



Faculty of Electrical Engineering

**DISPARITY MAP ALGORITHM FOR STEREO MATCHING PROCESS
USING LOCAL BASED METHOD**



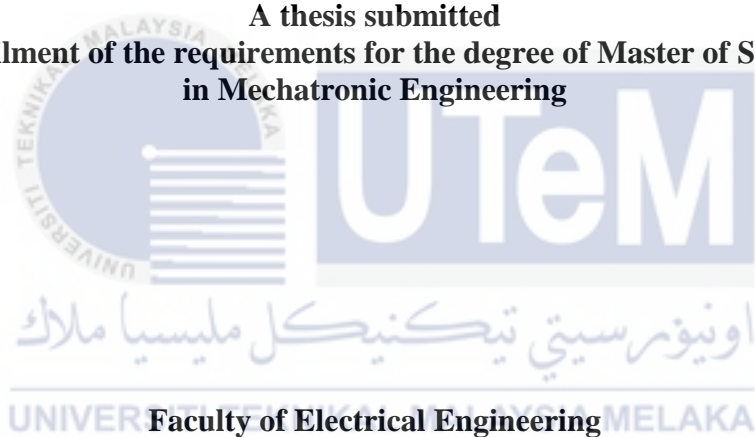
Master of Science in Mechatronic Engineering

2021

**DISPARITY MAP ALGORITHM FOR STEREO MATCHING PROCESS USING
LOCAL BASED METHOD**

MELVIN GAN YEOU WEI

**A thesis submitted
in fulfilment of the requirements for the degree of Master of Science
in Mechatronic Engineering**



UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2021

DECLARATION

I declare that this thesis entitles “Disparity Map Algorithm for Stereo Matching Process Using Local Based Method” 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

MELVIN GAN YEOU WEI
.....

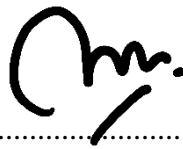
Date

12/12/2021
.....

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

APPROVAL

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



Signature :

Supervisor's Name : TS. DR. ROSTAM AFFENDI BIN HAMZAH

Date : 12/12/2021



DEDICATION

To my late father and beloved mother



ABSTRACT

The aim of Stereo Vision Disparity Map (SVDM) algorithm is to obtain the disparity map from two images. These input images have different viewpoints that corresponds with each other, forming a two-dimensional mapping of matching pixels and is known as disparity map. The SVDM algorithm can be categorized into local, semi global and global methods. Global method performs a matching process using global energy or a probability function over the whole image. This method involves high computational complexity and slow implementation. Therefore, it is not suitable for real-time applications. However, local method solves the matching problem via a local analysis and aggregating matching costs over a support region at each pixel in the images. The local method delivers fast execution and low computational requirement. Semi global method is the combination of both methods which make this method is more complex. However, the computation for the development of SVDM algorithm is more challenging especially for the images with complex scenes. There are several factors such as low texture region, repetitive patterns, illumination different, discontinuity, and occlusion. Hence, this thesis proposes a local-based SVDM algorithm that increases the accuracy on the complex scenes. The proposed SVDM algorithm involves four stages which starts from matching cost computation. At this stage, the proposed work uses the combination of Absolute Difference (AD) and Gradient Matching (GM) to produce the initial disparity map. Second stage, the Minimum Spanning Tree (MST) is utilised to remove noise from the initial disparity map. Then, the optimization stage uses a Winner-Take-All (WTA) strategy. The WTA strategy absorbs the minimal aggregated corresponding value for each valid pixel in disparity map. At the final stage, Bilateral Filter (BF) with Histogram Equalization and Weighted Median (WM) filter are proposed. These filters are capable to increase the accuracy and preserve the object edges. In this research, two standard online benchmarking database systems are used to measure the accuracy of the proposed algorithm. These systems are from the Middlebury Stereo for quantitative measurement and Karlsruhe Institute of Technology and Toyota Technological Institute (KITTI) for qualitative measurement. Once satisfactory results are obtained, there is demonstration on 3D surface reconstruction using disparity maps produced by the proposed SVDM algorithm. In conclusion, the proposed SVDM algorithm produces accurate results from the validation process. The final results are 7.55% for avg nonocc error and 10.6% for avg all error which are competitive with other established methods when compared in the standard benchmarking evaluation system from the Middlebury Stereo. The disparity map results on the complex scenes are also improved and fine quality of 3D surface reconstruction have been produced for both of the Middlebury and KITTI images.

