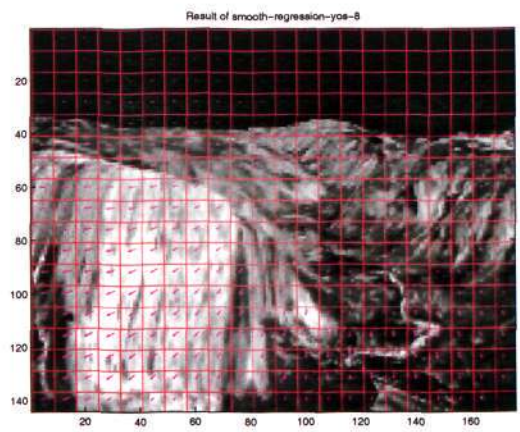
(a) Block Motion for Block-Size of 16 by 16



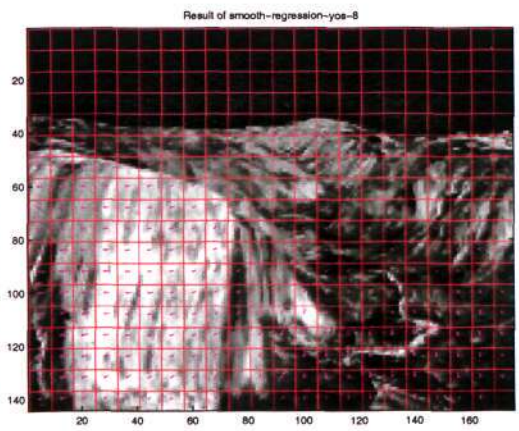
(b) Block Motion for Block-Size of 8 by 8



(c) Overlap Block Motion



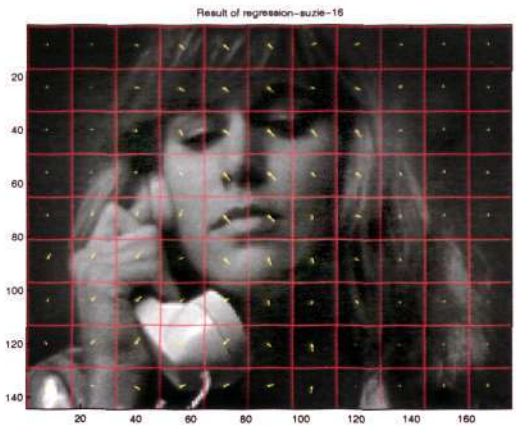
(d) Using Smoothness Prior Function



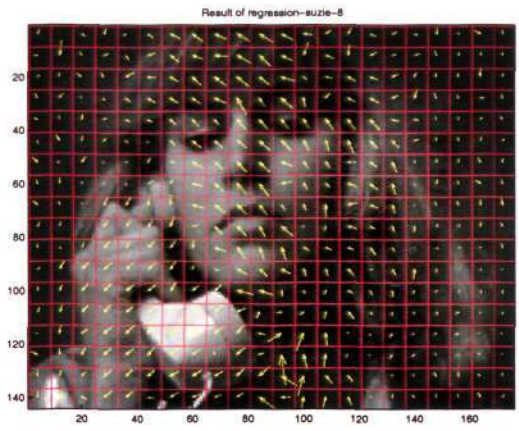
(e) Using Discontinuity Adaptive Prior Function

Figure 6.7: Results of Block Motion Model for Optical Flow Equation with the Yosemite Sequence.

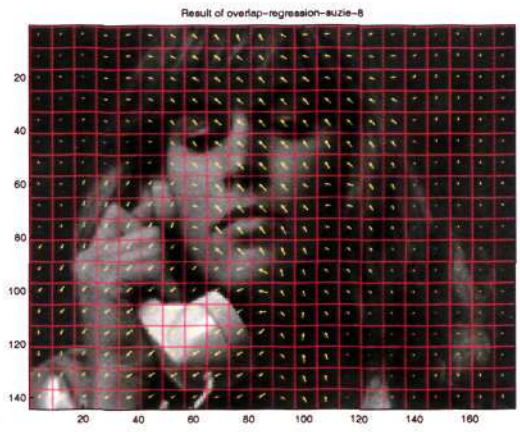
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



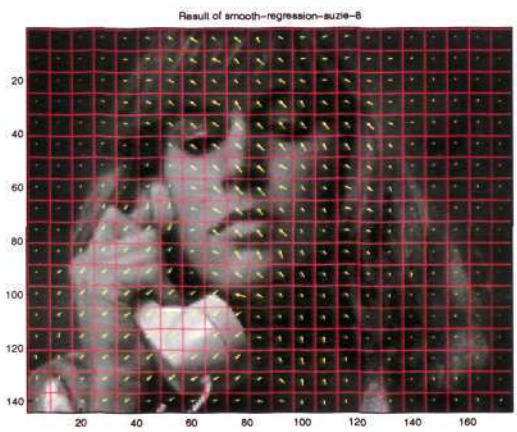
(a) Block Motion for Block-Size of 16 by 16



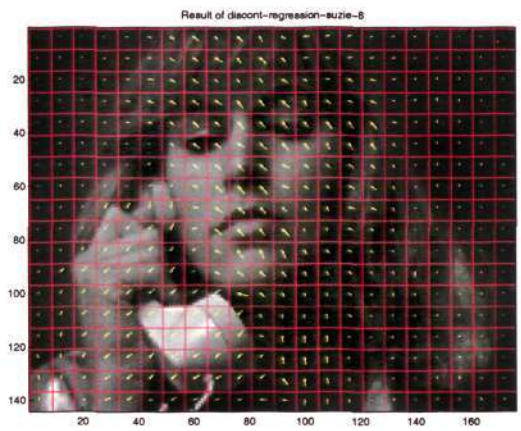
(b) Block Motion for Block-Size of 8 by 8



(c) Overlap Block Motion



(d) Using Smoothness Prior Function



(e) Using Discontinuity Adaptive Prior Function

Figure 6.8: Results of Block Motion Model for Optical Flow Equation with the Suzie Sequence.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

SEQUENCE	PSNR	ENTROPY
	dB	(bits)
Standard 16	29.82	1.75
Standard 8	30.55	1.65
Overlap	29.43	1.65
Prior	30.37	1.49
Discontinuity	30.13	1.29

Table 6.4: PSNR and Entropy for Foreman Sequence using Standard Block Model.

SEQUENCE	PSNR	Entropy
	dB	(bits)
Standard 16	34.47	0.50
Standard 8	36.25	0.77
Overlap	34.15	0.37
Prior	36.11	0.62
Discontinuity	35.88	0.54

Table 6.5: PSNR and Entropy for Salesman Sequence using Standard Block Model.

