

國立交通大學

資訊工程系

博士論文

由二維影像序列建構三維物體幾何模型

**Constructing 3D Object Models from Image
Sequences**

研究生：周宏隆

指導教授：陳 稔 教授

中華民國九十三年七月

由二維影像序列建構三維物體幾何模型
Constructing 3D Object Models from Image Sequences

研究生：周宏隆

Student : Hong-Long Chou

指導教授：陳 稔

Advisor : Zen Chen

國立交通大學
資訊工程系
博士論文



Submitted to Department of Computer and Information Science
College of Electrical Engineering and Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Computer Science and Information Engineering

July 2004

Hsinchu, Taiwan, Republic of China

中華民國九十三年七月

博碩士論文授權書

本授權書所授權之論文為本人在國立交通大學資訊工程學系所九十二學年度第二學期取得博士學位之論文。

論文名稱：由二維影像序列建構三維物體幾何模型

指導教授：陳稔 博士

1. 同意 不同意

本人具有著作財產權之上列論文全文(含摘要)資料，授予行政院國家科學委員會科學技術資料中心(或改制後之機構)，得不限地域、時間與次數以微縮、光碟或數位化等各種方式重製後散布發行或上載網路。

本論文為本人向經濟部智慧財產局申請專利(未申請者本條款請不予理會)的附件之一，申請文號為： ，註明文號者請將全文資料延後半年再公開。

2. 同意 不同意

本人具有著作財產權之上列論文全文(含摘要)資料，授予教育部指定送繳之圖書館及國立交通大學圖書館，基於推動讀者間「資源共享、互惠合作」之理念，與回饋社會及學術研究之目的，教育部指定送繳之圖書館及國立交通大學圖書館得以紙本收錄、重製與利用；於著作權法合理使用範圍內，不限地域與時間，讀者得進行閱覽或列印。

本論文為本人向經濟部智慧財產局申請專利(未申請者本條款請不予理會)的附件之一，申請文號為： ，註明文號者請將全文資料延後半年再公開。

3. 同意 不同意

本人具有著作財產權之上列論文全文(含摘要)，授予國立交通大學與台灣聯合大學系統圖書館，基於推動讀者間「資源共享、互惠合作」之理念，與回饋社會及學術研究之目的，國立交通大學圖書館及台灣聯合大學系統圖書館得不限地域、時間與次數，以微縮、光碟或其他各種數位化方式將上列論文重製，並得將數位化之上列論文及論文電子檔以上載網路方式，於著作權法合理使用範圍內，讀者得進行線上檢索、閱覽、下載或列印。
論文全文上載網路公開之範圍及時間 -

本校及台灣聯合大學系統區域網路： 年 月 日公開

校外網際網路： 年 月 日公開

上述授權內容均無須訂立讓與及授權契約書。依本授權之發行權為非專屬性發行權利。依本授權所為之收錄、重製、發行及學術研發利用均為無償。上述同意與不同意之欄位若未鈎選，本人同意視同授權。

研究生簽名：

(親筆正楷)

周宏隆

學號：8517801

(務必填寫)

日期：民國 93 年 7 月 27 日

1. 本授權書請以黑筆撰寫並影印裝訂於書名頁之次頁。

國家圖書館博碩士論文電子檔案上網授權書

本授權書所授權之論文為本人在國立交通大學資訊工程學系所九十二學年度第二學期取得博士學位之論文。

論文名稱：由二維影像序列建構三維物體幾何模型

指導教授：陳 稔 博士

同意 不同意

本人具有著作財產權之上列論文全文(含摘要)，以非專屬、無償授權國家圖書館，不限地域、時間與次數，以微縮、光碟或其他各種數位化方式將上列論文重製，並得將數位化之上列論文及論文電子檔以上載網路方式，提供讀者基於個人非營利性質之線上檢索、閱覽、下載或列印。

上述授權內容均無須訂立讓與及授權契約書。依本授權之發行權為非專屬性發行權利。依本授權所為之收錄、重製、發行及學術研發利用均為無償。上述同意與不同意之欄位若未鈎選，本人同意視同授權。

研究生簽名：
(親筆正楷)

周宏隆

學號：8517801
(務必填寫)

日期：民國 93 年 7 月 27 日

1. 本授權書請以黑筆撰寫，並列印二份，其中一份影印裝訂於附錄三之一(博碩士論文授權書)之次頁；另一份於辦理離校時繳交給系所助理，由圖書館彙總寄交國家圖書館。

國立交通大學
資訊工程系博士班

論文口試委員會審定書

本校 資 訊 工 程 系 周宏隆 君

所提論文 由二維影像序列建構三維物體幾何模型

合於博士資格水準、業經本委員會評審認可。

口試委員：貝 蘇 亭 韓 裕

洪一平

范國清

賴宗濤

李弘源

李淑望

范榮宏

指導教授：韓 裕

系主任：張明峰

中華民國九十三年七月十一日

Department of Computer Science and Information Engineering
College of Electrical Engineering and Computer Science
National Chiao Tung University
Hsinchu, Taiwan, R.O.C.

Date: Jul. 11, 2004

*We have carefully read the dissertation entitled **3D Object Reconstruction from a Sequence of Images** submitted by **Hong-Long Chou** in partial fulfillment of the requirements of the degree of Doctor of Philosophy and recommend its acceptance.*

Son-kyun

Sen Chen

Wing-Kit

Kuo-Chin Jan

Wang

Muh-Lian

Hao-jian Lee

Jun-Hay Chang

Thesis Advisor: _____

Sen Chen

Chairman: _____

Wing-Kit Chang

由二維影像序列建構三維物體幾何模型

研究生：周宏隆

指導教授：陳稔博士

國立交通大學資訊工程學系

摘要

本論文提出四個利用二維影像序列建構三維物體幾何模型的方法。首先，我們透過物體輪廓線及相機成像中心的連線，產生不同視角的成像空間。透過成像空間的交集，我們可以產生物體的八分樹表示式。因為有限的時間及空間資源限制，傳統上使用者透過指定最大八分樹細分階層來控制八分樹建構程序終止於資源殆盡之前。此方式之缺點在於使用者無法由所設定之參數推估建構出的模型和實際物體的差異。再者，八分樹節點無論投影誤差大小，皆須細分到指定的最大八分樹細分階層。在此論文中，我們提出一個投影誤差上限的概念。在前兩個方法，透過投影誤差的計算及比較，我們分別定義新的灰色八分樹節點。利用此新定義的灰色八分樹節點，我們提出兩個投影誤差控制的快速八分樹建構程序。不只是以實驗證明，在相同的建構模型品質下，我們所提出的方法在空間需求及運算時間上皆優於傳統的八分樹建構方法，我們亦在理論上得到相同的驗證。

在第三個方法中，我們提出一個最佳八分樹遊走演算法。我們利用八分樹節點投影誤差做排序，如此在建構八分樹模型時，可以最大誤差之節點為最優先處理之節點。實驗及理論上均證明所提出之方法在每單位八分樹節點的處理之誤差改進效能上皆能比傳統利用深度優先或橫向優先有較佳之表現。

在第四個方法中，我們提出一個未校正相機的平面物體模型建構方法。相較於前三種方法需要獲知相機內外部參數方可建構物體三維模型，在此方法中，我們提出一個不需事先校正相機參數的物體模型建構方法。並且，針對平面物體在建構，我們捨棄利用易受影像雜訊干擾之傳統點特徵的程序。我們透過平面轉換

矩陣，直接推導相當的投影矩陣。進而推導出物體組成平面的平面方程式。最後再利用平面轉換矩陣及平面方程式將不同的建構結果統合在同一個座標系統中。大量的實驗及雜訊影響分析證明我們的方法是有效且強健的。



Constructing 3D Object Models from Image Sequences

Student : Hong-Long Chou

Advisor : Dr. Zen Chen

Department of Computer Science and Information Engineering
National Chiao Tung University

ABSTRACT

Constructing 3D object models from image sequences has gained tremendous attraction in computer vision for the past few decades. The constructed object geometric model finds many applications including robotic system, computer aided design, virtual reality, digital entertainment and target recognition. There are many methods for constructing the object geometric model. Two types of approaches will be addressed in this dissertation: shape from silhouettes and shape fitting with planes. The shape-from-silhouette method is one of the popular techniques used to construct the object shape model from a sequence of object silhouette images. The object silhouette provides important clues of the object shape for human visual system. From the object silhouette and its corresponding viewing position, one can generate a viewing volume encapsulating the object in the particular view. By intersecting the viewing volumes obtained from all viewing directions, a volumetric representation for the object can be generated. However, it often requires a huge memory space to store the 3D object volume even using a hierarchical data structure like octree. Besides, a long computer time is taken to intersect all the viewing volumes to generate the final octree. Another type of construction methods is to fit the 3D object surface with planes. The feature correspondence and robust estimation of the 3D shape of the object are two of the main problems in this type of approaches.

In this dissertation, four different methods for constructing the 3D model of a real object from image sequences are proposed. In the first method, a new octree-based

subdivision strategy of two novel types of “grey” octants is proposed to speed up the octant subdivision process under a construction quality control. To further expedite the construction process a fast way for computing the 3D-to-2D projection of octant vertices using the information of vanishing point and cross ratio is proposed. In the second method, the octant whose image touches the silhouette is further classified into three types of grey octants: grey-grey, grey-black and grey-white. Then those octants having little intersection with object silhouettes will not be subdivided. This method has a great improvement on the computer processing time and memory storage space for a given construction quality specification. In the third method, a progressive mode instead of a recursive mode is proposed for the octree construction. The octree generation will be implemented using a best-first tree traversal scheme instead of a conventional depth first or breadth first tree traversal scheme. The precedence of octants for the arrangement of subdivision is ranked according to the XOR error between the projected octant image and the object silhouette. The progressive octree construction method generally gives the better visual quality rate of the object rendering effect, compared to the conventional recursive construction method.

The fourth construction method is to fit the object surface with planes. The 3D object model construction usually requires camera calibration beforehand. However, it is not easy for a user to calibrate the camera in circumstances, in particular, in the case of a hand-held camera. The fourth method for constructing the object from an uncalibrated image sequence is proposed. Instead of using the point based feature to construct the 3D information, the planar homography defined over an object planar surface is used to derive the projective geometric model of the object.

All of the four construction methods are tested on real and synthetic images. The performance of the methods is evaluated in terms of visual projection error, memory space and processing time. Analytical analysis on each method is also given. The four methods offer the user more choice in constructing the object geometric model to meet the different application requirements.

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to my advisor, Dr. Zen Chen, for his kind patience, constant encouragement, helpful guidance and inspiration throughout the course of this dissertation. In these years, he has stimulated the research work and also provided an excellent research environment at the Pattern Analysis and Intelligent System (PAIS) Laboratory. He also teaches me how to learn and how to think independently. Especially, he encourages me to challenge what we are used to be. I also express my sincere gratitude to the members of my thesis committee, Professor Lee and Professor Chuang, for their valuable suggestions and comments.

I am grateful to my fellow classmates in the PAIS Laboratory for their assistance and discussion, especially Mr. Y. H. Fang. Special thanks are also due to the alumni of PAIS Lab. It is a pleasure to be with you guys staying up for writing codes and playing baseball or basketball.

A special thank is given to my sister, brother-in-law and brothers for their constant encouragement. Finally, I would like to dedicate this dissertation to my parents. Mr. S. S. Chou and Mrs. M. Hsiao for their love, encouragement and support.