ALGORITMA PETA PERBEZAAN UNTUK PROSES PADANAN STEREO MENGUNAKAN KAEDAH SETEMPAT

ABSTRAK

Tujuan algoritma Peta Perbezaan Penglihatan Stereo (PPPS) adalah untuk mendapatkan peta perbezaan dari dua imej. Imej-imej ini mempunyai sudut pandang berbeza yang saling berkaitan, membentuk pemetaan piksel dua dimensi yang sepadan dikenali sebagai peta perbezaan. Algoritma PPPS boleh dikategorikan kepada kaedah tempatan, global dan separa global. Kaedah global melakukan proses padanan menggunakan tenaga global atau fungsi kebarangkalian pada keseluruhan imej. Kaedah ini melibatkan kekompleksan pengkomputeran yang tinggi dan pelaksanaan yang perlahan. Oleh itu, ia tidak sesuai untuk aplikasi-aplikasi masa nyata. Walau bagaimanapun, kaedah tempatan menyelesaikan masalah pemadanan melalui analisis tempatan dan mengagregatkan kos pemadanan di kawasan sokongan pada setiap piksel dalam imej. Kaedah tempatan memberikan pelaksanaan cepat dan keperluan pengkomputeran yang rendah. Kaedah separa global adalah gabungan kedua-dua kaedah yang menjadikan kaedah ini lebih kompleks. Walau bagaimanapun, pengiraan untuk pembangunan algoritma PPPS lebih mencabar terutamanya untuk imej dengan pemandangan yang kompleks. Terdapat beberapa faktor seperti kawasan bertekstur rendah, corak berulang, pencahayaan berbeza, ketakselajaran, dan oklusi. Oleh itu, tesis ini mencadangkan algoritma PPPS berasaskan tempatan yang meningkatkan ketepatan pada pemandangan kompleks. Algoritma PPPS yang dicadangkan melibatkan empat peringkat yang bermula dari pengiraan kos yang sepadan. Pada tahap ini, algoritma yang dicadangkan menggunakan kombinasi Perbezaan Mutlak (PM) dan Padanan Gradient (PG) untuk menghasilkan peta perbezaan awal. Tahap kedua, Minimum Spanning Tree (MST) digunakan untuk menghilangkan hingar dari peta perbezaan awal. Kemudian, tahap pengoptimuman menggunakan strategi Winner-Take-All (WTA). Strategi WTA menyerap nilai sepadan minimum terkumpul untuk setiap piksel yang sah dalam peta perbezaan. Pada peringkat akhir, penapis Bilateral, pemerataan histogram dan penapis Pemberat Pertengahan (PP) dicadangkan. Penapis ini mampu meningkatkan ketepatan dan memelihara tepi objek. Dalam penyelidikan ini, dua sistem pangkalan data penandaarasan standard digunakan untuk mengukur ketepatan algoritma yang dicadangkan. Sistem ini adalah daripada Stereo Middlebury untuk pengukuran kuantitatif dan Institut Teknologi Karlsruhe dan Institut Teknologi Toyota (KITTI) untuk pengukuran kualitatif. Setelah keputusan yang memuaskan diperolehi, terdapat demonstrasi pada pembinaan semula permukaan 3D menggunakan peta perbezaan yang dihasilkan oleh algoritma PPPS yang dicadangkan. Kesimpulannya, algoritma PPPS yang dicadangkan menghasilkan keputusan yang jitu daripada proses validasi. Keputusan akhir ialah 7.55% untuk purata ralat bukan "nonocc" dan 10.6% untuk purata ralat "all" di mana ianya berdaya saing dengan kaedah-kaedah yang sudah teruji jika dibandingkan dalam sistem standard penilaian penanda aras daripada Middlebury Stereo. Keputusan peta perbezaan pada pemandangan yang kompleks juga bertambah baik dan pembinaan semula permukaan 3D dengan kualiti yang jelas dihasilkan untuk kedua-dua imej Middlebury dan KITTI.