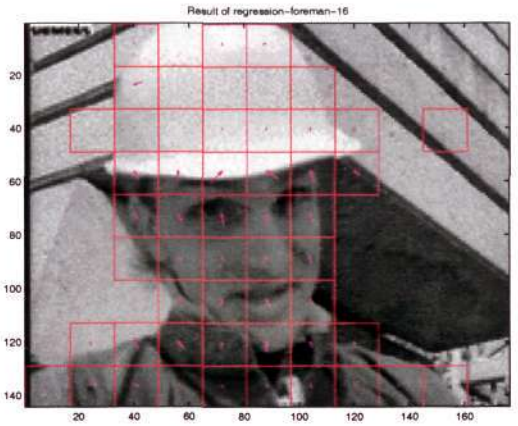
6.4.3.4 Foreman Sequence

In this sequence (Figure 6.9), we notice wrong vectors at the edge of the helmet with small block-size. The overlap block motion model cannot improve the result. Using the smoothness prior function, the result is improved. Using this technique, the white area of the helmet that is supposed to be a moving object cannot be detected. We can compare this to the result from the previous chapter. This is caused by the lack of gradient on the white area of helmet.

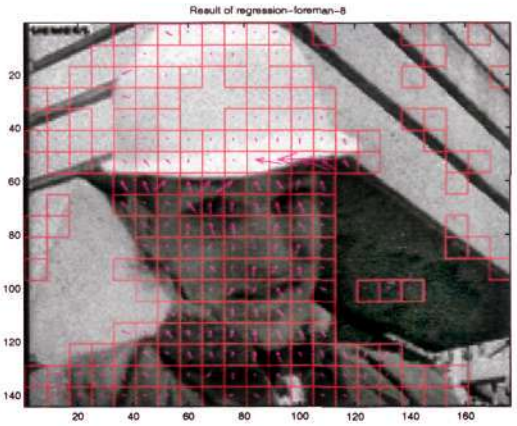
6.4.3.5 Salesman Sequence

In Figure 6.10, the Salesman Sequence has large motion for the small moving object (the box). We notice several wrong motion vectors at the box and the salesman's left hand. By the using smoothness prior function, we expect the result to improve. From the diagram of PSNR vs weight, we notice a small increase of PSNR with the use of the smoothness prior function.

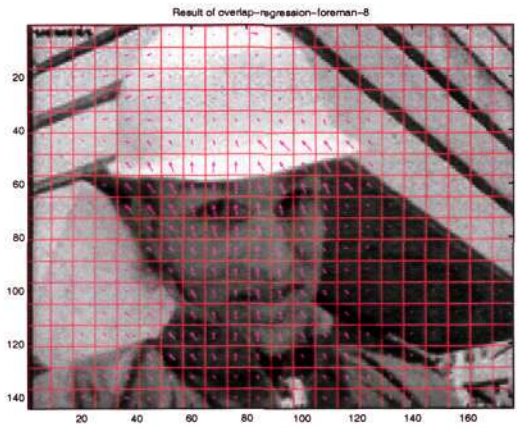
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



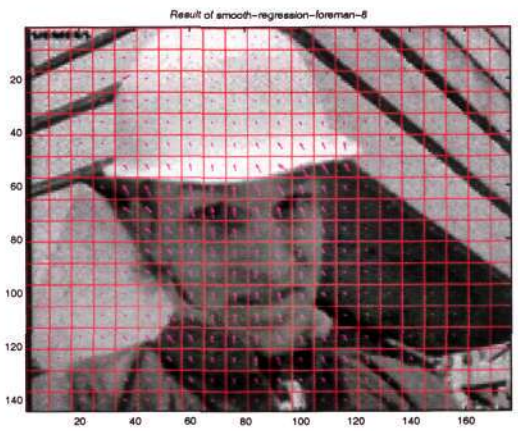
(a) Block Motion for Block-Size of 16 by 16



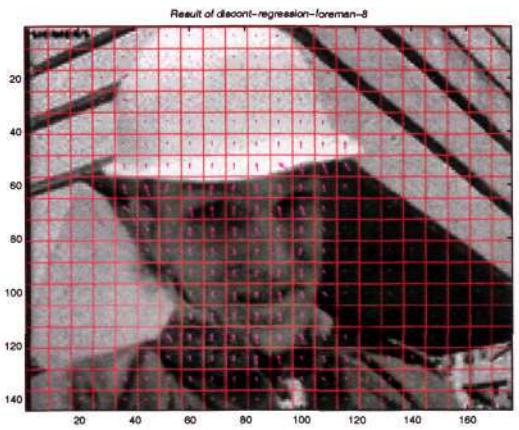
(b) Block Motion for Block-Size of 8 by 8



(c) Overlap Block Motion



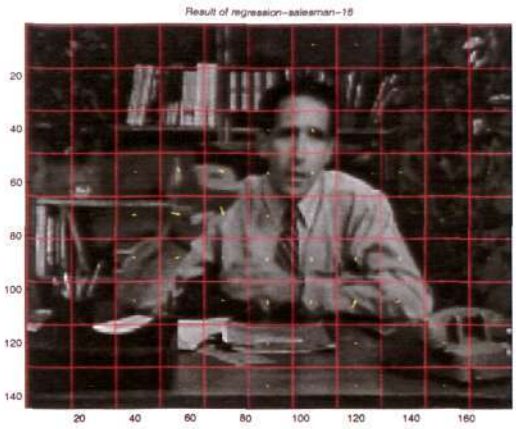
(d) Using Smoothness Prior Function



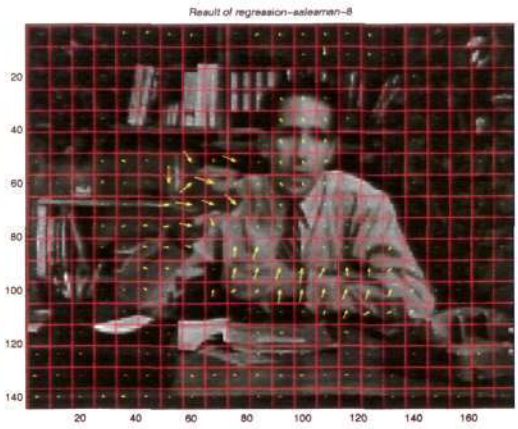
(e) Using Discontinuity Smoothness Prior Function

Figure 6.9: Result of Block Motion Model for Optical Flow Equation with the Foreman Sequence.

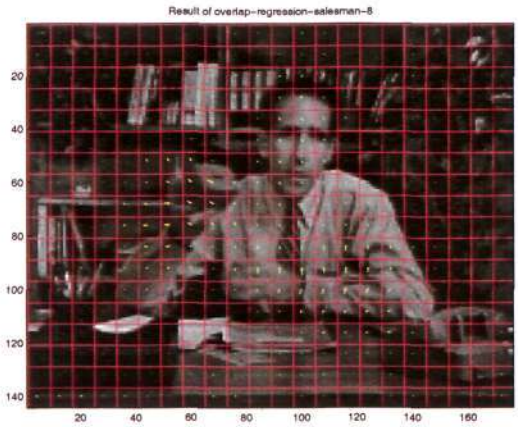
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



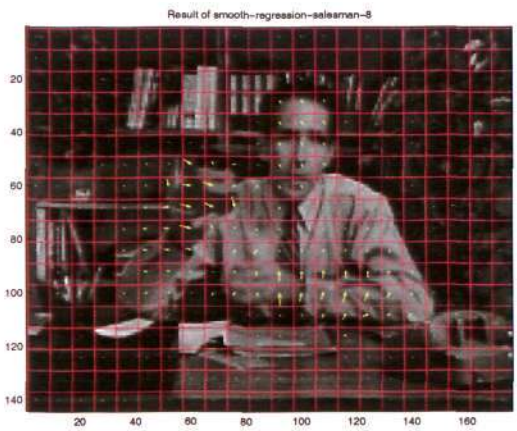
(a) Block Motion for Block-Size of 16 by 16



