



**Faculty of Electronics and Computer Engineering**

**DEVELOPMENT OF DOUBLE STAGE FILTER (DSF) ON STEREO  
MATCHING ALGORITHM FOR  
3D COMPUTER VISION APPLICATIONS**

**Teo Chee Huat**

**Master of Science in Electronic Engineering**

**2016**

**DEVELOPMENT OF DOUBLE STAGE FILTER (DSF) ON STEREO MATCHING  
ALGORITHM FOR 3D COMPUTER VISION APPLICATIONS**

**TEO CHEE HUAT**

**A thesis submitted  
in fulfillment of requirements for the degree of Master of Science  
in Electronic Engineering**

**Faculty of Electronics and Computer Engineering**

**UNIVERSITI TEKNIKAL MALAYSIA MELAKA**

**2016**

## DECLARATION

I declare that this thesis entitled “Development of Double Stage Filter (DSF) on Stereo Matching Algorithm for 3D Computer Vision Applications” is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature : .....

Name : Teo Chee Huat  
.....

Date : .....

## **APPROVAL**

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of Master of Science in Electronic Engineering.

Signature : .....

Supervisor Name : Nurulfajar bin Abd Manap

Date : .....

## **DEDICATION**

To my beloved family

## ABSTRACT

In the field of stereo vision, some of existing stereo matching algorithms are designed with less accuracy of algorithm. Thus, a new hybrid algorithm with higher accuracy of computation is developed in this project. This thesis will present the design, development and analysis of performance on a developed Double Stage Filter (DSF) algorithm and other existing stereo matching algorithms. DSF algorithm is a hybrid stereo matching algorithm which divided into two phases. Phase 1 is consists of the part on Sum of Absolute Differences from basic block matching and the part of Scanline Optimization (SO) from dynamic programming approaches while phase 2 includes segmentation, merging and basic median filter process. The main feature of DSF algorithm is mainly on the phase 2 or as post-processing in which to remove the unwanted aspects like random noises and horizontal streaks, which is obtained from the raw disparity depth map on the step of optimization. In order to remove the unwanted aspects, two stages filtering process are needed along with the developed approaches in the phase 2 of DSF algorithm. There are two categorized evaluations done on the disparity maps obtained by the algorithms : objective evaluation and subjective evaluation. The objective evaluation includes the evaluation system of Middlebury Stereo Vision website page, computation analysis and traditional methods of Mean Square Errors (MSE), Peak to Signal Noise Ratio (PSNR) and Structural Similarity Index Metric (SSIM). Besides, for subjective evaluation, the datasets are captured from LNC IP camera and the results obtained by the selected algorithms are evaluated by human's eyes perception. Based on the results of evaluations, the results obtained by DSF is better than the algorithms, basic block matching and dynamic programming.

## **ABSTRAK**

*Dalam bidang visio stereo, terdapat banyak rekaan algoritma padanan stereo yang kurang ketepatan. Oleh itu, satu algoritma hibrid baru dengan komputasi yang lebih ketepatan akan direka dalam projek ini. Tesis ini membentangkan reka bentuk, pembangunan dan analisis prestasi algoritma Penapis Gandaan Peringkat (DSF) dengan algoritma padanan stereo yang sedia ada. Algoritma DSF merupakan satu algoritma hibrid padanan stereo yang terbahagi kepada fasa 1 dan fasa 2. Fasa 1 adalah terdiri daripada Sum of Absolute Differences (SAD) iaitu sebahagian daripada padanan blok asas dan Scanline Optimization (SO), sebahagian daripada teknik pengaturcaraan dinamik manakala fasa 2 terdiri daripada proses segmentasi, penggabungan dan penapis median asas. Ciri utama algoritma DSF adalah pada fasa 2 atau sebagai pasca-pemprosesan untuk menghilangkan aspek-aspek yang tidak diinginkan seperti hingar rawak dan jalur-jalur mendatar yang terhasil daripada peta perbezaan kasar daripada langkah pengoptimuman. Dalam usaha untuk menghapuskan aspek-aspek yang tidak diinginkan, dua peringkat proses penapisan diperlukan bersama teknik yang direka dalam pasca-pemprosesan algoritma DSF. Terdapat dua kategori penilaian yang dilakukan pada hasil peta perbezaan daripada algoritma-algoritma : penilaian objektif dan penilaian subjektif. Penilaian objektif termasuk sistem penilaian yang ada dalam laman web Middlebury Stereo Visio, analisis komputasi dan kaedah-kaedah tradisional seperti Mean Square Error (MSE), Peak to Signal Noise Ratio (PSNR) dan Similarity Structural Index Metric (SSIM). Selain itu, penilaian subjektif adalah menggunakan dataset yang dirakam daripada kamera LNC IP dan hasil daripada beberapa algoritma yang dipilih dinilai oleh penglihatan manusia. Berdasarkan hasil penilaian, keputusan yang dihasil oleh algoritma DSF adalah lebih baik berbanding algoritma Basic Block Matching dan Dynamic Programming.*

