UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Faculty of Electronic and Computer Engineering

DEVELOPMENT OF STEREO-MATCHING ALGORITHM BASED ON ADAPTIVE WEIGHTED PREDICTION

Siti Safwana binti Abd Razak

Master of Science in Electronic Engineering

2019

DEVELOPMENT OF STEREO-MATCHING ALGORITHM BASED ON ADAPTIVE WEIGHTED PREDICTION

SITI SAFWANA BINTI ABD RAZAK

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

Faculty of Electronic and Computer Engineering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2019

DECLARATION

I declare that this thesis entitled "Development of Stereo-Matching Algorithm Based on Adaptive Weighted Prediction." 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 the candidature of any other degree.

Signature	: ………………………………………….......
Name	: SITI SAFWANA BINTI ABD RAZAK
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 : DR. MOHD AZLISHAH BIN OTHMAN

Date : …………………………………………...

DEDICATION

Dedicated to ALLAH Almighty, my loving son Danial Haziq and my family for your infinite and unfading love, sacrifice, patience, encouragement and best wishes.

ABSTRACT

In computer vision field, one of most active topics in research is estimation of depth map which is Stereo Matching (SM) process and it's also known as Stereo Vision Disparity Map (SVDM). The real challenge in SM is to get high accuracy of disparity map. Matching cost computation produces high noise of disparity map and input images contains low texture area and repetitive pattern also leads to high error on development of disparity map. Other than that, wrong information of each pixel of disparity map which can affect the accuracy of SVDM. Therefore, to overcome the causes of effected accuracy, new Stereo Matching Algorithm (SMA) based on Adaptive Weighted Bilateral Filter (AWBF) was introduced together with characterize the SMA based on quantitative and qualitative measurements and produced SMA performance were evaluate using standard taxonomy of SM. This thesis proposes an algorithm to handle the limitations. Firstly, pre-processing used Sobel Filter was added at initial step to compensate photometric distortion of input images. Then for matching cost, the proposed SMA combine one of matching cost method with some threshold adjustment to reduces the radiometric distortions. The Sum of Absolute Different (SAD) is the matching cost's method and some threshold adjustment were used in this thesis where the SAD's window size of 11 and threshold value of 0.8 was selected based on the experimental results. Secondly, to overcome low texture area and repetitive pattern limitation, this thesis present AWBF at cost aggregation stage where AWBF's radius or window size of 19, spatial adjustment of 17 and disparity similarity value of 0.3 was selected. This process is introduced to preserve and improve the object boundaries. Finally, this thesis present Outlier Detection, window size of 5 for Median Filter and fill-in invalid disparity to utilize the last stage of SMA to handle the accuracy of disparity map in regions of occluded, repetitive and low texture. The experimental result on the proposed algorithm is able to reduce 17.4% of weighted average error for *all* and 9.62% of weighted average error for *nonocc* (nonoccluded) compared to others Stereo Matching Algorithm without the proposed framework. This framework experimental result was also compared with other methods which located in standard benchmarking dataset from the Middlebury. New SMA was developed in this thesis based on AWBF. The weighted average error of disparity map in both *all* and *nonocc* attribute are reduced based on the quantitative and qualitative measurements. Comparison of this framework continue in some of the state-of-the-arts algorithm in the literature and the results is outperformed based on the proposed algorithm experimental result.