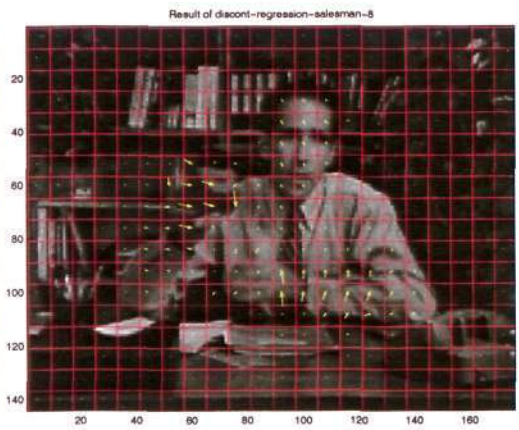
(b) Block Motion for Block-Size of 8 by 8



(c) Overlap Block Motion



(d) Using Smoothness Prior Function



(e) Using Discontinuity Adaptive Prior Function

Figure 6.10: Results of Block Motion Model for Optical Flow Equation with the Salesman Sequence.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

SEQUENCE	PSNR dB	Entropy (bits)
Standard 16	37.62	0.44
Standard 8	38.43	0.88
Overlap	37.32	0.33
Prior	38.01	0.28
Discontinuity	37.57	0.12

Table 6.6: PSNR and Entropy for Hamburg Taxi Sequence using Standard Block Model.

6.4.3.6 Hamburg Taxi Sequence

Figure 6.11 shows the Taxi Sequence that has several small objects with small movements. With this sequence we can show the advantage of using small block-size. Small block-size gives more vectors, in effect it gives flexibility and higher PSNR as shown in Table 6.6.

6.5 Conclusion

In this chapter, the smoothness prior function controls the smoothness between blocks in a frame. The smoothness prior function is added to the optical flow equation that is over the block of pixels. The energy function or the error definition for each block is solved one by one until the whole frame is completed. It is iterated a few times until it reaches the convergence condition.

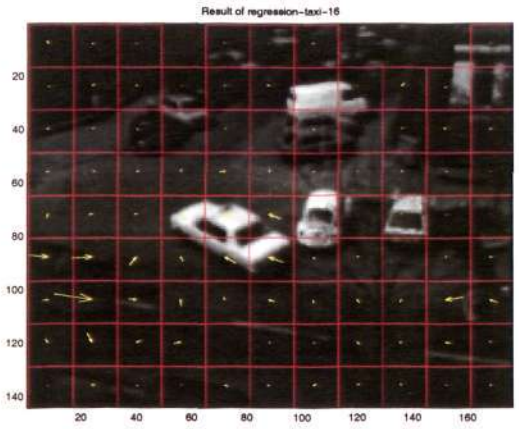
The experiments have demonstrated that the algorithm is stable and consistent. The predictable result is achieved if we set the weight and the discontinuity parameter within the allowable range.

The contribution of the work in this chapter are as following:

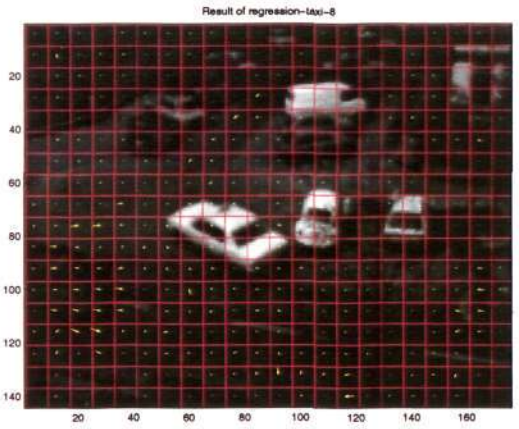
- We introduce smoothness prior function to the block motion model that is proposed by Lucas and Kanade. It is shown experimentally that this gives better results than the original design.

- The addition of the smoothness prior function is proven to be useful as evident from the present of a peak in the PSNR/Entropy versus weight diagram.

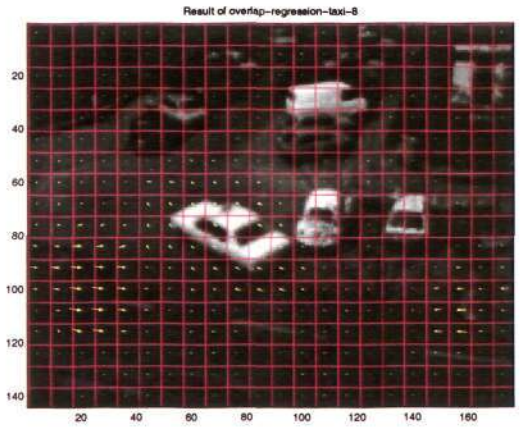
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



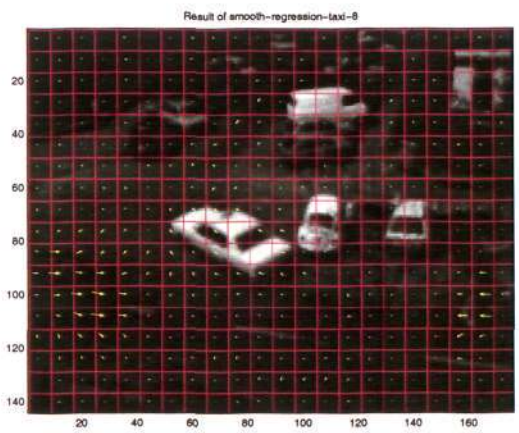
(a) Block Motion for Block-size of 16 by 16



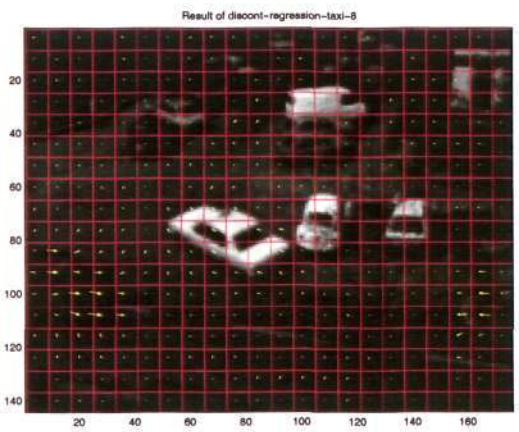
(b) Block Motion for Block-size of 8 by 8



(c) Overlap Block Motion



(d) Using Smoothness Prior Function



(e) Using Discontinuity Adaptive Prior Function

Figure 6.11: Results of Block Motion Model for Optical Flow Equation with the Hamburg Taxi Sequence.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

- Discontinuity adaptive function is successfully implemented.ub

The experiments are made in parallel with the previous chapter. However, we do not design re-sizing and multiresolution as in the previous chapter. The main reason is that the initialization values do not play too much a part in this design. Though it may help to improve the results of the Foreman Sequence by modifying the design and giving a different initialization value.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 7

Conclusions

With GMRF modeling tools and Bayesian framework, we have introduced many of the available motion estimators in a unified way. This framework is also selected to model all the three motion estimation methods proposed in this thesis. In this framework, each method consists of three parts: likelihood energy, prior energy and the optimization of the total energy (posterior energy). All of the three proposed methods try to estimate the "true" motion field, which is useful in application such as motion segmentation and motion in-betweening.