## **ACKNOWLEDGEMENTS**

I would like to express my highly appreciation to my main supervisor Dr. Nurulfajar bin Abd Manap who had guided, motivated me throughout the research on image processing and his advice had greatly improved my thesis writing. I also would like to thank my co-supervisor, Dr. Masrullizam bin Mat Ibrahim for his advices. Without their moral support and concern, this thesis would not have been completed and presented here.

Furthermore, I want to thank my lecturers, specially Dr. Abd Majid bin Darsono, Mr. Khairul Muzzammil bin Saipullah, Mr. Sani Irwan bin Md Salim and Dr. Lim Kim Chuan for their advice and motivation throughout my research work.

At last, I would like to extend my gratitude to my parents, Mr. Teo Koon Hock and Mrs. Lam Fong Kuen, my siblings and my friends for their encouragement, love and support. I am grateful to have all these people around me throughout the research. Once again, thank you so much.



## TABLE OF CONTENTS

	<b>PAGE</b>
<b>DECLARATION</b>	
<b>APPROVAL</b>	
<b>DEDICATION</b>	
<b>ABSTRACT</b>	<b>i</b>
<b>ABSTRAK</b>	<b>ii</b>
<b>ACKNOWLEDGEMENTS</b>	<b>iii</b>
<b>TABLE OF CONTENTS</b>	<b>iv</b>
<b>LIST OF TABLES</b>	<b>vi</b>
<b>LIST OF FIGURES</b>	<b>vii</b>
<b>LIST OF APPENDICES</b>	<b>xi</b>
<b>LIST OF ABBREVIATIONS</b>	<b>xii</b>
<b>LIST OF SYMBOLS</b>	<b>xiv</b>
<b>LIST OF PUBLICATIONS</b>	<b>xvi</b>
<b>CHAPTER</b>	
<b>1. INTRODUCTION</b>	<b>1</b>
1.1 Background	1
1.2 Problem Statement	2
1.3 Aims and Objectives	3
1.4 Scope of Work	4
1.5 Research Contributions	5
1.6 Thesis Organization	6
<b>2. LITERATURE REVIEW</b>	<b>8</b>
2.1 Introduction	8
2.2 Basic Concept of Stereo Matching	9
2.2.1 Matching Cost	10
2.2.2 Cost Aggregation	12
2.2.3 Disparity Computation And Optimization	15
2.2.4 Disparity Refinement	16
2.3 Block Matching Techniques	18
2.3.1 Winner-Update-Strategy (WUS)	19
2.3.2 Hierarchical Block Matching	20
2.3.3 Active Colour Illumination	21
2.3.4 Comparison of Block Matching Approaches	22
2.4 Dynamic Programming Approaches	23
2.4.1 One Dimension Optimization (1D)	24
2.4.2 Two Dimension Optimization (2D)	25
2.4.3 Dynamic Programming on Tree	25
2.4.4 Comparison of Dynamic Programming Approaches	28
2.5 Segmentation	29
2.5.1 Fast 3D Stereo Vision	30
2.5.2 Segment-Based Stereo Matching	32
2.6 Comparison of Stereo Matching Algorithms	36
2.7 Image Filtering	37