ACKNOWLEDGEMENTS

First, I would like to express my sincere acknowledgment and gratitude toward my main supervisor, Ts. Dr. Rostam Affendi bin Hamzah, senior lecturer in Fakulti Teknologi Kejuruteraan Elektrik & Elektronik (FTKKEE) of Universiti Teknikal Malaysia Melaka (UTeM), for his guidance, patience, encouragement, time, knowledge and wisdom. With all these, I was able to enrich my knowledge, especially on Stereo Vision Disparity Map (SVDM) technology, which also help me in completing this thesis. Next, I would also like to express my gratitude towards my co-supervisor, Dr. Nik Syahrim bin Nik Anwar, senior lecturer from Fakulti Kejuruteraan Elektrik (FKE) for his guidance, encouragement, knowledge and time, as it also helps me to completing this thesis. Then, I would also like to express my gratitude towards the panels during my proposal defence, Prof. Madya Ir. Dr. Abdul Rahim bin Abdullah, Prof. Madya Ts. Dr. Muhammad Fahmi bin Miskon and Prof. Madya Dr. Herman bin Jamaluddin for providing me the insight in improving my research. Finally, to my late father, my beloved mother, my sibling; my younger sister and younger brother, family and friends, I would like to acknowledge their continuous support throughout the project.

TABLE OF CONTENTS

	PAGE
DECLARATION	
APPROVAL	
DEDICATION	
ABSTRACT	i
ABSTRAK	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
LIST OF TABLES	vi
LIST OF FIGURES	vii
LIST OF APPENDICES	x
LIST OF ABBREVIATIONS	xi
LIST OF SYMBOLS	xiv
LIST OF PUBLICATIONS	xvi
CHAPTER	
1. INTRODUCTION	1
1.1 Background	1
1.2 Application	4
1.3 Problem statement	7
1.4 Objectives	12
1.5 Scope	12
1.6 Outline of thesis	13
2. LITERATURE REVIEW	15
2.1 Introduction	15
2.2 Matching cost computation	21
2.3 Cost aggregation	34
2.4 Disparity optimization	37
2.5 Disparity refinement	38
2.6 3D surface reconstruction	42
2.7 Standard online benchmarking database	43
3. METHODOLOGY	45
3.1 Methodology for research	45
3.2 Methodology for proposed SVDM algorithm	49
3.2.1 Matching cost computation	51
3.2.2 Cost aggregation	55
3.2.3 Disparity optimization	56
3.2.4 Disparity refinement	57
3.3 3D surface reconstruction	59
3.4 Evaluation of experimental results	60
4. RESULT AND DISCUSSION	62
4.1 Parameters selection	62
4.1.1 Parameters selection on α	65
4.1.1.1 Quantitative results for parameters selection on α	65
4.1.1.2 Qualitative results for parameters selection on α	69

4.1.2	Parameters selection on σ_s	70
4.1.2.1	Quantitative results for parameters selection on σ_s	70
4.1.2.2	Qualitative results for parameters selection on σ_s	74
4.1.3	Parameters selection on σ_c	75
4.1.3.1	Quantitative results for parameters selection on σ_c	75
4.1.3.2	Qualitative results for parameters selection on σ_c	79
4.1.4	Parameters selection on w_p	81
4.1.4.1	Quantitative results for parameters selection on w_p	82
4.1.4.2	Qualitative results for parameters selection on w_p	86
4.1.5	Summary for parameters selection process	88
4.1.5.1	Summary for parameters selection (quantitative results)	89
4.1.5.2	Summary for parameters selection (qualitative results)	90
4.2	Disparity maps for proposed SVDM algorithm (Middlebury Stereo)	93
4.3	Accuracy analysis for proposed SVDM algorithm (Middlebury Stereo)	96
4.4	Comparison analysis for proposed algorithm (Middlebury stereo)	104
4.4.1	Comparison for low texture region	105
4.4.2	Comparison for repetitive patterns	111
4.4.3	Comparison for depth discontinuity	117
4.4.4	Comparison for illumination differences	122
4.4.5	Comparison for occlusion	126
4.5	Disparity maps for using other input images	131
4.6	Demonstration for 3D surface reconstruction (Depth map)	134
5	CONCLUSION AND FUTURE WORK	137
5.1	Conclusion	137
5.2	Future works	140
	REFERENCES	142
	APPENDICES	156

LIST OF TABLES

TABLE	TITLE	PAGE
1.1	Taxonomy for local based SVDM method	3
2.1	Summary of SVDM algorithms framework cited in this thesis	17
2.2	Characteristic of matching cost computation approaches	21
2.3	Brief comparison of matching cost computation methods	32
2.4	Brief comparison of segmentation techniques	36
4.1	Parameter Selection for proposed method	64
4.2	Results for the parameter selection on α	68
4.3	Result for the parameter selection on σ_s	73
4.4	Results for the parameter selection on σ_c	78
4.5	Results for the parameter selection on w_p	85
4.6	Summary parameters for the proposed method	89
4.7	The performance based on each stage of the proposed algorithm	96
4.8	Taxonomy of proposed SVDM algorithms and compared SVDM algorithms in Middlebury Stereo	97
4.9	The accuracy analysis for average nonocc error	102
4.10	The accuracy analysis for avg all error	102
4.11	Results based on different types of algorithms	104