In Chapter 4, we deal with the motion discontinuity adaptive function. This function is the prior energy for motion estimation algorithms under the GMRF framework. This approach has not been formally reported in our existing publications. The motion estimation is based on the optical flow equation and the explicit smoothness constraint. A detailed investigation shows that the DA prior is useful in maintaining the motion discontinuities. Hence it is also used in Chapters 5 and 6. In Chapters 5 and 6, motion estimation is based on the block motion model, which assumes that an implicit smoothness constraint has automatically been applied. Our interest is to reveal that, if we further apply explicit smoothness constraint, a better motion field could be obtained. To evaluate the quality of motion fields, we calculate a measure of $PSNR/(Entropy^{\kappa})$, where κ is the influence strength of the entropy. The measure encapsulates the requirement of higher PSNR and lower Entropy. Further investigation of the measure will be useful for automatic parameter setting. Our study shows that the coexistence of the implicit and explicit constraints

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chap.	Motion repr.	Motion field	Motion vector	Constraints
4	optical flow	dense	real number	explicit
5	block matching	sparse	integer	implicit and explicit
6	optical flow	sparse	real number	implicit and explicit

Table 7.1: Characteristics of Three Proposed Methods. (Chap.= Chapter, repr.= representation)

gives a better quality measure in both block matching based and optical flow based models. This also gives us more flexibility to choose the motion field with higher-PSNR-higher-entropy or lower-PSNR-lower-entropy, according to different needs in different applications.

The three proposed methods can be further characterized as in Table 7.1. The novelties of each method will be further explained in the following section. This is followed by a plan in the future works.

7.1 Thesis Contributions

In Chapter 4, we have implemented discontinuity adaptive (DA) MRF for the optical flow problem. This is a simple but novel extension to the modeling using GMRF. This model is comparable to robust technique in handling the motion discontinuity.

The derivation of optical flow equation follows the Horn-Schunck's approach. But our modified Horn-Schunck's iterative scheme has the capability of preserving the discontinuity of the motion field. We have compared the scheme with the original iterative scheme and our modified algorithm has shown improvement in PSNR and subjective judgment.

Block matching, usually used for compression technologies, has a mathematical sound basis. In Chapter 5, block matching is designed within the MRF framework, which offers a new way of thinking. In such a scheme, the conventional block matching algorithm using a full search strategy is only seen as a maximum likelihood estimator. No explicit prior knowledge is present in this strategy. Moreover, when we reduce the block-size, the motion field is found to be in chaotic condition. If we implement the explicit prior knowledge, we could improve the quality of the motion field. This is verified by our experiments.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

We have proposed several procedures that lead to the modification of the standard block matching algorithm. Firstly likelihood values for each candidate labels are kept in the memory. The reason is that on every iteration, the likelihood value of the candidate label is needed. Secondly likelihood values are computed for the smallest block-size from which the likelihood values for bigger block-sizes are easy to synthesize. If those procedures are available, the iterative scheme can be efficiently implemented.

As our algorithm uses a local minimization technique, a bigger block-size is usually fed into the algorithm first to generate the initialization values. This initialization step improves the result for the smaller block-size.

Synthetic and real image sequences are tested. The experiments show that better result is achieved if we compare them with the conventional block matching algorithm. We have further flexibility to balance the PSNR and Entropy.

Chapter 6 is parallel to Chapter 5. The difference is that the likelihood function is a block of optical flow equations instead of block differences in this algorithm. We add the explicit smoothness constraint into the equations. Using this approach result could be smoother according to the weights of the explicit smoothness. Some iterations are needed to update the motion fields.

Similar to Chapter 5, the algorithm is tested with synthesized and real image sequences. The experiments show that better results are achieved.

7.2 Topics for Future Research

7.2.1 Comparison of the Algorithms

In the literature, Bouthemy [66] uses the affine robust function. However, the software given is a type of global motion. This prevents us from using it as a comparison. We need to reprogram them for the same purpose, the block motion estimation.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

7.2.2 Fast Block Matching using Smoothness Prior Constraints

Many of the available fast block matching schemes take the mathematical strategy or reduce the searching area to decrease the computation tasks. If we can combine our proposal with the available fast block matching algorithm, we will have fast block matching using smoothness prior constraint.

7.2.3 Multiscale Motion Estimation

Another area using GMRF and the smoothness constraints is developed by the researchers from control engineering [78–80]. The result is found to be appealing. We have tested and used their multiscale MRF with smoothness constraint for our matching problem [81]. We have also implemented the multiscale framework to solve an optical flow problem. However, we cannot embed the result in this thesis because it is still in infancy.

7.2.4 Ratio Measure

The influence of likelihood and the prior function to the energy function is adjusted by a weight constant. The value is selected where the highest PSNR/Entropy ratio is achieved. However, the degree of entropy influence and the interpretation of ratio value need further research.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Appendix A

Basic Iterative Methods

This appendix explains the basic iterative method. The explanation is summarized from [82].

A.1 Point Jacobi Method

In matrix notation, the linear equations are written

$$A\mathbf{x} = \mathbf{k}, \tag{A.1.1}$$

where $\mathbf{k}$ is a given column vector. The solution vector $\mathbf{x}$ exists and it is unique if and only if A is also nonsingular. The solution vector is

$$\mathbf{x} = A^{-1}\mathbf{k}. \tag{A.1.2}$$

Instead of solving the equation using a direct inverse matrix, we can also solve the equation in an iterative way. Matrix A is rearranged as

$$A = D - E - F, \tag{A.1.3}$$

where $D = \mathtt{diag}\{a_{1,1}, a_{2,2}, a_{3,3}, \ldots, a_{n,n}\}$, and E is lower triangular and F is upper triangular $n \times n$ matrices.

The A.1.1 can be written as

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

$$D\mathbf{x} = (E + F)\mathbf{x} + \mathbf{k}. \tag{A.1.4}$$

The iterative method is

$$D\mathbf{x}^{(t+1)} = (E + F)\mathbf{x}^{(t)} + \mathbf{k}, \qquad t \geq 0, \tag{A.1.5}$$

and since D is a nonsingular matrix, it becomes

$$\mathbf{x}^{(t+1)} = D^{-1}(E + F)\mathbf{x}^{(t)} + D^{-1}\mathbf{k}, \qquad t \geq 0. \tag{A.1.6}$$

A.2 Point Gauss-Seidel Method

The use of the latest estimates $x_i^{(t)}$ to estimate $x_i^{(t+1)}$, Equation A.1.5 becomes

$$(D - E)\mathbf{x}^{(t+1)} = F\mathbf{x}^{(t)} + \mathbf{k}, \qquad t \geq 0, \tag{A.2.7}$$

and as $D-E$ is a nonsingular lower triangular matrix, we can write A.2.7 equivalently as

$$\mathbf{x}^{(t+1)} = (D - E)^{-1}F\mathbf{x}^{(t+1)} + (D - E)^{-1}\mathbf{k}, \qquad t \geq 0. \tag{A.2.8}$$