ABSTRAK

Dalam bidang penglihatan komputer, salah satu topik yang paling aktif dalam penyelidikan ialah penaksiran peta kedalaman yang merupakan proses Pemadanan Stereo (SM) dan ia juga dikenali sebagai Peta Perbezaan Penglihatan Stereo (SVDM). Cabaran sebenar di dalam SM adalah untuk mendapatkan ketepatan peta perbezaan yang tinggi. Pemadanan kos pengiraan menghasilkan peta perbezaan yang tinggi hingar and gambar-gambar input yang mengandungi kawasan tekstur rendah dan corak berulang juga menerajui kepada ralat yang tinggi dalam penghasilan peta perbezaan. Selain daripada itu, informasi salah dari setiap piksel dari peta perbezaan boleh menjejaskan ketepatan SVDM. Lantaran itu, untuk mengatasi penyebab kesan ketepatan, Algoritma Pemadanan Stereo (SMA) baru berdasarkan Penapisan Bilateral Berasaskan Penyesuaian (AWBF) telah diperkenalkan bersama ciri-ciri SMA berdasarkan pengukuran kuantitatif dan kualitatif dan prestasi SMA yang di peroleh telah di nilai mengunakan SM piawai taksonomi. Tesis ini mencadangkan satu algoritma untuk mengendali batasan-batasan tersebut. Pertama, pra-pemprosesan mengunakan Penapisan Sobel yang telah ditambah di permulaan langkah untuk mengimbangi herotan fotometrik dari gambar-gambar input. Kemudian untuk pemadanan kos, SMA yang dicadangkan menggabungkan satu daripada kaedah pemadanan kos bersama sedikit pelarasan ambang untuk mengurangkan herotan radiometric. Jumlah Perbezaan Mutlak (SAD) merupakan kaedah kos padanan dan pelarasan ambang yang telah digunakan di dalam tesis ini di mana saiz tingkap SAD ialah 11 dan 0.8 dari nilai ambangan telah dipilih berdasarkan hasil eksperimen. Kedua, untuk mengatasi batasan kawasan tekstur rendah dan corak berulang juga, tesis ini mempersembahkan AWBF di peringkat kos agregasi di mana dari jejari atau saiz tingkap AWBF ialah 19, pelarasan spatial ialah 17 dan persamaan nilai perbezaan ialah 0.8 telah dipilih. Proses ini diperkenalkan untuk mengekalkan dan memperbaiki sempadan objek. Akhir sekali, tesis ini membentangkan Pengesanan Lebihan, saiz tingkap 5 untuk Penapis Median dan mengisi kekurangan perbezaan yang tidak sah untuk menggunakan peringkat akhir SMA untuk mengendalikan ketepatan peta perbezaan di kawasan-kawasan sempit, berulang dan kawasan bertekstur rendah. Hasil keputusan eksperimen pada algoritma yang dicadangkan dapat mengurangkan 17.4% purata ralat wajaran untuk semua dan 9.62% purata ralat wajaran untuk tidak bertutup dibandingkan dengan Stereo Pemadanan Algoritma tanpa rangka kerja yang dicadangkan. Rangka kerja keputusan eksperimen juga dibandingkan dengan kaedah-kaedah lain dimana ia terletak di penanda aras piawai set data dari Middlebury. SMA baru yang telah dibangunkan di dalam tesis ini berdasarkan AWBF. Purata ralat wajaran dari peta perbezaan di dalam kedua-dua atribut semua dan tidak bertutup telah berkurang berdasarkan pengukuran kuantitatif dan kualitatif. Perbandingan rangka kerja ini berterusan dalam beberapa algoritma terkini di dalam lateratur dan hasil lebih baik berdasarkan hasil eksperimen algoritma yang dicadangkan.

ACKNOWLEDGEMENTS

Alhamdulillah, Here, I was grateful and like to express my heartfelt gratitude and appreciation towards some individuals who have been assisted me throughout my project, Development of Spatiotemporal Stereo-Matching Algorithm Based on Adaptive Weighted Prediction. The completion who made it possible and who constant support has crowned our efforts with success. First and foremost, I would love to convey my sincere acknowledgement to my supervisor Dr. Mohd Azlishah Bin Othman from the Faculty of Electronic and Computer Engineering Universiti Teknikal Malaysia Melaka and En. Ahmad Fauzan Bin Kadmin from Faculty of Engineering Technology Universiti Teknikal Malaysia Melaka for their vital supervision, advice, support, suggestion and encouragement towards the completion of this thesis. Million thanks to Allah, son, my beloved father, mother and siblings for their constant moral supports throughout the completion of this thesis. Lastly, thank you to everyone who had been of the utmost parts of this project.

Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

1D	-	One-dimensional
2D	-	Two-dimensional
3D	-	Three-dimensional
AD	-	Absolute Differences
ASW	-	Adaptive Support Weight
AW	-	Adaptive Window
BP	-	Belief Propagation
CPU	-	Computer Processor Unit
CT	-	Census Transforms
DP	-	Dynamic Programming
DSI	-	Disparity Space Image
FPGA	-	Field Programmable Gate Array
FW	-	Fix Window
GC	-	Graph Cut
GCP	-	Ground Control Points
GCSF	-	Growing Scene Flow
GD	-	Gradient Matching
GF	-	Guided Filter
GPU	-	Graphic Processor Unit

HBDS	-	Hierarchical Bilateral Disparity Structure
LPS	-	Low Pass Filter
LR	-	Left Right
MeanShiftSeg	-	Shift Image Segmentation
MF	-	Median Filter
MRF	-	Markove Random field
MVS	-	Multi View Stereo
MW	-	Multiple Window
NCC	-	Normalsed Cross Correlation
RANSAC	-	Random Sample of Consensus
RGB	-	Red Green Blue
RT	-	Rank Transforms
SAD	-	Sum Absolute Differences
SD	-	Squared Differences
SGM	-	Semi Global
SMA	-	Stereo Matching Algorithm
SSD	-	Sum Square Differences
ToF	-	Time-of-Flight
V2	-	Version Two
WMF	-	Weighted Median Filter
WTA	-	Winner Take All

LIST OF PUBLICATIONS

Abd Razak, S.S., Othman, M.A., and Kadmin, A.F., 2019. The Effect of Adaptive Weighted Bilateral Filter on Stereo Matching Algorithm. *International Journal of Engineering and Advanced Technology (IJEAT)*, 8(3), pp.284-287.

CHAPTER 1

INTRODUCTION

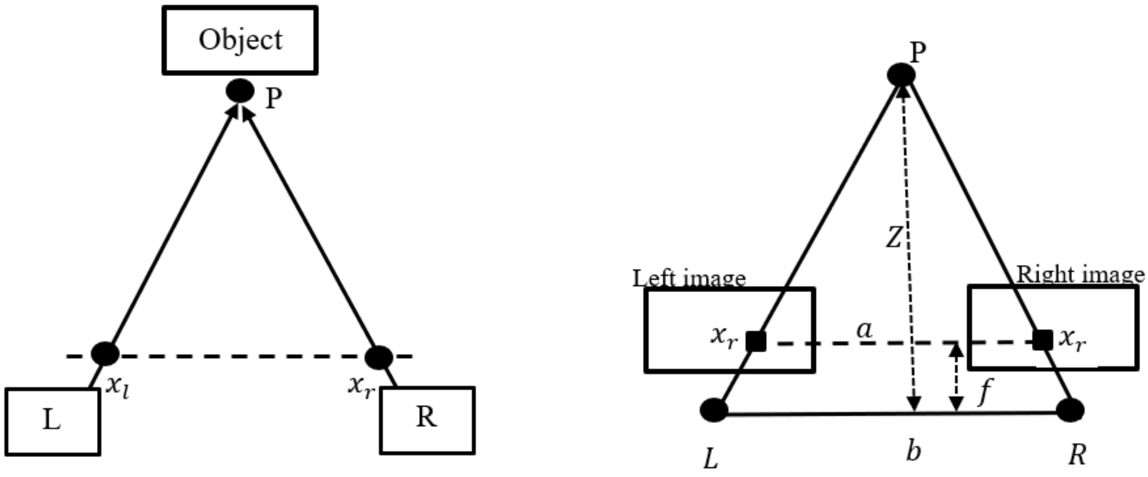
This chapter is divided into seven main sections. The fundamentals of the mathematical model will be explained in Section 1.1, along with the background of stereo vision system. Section 1.2 will be the overview and application examples of stereo matching system. Then, Section 1.3 introduces the research challenges and Section 1.4 provides the problem statement. Section 1.5 presents the objectives of the research while Section 1.6 will be about the scopes of the research and thesis. Section 1.7 discusses on the structure of the thesis.

1.1 Background of stereo matching

Stereo Matching is one of Computer Vision field. Estimating disparity maps' depth one of most active research conducted in the past years under Stereo Matching. New Stereo Matching (SMA) was develop in order to fulfil the main objective SMA which are to develop the most accurate disparity maps with lowest error as possible and minimum processing time. In this thesis, we share the fundamental objectives of SMA with other researchers.