TABLE OF CONTENTS

ABSTRACT (in Chinese).....	vii
ABSTRACT (in English).....	ix
ACKNOWLEDGEMENT.....	xi
TABLE OF CONTENTS.....	xii
LIST OF FIGURES.....	xiv
LIST OF TABLES.....	xvii
CHAPTER 1 INTRODUCTION	1
1.1 Statement of the Problem.....	1
1.2 Survey of Related Research.....	1
1.3 Sketch of the Work	4
1.4 Contribution of the Work.....	5
1.5 Dissertation Organization	6
CHAPTER 2 FAST OCTREE CONSTRUCTION ENDOWED WITH AN ERROR BOUND CONTROLLED SUBDIVISION SCHEME.....	7
2.1 Introduction.....	7
2.2 Fast Octant Projection Computation Using Cross Ratio and Vanishing Point	8
2.3 2D Intersection Test Using the Precomputed Distance Maps of Silhouette Images and Approximation of an Octant Image by a Circle	12
2.4 New Subdivision Strategy and Construction Quality Measure	14
2.5 Experimental Results	15
2.6 Analysis on the New Octree Construction Method	33
2.7 Conclusions and Future Work	38
CHAPTER 3 AN OCTREE CONSTRUCTION METHOD WITH THREE TYPES OF GREY OCTANTS	39
3.1 Introduction.....	39
3.2 Types of New Grey Octants.....	39
3.3 The Octant Subdivision Algorithm of the Second Construction Method.....	40
3.4 The Relation between Object Spatial Resolution and Projection Error Upper Bound.....	40
3.5 Experimental Results	41
3.6 Analytical Analysis.....	53
3.7 Conclusion	60
CHAPTER 4 A PROGRESSIVE OCTREE CONSTRUCTION METHOD WITH A PROJECTION ERROR MEASURE	61
4.1 Introduction.....	61

4.2 Depth-First Tree Traversal Scheme vs. Breadth-First Tree Traversal Scheme	62
4.3 Best-First Tree Traversal Scheme with Octant Sorting Based on Projection Error	64
4.4 Experimental Results	67
4.5 Analytical Analysis on the Performance of the Proposed Method	73
4.6 Conclusion	79
CHAPTER 5 A NOVEL 3D PLANAR OBJECT RECONSTRUCTION FROM MULTIPLE UNCALIBRATED IMAGES USING THE PLANE-INDUCED HOMOGRAPHIES	80
5.1 Introduction.....	80
5.2 Preliminaries and Mathematical Notations for Projective Reconstruction....	82
5.3 Reconstruction of All Visible Planes from a Given Image Pair	85
5.4 Integration of Planes Reconstructed from Different Image Pairs	87
5.5 Computation of Homographies.....	89
5.6 Experimental Results	90
5.7 Conclusions.....	97
CHAPTER 6 SUMMARY AND FUTURE RESEARCH.....	98
6.1 Summary and Conclusions	98
6.2 Topics for Future Research.....	99
REFERENCES	101
VITA.....	106
PUBLICATIONS.....	107

LIST OF FIGURES

Fig. 2.1. The 27 distinct vertices involved in the subdivision of a parent octant including eight A-type vertices (A_0, A_1, \dots, A_7), twelve B-type vertices (B_0, B_1, \dots, B_{11}), six C-type vertices (C_0, C_1, \dots, C_5), and one D-type vertex D_0	9
Fig. 2.2. The four collinear points, p_1, p_2, p_3 and p_∞ , together with their 3D corresponding vertices P_1, P_2, P_3 and P_∞ , define a common cross ratio value.....	10
Fig. 2.3. The illustration of octant subdivision strategy based on the spatial relation between the circle containing the projected octant and the object silhouette.	14
Fig. 2.4. The hardware setup of the system.	15
Fig. 2.5. The schematic diagram of the right cube. The dimensions of the cube are 20 inches in depth (the x-direction), 20 inches in width (the z direction) and 20 inches in height (the y direction).	16
Fig. 2.6. Three distinct images I_0, I_3 and I_7 of the cube taken at 0, 90 and 180 degrees.	16
Fig. 2.7. The comparison between construction results of the synthetic cube obtained by the conventional method and the new method.....	18
Fig. 2.8. The schematic diagram of the tilted cube. The dimensions of the cube are 20 inches in depth (the x-direction), 20 inches in width (the z direction) and 20 inches in height (the y direction).	19
Fig. 2.9. Three distinct images I_0, I_3 and I_7 of the tilted cube taken at 0, 120 and 280 degrees.	19
Fig. 2.10. The comparison between the construction results of the synthetic cube obtained by the conventional method and the new method.	21
Fig. 2.11. Three real objects used in the experiments: (a) a cone, (b) a vase, and (c) a boy sculpture.	21
Fig. 2.12. New views generated from the constructed octree models for (a) a cone, (b) a vase, and (c) a boy sculpture. The octree models are converted to VRML format and rendered using Cosmo© player.	22
Fig. 2.13. The comparison between construction results of the cone obtained by the conventional method and the new method.....	26
Fig. 2.14. The comparison between construction results of the vase obtained by the conventional method and the new method.....	27
Fig. 2.15. The comparison between construction results of the boy sculpture obtained by the conventional method and the new method.....	28
Fig. 2.16. The total leaf node number vs. XOR error plot of the construction results obtained by the new method and the conventional method.	31

Fig. 2.17. The computation time vs. XOR error plot of the construction results obtained by the new method and the conventional method.....	32
Fig. 3.1. The depiction of the relation between the object spatial resolution and the projection error upper bound.	41
Fig. 3.2. The comparison between construction results of the synthetic cube obtained by method 1 and method 2.....	43
Fig. 3.3. The comparison between construction results of the cone obtained by the conventional method and the new method.....	49
Fig. 3.4. The comparison between construction results of the vase obtained by the conventional method and the new method.....	50
Fig. 3.5. The comparison between construction results of the boy sculpture obtained by the conventional method and the new method.....	51
Fig. 3.6. (a) One of the input image to the second method. (b) The new view generated from the constructed octree model.	52
Fig. 3.7. (a) One of the input image to the second method. (b) The new view generated from the constructed octree model.	52
Fig. 3.8. The depiction of centroids of the bounding circles for an octant and its parent octant and child octant. (Assume $r_L < p < r_{L-1} = 2r_L$).....	56
Fig. 3.9. The spatial relation depiction of bounding circles for an octant and its two sub-octants	58
Fig. 4.1. A typical octree structure.....	63
Fig. 4.2. The visiting ordering of the black nodes by the depth-first tree traversal scheme for the octree in Fig. 4.1.....	63
Fig. 4.3. The visiting ordering of the black nodes by the breadth-first tree traversal scheme for the octree in Fig. 4.1.....	64
Fig. 4.4. The modified insertion process which avoids inserting GB^1 octant into the priority queue.	65
Fig. 4.5. The modified insertion process which avoids inserting GB^2 or GW^2 octant into the priority queue.	65
Fig. 4-6. The algorithm of the proposed progressive octree construction method.	66
Fig. 4.7. The constructed octree model of the first 11 octants processed by the DFS method when viewed from the 1 st viewing direction.....	68
Fig. 4.8. The constructed octree model of the first 11 octants processed by the BFS method when viewed from the 1 st viewing direction.....	68
Fig. 4.9. The constructed octree model of the first 11 octants processed by the BestFS method when viewed from the 1 st viewing direction.....	68
Fig. 4.10. The plot of XOR vs. subdivided octant number of the octree model for the vase constructed from the DFS, BFS and BestFS methods.	69

Fig. 4.11. The plot of XOR vs. Subdivided octant number of the octree model for the boy sculpture constructed from the DFS, BFS and BestFS methods.	70
Fig. 4.12. The priority queue length when constructing the model of the vase using the BestFS_m1 method.	71
Fig. 4.13. The priority queue length when constructing the model of the vase using the BestFS_m2 method.	72
Fig. 4.14. The priority queue length when constructing the model of the boy sculpture using the BestFS_m1 method.	72
Fig. 4.15. The priority queue length when constructing the model of the boy sculpture using the BestFS_m2 method.	73
Fig. 4.16. The relation of radii of the bounding circles for octant O_i and O_j and their $max_proj_error(O_i)$ and $max_proj_error(O_j)$	75
Fig. 5.1. The flow diagram of our reconstruction method.	82
Fig. 5.2. The schematic diagram of the tower. The dimensions of the tower are 40 inches in depth (the x-direction), 40 inches in width (the z direction) and 180 inches in height (the y direction).	91
Fig. 5.3. Three distinct images I_1 , I_2 and I_3 taken at a distance of about 500 inches. The visible planes in the three images are Π_{Gr} , Π_A , Π_B , Π_E , Π_F in I_1 and I_2 , and Π_{Gr} , Π_B , Π_F , Π_C , Π_G in I_3	91
Fig. 5.4. The effects of the uniformly generated reference plane coefficient vectors and the noise at different levels on the reconstruction result.	93
Fig. 5.5. The indices of the vertices and planes of the object.	95
Fig. 5.6. New views of the reconstructed object with texture mapping.	96

LIST OF TABLES

Table 2.1. The comparison of projection computation of octant vertices between the new method and the conventional method.....	12
Table 2.2 Intersection relation between the projected octant image and a silhouette image.....	13
Table 2.3. Numbers of black, grey-black, grey-grey, and white nodes of the cube generated by our octree construction method with a specified projection error bound.....	17
Table 2.4. Numbers of black, grey, and white nodes of the cube generated by the conventional octree construction method with a fixed subdivision level.....	17
Table 2.5. Numbers of black, grey-black, grey-grey, and white nodes of the cube generated by our octree construction method with a specified projection error bound.....	18
Table 2.6. Numbers of black, grey, and white nodes of the cube generated by the conventional octree construction method with a fixed subdivision level.....	20
Table 2.7. Numbers of black, grey-black, grey-grey, and white nodes of the cube generated by our octree construction method with a specified projection error bound.....	20
Table 2.8. Numbers of black, grey, and white nodes of the cone generated by the conventional octree construction method with a fixed subdivision level.....	22
Table 2.9. Numbers of black, grey-black, grey-grey, and white nodes of the cone generated by our octree construction method with a specified projection error bound.....	23
Table 2.10. Numbers of black, grey, and white nodes of the vase generated by the conventional octree construction method with a fixed subdivision level.....	23
Table 2.11. Numbers of black, grey-black, grey-grey, and white nodes of the vase generated by our octree construction method with a specified projection error bound.....	24
Table 2.12. Numbers of black, grey, and white nodes of the boy sculpture generated by the conventional octree method with a fixed subdivision level.....	24
Table 2.13. Numbers of black, grey-black, grey-grey, and white nodes of the boy sculpture generated by our octree construction method with a specified projection error bound.....	25
Table 3.1. Number of the black, grey-black, grey-grey, grey-white and white octants of the constructed cube generated by method 2.....	42
Table 3.2. Number of black, grey-black, grey-grey and white octants of the cube generated by method 1.....	42

Table 3.3. The radius range of the circle containing the octant projection at different level.....	44
Table 3.4. Number of the black, grey-black, grey-grey, grey-white and white octants of the constructed cone generated by the second construction method.	44
Table 3.5. Number of black, grey-black, grey-grey and white octants of the cone generated by the first construction method.....	45
Table 3.6. Number of grey-black and grey-white leaf octants of the cone generated by applying the second method octant type evaluation criteria on grey-black octants in Table 3.3.	45
Table 3.7. Number of the black, grey-black, grey-grey, grey-white and white octants of the constructed vase generated by the second construction method.	46
Table 3.8. Number of black, grey-black, grey-grey and white octants of the vase generated by the first construction method.....	46
Table 3.9. Number of grey-black and grey-white leaf octants of the vase generated by applying the second method octant type evaluation criteria on grey-black octants in Table 3.6.	47
Table 3.10. Number of the black, grey-black, grey-grey, grey-white and white octants of the constructed boy sculpture generated by the second construction method.	47
Table 3.11. Number of black, grey-black, grey-grey and white octants of the boy sculpture generated by the first construction method.	48
Table 3.12. Number of grey-black and grey-white leaf octants of the boy sculpture generated by applying the second method octant type evaluation criteria on grey-black octants in Table 3.9.	48
Table 5.1. 3D object centered coordinates of the tower feature points.....	92
Table 5.2. The statistics of the distance errors of the reconstruction results.	92
Table 5.3. The mean errors of the reconstructed Euclidean point positions for different setups.....	94
Table 5.4. The estimated angles between the object planes and the ground plane.....	96

CHAPTER 1

INTRODUCTION

1.1 Statement of the Problem

The inference of 3D object geometric model from multiple object images is a key problem in computer vision which finds many applications including robotic system, computer aided design, virtual reality, digital entertainment and target recognition. There are a variety of methods in the literature to construct such a 3D object model. They are characterized by different features such as:

- (1) image data acquisition mode: active, passive, ...
- (2) image processing unit: point, edge, region, contour, texture, shading...
- (3) object geometric model: voxel-based, surface-based, point-based, ...
- (4) number of images needed for construction: stereo pairs, triplets, multiple, ...
- (5) camera parameter setting: calibrated, uncalibrated, ...