LIST OF FIGURES

FIGURE	TITLE	PAGE
1.1	Stereo vision system	2
1.2	Stereo vision basics	2
1.3	Object with low texture region and repetitive region	8
1.4	Example of depth discontinuity	9
1.5	Input images with illumination difference	10
1.6	Causes of occlusion	11
2.1	Block Diagram for the proposed SVDM algorithm	16
2.2	3x3 pixel square window	23
2.3	SVDM and 3D surface reconstruction process	43
3.1	The flowchart for the research	48
3.2	Flowchart for proposed algorithm	51
3.3	Flowchart for stage one of the proposed algorithm; matching cost computation	54
3.4	Flowchart for stage two of the proposed algorithm; cost aggregation	56
3.5	Flowchart for stage four of the proposed algorithm; disparity refinement stage	59
4.1	Line graphs based on the results from parameter selection process of α	68
4.2	Disparity maps for parameter selection on α	70
4.3	Line graphs based on the results from parameter selection process of σ_s	73

4.4	Disparity maps for parameter selection on σ_s	75
4.5	Line graphs based on the results from parameter selection process of σ_c	78
4.6	Disparity maps - Playroom for parameter selection on σ_c	80
4.7	Disparity maps - Piano for parameter selection on σ_c	81
4.8	Line graphs based on the results from parameter selection process of w_p	85
4.9	Disparity maps - Vintage for parameter selection on w_p	87
4.10	Disparity maps - Pipes for parameter selection on w_p	88
4.11	Disparity maps (MotorcycleE) for parameter selection process	91
4.12	Disparity maps (PlaytableP) for parameter selection process	92
4.13	The final disparity maps for the proposed SVDM algorithm	94
4.14	Input images for low texture region comparison	106
4.15	Comparison on low texture region for Adirondack	107
4.16	Comparison on low texture region for Recycle	109
4.17	Comparison on low texture region for Vintage	110
4.18	Input images for repetitive patterns comparison	112
4.19	Comparison on repetitive patterns for Jadeplant	113
4.20	Comparison on repetitive patterns for Playroom	115
4.21	Comparison on repetitive patterns for Playtable	117
4.22	Input images for discontinuity region comparison	118
4.23	Comparison on depth discontinuity for Pipes	119
4.24	Comparison on depth discontinuity for Playroom	121
4.25	Input images for illumination differences comparison	123
4.26	Comparison on illumination differences for ArtL	124
4.27	Comparison on illumination differences for PianoL	125

4.28	Input images for occlusion comparison	127
4.29	Comparison on occlusion for Shelves	129
4.30	Comparison on occlusion for Teddy	130
4.31	Final disparity maps for KITTI	132
4.32	3D surface reconstruction for final disparity maps (Middlebury Stereo)	135
4.33	3D surface reconstruction for final disparity maps (KITTI)	136
A.1	Disparity maps (Adirondack) results for parameter selection process	156
A.2	Disparity maps (ArtL) results for parameter selection process	157
A.3	Disparity maps (Jadeplant) results for parameter selection process	158
A.4	Disparity maps (Motorcycle) results for parameter selection process	159
A.5	Disparity maps (MotorcycleE) results for parameter selection process	160
A.6	Disparity maps (Piano) results for parameter selection process	161
A.7	Disparity maps (PianoL) results for parameter selection process	162
A.8	Disparity maps (Pipes) results for parameter selection process	163
A.9	Disparity maps (Playroom) results for parameter selection process	164
A.10	Disparity maps (Playtable) results for parameter selection process	165
A.11	Disparity maps (PlaytableP) results for parameter selection process	166
A.12	Disparity maps (Recycle) results for parameter selection process	167
A.13	Disparity maps (Shelves) results for parameter selection process	168
A.14	Disparity maps (Teddy) results for parameter selection process	169
A.15	Disparity maps (Vintage) results for parameter selection process	170

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
A	Final results for parameter selection process	156