A.3 Successive Overrelaxation Iterative Method

This is a method which is closely related to the point Gauss-Seidel iterative method. Starting with the non matrix Gauss-Seidel method, we define the components of the auxiliary vector iterates $\widetilde{x}^{(t)}$ from

$$a_{i,i}\widetilde{x}_i^{(t+1)} = -\sum_{j=1}^{i-1} a_{i,j}x_j^{(t+1)} - \sum_{j=i+1}^{n} a_{i,j}x_j^{(t)} + k_i, \tag{A.3.9}$$

with

$$1 \leq i \leq n,\ t \geq 0.$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

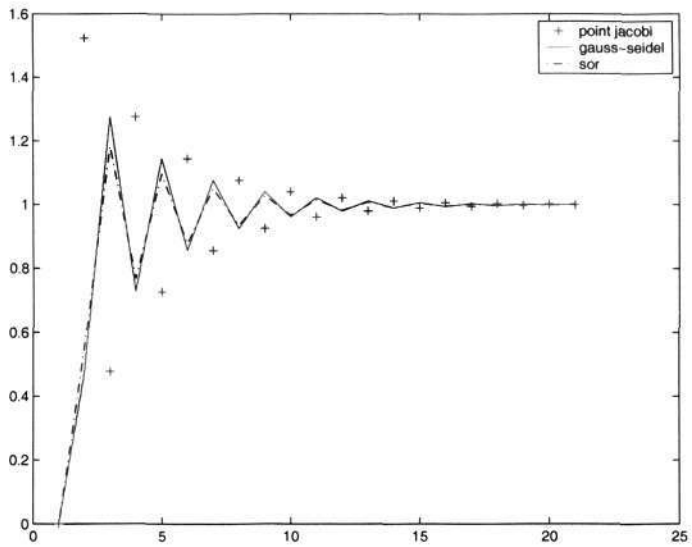


Figure A.1: Iteration Process for Three Methods

The actual components $x_i^{(t+1)}$ of this iterative method are then defined from

$$x_i^{(t+1)} = x_i^{(t)} + \omega\{\widetilde{x}_i^{(t+1)} - x_i^{(t)}\}. \tag{A.3.10}$$

The variable ω is called the *relaxation factor*. The value of this variable should be $0 \leq \omega \leq 1$. When $\omega > 1$, it corresponds to overrelaxation and $\omega < 1$, it corresponds to *underrelaxation*. The matrix form is

$$(D - \omega E)\mathbf{x}^{(t+1)} = \{(1-\omega) + \omega F\}\mathbf{x}^{(t)} + \mathbf{k}, \qquad t \geq 0, \tag{A.3.11}$$

and as $D - \omega E$ is nonsingular for any choice of ω, then with $L \equiv D^{-1}E$, $U \equiv D^{-1}F$, this takes the form

$$\mathbf{x}^{(t+1)} = (I - \omega L)^{-1}\{(1-\omega)I + \omega U\}\mathbf{x}^{(t)} + \omega(I - \omega L)^{-1}D^{-1}\mathbf{k}. \tag{A.3.12}$$

The example of the iteration process for the three methods using $\omega = 0.7$ is shown in Figure A.1.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Appendix B

Adaptive MRF

There are two topics: regularization model and the guideline for selecting g. This appendix is mostly summarized from the third chapter of [6].

B.1 Regularization and Discontinuities

A regularization model is a special case of the MRF model. Whereas this MRF model can encode prior constraints other than the smoothness, an analytic regularization model provides a convenient platform for the study of smoothness priors. and allows arbitrary neighborhood systems other than the nearest ones. It is because of the close relationship between smoothness and analytical continuity.

When we have an energy function in the regularization framework

$$E(f) \doteq U(f|d) = U(d|f) + U(f) \tag{B.1.1}$$

with $U(d|f)$ is the *closeness* term and $U(f)$ is the *smoothness* term, it corresponds to the energies in the posterior, the likelihood and the prior of Gibbs distribution of an MRF.

A smoothness term $U(f)$ is called *a regularizer* when it penalizes the irregularities according to a priori smoothness constraint encoded in it. Generally, it is defined as

$$U(f) \doteq \sum_{n=1}^{N} U_n(f) = \sum_{n=1}^{N} \lambda_n \int_a^b g(f^{(n)}(i))di \tag{B.1.2}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

where $U_n(f)$ is the n^{th} order of regularizer, N is the highest order and $\lambda \geq 0$ is a weighting factor. A *potential function* $g(f^{(n)}(i))$ is the penalty against irregularity in $f^{(n)}(i)$. This function corresponds to prior clique potentials in MRF models.

Standard regularizer takes the pure quadratic form $g_q(\eta) = \eta^2$. The consequence of this form is the more irregular $f^{(n)}$, the larger $|f^{(n)}|$ and the larger the potential $g(f^{(n)})$ contributed to $U_n(f)$. The general form of this standard regularizer is

$$U_n(f, \omega_n) = \lambda_n \int_a^b w_n(i)[f^{(n)}(i)]^2 di \tag{B.1.3}$$

where $\omega_n(i)$ are pre-specified non-negative continuous functions. If the function $\omega_n(i)$ can be pre-specified that $\omega_n(i) = 0$ at (i) where $f^{(n)}(i)$ is infinite, then oversmoothing will not happened. A line process is introduced to take the task of this $\omega_n(i)$.

The labeling problem regularizer with line process is

$$U(f, l) = \lambda \sum_{i=2}^{m} [f_i - f_{i-1}]^2 [1 - l_i] + \lambda\alpha \sum_{i=2}^{m} l_i \tag{B.1.4}$$

where l is the line field, related to $l_i = 1 - \omega_i$, with $l_i = 1$ is the on-state and $l_i = 0$ is the off-state, λ are the weighting function between the two function and $\lambda\alpha$ penalized the on-state line process. The line process variable can be set based on a certain threshold proportional to $|f_i - f_{i-1}|$.

Blake and Zisserman [71] convert the line process into function containing only real variables

$$U(f) = \lambda \sum_{i=2}^{m} g_\alpha(f_i - f_{i-1}) \tag{B.1.5}$$

where the truncated quadratic function is $g_\alpha(\eta) = \min\{\eta^2, \alpha\}$. Furthermore, Blake and Zisserman introduces a parameter p into this truncated quadratic function to control the convexity of E. The equation becomes $g_\alpha^p(\eta)$ with p varies from 1 to 0.

Li proposes a continuous adaptive regularizer model, the line field is not used, instead smoothness constraint has its effects as a function of the derivative magnitude. The larger the derivative the more it influence the interaction.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

B.2 Discontinuity Adaptive (DA) Model

A necessary condition for regularization model to be defined as DA is

$$\lim_{\eta\to\infty} |g'(\eta)| = \lim_{\eta\to\infty} |2\eta h(\eta)| = C \tag{B.2.6}$$

where $C \in [0, \infty)$ is a constant. When $C = 0$, this condition prohibits smoothing at discontinuities, while $C > 0$ allows limited smoothing.