The methods of concern in this dissertation fall under the category characterized by

- (1) passive sensing,
- (2) contour- or region-based,
- (3) voxel- or surface-based object model,
- (4) multiple images required, and
- (5) camera-calibrated or -uncalibrated.

The existing methods will be reviewed first in order to point out the drawbacks of these methods. Then new methods will be described which overcome most of the drawbacks discussed, if not all of the drawbacks are removed. Theoretical analysis on the methods will be provided.

1.2 Survey of Related Research

Many 3D structure construction methods have been proposed. They can be classified into:

- (A) Active Vision

(B) Passive Vision

The techniques proposed in the active vision class project light pattern onto the object. Then the 3D triangulation or time of light is calculated to estimate the depth information [1]-[10]. Active vision 3D construction methods provide high accuracy results. However, additional procedures are required to set up and calibrate the pattern projection device. Besides, the range of working area is limited to the brightness of the projected light pattern onto the object or scene.

On the contrary, passive vision techniques do not require to project light pattern on the surface to infer the 3D information. Features from the images of the object surfaces are extracted and then used to infer the 3D information. The object silhouette provides an important visual clue which is used to construct the 3D structure of the object. Besides, in the physical world (especially the man-made world) planar surfaces such as walls, windows, table, roof, road, and terrace can be found in the indoor as well as the outdoor scenes. Another task is to reconstruct the 3D planar surfaces in a scene from multiple uncalibrated images taken by a camera placed at different viewpoints. In the following paragraphs, existing methods using the silhouette or planar surfaces information for inferring 3D information are reviewed.

The reconstruction methods using the silhouette information, usually named the silhouette from silhouette (SFS) method, can be further divided into two groups. The first group of methods [11]-[12] applies the differential geometry constraints onto the extracted object silhouettes and generates 3D curves corresponding to the object silhouettes. However, at least three consecutive silhouette images are required to construct the 3D curves. Besides, the change of the object silhouettes should be slightly different, thus a large number of views are required to construct a complete object. Another group of the SFS methods obtain the 3D geometry of the object using the volume intersection techniques. Different structures [13]-[30] have been proposed to represent the initial volume used to intersect with the space that the object occupies. Recently an efficient image-based approach was proposed [66] to compute the so called visual hull of an object to represent the volume where the object occupies. The operation of the method was based on the intersection of the epipolar lines with the silhouette images one by one. The performance of the method is highly dependent on the resolution of the image. Besides, additional rendering techniques, like pixel splating, are required for rendering the new images. In [27][65] methods are proposed to represent the $N \times N \times N$ voxels to represent the initial volume. They then projected

each voxel onto the images and applied the color consistency check to determine whether the voxel belongs to the object surface or not. The number of N will affect the accuracy of the final result of the constructed model. The larger the N , the better the reconstructed model. However, it will take much time to do the voxel projection and color consistency check. An octree is one of the volumetric representations of an object geometric model. It is well known in the field of computer vision that an octree can be constructed from multiple silhouettes taken from different viewpoints of an object [13]-[24]. They can be divided into two classes: 3D space approach [13]-[16] and 2D space approach [17]-[24]. Both approaches recursively subdivide a partially occupied octree octant into smaller octants until all the generated octants are entirely inside or outside the object. The overlapping between an octree octant and the object is determined by an intersection test. The above two approaches differ in their intersection tests. In the 3D space approach the octant under examination is tested against the conic view volume formed by the individual silhouette and the center of projection for each viewpoint. In the 2D space approach the projection image of the octant is tested against the silhouette for each viewpoint. Generally speaking, the 2D space approach is performed in a space with a lower dimension, so it is more efficient than the 3D space approach [17]. Since the octant number grows exponentially with the subdivision level, an upper bound is usually imposed on the number of subdivision levels to avoid the insufficient memory space problem. However, a larger value of subdivision levels leads to a better construction result. It is generally difficult to make a good balance between memory space and construction quality. In this dissertation, we propose a new subdivision strategy which is governed by the degree of overlapping between the generated octant and the object. We shall consider making effective use of the octant subdivisions to improve the overall system performance.

Recently, much research efforts are involved in developing the 3D geometry reconstruction of an object without calibrating the camera a priori. In general, the methods for 3D projective or uncalibrated reconstruction [33][36][37][40][41][50] are point-based. They estimate the fundamental matrix from a sufficient number of corresponding point pairs first, and then derive the epipole and the canonical geometric representation for projective views using the fundamental matrix. Then, for each pair of corresponding points, they use a triangulation technique or bundle adjustment technique to compute the 3D point coordinates in the projective space. Finally, for the determination of the uncalibrated planar scene structure

[34][39][46][47][49][54][56][60], the 3D points found are fitted by planes. However, it is desirable to derive the 3D planar scene structure in terms of plane features in the images directly, for these features are more reliable than the point or line features [49]. The estimation of the 3D projective planar structure based on the projected plane feature information exclusively has not yet received much attention, although it is known that the corresponding projected plane regions in a pair of stereo images induce a homography. It is also known that homographies are useful to many other practical applications including:

- (a) Fundamental matrix estimation or canonical projective geometry representation [48][49].
- (b) 2D image mosaicing or view synthesis [43].
- (c) Plane + parallax analysis [34][36][56].
- (d) Planar motion estimation and ego-motion [45][54][60].

Recently, two methods have been proposed for the 3D projective reconstruction of planes and cameras. The first method assumes all planes are visible in all images and the second method assumes a reference plane is visible in all images [51][52]. In practice, it is not realistic to have all planes or even one plane visible in all images unless a very large ground plane is available. When there is no reference plane visible in all images, the reconstruction problem cannot be formulated within a common projective space and the reconstruction results will be inevitably obtained in different projective spaces.

1.3 Sketch of the Work

In this dissertation, four different methods for constructing the 3D model of a real object from image sequences are proposed. In the first method, a close look at the octant subdivisions reveals that the subdivided octants have varying projection errors; some may be very large and some may be rather small. We shall consider making effective use of the octant subdivisions to improve the overall system performance. In this method, we propose a new subdivision strategy which is governed by the degree of overlapping between the generated octant and the object. This degree of overlapping will be measured by the maximum 2D projection error of the projected octant image relative to the object silhouette for all viewpoints. Those octants with a projection error exceeding the pre-specified upper bound will be subdivided

recursively until all the octants satisfy the maximum constraint imposed on the projection error. Furthermore, we will also present a fast computation of the 2D projection and a new intersection test to reduce the computer processing time.

An extension to the first method is proposed next in the second method. In this method the octant whose image touches the silhouette is further classified into three types of grey octants: grey-grey, grey-black and grey-white. Then those octants having little intersection with object silhouettes will not be subdivided. This method has a great improvement on the computer processing time and memory storage space.

A progressive mode instead of a recursive mode is proposed in the third method for the octree construction. The octree generation will be implemented using a best-first tree traversal scheme instead of a conventional depth first or breadth first tree traversal scheme. The precedence of octants for subdivision is ranked according to the XOR error between the projected octant image and the object silhouette. The progressive octree model generally gives the best visual quality rate of the object rendering effect.

The fourth construction method is to fit the object surface with planes. The 3D object model construction usually requires camera calibration beforehand. However, it is not easy for a user to calibrate the camera in circumstances, in particular, in the case of a hand-held camera. The fourth method for constructing the object method from an uncalibrated image sequence is proposed. Instead of using the point based feature to construct the 3D information, the planar homography defined over an object planar surface is used to derive the projective geometric model of the object.

1.4 Contribution of the Work

The contributions of the three proposed octree construction methods are:

- (1) The first construction method categorizes the conventional grey nodes into two types and controls the node or octant subdivision by measuring an exclusive OR projection error. Therefore, the ultimate octree subdivision level of the first construction method is not fixed as in the conventional method.
- (2) The first construction method outperforms the conventional method in terms of memory space and computation time, given that they have compatible projection errors.
- (3) Theoretical analysis on the space and time complexities of the new method is

given.

- (4) A further improvement over the first construction method is achieved in the second construction method by categorizing the conventional grey nodes into more types of grey nodes.
- (5) A progressive, instead of a recursive, reconstruction method achieves with a fast error rate reduction is made possible with a subdivision scheme using the octant sorting based on the projection error.

The contributions of the proposed planar object reconstruction method from uncalibrated images are:

- (1) A novel planar object reconstruction method from uncalibrated images is proposed
- (2) The estimation of 3D equations of the object planes is directly derived with respect to the assigned reference plane equations and plane-induced homographies.
- (3) The integration of the reconstruction results from different images pairs is introduced using the plane-induced homography information.
- (4) A performance comparison between our method and the point-based method employing the fundamental matrix with the aid of hallucinated points is provided.

1.5 Dissertation Organization

This dissertation is organized as follows: In Chapter 2, we propose a fast octree construction method with quality control for multiple object silhouette images. Chapter 3 presents the second octree construction method which allows stopping of an octant subdivision if the octant is classified as a “grey-black” octant. In Chapter 4, we propose the third octree construction method which is a progressive method using a priority queue for sorting octants based on the projection errors. In Chapter 5, we propose a novel 3D planar object reconstruction method from multiple uncalibrated images using the plane-induced homographies Chapter 6 is the summary and future work.

CHAPTER 2

FAST OCTREE CONSTRUCTION ENDOWED WITH AN ERROR BOUND CONTROLLED SUBDIVISION SCHEME

2.1 Introduction

An octree is a volumetric representation of an object geometric model. It is well known in the field of computer vision that an octree can be constructed from multiple silhouettes taken from different viewpoints of an object [1]-[18]. The constructed object geometric model finds many applications including robotic system, computer aided design, virtual reality, and object tracking. The conventional construction methods can be categorized into two classes: 3D space approach [3]-[6] and 2D space approach [7]-[18]. Both approaches recursively subdivide a partially occupied octree octant into smaller octants until all the generated octants are entirely inside or outside the object. The overlapping between an octree octant and the object is determined by an intersection test. The above two approaches differ in their intersection tests. In the 3D space approach the octant under examination is tested against the conic view volume formed by the individual silhouette and the center of projection for each viewpoint. In the 2D space approach the projection image of the octant is tested against the silhouette for each viewpoint. Generally speaking, the 2D space approach is performed in a space with a lower dimension, so it is more efficient than the 3D space approach [7]. Since the octant number grows exponentially with the subdivision level, an upper bound is usually imposed on the number of subdivision levels to avoid the insufficient memory space problem. However, a larger value of subdivision levels leads to a better construction result. It is generally difficult to make a good balance between memory space and construction quality.