LIST OF ABBREVIATIONS

1D	-	One Dimensional
2D	-	Two Dimensional
3D	-	Three Dimensional
AD	-	Absolute Differences
AR	-	Augmented Reality
AMF	-	Adaptive Median Filter
ASW	-	Adaptive Support Weight
ASMF	-	Adaptive Switching Median Filter
AW	-	Adaptive Window
BF	-	Bilateral Filter
BP	-	Belief Propagation
CLAHE	-	Contrast Limited Adaptive Histogram Equalization
CPU	-	Central Processing Unit
CN	-	Census Transform
CNN	-	Convolutional Neural Network
DAMF	-	Different Applied Median Filter
DP	-	Dynamic Programming
FM	-	Feature Matching
FW	-	Fixed Window
GC	-	Graph Cut

GF	-	Guided Filter
GM	-	Gradient Matching
GPU	-	Graphical Processing Unit
LPS	-	Local Plane Sweep
LR	-	Left Right
LRC	-	Left Right Consistency
MF	-	Median Filter
MRF	-	Markov Random Field
MSDF	-	Modified Switching Median Filter
MST	-	Minimum Spanning Tree
NCC	-	Normalised Cross Correlation
PP	-	Post Processing
PSMF	-	Progressive Switching Median Filter
RANSAC	-	Random Sample Consensus
RT	-	Rank Transform
SAD	-	Sum Absolute Differences
SCL	-	Scattered Control Landmarks
SD	-	Squared Differences
SGM	-	Semi Global Method
SLIC	-	Simple Linear Iterative Clustering
SPP	-	Spatial Pyramid Pooling
SSD	-	Sum Squared Differences
SVDM	-	Stereo Vision Disparity map
TAD	-	Truncated Absolute Differences
WTA	-	Winner Take All

- WGIF - Weighted Guided Image Filtering
- ZNCC - Zero Normalised Cross Correlation



LIST OF SYMBOLS

b	-	Baseline of stereo camera
d	-	Disparity
f	-	Focal length of stereo camera
m	-	Gradient magnitude
p	-	Targeted Pixel
q	-	Neighbouring Pixels
w	-	Window size
z	-	Information of depth
E	-	Edge connects the vertices
I	-	Image
V	-	A set of vertices
(i, j)	-	Coordinates for neighbouring pixels
(x, y)	-	Coordinate of targeted pixel
$ w $	-	number of pixels in support window
m_l	-	Gradient magnitude on left image
m_r	-	Gradient magnitude on right image
w_k	-	size for the support window
G_x	-	Gradient in horizontal direction
G_y	-	Gradient in vertical direction
I_l	-	Left image

- I_p - pixel of interest at reference image
- I_q - neighbouring pixels at reference image
- I_r - Right image
- μ_k - Mean value
- σ_c - Colour similarity parameters
- σ_s - Spatial distance
- σ_k - Variance value



LIST OF PUBLICATIONS

Journal (Scopus Index)

1. Hamzah, R. A., Wei, M. Y., and Anwar, N. S. N., 2020. Development of Stereo Matching Algorithm based on Sum of Absolute RGB Color Differences and Gradient Matching. *International Journal of Electrical and Computer Engineering*, 10(3), pp.2375 - 2382.
2. Hamzah, R. A., Wei, M. G. Y., and Anwar, N. S. N., 2019. Stereo Matching based on Absolute Differences for Multiple Objects Detection. *Telkomnika*, 17(1), pp.261 - 267.

Proceeding (Scopus Index)

1. Hamzah, R. A., Wei, M. G. Y. and Anwar, N. S. N., 2020. Depth Estimation Based on Stereo Image Using Passive Sensor. In *Advances in Electronics Engineering*, Springer, pp.127 - 136.
2. Gan, Y., Hamzah, R. A., & Anwar, N. N. (2018). Local Stereo Matching Algorithm based on Pixel Difference Adjustment, Minimum Spanning Tree and Weighted Median Filter. In *2018 IEEE Conference on Systems, Process and Control (ICSPC)*, pp. 39 - 43.

Journal (Non-Scopus Index)

1. Gan, Y., Hamzah, R. A., and Anwar, N. N., 2019. A New Structure of Stereo Algorithm Using Pixel Based Differences and Weighted Median Filter. *International Journal of Human and Technology Interaction (IJHaTI)*, 3(2), pp. 23 - 28.



CHAPTER 1

INTRODUCTION

This chapter provides a brief introduction of stereo vision technology and is divided into six sections. Here, section 1.1 introduces the background of stereo vision technology. Next, section 1.2 briefly explains some applications for stereo vision technology. Then, section 1.3 explains the problem statements and section 1.4 presents the objectives of this research. After that, section 1.5 describes the scope of this research and finally, section 1.6 provides the overall outline of the thesis.