The guideline for selecting g and f is defined here:

Definition. 1. An adaptive interaction function (AIF) h_γ parameterized by $\gamma(> 0)$ is a function that satisfies:

1. $h_\gamma \in C^1$. C is a continuity requirement.
2. $h_\gamma(\eta) = h_\gamma(-\eta)$
3. $h_\gamma(\eta) > 0$
4. $h'_\gamma(\eta) < 0$ $(\forall \eta > 0)$
5. $\lim_{\eta\to\infty} |\eta h_\gamma(\eta)| = C > \infty$

The continuity requirement (1) guarantees the twice differentiable of the integrand $u(f|d)$ with respect to f', a condition for the solution f to exist. The evenness (2) is assumed for spatially unbiased smoothing. The positive definiteness of (3) keeps the interaction positive. The monotony of (4) leads to decreasing interaction as the magnitude of the derivative increases. The bounded asymptotic property of (5) provides the adaptive discontinuity control.

Definition 2. The DA solution f is defined by Euler equation constrained by $h = h_\gamma \in \mathbf{H}_\gamma$.

Definition 3. The *adaptive potential function* (APF) is defined by

$$g_\gamma(\eta) = 2\int_0^\eta \eta' h_\gamma(\eta') d\eta \tag{B.2.7}$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Some properties of g_γ: one order higher than h_γ, it is even, its derivative function is odd, it is not necessary for $g_\gamma(\infty)$ to be bounded, it is strictly monotonically increasing.

For a given $g_\gamma(\eta)$, there exists a region of η within which the smoothing strength increases monotonically and the function g_γ is convex:

$$B_\gamma = \left\{\eta | g''_\gamma > 0\right\} \doteq (b_L, b_H) \tag{B.2.8}$$

This region is referred to as the *band of convexity*. The lower and upper bounds b_L, b_H correspond to the two extrema of $g'_\gamma(\eta)$, which can be obtained by solving $g'_\gamma(\eta) = 0$.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Appendix C

Measurement

In this thesis, we use PSNR and Entropy to evaluate the motion field. They are explained in this appendix.

C.1 Peak Signal-to-Noise Ratio

The peak signal-to-noise ratio (PSNR) equation is

$$\begin{aligned} PSNR_t &= 10 \log \frac{255^2}{MSE_t} \\ MSE_t &= \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} \left(d_{i,j,t} - d_{i+f_i,j+f_j,t+1} \right)^2}{I \times J} \end{aligned} \tag{C.1.1}$$

where $d_{i,j,t}$ is intensity at position (i, j) and time t. The displacement vector is (f_i, f_j) and I, J are the size of row and column of image respectively.

C.2 Entropy

The entropy equation is

$$H(u) = -\sum_{u} p(u) \log_2 p(u) \tag{C.2.2}$$

$$H(v) = -\sum_{v} p(v) \log_2 p(v) \tag{C.2.3}$$

$$H(u, v) = H(v) + H(v) \tag{C.2.4}$$

where u and v are all motion vectors in a frame.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

C.3 Average Angular Error

We use an angular measure of error [38]. The velocity $f = (u, v)$ may be written as space-time direction vector $(u, v, 1)$ in units of $(pixel, pixel, frame)$. It represents 3-d direction vector. The angular error between the correct velocity $\overrightarrow{f_c}$ and an estimate $\overrightarrow{f_e}$ is

$$\psi_E = \arccos\left(\overrightarrow{f_c} \cdot \overrightarrow{f_e}\right).$$

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Appendix D

Author's Publications

Main Author

Motion Estimation

1. **S. S. Tandjung**, T. S. Gunawan, and M. N. Chong. Combination Of Spatial and Multiscale Markov Random Field Modeling for Motion Estimation. *Int. Conf. on Control, Automation, Robotic and Vision*, 6th, Dec. 2000.

2. **S. S. Tandjung**, T. S. Gunawan, and M. N. Chong. Motion Estimation using Adaptive Block Size Observation Model and Efficient Multiscale Regularization. *IEEE Int. Conf. on Image Processing*, pages 566-569, vol. 2, Sept. 2000.

3. **S. S. Tandjung**, T. S. Gunawan, and M. N. Chong. Motion Estimation using Adaptive Matching and Multiscale Method. *SPIE Conf. on Visual Communication and Image Processing*, June 2000.

4. **S. S. Tandjung**, T. S. Gunawan, and M. N. Chong. Multiscale Motion Estimation. *IECI Conf. on Multimedia and Network*, pages 260-264, Mar. 2000.

5. **S. S. Tandjung**, M. N. Chong, and K. K. Wong. Adaptive Blocksize Observation Model for MRF Modeling of Motion Field. *IEEE Int. Conf. on Information, Communications and Signal Processing*, 2nd, Dec. 1999.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Motion Tracking System

1. **S. S. Tandjung**, T. Adiono, C. Machbub, and T. L. R. Mengko. Camera Object Tracking System. *IEEE Int. Conf. on Information, Communications and Signal Processing*, 1st, pages 1557-1561, vol. 3, Sept. 1997.

2. **S. S. Tandjung**, D. Astharini, , H. R. Pranoto, T. Adiono, C. Machbub and T. L. R. Mengko,. Real Time Visual Servoing System. *Int. Conf. on Microelectronics*, 9th, Oct. 1997.

Co-author

Parallel Implementation of Motion Estimation

1. T. S. Gunawan, **S. S. Tandjung**, and M. N. Chong. An Efficient Parallelization Scheme of Motion Estimation Algorithms on Myrinet-Connected Workstation. *Int. Conf. on Automation, Robotics, Control and Vision*, 6th, Dec. 2000.

2. T. S. Gunawan, **S. S. Tandjung**, and M. N. Chong. Parallel Implementation of Adaptive Bi-directional Motion Estimation based on Texture Analysis. *Int. Conf. on Automation, Robotics, Control and Vision*, 6th, Dec. 2000.

3. T. S. Gunawan, **S. S. Tandjung**, and M. N. Chong. Testing The Myrinet Reliability on The Cluster of Linux Systems. *First Myrinet User Group Conf.*, Sept. 2000.

4. T. S. Gunawan, **S. S. Tandjung**, and M. N. Chong. Performance of The Communication Layer with The Myrinet Gigabit LAN for HDTV Images on Pentium-Linux Cluster using MPI. *Proc. IECI MULNET*, pages 255-259, April 2000.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Robot Vision

1. Y. Pribadi, **S. S. Tandjung**, A. Setiadji, C. Machbub and T. L. R. Mengko. Mobile Robot System with Visual Feedback. *Int. Conf. on Microelectronics*, 9th, Oct. 1997.

All of the publications are related to the motion estimation directly and indirectly. They are available at

- http://sentosa.sas.ntu.edu.sg:8000/~steph,
- http://www.angelfire.com/journal2/stephanusst,
- http://groups.yahoo.com/group/motionestimation/,
- http://www.vlsi.ss.titech.ac.jp/~steph/thesis.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Bibliography

[1] C. Machbub, P. Yusuf, S. S. Tandjung, Ari Setiadji, and Tati L. R. Mengko, "Mobile robot system with visual feedback," in *Ninth Int. Conf. on Microelectronics*, Oct. 1997.