A close look at the octant subdivisions reveals that the subdivided octants have varying projection errors; some may be very large and some may be rather small. We shall consider making effective use of the octant subdivisions to improve the overall system performance. In this chapter, we propose a new subdivision strategy which is governed by the degree of overlapping between the generated octant and the object. This degree of overlapping will be measured by the maximum 2D projection error of

the projected octant image relative to the object silhouette for all viewpoints. Those octants with a projection error exceeding the pre-specified upper bound will be subdivided recursively until all the octants satisfy the maximum constraint imposed on the projection error. Furthermore, we will also present a fast computation of the 2D projection and a new intersection test to reduce the computer processing time. A summary of contributions of the chapter is given below:

- (1) The ultimate octree subdivision level of the new construction method is not fixed. Instead, it is controlled by a projection error bound.
- (2) The construction quality can be evaluated via an exclusive OR projection error.
- (3) Fast computation of 2D octant projections to the image planes is provided.
- (4) Fast 2D intersection test utilizes the distance maps generated for the multiple silhouettes in advance.
- (5) Analysis on the properties of the proposed method reveals its superiority over the conventional method in terms of memory space and computation time.

The chapter is organized as follows. Section 2 presents a fast 2D octant projection computation. Section 3 describes the use of distance maps generated from the multiple silhouettes and the approximation of an octant projection image by a circle. Section 4 describes the new subdivision process and the evaluation of the quality of the final octree constructed for the object. Section 5 is the experimental results. Section 6 is the analysis on the new method. Section 7 is the conclusion remarks.

2.2 Fast Octant Projection Computation Using Cross Ratio and Vanishing Point

2.2.1 Computation of perspective projections of octant vertices

When a parent octant is subdivided into 8 child octants, the projection images of these 8 child octants involves $8 \times 8 = 64$ vertices in the recursive construction process. However, each child octant shares some of its 8 vertices with its siblings. There are 27 (instead of 64) distinct vertices, as shown in Fig. 2.1. We shall compute the perspective projections of the 27 distinct vertices onto the image plane before dividing them into 8 octant projection images.

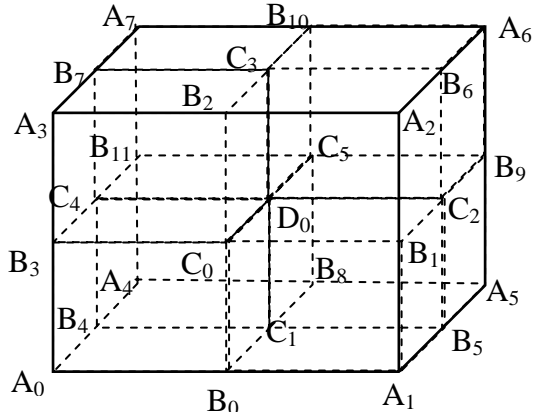


Fig. 2.1. The 27 distinct vertices involved in the subdivision of a parent octant including eight A-type vertices (A_0, A_1, \dots, A_7), twelve B-type vertices (B_0, B_1, \dots, B_{11}), six C-type vertices (C_0, C_1, \dots, C_5), and one D-type vertex D_0 .

There are four types of sub-octant vertices in regard to their parent octant vertices:

(i) The A-type vertices in $\{A_0, A_1, \dots, A_7\}$ are the vertices belonging to the parent octant. Their perspective projections on the image plane are passed down to the next round of octant subdivision.

(ii) The B-type vertices in $\{B_0, B_1, \dots, B_{11}\}$ are the vertices located on the separate lines defined by pairs of A-type vertices. Their perspective projections on the image plane can be calculated more rapidly from A-type vertices through the use of the cross ratio and vanishing points, as described later.

(iii) The C-type vertices in $\{C_0, C_1, \dots, C_5\}$ are the vertices falling on the separate lines defined by pairs of B-type vertices. Their perspective projections on the image plane can be calculated more rapidly from B-type vertices through the use of the cross ratio and vanishing points, too.

(iv) The D-type vertex D_0 falls on the line defined by two C-type vertices. Its perspective projections on the image plane can be calculated from C-type vertices through the use of the cross ratio and vanishing point.

In the following the concepts of cross ratio and vanishing point are used to derive the 2D projections of the octant vertices on the image plane.

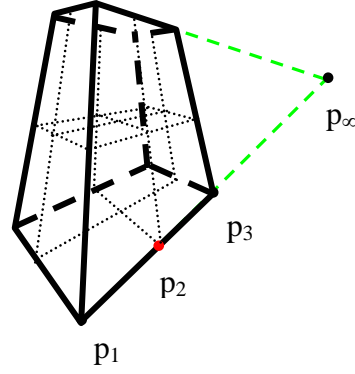


Fig. 2.2. The four collinear points, p_1 , p_2 , p_3 and p_∞ , together with their 3D corresponding vertices P_1 , P_2 , P_3 and P_∞ , define a common cross ratio value.

We only consider the 2D perspective projection computation for a B-type vertex. The other types of octant vertex can be computed similarly. Let a B-type vertex be denoted by P_2 and its projected point by p_2 and assume it lies on the line between two A-type vertices P_1 and P_3 whose projected points are p_1 and p_3 , respectively. Let the point at the infinity along the direction from P_1 to P_3 be P_∞ whose vanishing point is denoted by p_∞ . Figure 2.2 is the depiction of 2D projections of P_1 , P_2 , P_3 and P_∞ on the image plane. Then the four collinear points, p_1 , p_2 , p_3 and p_∞ , together with their 3D corresponding vertices P_1 , P_2 , P_3 and P_∞ , define a common cross ratio value [19], which is invariant under the perspective projection.

$$cr(P_1, P_2, P_3, P_\infty) = cr(p_1, p_2, p_3, p_\infty), \text{ or}$$

$$\frac{\overline{P_1 P_2} \cdot \overline{P_3 P_\infty}}{\overline{P_1 P_3} \cdot \overline{P_2 P_\infty}} = \frac{\overline{p_1 p_2} \cdot \overline{p_3 p_\infty}}{\overline{p_1 p_3} \cdot \overline{p_2 p_\infty}} \quad (2-1)$$

Since $\overline{P_3 P_\infty}$ is equal to $\overline{P_2 P_\infty}$,

$$\frac{\overline{P_1 P_2}}{\overline{P_1 P_3}} = \frac{\overline{p_1 p_2} \cdot \overline{p_3 p_\infty}}{\overline{p_1 p_3} \cdot \overline{p_1 p_3} - \overline{p_1 p_3} \cdot \overline{p_1 p_2} + \overline{p_1 p_3} \cdot \overline{p_3 p_\infty}}$$

Let $r = \frac{\overline{P_1 P_2}}{\overline{P_1 P_3}}$ ($=1/2$, as explained below)

$$r = \frac{\overline{p_1 p_2} \cdot \overline{p_3 p_\infty}}{\overline{p_1 p_3} \cdot \overline{p_1 p_3} - \overline{p_1 p_3} \cdot \overline{p_1 p_2} + \overline{p_1 p_3} \cdot \overline{p_3 p_\infty}}$$

$$\overline{p_1 p_2} = \frac{r \overline{p_1 p_3} \cdot \overline{p_1 p_\infty}}{(r \overline{p_1 p_3} + \overline{p_3 p_\infty})} \quad (2-2)$$

A parametric form of the projected line containing the four collinear points, p_1 , p_2 , p_3 and p_∞ is given by

$$p_2 = p_1 + t \cdot (p_3 - p_1), p_\infty = p_1 + k \cdot (p_3 - p_1) \quad \text{with} \quad k = \frac{\overline{p_1 p_\infty}}{\overline{p_1 p_3}} \quad \text{and} \quad t = \frac{\overline{p_1 p_2}}{\overline{p_1 p_3}} \quad (2-3).$$

Substituting these Equations into Eq. (2-2), we obtain

$$t = \frac{r \cdot k}{(r + k - 1)}.$$

Since the side length of the child octant equals one half of the side length of the parent octant, $r = 1/2$. It follows that

$$t = \frac{k}{(2 \cdot k - 1)} \quad (2-4)$$

The locations of the following three projected points can be read from the image:

$p_1 = (u_1, v_1)^T$, $p_3 = (u_3, v_3)^T$, $p_\infty = (u_\infty, v_\infty)^T$. Also the values of k and t can be found from Eqs. (2-3) and (2-4). Then the projection of the B-type vertex $p_2 = (u_2, v_2)^T$ can be computed from the projections of the two A-type vertices:

$$p_2 = p_1 + t \cdot (p_3 - p_1) \quad (2-5)$$

2.2.2 Time Complexity Analysis

Before the analysis of the time complexity of the computation of the projections of octant vertices, recall that in the conventional octree construction method the projection of a 3D octant vertex onto the 2D image is done via a projection matrix $H_{3 \times 4}$:

$$\begin{bmatrix} u \cdot w \\ v \cdot w \\ w \end{bmatrix} = H_{3 \times 4} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

After eliminating the parameter w , the coordinates of the 2D projected point are described by

$$\begin{cases} u = \frac{h_{11} \cdot X + h_{12} \cdot Y + h_{13} \cdot Z + h_{14}}{h_{31} \cdot X + h_{32} \cdot Y + h_{33} \cdot Z + h_{34}} \\ v = \frac{h_{21} \cdot X + h_{22} \cdot Y + h_{23} \cdot Z + h_{24}}{h_{31} \cdot X + h_{32} \cdot Y + h_{33} \cdot Z + h_{34}} \end{cases} \quad (2-6)$$

Thus, 9 multiplications, 2 divisions and 9 additions are performed to obtain the 2D

projected point on the image plane by Eq. (2-6). But in the new method based on the cross ratio and vanishing point, it requires 3 subtractions and 1 division to compute the parameter k by Eq. (2-3); 1 multiplication, 1 division and 1 subtraction to compute the parameter t by Eq. (2-4); and 2 multiplications and 2 additions to compute the 2D coordinates of the projected point by Eq. (2-5). Table 2.1 is the comparison of the new method with the conventional 3D projection method. Roughly speaking the new method requires less than one half of the computation time of the conventional method. In the experimental results, we will show the new method does have better performance than the conventional 3D-to-2D projection computation based on the projection matrix H .

Table 2.1. The comparison of projection computation of octant vertices between the new method and the conventional method.

	Our method		Direct 3D projection method	
	+/-	×/÷	+/-	×/÷
Operation number	6	5	9	11

2.3 2D Intersection Test Using the Precomputed Distance Maps of Silhouette Images and Approximation of an Octant Image by a Circle

2.3.1 Distance maps of silhouette images

After obtaining the 2D projections of the octant surface, which is generally a hexagon, we need to test this hexagon against the object silhouette images to determine whether it is completely inside or outside the object, or it intersects the object. We give up the conventional intersection test through first approximating each silhouette by a set of convex sub-polygons and then examining the intersection between two convex polygons. We generate beforehand the signed distance map for each silhouette image using a chess board distance definition given by [32]:

$$I_0(u, v) \equiv I(u, v) \quad \text{and}$$

$$I_k(u, v) = I_0(u, v) + \min_{\Delta(u, v; i, j)} \{I_{k-1}(i, j); ((i, j) : \Delta(u, v; i, j) \leq 1)\}$$

where $\Delta(u,v;i,j)$ is the distance between (u,v) and (i,j) and $I(u,v)$ is the input binary silhouette image.

Here the map distance value is positive at a pixel located inside the silhouette, negative at a pixel outside the silhouette, and 0 at a pixel on the silhouette boundary. The absolute distance value at each pixel in the distance map represents the shortest distance from the pixel to the silhouette boundary.

2.3.2 Intersection test using the circular approximation of the octant image

Next, we find the center (or centroid) of the 8 projected octant vertices and estimate the radius of the bounding circle of the projected octant image. The radius is roughly taken to be the largest distance among the center and the eight projected octant vertices. The distance map value at the circle center c in the distance map $DistMap$ and the radius r of the bounding circle are then used to determine the intersection between the silhouettes and the projected octant image. Table 2.2 indicates the spatial relationship between the projected octant image and a silhouette image. The spatial relationship must be carried out for all the views one at a time. However, it is well known that if the octant image is outside any single silhouette, then the octant is definitely outside the object and classified as a white node; no view is needed for further examination. Also, an octant is classified as a black node if it is marked black in all views, and an octant is classified as a grey node if it is grey in some view.

Table 2.2 Intersection relation between the projected octant image and a silhouette image.