2.7.1	Low Pass and High Pass Filter	37
2.7.2	Median Filtering	39
2.7.3	Average Filtering	40
2.7.4	Adaptive Wiener Filtering	41
2.7.5	Comparison of Filtering Techniques	42
2.8	Summary	43
<b>3.</b>	<b>METHODOLOGY</b>	<b>45</b>
3.1	Introduction	45
3.2	Double Stage Filter (DSF) Design Structure	47
3.3	Phase 1: Obtaining the Depth Map	50
3.3.1	Matching Cost	50
3.3.2	Cost Aggregation	53
3.3.3	Disparity Computation (Optimization)	54
3.4	Phase 2: Disparity Depth Map Refinement Stage	57
3.4.1	Segmentation	58
3.4.2	Median filter	62
3.4.3	Merging	66
3.5	Evaluation Methods	69
3.5.1	Objective Evaluation	69
3.5.1.1	Evaluation from Middlebury System	69
3.5.1.2	Common Objective Evaluation	70
3.5.2	Subjective Evaluation : LNC IP Camera	73
3.6	Summary	75
<b>4.</b>	<b>RESULTS AND DISCUSSIONS</b>	<b>76</b>
4.1	Introduction	76
4.2	Results in Middlebury	77
4.2.1	Disparity Depth Maps	77
4.2.2	Ranking in Middlebury	79
4.3	Results Analysis for Stereo Pairs	84
4.3.1	Results Analysis for Tsukuba Stereo	84
4.3.2	Results Analysis for Venus Stereo	88
4.3.3	Results Analysis for Teddy Stereo	92
4.3.4	Results Analysis for Cones Stereo	96
4.4	Subjective Evaluations	102
4.5	Summary	104
<b>5.</b>	<b>CONCLUSION AND FUTURE WORK</b>	<b>105</b>
5.1	Conclusion	105
5.2	Future Work	107
	<b>REFERENCES</b>	<b>109</b>
	<b>APPENDICES</b>	<b>131</b>

## LIST OF TABLES

<b>TABLE</b>	<b>TITLE</b>	<b>PAGE</b>
2.1	Descriptions and equations of Absolute Difference (AD), Squared Difference (SD) and Truncated Absolute Difference (TAD)	11
2.2	Equations of area-based matching cost functions	13
2.3	Calculation of penalty score for each player	20
2.4	Comparison of segmentation methods	29
2.5	Comparison of steps among stereo matching algorithms	37
2.6	Comparison of image filtering techniques	43
3.1	Suitable disparity range for stereo pairs and extracted segments	62
3.2	Proper scales for stereo pairs by Middlebury	65
3.3	Mask sizes for median filter on stereo pairs	65
3.4	Specifications of LNC IP Camera	74
4.1	Middlebury ranking for DSF algorithm	81

## LIST OF FIGURES

<b>FIGURE</b>	<b>TITLE</b>	<b>PAGE</b>
2.1	Basic steps of stereo matching algorithm	9
2.2	Overview of pixel based matching	11
2.3	Disparity selection by Winner Takes All (WTA)	13
2.4	Streaks on Tsukuba obtained by DP algorithm	16
2.5	Three matching costs with parabola function	17
2.6	(a) Depth obtained by non-hierarchical approach ; (b) Improved depth map using hierarchical approach	21
2.7	System of active colour illumination	22
2.8	Disparity Space Image (DSI)	24
2.9	Dynamic programming on tree	27
2.10	Block diagram design of fast 3D stereo vision algorithm	30
2.11	(a) Original image ; (b) Pixel disparity map	31
2.12	Filtered disparity map	32
2.13	Design overview of segment-based stereo matching	32
2.14	Disparity planes	35
2.15	Low pass filter	38
2.16	High pass filter	38

2.17	Sample results of low pass and high pass filter :	39
	(a) Original image ; (b) Low pass filter ; (c) High pass filter	
2.18	Median filtering : (a) Source of noisy image ;	40
	(b) Image after median filtering	
2.19	Sample results of average filtering : (a) Original image;	41
	(b) Results after filtering	
2.20	AW filtering : (a) Source of noisy image ; (b) Image after filtering	42
3.1	General design of block diagram for DSF	48
3.2	(a) Raw disparity map ; (b) Refined disparity map	48
3.3	Original stereo pair of Tsukuba image : (a) Left image ;	51
	(b) Right image	
3.4	Flowchart for Sum of Absolute Differences (SAD)	52
3.5	Depth map by SAD approach	52
3.6	Cost aggregation : (a) Reference image ; (b) Target image	54
3.7	Flowchart of Scanline Optimization (SO)	56
3.8	(a) Disparity map by dynamic programming on Teddy ;	57
	(b) Disparity map by DSF	
3.9	Flowchart of segmentation approach	59
3.10	Flow of segmentation process	59
3.11	Segmentation process on stereo pairs	60
3.12	Pseudocode for segmentation	61
3.13	Filter mask of $3 \times 3$ region	62
3.14	Flowchart for median filtering	64
3.15	Pseudocode for median filter	65