[2] T. L. R. Mengko, D. Astharini, S. S. Tandjung, H. R. Pranoto, T. Adiono, and C. Machbub, "Real time visual servoing system," in *Ninth Int. Conf. on Microelectronics*. ITB, Oct. 1997.

[3] J. Konrad, *Image and Video Processing Handbook*, chapter 3.8, Motion detection and estimation, Academic Press, 1999.

[4] J. Konrad and E. Dubois, "Estimation of image motion fields: Bayesian formulation and stochastic solution," in *International Conf. Acoustics, Speech and Signal Processing*. IEEE, 1988, pp. 1072–1075.

[5] P. Bouthemy and P. Lalande, "Motion detection in an image sequence using Gibbs distributions," in *Int. Conf. Acoust., Speech, Signal Processing*. 1989, pp. 1651–1654, IEEE.

[6] S. Z. Li, *Markov Random Field Modeling in Computer Vision*, Springer-Verlag, Hongkong, 1995.

[7] M. J. Black, *Robust Incremental Optical Flow*, Ph.D. thesis, University of Cambridge, Sept. 1992.

[8] M. J. Black and P. Anandan, "A framework for the robust estimation of optical flow," in *Int. Conf. on Computer Vision*, May 1993.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[9] M. J. Black and P. Anandan, "The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields," *Computer Vision and Image Understanding*, vol. 63, no. 1, pp. 75–104, January 1996.

[10] S. Z. Li, "Discontinuity-adaptive MRF prior and robust statistics: A comparative study," *Image and Vision Computing*, vol. 13, no. 4, pp. 227–223, April 1995.

[11] E. T. Jaynes, *Probability Theory: The Logic of Science*, Cambridge University, 2003.

[12] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 6, pp. 721–741, Nov. 1984.

[13] P. Perez and F. Heitz, "Restriction of a markov random field on a graph and multiresolution image analysis," *IEEE Trans. Information Theory*, vol. 42, pp. 180–190, Jan. 1996.

[14] C. C. Chen, *Markov Random Fields in Image Analysis*, Ph.D. thesis, Thesis, Michigan State University, USA, 1988.

[15] C. A. Bouman, "Tutorial: Markov random fields and stochastic image models," IEEE International Conference on Image Processing, October 1995.

[16] A. K. Jain and J. Mao, "Texture classification and segmentation using multiresolution simultaneous autoregressive models," *Pattern Recognition*, vol. 25, no. 2, pp. 173–188, 1992.

[17] R. W. Picard and I. M. Elfadel, "Gibbs random fields, coocurrences and texture modeling," *Pattern Analysis and Machine Intelligence*, vol. 16, no. 1, pp. 24–37, January 1994.

[18] J. Zhang, Modestion J. W., and Langan D. A., "Maximum-likelihood parameter estimation for unsupervised stochastic model-based image segmentation," *Image Processing*, vol. 20, no. 1, pp. 128–140, February 1994.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[19] R. Chellapa and J Zerubia, "Mean field annealing using compound Gauss-Markov random fields for edge detection and image estimation," *Neural Networks*, vol. 4, no. 4, pp. 703–709, July 1993.

[20] J. Zhang, "The mean field theory in EM procedures for blind Markov random field image restoration," *Image Processing*, vol. 2, no. 1, pp. 27–40, January 1993.

[21] J. Zhang, "The application of mean field theory to image motion estimation," *Image Processing*, vol. 4, no. 1, pp. 19–33, January 1995.

[22] A. Caplier and F. Luthon, "Spatiotemporal multiresolution associated to MRF modelling for motion detection," in *IEE Conference on Image Processing and Its Application*, July 1995, pp. 158–162.

[23] C. Stiller and J. Konrad, "Estimating motion in image sequences, a tutorial on modeling and computation of 2-D motion," *IEEE Signal Processing Magazine*, pp. 70–98, July 1999.

[24] A. M. Tekalp, *Digital Video Processing*, Prentice Hall, 1995.

[25] D. P. Elias, *The Motion-Based Segmentation of Image Sequences*, Ph.D. thesis, University of Cambridge, Signal Processing and Communications Lab., Dept. of Eng., Aug. 1998.

[26] C. H. Yang and J. Konrad, "Motion-based video segmentation using continuation method and robust cost functions," in *SPIE Proc. Visual Communications and Image Process.*, 1998.

[27] A. Benzougar, P. Bouthemy, and R. Fablet, "MRF-based moving object detection from MPEG CODEC video," in *Int. Conf. of Image Processing*. IEEE, Sept. 2001, vol. 2, pp. 402–980.

[28] Y. Chen and T. S. Huang, "Hierarchical MRF model for model-based multi-objet tracking," *Int. Conf. of Image Processing*, pp. 385–388, Sept. 2001.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[29] M. Silveira and M. Piedade, "MRF-motion segmentation based on dominant motion estimation and the detection of uncovered regions," in *Int. Conf. on Image Processings*, Sept. 2001, pp. 373–376.

[30] S. H. Lee, Kanatsugo Y., and J.-I. Park, "Object segmentation in stereo image usin cooperative line field in stochastic diffusion," *Int. Conf. on Image Processing*, 2001.

[31] E. Dubois and J. Konrad, "Bayesian estimation of motion vector fields," *IEEE Trans. on Pattern Analysis and Machine Intelligent*, vol. 14, pp. 910–927, Sept. 1992.

[32] E. Dubois and J. Konrad, *Motion Analysis and Image Sequence Processing*, chapter 3, Estimation of 2-D motion fields from image sequences with application to motion-compensated processing, Kluwer Academic Publishers, 1993.

[33] F. Heitsz and P. Bouthemy, "Motion estimation and segmentation using a global Bayesian approach," in *Proc. Int. Conf. Acoust., Speech, Signal Processing.* IEEE, 1990, pp. 2305–2307.

[34] F. Heitz and P. Bouthemy, "Multimodal estimation of discontinuous optical flow using Markov random fields," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 15, no. 12, pp. 1217–1232, Dec. 1993.

[35] R. D. Morris, *Image Sequence Restoration Using Gibbs Distributions*, Ph.D. thesis, University of Cambridge, May 1995.

[36] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artifical Intelligence*, vol. 17, pp. 185–203, 1981.

[37] B. Galvin, B. McCane, K. Novins, D. Mason, and S. Mills, "Recovering motion fields: An evaluation of eight optical flow algorithms," in *Ninth British Machine Vision Conference*, Sept. 1998, vol. 1, pp. 195–204.

[38] J. L. Barron, D. J. Fleet, S. S. Beauchemin, and T. A. Burkitt, "Performance of optical flow techniques," *CVPR*, vol. 92, pp. 236–242, 1992.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[39] J. Weng, N. Ahuja, and Thomas S. Huang, "Matching two perspective views," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 14, no. 8, pp. 806–824, August 1992.

[40] J. Weng, N. Ahuja, and Thomas S. Huang, "Optimal motion and structure estimation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, no. 9, pp. 864–884, September 1993.

[41] H. S. Seah, *From Image Sequence Correspondences to Structure and Motion using Active Template*, Ph.D. thesis, School of Applied Science, Nanyang Technological University, 2000.