$DistMap(c) \geq 0$	$r > DistMap(c)$	across the silhouette boundary (indicating a grey node)
	$r \leq DistMap(c)$	inside the silhouette (indicating a black node)
$DistMap(c) < 0$	$r > -DistMap(c)$	across the silhouette boundary (indicating a grey node)
	$r \leq -DistMap(c)$	outside the silhouette (indicating a white node)

2.4 New Subdivision Strategy and Construction Quality Measure

2.4.1 New subdivision strategy

Suppose we use N views to construct the octree for the object. Let A_i be the projection image of any grey octant A and S_i be the object silhouette for $i = 0, 1, 2, \dots, N$. Let $DistMap_i$ be the signed distance map of S_i defined previously. We find the center c_i and the radius r_i of the bounding circle of A_i , as shown in Fig. 3. Then we define the white portion (or white extent) of a grey octant to be $\{r_i - DistMap_i(c_i)\}$.

Now we introduce two types of grey nodes in the following:

Let “Error_bound” be a pre-specified projection error upper bound.

(1) A grey node is said to be a “grey-grey” node if $\{r_i - DistMap_i(c_i)\} > Error_bound$ for some $i, 0 \leq i \leq N-1$.

(2) A grey node is defined as a “grey-black” node, if $\{r_i - DistMap_i(c_i)\} > Error_bound$ for all $i, 0 \leq i \leq N-1$.

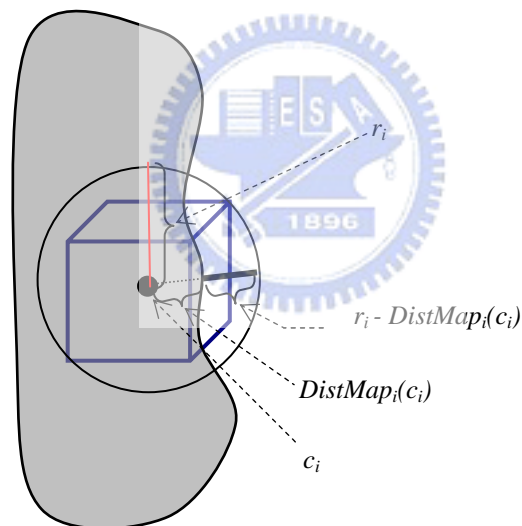


Fig. 2.3. The illustration of octant subdivision strategy based on the spatial relation between the circle containing the projected octant and the object silhouette.

Based on the new types of grey nodes a new subdivision scheme for grey nodes is proposed below:

The new subdivision scheme for grey nodes:

(1) If a node is classified as a “grey-grey” node it is subdivided recursively until its descendant nodes contain no “grey-grey” descendant nodes.

(2) If a node is classified as a “grey-black” node it is retained without the need of subdivision.

2.4.2 Construction quality measure

To measure the construction quality of the constructed model, we compute the exclusive OR between the binary silhouette image S_j^C of the constructed object C and the actual object silhouette S_j^O for all views ($j=0, 1, \dots, N-1$). The index of construction quality is the sum of the exclusive OR projection errors over all the views:

$$Quality_Index = \sum_{j=0}^{N-1} S_j^C \otimes S_j^O$$

2.5 Experimental Results

In this section, experiments are conducted to evaluate the performance of the method we proposed. Fig. 2.4 is the setup of our hardware system including a CCD, a turntable and a PC with the Pentium-IV 3G and 768M RAM. The camera is triggered by the PC to capture the images of the real object resting on the rotating turntable which is controlled by the PC. The whole reconstruction program is coded with Borland C++ Builder under the Windows environment. We also choose an octree construction method recently developed in [20] as the conventional method in order to make comparisons with our new construction method.



Fig. 2.4. The hardware setup of the system.

Experiment 1:

In the experiment 1, we first use a synthetic right cube whose schematic diagram is given in Fig. 2.5. We align the cube with the root octant in three surface normal directions. The volume of the synthetic cube is 8000 inch^3 . We take four pictures from the viewing directions parallel to the normals of the cube surface. The image resolution is 640 pixel x 480 pixel. Three selected images of the sequence, I_0 , I_1 , and I_2 , are shown in Fig. 2.6. We apply the octree reconstruction process to this data set. In Table 2.3, one can find that if the error bound (P) is set to 1, the constructed octree model is the same as the one constructed from the conventional method. That is, additional efforts are required to perform subdivision check in our method. However, if the error bound (P) is larger than 1, our method requires less memory space to represent an object. We also list the numbers of all types of octants generated at each subdivision level, together with the 3D volume of the final constructed models in Tables 2.4 and 2.5 under different error bound and maximum subdivision level of our method and the conventional method. In addition, the graphical display of the constructed object models obtained by the conventional method and the new method is shown in Fig. 2.7.

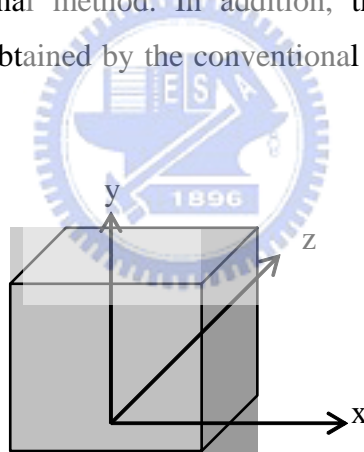


Fig. 2.5. The schematic diagram of the right cube. The dimensions of the cube are 20 inches in depth (the x-direction), 20 inches in width (the z direction) and 20 inches in height (the y direction).

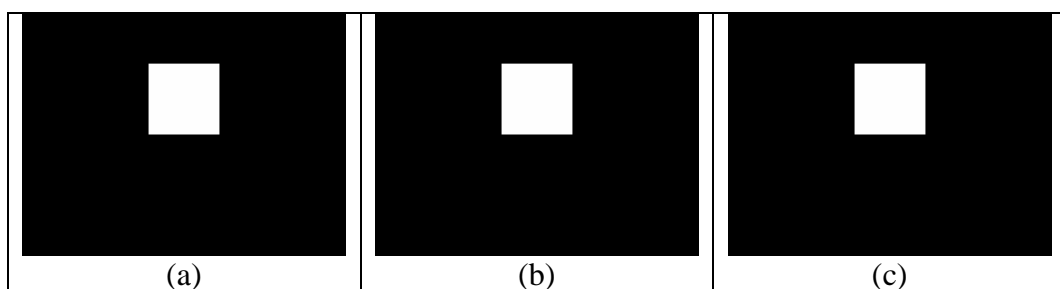


Fig. 2.6. Three distinct images I_0 , I_3 and I_7 of the cube taken at 0, 90 and 180 degrees.

Table 2.3. Numbers of black, grey-black, grey-grey, and white nodes of the cube generated by our octree construction method with a specified projection error bound.

	Protrusion=1				L=8			
	Volume = 8852.661				Volume = 8852.661			
	B	GB	GG	W	B	GB	GG	W
0	0	0	1	0	0		1	0
1	0	0	8	0	0		8	0
2	0	0	48	16	0		48	16
3	8	0	171	204	8		171	204
4	236	0	600	540	236		600	540
5	1204	0	1940	1656	1204		1940	1656
6	5496	0	5644	5280	5496		5644	5280
7	15460	0	17808	11884	15460		17808	11884
8	67048	0	0	75416	67048		0	75416

Table 2.4. Numbers of black, grey, and white nodes of the cube generated by the conventional octree construction method with a fixed subdivision level.

	Level=4				Level=5				Level=6				Level=7			
	Volume= 14062.5				Volume = 10828.13				Volume = 9758.789				Volume = 9396.118			
	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W
0	0		1	0	0		1	0	0		1	0	0		1	0
1	0		8	0	0		8	0	0		8	0	0		8	0
2	0		48	16	0		48	16	0		48	16	0		48	16
3	8		171	204	8		171	204	8		171	204	8		171	204
4	236		600	540	236		600	540	236		600	540	236		600	540
5					1204		1940	1656	1204		1940	1656	1204		1940	1656
6									5496		5644	5280	5496		5644	5280
7													15460		17808	11884
8																

(B for Black, GB for Grey-Black, GG for Grey-Grey, and W for White)

Table 2.5. Numbers of black, grey-black, grey-grey, and white nodes of the cube generated by our octree construction method with a specified projection error bound.

	Protrusion=30				Protrusion=15				Protrusion=7				Protrusion=4			
	Volume = 15062.5				Volume = 10828.13				Volume = 9522.827				Volume = 9185.425			
	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W
0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
1	0	0	8	0	0	0	8	0	0	0	8	0	0	0	8	0
2	0	8	40	16	0	0	48	16	0	0	48	16	0	0	48	16
3	8	0	116	204	8	56	116	204	8	20	152	204	8	4	168	204
4	0	388	0	540	0	4	384	540	108	180	388	540	204	124	746	540
5					0	1416	0	1656	16	340	1092	1656	376	372	1404	1656
6									240	3140	76	5280	1424	1688	2840	5280
7									0	76	0	532	144	10988	0	11588
8																



Conventional method	Level = 4	Level = 5	Level = 6	Level = 7
Image of the constructed model				
New method	Protrusion= 30	Protrusion= 15	Protrusion= 7	Protrusion= 4
Image of the constructed model				

Fig. 2.7. The comparison between construction results of the synthetic cube obtained by the conventional method and the new method.

We then slightly rotate the cube whose schematic diagram is given in Fig. 2.8. Nine pictures are taken from different angles to cover all aspects of the cube. Three selected images of the sequence, I_0 , I_3 , and I_7 , are shown in Fig. 2.9. We apply the octree reconstruction process to this data set. We list the numbers of all types of octants generated at each subdivision level, together with the 3D volume of the final constructed models in Tables 2.6 and 2.7 under different error bound and maximum subdivision level of our method and the conventional method. One can find that under the same reconstruction volume generated, our method requires less octant nodes to represent the object. In addition, the graphical display of the constructed object models obtained by the conventional method and the new method is shown in Fig. 2.10.

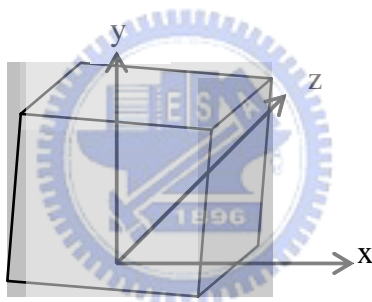


Fig. 2.8. The schematic diagram of the tilted cube. The dimensions of the cube are 20 inches in depth (the x-direction), 20 inches in width (the z direction) and 20 inches in height (the y direction).

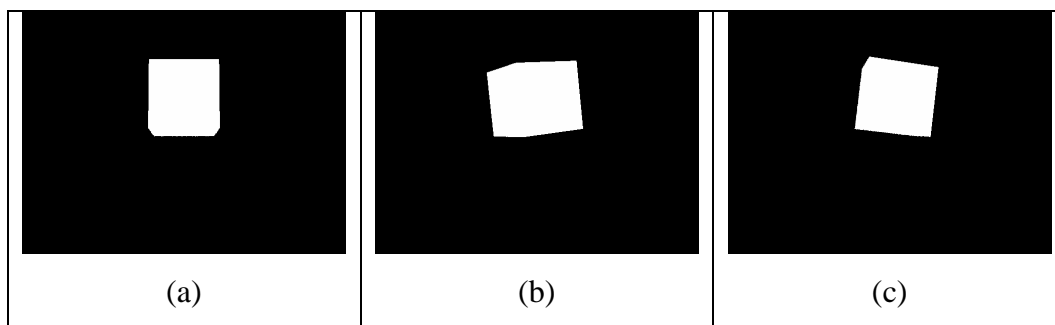


Fig. 2.9. Three distinct images I_0 , I_3 and I_7 of the tilted cube taken at 0, 120 and 280 degrees.

Table 2.6. Numbers of black, grey, and white nodes of the cube generated by the conventional octree construction method with a fixed subdivision level.