Humans were gifted with stereoscopic fusion vision from the eyes. We are able to recognise depth from our view easily. The input from both eyes were automatically processed by the brain and the depth of surrounding area were realised instantly. To mimic the depth-recognition the function of human vision processed by the brain, the depth estimation can be modelled mathematically from a scene of stereoscopic fusion (Bhatti, 2012). This model is called as stereo vision system, which is one of the most popular and active research areas in computer vision field. Stereo vision is practically an experiment of stereoscopic view which utilises two stereo cameras (i.e., left and right) to perceive one scene from two different point of view. From the result, it produces mathematical algorithm called stereo matching algorithm. This algorithm processes the two viewpoints to extract visual depth data. Two-dimensional (2D) data from stereo vision scene can be processed into Three-dimensional (3D) information. The matching process is the shifting method of the part or pixels to match with each the other scene which can produce the depth information value. This information value is known as disparity (Xu and Zhang, 1996). Depth of the images depends on the disparity value, where a higher disparity value indicates a closer distance from the object to the cameras while lower or near zero disparity value indicates the object is far from the cameras. The pixel location of left and right image was also located at same place.

The basic mathematical models of stereo vision system concept (Virgin and Wagg, 2012) is shown in Figure 1.1. Figure 1.1(a) shows the stereo sensor (i.e., L=left camera and R=right camera) detects an object at P point with the same viewpoint. The x_l and x_r is the location of left and right camera placement respectively. The horizontal dotted line is the plane projection of the stereo camera location. Figure 1.1(b) shows the geometry version of Figure 1.1(a) where x_l and x_r at the plane view produce left and right images respectively from left camera and right camera.



(a) Stereo vision sensor with an object detection at point P

(b) Stereo vision geometry

Figure 1.1: A basic concept of stereo vision system which contains a point detection and its translation model

The distance between L and R represent by b and matching pixels distance located from x_l and x_r is represented by a. Essentially, depth computation is based on the triangular principle angle $(\angle L, P, R)$ and $(\angle x_l, P, x_r)$. The angle must be similar to enable disparity to be produced. The relation between depth and baseline is explained in Equation (1.1):

$$\frac{b}{Z} = \frac{a}{Z-f} = \frac{(b-x_l)+x_r}{Z-f} \tag{1.1}$$

In the equation above, b represents the baseline of the stereo camera and Z denotes the depth or distant. The x_l and x_r represent the coordinates of plane projection on matching pixel and f represent focal length of the stereo camera. The final depth estimation Z, is gained through Equation (1.2):

$$Z = \frac{bf}{x_l - x_r} = \frac{bf}{d} \tag{1.2}$$

Disparity value is represented by $d = x_l - x_r$. This value is known as disparity map where it can be plotted in 2D format. The disparity map contains important information such as depth value for stereo vision applications. This map is the result of an algorithm called Stereo Matching Algorithm based on Scharstein and Szeliski (2002) taxonomy. In their taxonomy, there are three major categories of methods of stereo matching algorithm (i.e., global, Semi Global (SGM) and local method). They also emphasise on four main frameworks for stereo matching algorithm which are:

- Step 1: Matching Cost Computation
- Step 2: Cost Aggregation
- Step 3: Disparity Selection
- Step 4: Disparity Refinement

These four steps will be extensively discussed in Chapter 2.

1.2 Applications of stereo vision

There are multiple-applications of stereo vision system which covers a wide range of filed such as:

i. Surveillance

   Vehicle detection performed by images acquired from stereo matching/vision system implemented by Bertozzi et. al. (2000).

ii. Medical

   Ability to derive depth information are an advantage of stereo matching which could support neuronal mechanisms. Read (2014) mentioned how it disrupted in strabismus.

iii. Robotic and automation

Both robotic and automotive field requires an embedded system of depth and disparity map information which was proven by Humenberger et. al. (2010).

iv. 3D surface reconstruction

Dellepiane et. al. (2012) stated that analysis of 3D surface reconstruction as one of the important processes to determine the status and objects' condition for example in archaeological artefact observation.

1.3 Research challenges

The main purposed of stereo matching experiments conducted widely by researchers are to achieve high accuracy of disparity map and produced minimum processing time to enable it to run in real-time system. There are several factors that may contribute to the effectiveness of stereo matching algorithm disparity map. The challenges are explained as follows by using images taken from Unger and Navab (2018):

i. Occlusions

The most difficult challenge of stereo matching algorithm is occlusions. From Figure 1.2(a) in the left image, the region of divider is widely visible compared to Figure 1.2(b) where occlusions occurred in the red box. This happened because of geometric displacement between the two cameras. Occluded region is extremely hard to estimate or filled in disparity map because of unknown shapes and objects behind the occluded area.