[42] H. S. Seah and F. Tian, "Computer-assisted coloring by matching line drawings," *The Visual Computer, Springer-Verlag*, vol. 16, pp. 289–304, 2000.

[43] B. K. P. Horn, *Robot Vision*, The MIT Press, McGraw-Hill Book Company, 1986.

[44] H. H. Nagel and W. Enkelmann, "An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 565–593, Sept. 1986.

[45] S. V. Fogel, "Estimation of velocity vector fields from time-varying image sequences," *CVGIP: Image Understanding*, vol. 53, pp. 253–287, 1991.

[46] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," *Proc. DARPA Image Understanding Workshop*, pp. 121–130, 1981.

[47] J. Jain and A. Jain, "Displacement measurement and its application in interframe coding," *IEEE Trans. on Communications*, vol. 29, pp. 1799–1804, Dec. 1981.

[48] J. Zhang and G. G. Hanauer, "The application of mean field theory to image motion estimation," *IEEE Trans. Image Processing*, vol. 4, no. 1, pp. 19–33, Jan. 1995.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[49] M. Ghanbari, "The cross-search algorithm for motion estimation," *IEEE Trans. on Communications*, vol. 38, no. 950-953, July 1990.

[50] G. de Hann, P. W. Biezen, H. Huijgen, and O. A. Ojo, "True-motion estimation with 3-D recursive search block matching," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 3, pp. 368–379, Oct 1993.

[51] J. C. H. Ju, Y. K. Chen, and S. Y. Kung, "Rate optimization by true motion estimation," *IEEE Workshop Multimedia Signal Processing*, Jun. 1997.

[52] J. C. H. Ju, Y. K. Chen, and S. Y. Kung, "A fast algorithm for rate optimized motion estimation," in *International Symposium on Multimedia Information Processing*, Dec. 1997, pp. 472–477.

[53] K. W. Wong, K. M. Lam, and W. C. Siu, "An efficient low bit-rate video-coding algorithm focusing on moving regions," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11, no. 10, pp. 1128–1134, October 2001.

[54] Y. L. Chan and W. C. Siu, "An efficient search strategy for block motion estimation using image features," *IEEE Trans. on Image Processing*, vol. 10, no. 8, pp. 1223–1238, August 2001.

[55] W. Li and E. Salari, "Successive elimination algorithm for motion estimation," *IEEE Trans. Image Processing*, vol. 4, no. 1, pp. 105–107, 1995.

[56] H. S. Wang and R. M. Mersereau, "Fast algorithms for the estimation of motion vectors," *IEEE Trans. Image Processing*, vol. 8, no. 3, pp. 435–437, 1999.

[57] G. Heising, D. Marpe, and A. P. Petuknov, "Wavelet-based very low bit-rate video coding using image warping and overlapped block motion compensation," *IEE Vis. Image Signal Process.*, vol. 148, no. 2, pp. 93–101, 2001.

[58] M. Orchard, "Predictive motion-field segmentation for image sequence coding," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 3, pp. 54–70, Feb. 1993.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[59] H. Derin, H. H. Li, and S. Sun, *Video Data Compression for Multimedia Computing*, Kluwer Academic, 1997.

[60] M. I. Sezan and R. L. Lagendijk, *Motion Analysis and Image Sequence Processing*, Kluwer Academic, 1993.

[61] N. Paragios and G. Tziritas, "Adaptive detection and localization of moving objects in image sequences," *ELSEVIER Signal Processing Image Communication*, Feb. 1999.

[62] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. on Communications*, vol. 31, pp. 532–540, Apr. 1983.

[63] T. Lindenberg, *Scale space Theory in Computer Vision*, Kluwer, Netherland, 1994.

[64] S. Kruger, *Motion Analysis and Estimation using Multiresolution Affine Models*, Ph.D. thesis, University of Bristol, 1998.

[65] Y. Weis, *Bayesian motion estimation and segmentation*, Ph.D. thesis, Massachusetts Institute of Technology, 1998.

[66] J. M. Odobez and P. Bouthemy, "Robust multiresolution estimation of parametric motion models applied to complex scenes," Publication intern, no. 788, Janvier, Inria, France, 1994.

[67] M. Chan, Y. Yu, and A. Constanides, "Variable size block matching motion compensation with appications to video coding," *IEE Proc. Part I Communications, Speech and Vision*, vol. 137, pp. 205–212, Aug. 1990.

[68] F. Dufaux and M. Kunt, "Multigrid glock matching motion estimation with adaptive local mesh refinement," *SPIE in Visual Communications and Image Processing*, pp. 97–109, 1992.

[69] J. Zan, M. O. Ahmad, and M. N. S. Swamy, "Neighbourhood-blocks motion vector estimation technique using pyramidal data structure," *IEE Proc. Vis. Image Signal Process.*, vol. 149, no. 3, pp. 140–151, June 2002.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[70] M. Bierling, "Displacement estimation by hierarchical block matching," in *SPIE on Visual Communications and Image Processing*, 1988, pp. 942–951.

[71] A. Blake and A. Zisserman, "Visual reconstruction," *MIT Press, Cambridge, MA*, 1987.

[72] E. Memin, P. Perez, and D. Machecourt, "Dense estimation and object-oriented segmentation of the optical flow with robust techniques," Tech. Rep., INRIA, March 1996.

[73] M. J. Black and A. Rangarajan, "On the unification of line processes, outlier rejection, and robust statistics with applications in early vision," *International Journal of Computer Vision*, vol. 19, no. 1, pp. 57–92, July 1996.

[74] S. Z. Li, "On discontinuity-adaptive smoothness priors in computer vision," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 17, no. 6, pp. 576–586, June 1995.

[75] S. Z. Li, "Roof-edge preserving image smoothing based on MRFs," *IEEE Trans. on Image Processing*, vol. 9, no. 6, pp. 1134–1138, June 2000.

[76] K. P. Lim, M. N. Chong, and A. Das, "Dense motion field and uncovered region estimation for low bit-rate video coding application," *SPIE Conf. Visual Communications and Image Processing*, pp. 547–558, Jan. 1999.

[77] B. G. Haskel, A. Puri, and A. N. Netravali, *Digital Video: An Introduction to MPEG-2*, Chapman & Hall, 1997.

[78] P. W. Fieguth, *Application of Multiscale Estimation to Large Scale Multidimensional Imaging and Remote Sensing Problems*, Ph.D. thesis, Massachusetts Institute of Technology, June 1995.

[79] M. R. Luettgen, W. C. Karl, A. S. Willsky, and R. R. Tenney, "Multiscale representations of Markov random fields," *IEEE Trans. Information Theory*, vol. 41, no. 12, pp. 3277–3395, Dec. 1993.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

[80] W. C. Karl, A. S. Willsky, and M. R. Luettgen, "Efficient multiscale regularization with applications to the computation of optical flow," *IEEE Trans. on Image Processing*, vol. 3, no. 1, pp. 41–64, Jan. 1994.

[81] S. S. Tandjung, T. S. Gunawan, and M. N. Chong, "Motion estimation using adaptive block size observation model and efficient multiscale regularization," in *Int. Conf. on Image Processing*. IEEE, Sept. 2000.

[82] R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall International, 1962.