	Level=4				Level=5				Level=6				Level=7			
	Volume=14140.63				Volume = 11689.45				Volume = 10389.4				Volume = 9799.713			
	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W
0	0		1	0	0		1	0	0		1	0	0		1	0
1	0		8	0	0		8	0	0		8	0	0		8	0
2	0		48	16	0		48	16	0		48	16	0		48	16
3	12		167	205	12		167	205	12		167	205	12		167	205
4	250		559	527	250		559	527	250		559	527	250		559	527
5					1089		2128	1255	1089		2128	1255	1089		2128	1255
6									4755		6944	5325	4755		6944	5325
7													17757		18472	19323
8																

(B for Black, GB for Grey-Black, GG for Grey-Grey, and W for White)

Table 2.7. Numbers of black, grey-black, grey-grey, and white nodes of the cube generated by our octree construction method with a specified projection error bound.

	Protrusion=30				Protrusion=15				Protrusion=7				Protrusion=4			
	Volume = 14109.38				Volume = 11755.86				Volume = 10333.71				Volume = 9861.328			
	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W
0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
1	0	0	8	0	0	0	8	0	0	0	8	0	0	0	8	0
2	0	7	41	16	0	0	48	16	0	0	48	16	0	0	48	16
3	1	19	103	205	12	42	125	205	12	16	151	205	12	8	159	205
4	0	295	2	527	32	177	264	527	130	115	436	527	186	74	485	527
5					0	891	2	1219	319	775	1139	1255	541	514	1570	1255
6					0	0	0	16	67	3473	300	5272	1450	2255	3530	5325
7									0	151	0	2249	1667	9269	0	17304
8																









Conventional method	Level = 4	Level = 5	Level = 6	Level = 7
Image of the constructed model				
New method	Protrusion= 30	Protrusion= 15	Protrusion= 7	Protrusion= 4
Image of the constructed model				

Fig. 2.10. The comparison between the construction results of the synthetic cube obtained by the conventional method and the new method.

Experiment 2:

In the experiment 2, we test our method on three real objects with different geometric complexities: a cone, a vase and a boy sculpture. Fig. 2.11 shows the three real objects. A number of views of each object are taken, while the turntable is rotated by a fixed degree each time. In the experiments the rotation angle is 36 degrees. Fig. 2.12 is the new views rendered from the final constructed objects obtained.

We list the numbers of all types of octants generated at each subdivision level, together with the projection error upper bound value and the quality index of the final constructed models in Table 2.8 to Table 2.13. In addition, the graphical display of the constructed object models and the XOR error images obtained by the conventional method and the new method are shown in Fig. 2.13 to Fig. 2.15.

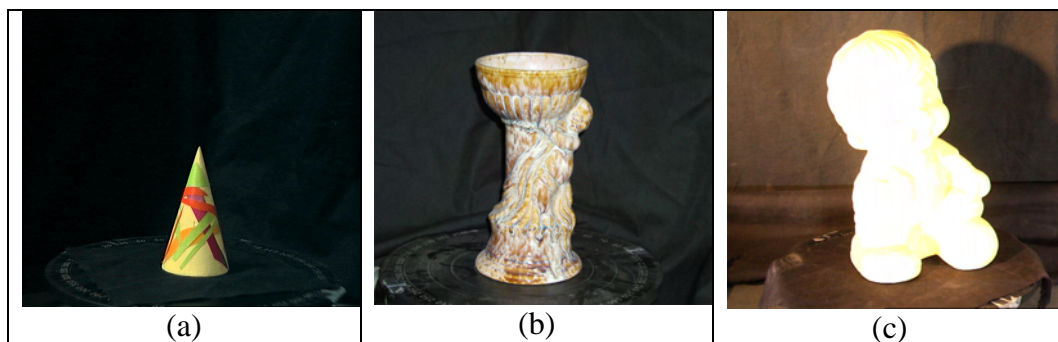


Fig. 2.11. Three real objects used in the experiments: (a) a cone, (b) a vase, and (c) a boy sculpture.

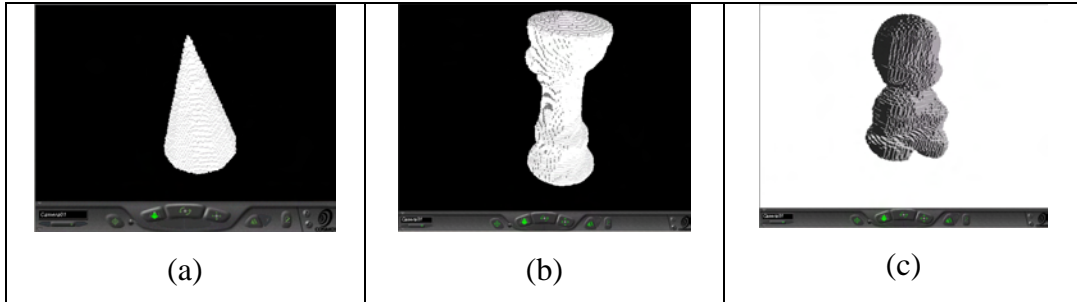


Fig. 2.12. New views generated from the constructed octree models for (a) a cone, (b) a vase, and (c) a boy sculpture. The octree models are converted to VRML format and rendered using Cosmo© player.

Table 2.8. Numbers of black, grey, and white nodes of the cone generated by the conventional octree construction method with a fixed subdivision level.

	Level=4				Level=5				Level=6				Level=7			
	Xor=116647				Xor= 53225				Xor=23682				Xor=9660			
	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W
0	0	-	1	0	0	-	1	0	0	-	1	0	0	-	1	0
1	0	-	7	1	0	-	7	1	0	-	7	1	0	-	7	1
2	0	-	8	48	0	-	8	48	0	-	8	48	0	-	8	48
3	0	-	32	32	0	-	32	32	0	-	32	32	0	-	32	32
4	4	-	108	144	4	-	108	144	4	-	108	144	4	-	108	144
5					84	-	393	387	84	-	393	387	84	-	393	387
6									609	-	1401	1134	609	-	1401	1134
7													3401	-	3349	4458
8																

(B for Black, GB for Grey-Black, GG for Grey-Grey, and W for White)

Table 2.9. Numbers of black, grey-black, grey-grey, and white nodes of the cone generated by our octree construction method with a specified projection error bound.

	Protrusion=30				Protrusion=18				Protrusion=10				Protrusion=5			
	Xor=67123				Xor=44365				Xor=23685				Xor=9283			
	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W
0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
1	0	0	7	1	0	0	7	1	0	0	7	1	0	0	7	1
2	0	0	8	48	0	0	8	48	0	0	8	48	0	0	8	48
3	0	1	31	32	0	0	32	32	0	0	32	32	0	0	32	32
4	2	54	48	144	4	26	82	144	4	9	99	144	4	4	104	144
5	0	61	0	323	1	213	55	387	31	137	237	387	54	59	332	387
6					0	7	0	433	7	762	0	1134	199	552	771	1134
7													122	1494	242	4310
8													1	0	0	1935



Table 2.10. Numbers of black, grey, and white nodes of the vase generated by the conventional octree construction method with a fixed subdivision level.

	Level=4				Level=5				Level=6				Level=7			
	Xor=289493				Xor=144377				Xor=68356				Xor=30069			
	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W
0	0	-	1	0	0	-	1	0	0	-	1	0	0	-	1	0
1	0	-	8	0	0	-	8	0	0	-	8	0	0	-	8	0
2	0	-	41	23	0	-	41	23	0	-	41	23	0	-	41	23
3	1	-	133	194	1	-	133	194	1	-	133	194	1	-	133	194
4	88	-	532	444	88	-	532	444	88	-	532	444	88	-	532	444
5	685	-	2074	1497	685	-	2074	1497	685	-	2074	1497	685	-	2074	1497
6									3811	-	7482	5299	3811	-	7482	5299
7													17526	-	21460	20870
8																

Table 2.11. Numbers of black, grey-black, grey-grey, and white nodes of the vase generated by our octree construction method with a specified projection error bound.

	Protrusion=30				Protrusion=18				Protrusion=10				Protrusion=5			
	Xor=178926				Xor=118666				Xor=68365				Xor=28185			
	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W
0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
1	0	0	8	0	0	0	8	0	0	0	8	0	0	0	8	0
2	0	0	41	23	0	0	41	23	0	0	41	23	0	0	41	23
3	1	17	116	194	1	8	125	194	1	1	132	194	1	0	133	194
4	12	284	188	444	40	151	365	444	81	76	455	444	88	32	500	444
5	0	316	0	1188	3	1137	283	1497	180	808	1155	1497	441	359	1703	1497
6					0	68	0	2196	1	3940	0	5299	1301	3100	3924	5299
7													496	8233	2489	20174
8													7	0	0	19905

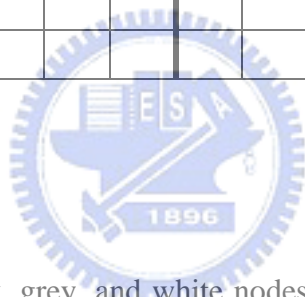


Table 2.12. Numbers of black, grey, and white nodes of the boy sculpture generated by the conventional octree method with a fixed subdivision level.

	Level=4				Level=5				Level=6				Level=7			
	Xor=366074				Xor=186064				Xor=89866				Xor=41534			
	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W
0	0	-	1	0	0	-	1	0	0	-	1	0	0	-	1	0
1	0	-	8	0	0	-	8	0	0	-	8	0	0	-	8	0
2	0	-	51	13	0	-	51	13	0	-	51	13	0	-	51	13
3	7	-	232	169	7	-	232	169	7	-	232	169	7	-	232	169
4	206	-	957	693	206	-	957	693	206	-	957	693	206	-	957	693
5					1483	-	3738	2435	1483	-	3738	2435	1483	-	3738	2435
6									7307	-	13437	9160	7307	-	13437	9160
7													31273	-	40955	35265
8																

Table 2.13. Numbers of black, grey-black, grey-grey, and white nodes of the boy sculpture generated by our octree construction method with a specified projection error bound.

	Protrusion=30				Protrusion=18				Protrusion=10				Protrusion=5			
	Xor=226886				Xor=156235				Xor=89872				Xor=37350			
	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W	B	GB	GG	W
0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
1	0	0	8	0	0	0	8	0	0	0	8	0	0	0	8	0
2	0	0	51	13	0	0	51	13	0	0	51	13	0	0	51	13
3	7	46	186	169	7	16	216	169	7	10	222	169	7	4	228	169
4	19	476	300	693	103	325	607	693	133	159	291	693	174	69	888	693
5	0	438	0	1962	6	1927	488	2435	415	1443	2035	2435	955	717	2997	2435
6					0	166	0	3738	13	7107	0	9160	2139	5531	6966	9160
7													822	14814	5743	34349
8													54	0	0	45890

















Conventional method	Level = 4	Level = 5	Level = 6	Level = 7
Image of the constructed model				
Xor error image of the image				
New method	Protrusion= 30	Protrusion= 18	Protrusion= 10	Protrusion= 5
Image of the constructed model				
Xor error image of the images				

Fig. 2.13. The comparison between construction results of the cone obtained by the conventional method and the new method.

















Conventional method	Level = 4	Level = 5	Level = 6	Level = 7
Image of the constructed model				
Xor error image of the image				
New method	Protrusion= 30	Protrusion= 18	Protrusion= 10	Protrusion= 5
Image of the constructed model				
Xor error image of the image				

Fig. 2.14. The comparison between construction results of the vase obtained by the conventional method and the new method.

















Conventional method	Level = 4	Level = 5	Level = 6	Level = 7
Image of the constructed model				
Xor error image of the image				
New method	Protrusion= 30	Protrusion= 18	Protrusion= 10	Protrusion= 5
Image of the constructed model				
Xor error image of the image				

Fig. 2.15. The comparison between construction results of the boy sculpture obtained by the conventional method and the new method.