3.16	Samples of disparity map ; (a) before median filtering ; (b) after median filtering	66
3.17	Operation of merging process	67
3.18	Pseudocode for merging algorithm	67
3.19	Sample of segment	68
3.20	Merged segments of first median filter	68
3.21	Second time of median filter with outliers reduced	69
3.22	Comparison of MSE and SSIM values for an image	72
3.23	Comparison of PSNR and SSIM values for an image	73
3.24	LNC IP Camera	74
4.1	Results of three stereo matching algorithms by applying the Middlebury benchmark datasets : Tsukuba, Venus, Teddy and Cones	78
4.2	Non-occluded areas (white) and occluded (black)	80
4.3	Depth discontinuities areas (white)	80
4.4	Non-occluded regions with occluded and border regions : (a) Samples; (b) BBM; (c) DP; (d) DSF	83
4.5	(a) Occluded and non-occluded regions Tsukuba of DP algorithm ; (b) Occluded and non-occluded regions Tsukuba of DSF algorithm	83
4.6	MSE value for Tsukuba stereo pair	85
4.7	PSNR value for Tsukuba stereo pair	86
4.8	SSIM value for Tsukuba stereo pair	87
4.9	SSIM index map for Tsukuba ; (a) Window size; (b) BBM; (c) DP; (d) DSF	88
4.10	MSE value for Venus stereo pair	89

4.11	PSNR value for Venus stereo pair	90
4.12	SSIM value for Venus stereo pair	91
4.13	SSIM index map for Venus ; (a) Window size; (b) BBM; (c) DP; (d) DSF	92
4.14	MSE value for Teddy stereo pair	93
4.15	PSNR value for Teddy stereo pair	94
4.16	SSIM value for Teddy stereo pair	95
4.17	SSIM index map for Teddy ; (a) Window size; (b) BBM; (c) DP; (d) DSF	96
4.18	MSE value for Cones stereo pair	97
4.19	PSNR value for Cones stereo pair	98
4.20	SSIM value for Cones stereo pair	99
4.21	SSIM index map for Cones ; (a) Window size; (b) BBM; (c) DP; (d) DSF	100
4.22	Results from dataset captured using LNC IP camera : (a) Original stereo images ; (b) BBM ; (c) DP ; (d) DSF	102
4.23	Results of datasets from Middlebury Webpage : (a) Original ; (b) BBM ; (c) DP ; (d) DSF	103

## LIST OF APPENDICES

<b>APPENDIX</b>	<b>TITLE</b>	<b>PAGE</b>
A	Datasheet of LNC IP Camera	131
B	User Interface	132
C	Flowchart for Reading and Converting Stereo Images	133



## LIST OF ABBREVIATIONS

1D	-	One Dimensional
2D	-	Two Dimensional
3D	-	Three Dimensional
AD	-	Absolute Difference
BM	-	Bidirectional Matching
BBM	-	Basic Block Matching
CUDA	-	Compute Unified Device Architecture
DP	-	Dynamic Programming
DSI	-	Disparity Space Image
DSF	-	Double Stage Filter
GPU	-	Graphic Processing Unit
IP	-	Internet Protocol
LNC	-	Low Noise Converter
MF	-	Matching Function
MSE	-	Mean Squared Error

- NCC - Normalized Cross Correlation
- PSNR - Peak Signal to Noise Ratio
- RMS - Root Mean Squared
- SD - Squared Difference
- SO - Scanline Optimization
- SAD - Sum of Absolute Difference
- SNR - Signal to Noise Ratio
- SSD - Sum of Squared Differences
- SSIM - Structural Similarity Index Metric
- WTA - Winner Take All

## LIST OF SYMBOLS

$x$	-	Pixel in column
$y$	-	Pixel in row
$I$	-	Image
$d$	-	Position increment of a pixel
$W$	-	Square window for aggregation
$\nabla$	-	Gradient
$w$	-	Absolute weighting
$*$	-	Multiply
$P$	-	Disparity plane
$S$	-	Segment
$\lambda_{disc}$	-	Common border lengths
$C_{census}$	-	Hamming distance between the correspondence pixels of left image and right image
$d_p^*$	-	Highest votes or bin value for the disparity
$S_p$	-	Total number of good pixels

$d_p$	-	Value of disparity map
$E$	-	Edges
$m$	-	Matching penalty
$t$	-	Threshold
$N$	-	Neighborhood
$y_k$	-	Data point upon the iteration
$a_x$	-	Number of pixels in the segment from data point of $x$
$A$	-	Aggregated area
Occ	-	Occluded pixels
$\sigma_i$	-	Input of noise variance
$f(\bar{n})$	-	Function of noise density