1.1 Background

With a pair of eyes, humans are able to estimate depth and identify obstacles. Then, animals with binocular vision, which are normally hunters, uses their eyes to estimate depth for hunting. For living organism, the ability to estimate depth is naturally implemented by the brain but not for a computer. However, the ability to recognize the depth can be modelled mathematically (Bhatti, 2012) using the stereo vision system. Stereo vision technology uses two cameras, where they are aligned horizontally, mimicking the human eyes, the left side and the right side, which are presented in Figure 1.1, obtained from Hariyama et al. (2008). When a scene is captured, two input images are produced, the left image and the right image. Then these input images also provided two different viewpoints from a single scene. In stereo vision, a minimum of two viewpoints are need and these inputs images (left image and right image), represents those viewpoints. By using the mathematical models and computer computation, the input images will correspond with each other and produce disparity map. In addition, the

disparity map is later used for three dimensional (3D) reconstruction to obtain depth map. Hence, a good accuracy disparity map is always preferable.

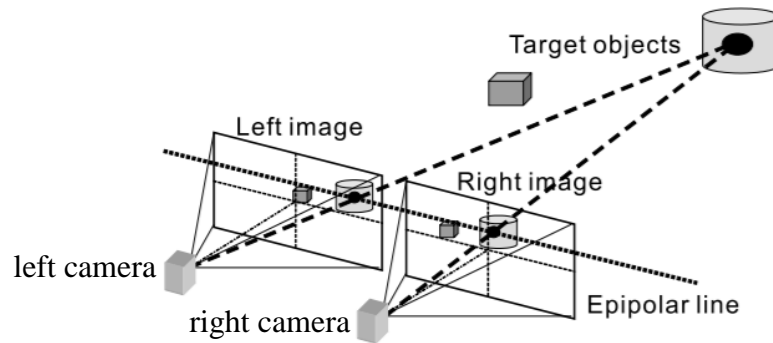


Figure 1.1: Stereo vision system

The basic of a stereo vision is presented in Figure 1.2, obtained from Fahmy et al. (2013). $P(x,y,z)$ represents Scene Point, P where is then observed at points $P_l(x_l, y_l)$ and $P_r(x_r, y_r)$ at the left image and right image, respectively. Then, f represents focal length and b represent baseline; the distance between the two cameras.

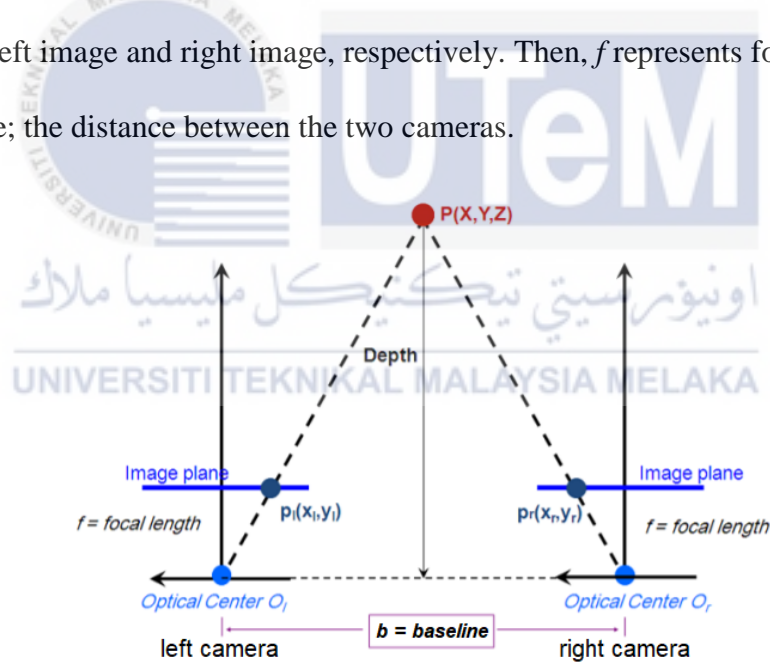


Figure 1.2: Stereo vision basics

The difference in position of the corresponding points for the respective images is defined as disparity, d and the respective images refer to the input images, left image and right image. The equation of disparity, d obtained from Adi and Widodo (2017), is presented in equation (1.1),

$$d = x_l - x_r \quad (1.1)$$