From the above tables the two methods can be compared in terms of memory storage, computation time, and the quality of the construction. Since the object model is represented by the leaf nodes of the constructed octree including black nodes and grey-black nodes, if available, the memory space required to store the constructed object model is proportional to the total number of black and grey-black nodes. On the other hand, the computer processing time required is mainly proportional to the entire set of leaf and internal nodes generated. Finally, the quality of the construction result is evaluated based on the XOR projection error given by the “*Quality_Index*”.

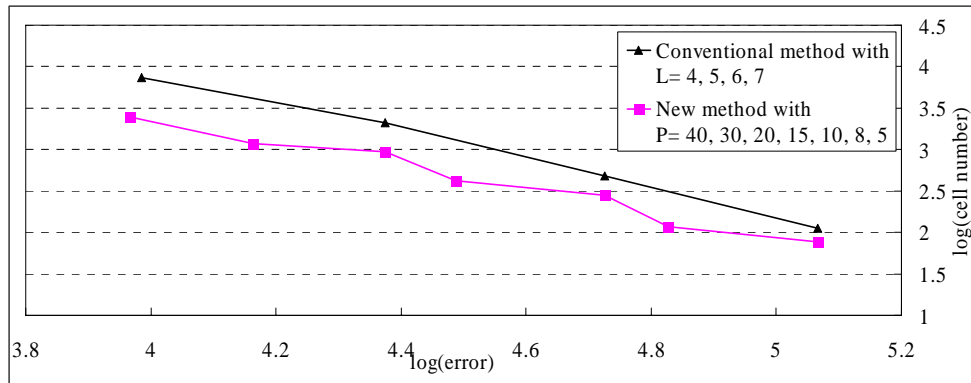
The following interesting observations can be made from the above experimental results:

- (1) Based on the quality of the constructed object models expressed by the XOR value we can rank the construction results obtained by both the conventional and the new methods. For instance, from Tables 2.8 and 2.9, the ascending quality order of the construction results can be found to be $V_4^C < V_{30}^N < V_5^C < V_{18}^N < V_6^C = V_{10}^N < V_7^C < V_5^P$, where the symbol V_L^C denotes the object model obtained by the conventional method with a maximum level parameter L , and the symbol V_P^N denotes the result obtained by the new method with a projection error upper bound parameter P .
- (2) During the construction process the generation patterns of the white nodes by the two methods are identical. To be more specific, the white nodes at a lower subdivision level must be completely generated before any white node can be generated at the next higher subdivision level. For instance, in Table 2.9 the full set of 1134 white nodes must be generated at level 6 before any white node at

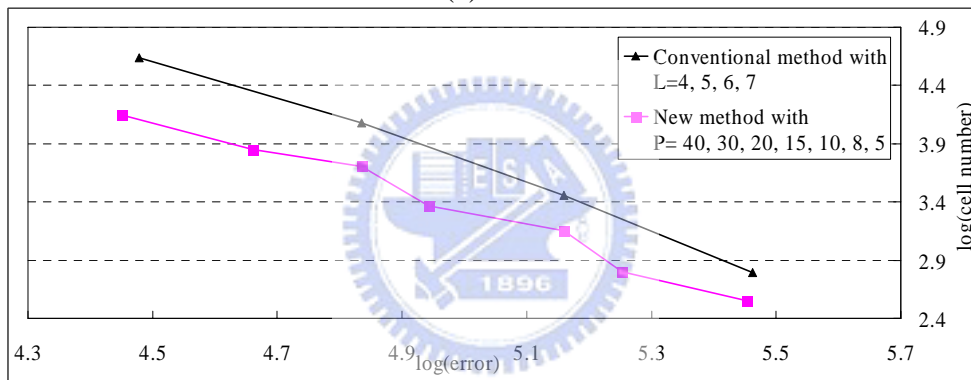
level 7 is generated when the error control parameter P decreases from 18, 10, to 5.

- (3) For the same quality of the construction results produced by the conventional and new methods the total leaf node number of the object model generated by the conventional method is greater than that generated by the new method. For instance, in Tables 2.8 and 2.9 the total leaf node number of V_{10}^N is smaller than that of V_6^C . The same observations can be also found in Tables 2.10 to 2.13.
- (4) Under the assumption of the same or equivalent construction quality, the total internal and leaf nodes is less in the new method.

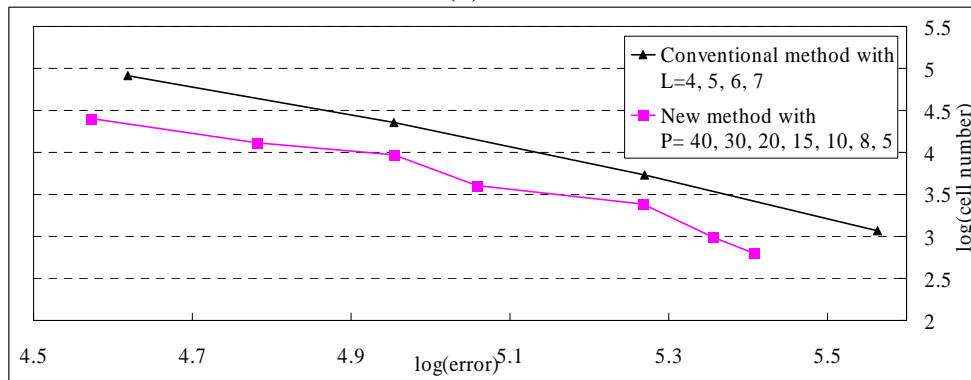
The observations (3) and (4) indicate that the new method takes less memory space and computation time to achieve the same construction quality when compared to the conventional method. These findings are plotted in Figures 2.16 to 2.17. In these plots a linear interpolation of the discrete set of construction results is made. Since the colors of the octants in the two methods are classified as three colors (White, Grey, and Black) and four colors (White, Grey-Grey, Grey-Black, and Black), respectively, it is sufficient to use two bits to represent the color of the octant in the two methods. Let num_c and num_n be the total numbers of octants generated in the conventional method and our new method. Then the total memory space in bits required for the two methods are $num_c * 2$ and $num_n * 2$, respectively. Since num_n is less than num_c , as shown in the later section. We can conclude that the total memory space requirement in bit representation for our new method is less than that of the conventional method.



(a) cone

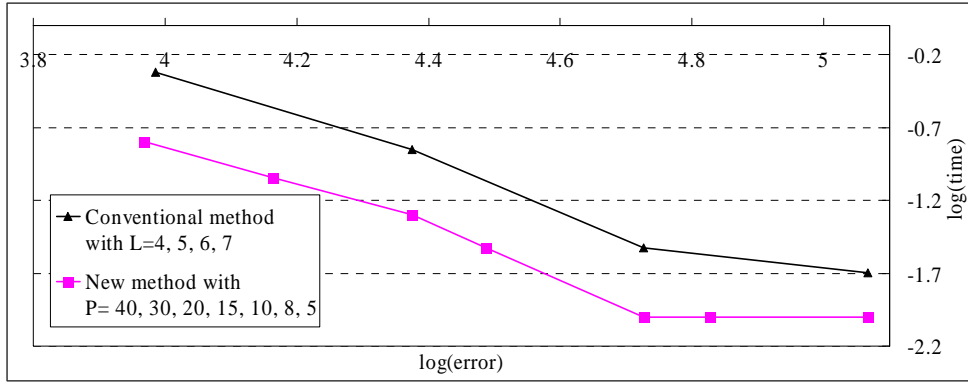


(b) vase

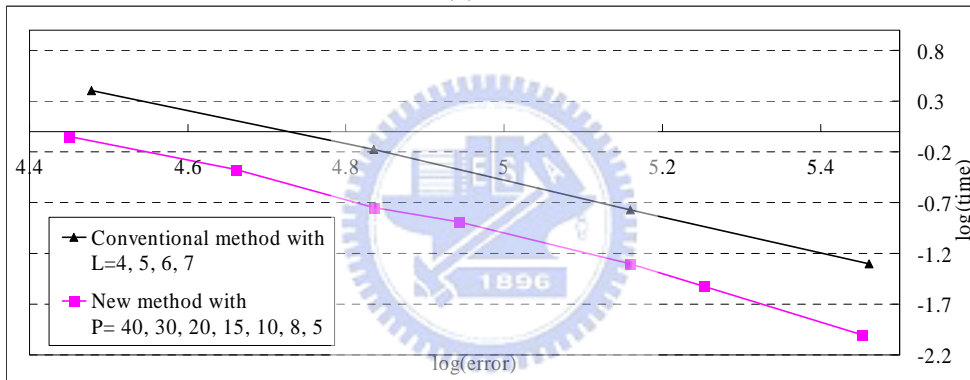


(c) boy sculpture

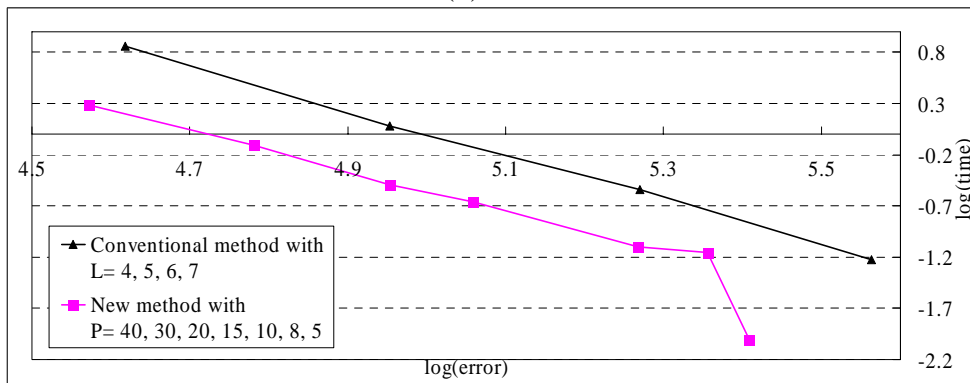
Fig. 2.16. The total leaf node number vs. XOR error plot of the construction results obtained by the new method and the conventional method.



(a) cone



(b) vase



(c) boy sculpture

Fig. 2.17. The computation time vs. XOR error plot of the construction results obtained by the new method and the conventional method.

2.6 Analysis on the New Octree Construction Method

In the following we shall define first the notations and symbols used in the forthcoming theoretical analysis on the performance of the new octree construction method:

- (1) $V_T, V_{Root}, V_L^C, V_P^N$: the volumes of the true object, the root node of the octree, the constructed object obtained by the conventional method with a maximum subdivision level L , and the constructed object obtained by the new method with a specified projection error upper bound P , respectively.
- (2) B_l^C, W_l^C, G_l^C : the sets of black, white, and grey nodes generated at level l in the conventional method.
- (3) $B_l^N, W_l^N, GB_l^N, GG_l^N$: the sets of black, white, grey-black, and grey-grey nodes generated at level l in the new method.
- (4) V_{N^m} (e.g., $V_{B_L^C}, V_{B_P^N}$): the volume of the set of a particular type of nodes $N \in \{B, GG, GB, W\}$ generated by the one of the two methods “ m ” $\in \{C(\text{conventional method}), N(\text{new method})\}$ with a “parameter” $\in \{L(\text{level value}), P(\text{the protrusion value})\}$.
- (5) I_V : the binary image obtained through the projection of a particular object or a particular node set to the image plane in a view, where $V \in \{V_T, V_{Root}, V_L^C, V_P^N, V_{B_L^C}, V_{B_P^N}, \dots\}$.
- (6) $E_{V_L^C}, E_{V_P^N}$: the total XOR projection errors of $I_{V_L^C}$ and $I_{V_P^N}$ with the true object silhouette over all views, respectively.

Next, we shall begin to derive the useful lemmas for the theoretical analysis on the performance of the two construction methods. These lemmas justify the observations made in the last section.