## LIST OF PUBLICATIONS

### Journals:

**Teo, C.H.**, Abd Manap, N., 2014. Evaluation of Stereo Matching Algorithms and Dynamic Programming for 3D Triangulation. *Advanced Computer and Communication Engineering Technology, Volume 315*, pp.641-650. (Scopus)

**Teo, C.H.**, Abd Manap, N., 2014. Evaluation of Dynamic Programming among the Existing Stereo Matching Algorithms. *International Conference on Mathematics Engineering and Industrial Applications (ICoMEIA)*. (Scopus)

**Teo, C.H.**, Abd Manap, N., Mat Ibrahim, M., 2015. Development of Double Stage Filter (DSF) for Stereo Matching Algorithms and 3D Vision applications. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*. (Scopus and Google Scholar) (Accepted)

**Articles:**

**Teo, C.H.**, Abd Manap, N., 2015. Computation and Performance Analysis of Double Stage Filter for Image Processing. *International Conference on Intelligent and Interactive Computing (IIC 2015)*. (Scopus) (Accepted)

**Teo, C.H.**, Abd Manap, N., 2015. Performance Evaluation of Double Stage Filter Algorithm for Stereo Vision Applications. In: IEEE, *International Conference on Signal and Image Processing Applications (ICSIPA)* . (Scopus and ISI) (Accepted)

Abd Manap, N., **Teo, C.H.**, 2015. Analysis on Segment-Based Double Stage Filter Algorithm for Stereo Matching. *International Conference on Telecommunication, Electronic and Computer Engineering (ICTEC 2015)* . (Scopus and ISI) (Accepted)

# CHAPTER 1

## INTRODUCTION

### 1.1 Background

Stereo matching is one of the popular topic in the field of computer vision and gain interest among researchers. Stereo vision is a technique on inferring depth from stereo camera with division of left and right side to form stereo pair images for stereo matching algorithm purpose (Bleyer & Gelautz, 2005). Stereo matching algorithms are the procedure to obtain a disparity depth map from the corresponding pixels from the pair of stereo images. It can be used in 3D reconstruction, scene analysis and other depth related usage. Most computer vision based approaches for extracting stereo information can mainly classified as monocular and dual-camera system. The analysis of depth details which from monocular and dual-camera are through its disparity and depth map (Scharstein & Szeliski, 2002). Disparity and depth map data can be gathered by capturing the stereo image simultaneously and calculated through stereo matching algorithm (Guillemaut & Hilton, 2012). Disparity depth map is used to synthesize the inter-view images based on the depth layer extraction (Bleyer & Gelautz, 2005). The depth information in a scene for a stereo image is used in many image analysis and 3D video processing applications.

There are many stereo matching algorithms has been developed, which can be found in the Middlebury Stereo Vision Website by (Scharstein & Szeliski, 2002). This

particular website provides various datasets of stereo images and standard evaluation for researchers to compare the results obtained from their proposed stereo matching algorithm with other. In general, most stereo matching algorithm consist of four common steps such as matching cost computation, cost aggregation, disparity computation optimization and disparity refinement according to (Scharstein & Szeliski, 2002). However, the four steps are not necessary to be included in a stereo matching algorithm as it is depending on the design developed by the researchers.

In the development of stereo matching algorithm, the most preferable step which attracts the interest from the researchers is on the post-processing step. Post-processing is the step of disparity refinement where the raw disparity map obtained from computation of correspondence and optimization. The raw disparity map contains some noises, mismatches and holes from occlusion, which need to be smoothed out and removed. There are various type of post-processing can performed such as sub-pixel estimation, median filter, cross-checking and surface fitting (Fua, 1993; Heo, Lee, & Lee, 2011; Li & Li, 2008; Ma & He, 2013). Some algorithms with high processing speed may obtain a relatively low precision of disparity depth map (Bobick & Intille, 1999; Calakli, Ulusoy, Restrepo, Taubin, & Mundy, 2012; Zhou, Troccoli, & Pulli, 2012).

## **1.2 Problem Statement**

There are many stereo matching algorithms can be found in the field of image processing. However, some of the stereo matching algorithms lack of accuracy in obtaining the disparity depth map. 3D triangulation and effects on image processing depending on the smoothness of the disparity depth map that obtained from stereo matching algorithms. The smoother the disparity depth map, the better effect of 3D applications (Gong & Yang,