Lemma 2-1: The new method always halts at a subdivision level with no remaining grey-grey node, while the conventional method generally terminates at a given maximum subdivision level with remaining grey nodes unless the maximum

level is large enough.

Proof:

This is the consequence of the different subdivision strategies adopted by the two methods. The conventional method terminates without any remaining grey nodes either when its octant resolution becomes so high that all the octant projections on the image plane have a unit area or when the object happens to be fitted nicely by an octree with a few resolution levels.

Lemma 2-2: If GG_L^N is not empty, then $W_l^N = W_l^C, l=1, 2, \dots, L-1$.

Proof:

We prove the lemma by contradiction. Assume $W_l^N \neq W_l^C, l=1, 2, \dots, L-1$.

Let $w_{k-1}, 1 \leq k \leq L$, be a node in the set W_{k-1}^C , not in the set W_{k-1}^N . Since w_{k-1} is not in the set W_{k-1}^N , its parent node gb_{k-2} should be in the set GB_{k-2}^N . Thus, we have

$$r_{k-2} - |d_{k-2}| \leq P$$

where r_{k-2} and d_{k-2} are the radius and the distance map value related to the bounding circle of the projected image of gb_{k-2} .

Since w_{k-1} is within gb_{k-2} , its projection should not exceed the white portion of the projection of its parent node. So

$$r_{k-1} + |d_{k-1}| \leq r_{k-2} - |d_{k-2}| \leq P$$

Also, since w_{k-1} is a white node, its entire projection should be outside the silhouette, so $|d_{k-1}| > r_{k-1}$.

Thus,

$$P \geq r_{k-1} + |d_{k-1}| > r_{k-1} + r_{k-1} > 2r_{k-1} \quad (2-7)$$

Given GG_L^N is not an empty set. There exists at least one grey-grey node gg_L belonging to GG_L^N . There are two possible cases for the location of this gg_L .

Case 1: The projection centroid of gg_L is outside the silhouette image.

Then $r_L + |d_L| > P$, where r_L is the radius of the bounding circle of the projection of gg_L , and $|d_L|$ is the shortest distance between the centroid

of the projection of gg_L and the silhouette and P is the pre-specified projection error upper bound used to control the node subdivision.

Since gg_L is partially overlapped with the silhouette, we have $|d_L| \leq r_L$.

Thus, $r_L + r_L > r_L + |d_L| > P$. That is, $2r_L > P$

Case 2: The projection centroid of gg_L is inside the silhouette image.

Then $r_L - |d_L| > P$. Thus, $2r_L > r_L + r_L - |d_L| > P$.

For the above two cases, we conclude

$$2r_L > P \quad (2-8)$$

From Equations (2-7) and (2-8), we have

$$r_L > r_{k-1} \text{ for } 1 \leq k \leq L$$

However, this is contradictory to the fact that $r_L < r_{k-1}$. Because the radius r_L of a projected node at level L should be smaller than the radius r_{k-1} of the bounding circle of its ancestor at level $k-1$. (Here the object rests on a turntable and the distance between the object and the camera is nearly constant for all viewpoints.) Thus, $W_l^N = W_l^C, l=1, 2, \dots, L-1$.

Lemma 2-3: If $|GG_L^N| = 0$ and $W_L^N = W_L^C$, then $E_{V_P^N} = E_{V_L^C}$ and the total number of

leaf nodes generated by the conventional method is greater than that in the new method. Also, the total number of leaf and internal nodes in the conventional method is greater.

Proof:

(a) Since $W_L^N = W_L^C$, GG_{L-1}^N is not empty, Lemma 2-2 implies that $W_l^N = W_l^C$, $l=1, 2, \dots, L-2$. We can show $W_{L-1}^N = W_{L-1}^C$ using the same technique used to prove Lemma 2-2.

Next, $V_L^C = V_{root} - \sum_{l=1}^L V_{W_l^C} = V_{root} - \sum_{l=1}^L V_{W_l^N} = V_P^N$. It implies that their

projected images of the two object models constructed are equal: $I_{V_L^C} = I_{V_P^N}$. Also, their exclusive OR images are equal: $I_{V_L^C} \otimes I_{V_T} = I_{V_P^N} \otimes I_{V_T}$. That is, $E_{V_L^C} = E_{V_P^N}$.

(b) From above $W_l^N = W_l^C, l=1, 2, \dots, L$. Also, $B_l^N \subseteq B_l^C, GB_l^N \subseteq G_l^C$ for $l=1,$

2, ..., L. To emulate the subdivision result of the conventional method, imagine that the grey-black nodes generated at any level lower than L are decomposed. Then the leaf nodes generated by the two methods satisfy the following inequality:

$$\sum_1^L |B_l^N| + \sum_1^{L-1} 8 \cdot |GB_l^N| + |GG_L^N| + |GB_L^N| + \sum_1^L |W_l^N| < \sum_1^L |B_l^C| + |G_L^C| + \sum_1^L |W_l^C|.$$

It implies that

$$\begin{aligned} & \sum_1^L |B_l^N| + \sum_1^{L-1} |GB_l^N| + |GB_L^N| + \sum_1^L |W_l^N| \\ & < \sum_1^L |B_l^N| + \sum_1^{L-1} 8 \cdot |GB_l^N| + |GG_L^N| + |GB_L^N| + \sum_1^L |W_l^N| \quad (\text{Note that } |GG_L^N| = 0) \\ & < \sum_1^L |B_l^C| + |G_L^C| + \sum_1^L |W_l^C| \end{aligned}$$

Therefore, the total numbers of leaf nodes generated by the two methods are not equal.

On the other hand, because the grey-black nodes are not decomposed in the new method, the total number of leaf and internal nodes generated by the conventional method is greater than that generated by the new method.

Lemma 2-4: If $|GG_{L+1}^N| = 0$ and $|W_L^N| < |W_L^C| < [\frac{|W_{L+1}^N|}{8} + |W_L^N|]$, then $0 < E_{V_L^C} - E_{V_P^N}$. If

$$\sum_{l=1}^{L-1} |GB_l^N| > |GG_L^N| \text{ holds additionally, then the total number of leaf nodes}$$

generated by the conventional method is greater than that generated by the new method. Also, the total number of leaf and internal nodes generated by the conventional method is also greater.

Proof:

(a) Since $|GG_L^N|$ is not zero, $W_l^N = W_l^C$, $l = 1, 2, \dots, L-1$ due to Lemma 2-2. Since the volume of each node in W_{L+1}^N is one eighth of that of W_L^N , so the condition

$$[\frac{|W_{L+1}^N|}{8} + |W_L^N|] > |W_L^C| \text{ indicates the total volume of } W_{L+1}^N \text{ and } W_L^N \text{ is greater}$$

than the volume of W_l^C . Therefore, the object volume obtained by the new method is smaller than that obtained by the conventional method. Namely,

$$V_P^N = V_{root} - \sum_{l=1}^{L+1} V_{W_l^N} < V_{root} - \sum_{l=1}^L V_{W_l^C} = V_L^C.$$

Consequently, the area of $I_{V_P^N}$ is smaller than that of $I_{V_L^C}$. Also, the area of $I_{V_P^N} - I_{V_T}$ is smaller than that of $I_{V_L^C} - I_{V_T}$. It can be readily shown that the XOR projection errors of the two methods satisfy the inequality:

$$I_{V_P^N} \otimes I_{V_T} < I_{V_L^C} \otimes I_{V_T}. \text{ That is, } E_{V_P^N} < E_{V_L^C}.$$

(b) As before, comparing the subdivision processes of the two methods reveals

$$\begin{aligned} & \sum_{l=1}^L |B_l^C| + |G_L^C| + \sum_{l=1}^L |W_l^C| \\ & > \sum_{l=1}^L |B_l^N| + 8 \sum_{l=1}^{L-1} |GB_l^N| + |GB_L^N| + \sum_{l=1}^L |W_l^N| + |GG_L^C| \\ & = \sum_{l=1}^L |B_l^N| + \sum_{l=1}^{L-1} |GB_{l=1}^N| + 7 \sum_{l=1}^{L-1} |GB_{l=1}^N| + |GB_L^N| + \sum_{l=1}^L |W_l^N| + |B_{L+1}^N| + |GB_{L+1}^N| + |W_{L+1}^N| - 7 |GG_L^N| \\ & > \sum_{l=1}^L |B_l^N| + \sum_{l=1}^{L-1} |GB_{l=1}^N| + |GB_L^N| + \sum_{l=1}^L |W_l^N| + |B_{L+1}^N| + |GB_{L+1}^N| + |W_{L+1}^N|, \text{ if } \sum_{l=1}^{L-1} |GB_l^N| > |GG_L^N|. \end{aligned}$$

Therefore, the total number of leaf nodes generated by the conventional method is greater than that generated by the new method.

A remark is in order here. When the subdivision level L is large enough, most nodes in the set of grey-grey nodes GG_L^N are mainly occupied by the white space due to the fact that the sets of black and grey-black nodes, B_L^N and GB_L^N are already extracted from GG_{L-1}^N . Thus, the nodes in GG_L^N will be likely decomposed

into white nodes belonging to W_{L+1}^N with $\frac{|W_{L+1}^N|}{8} \approx |GG_L^N|$. Under this circumstance

when the subdivision level is so high that the size of GG_L^N becomes generally small and the condition $\sum_{l=1}^{L-1} |GB_{l=1}^N| > |GG_L^N|$ is likely to be valid, too.

(c) Next, because the grey-black nodes are not decomposed in the new method, the total number of internal grey nodes generated by the conventional method is greater than the number of the internal grey-grey nodes generated by the new method. Besides, the total number of leaf nodes generated by the conventional method is already shown to be greater than that generated by the new method. Consequently, the total number of leaf and internal grey nodes generated by the

conventional method is greater than that in the new method.

2.7 Conclusions and Future Work

In this chapter we propose a fast and efficient octree construction method to obtain the object model from the multiple silhouettes. The computation of the projection of 3D octants onto the 2D image planes is reduced by using the invariant property of cross ratio. A maximum projection error is specified to decide whether an octant subdivision is needed. The experiments are conducted on three real objects to demonstrate the performance of the new method. The results show the improvement of the new method over the conventional method in terms of memory space, computation time, and quality of the construction result. Theoretical analysis on the new method is presented.

In the future, we plan to convert the reconstructed result to a polygonal representation. Then, we can extract textures from the real object images and map them onto the model to obtain the photo-realistic effect. Moreover, further investigations of other effective use of octant subdivisions are underway. Preliminary results of using both grey-grey and grey-white nodes indicate additional reduction on the memory space as well as the computation time is possible. Also, the progressive octree construction process instead of the recursive construction process is also under consideration.

CHAPTER 3

AN OCTREE CONSTRUCTION METHOD WITH THREE TYPES OF GREY OCTANTS

3.1 Introduction

In the previous chapter a new construction method with a new grey octant type, the so-called “grey-black” octant together with an Exclusive-OR projection error control is presented. This first construction method is to be extended to a method with an additional new grey octant type will be described in this chapter. We shall call this method the second construction method. The computer simulation of this second method is given and the analysis on the experimental results in the form of lemmas is given.

3.2 Types of New Grey Octants

There are five octant types in the second method: B_l , W_l , GB_l , GW_l , and GG_l , $l = 0, 1, 2, \dots, L$. As before, the bounding circle of the projected octant image has a radius $r_{l,v}$ and $d_{l,v}$, the distance map value at the circle center, where $v \in [1, N]$ is the view index and $l \in [0, L]$ is the level index. Assume at each level $l \in [0, L]$ the values of $r_{l,v}$ are nearly equal for $v \in [1, N]$ and the relation: $r_{l,v} = 2r_{l+1,v}$ roughly holds.. The definitions of these octants at a level $l \in [0, L]$ are given below:

- (1) For all views $v \in [1, N]$ if $d_{l,v} > 0$ and $r_{l,v} < d_{l,v}$, then the octant O_l is called a B_l octant.
- (2) For at least one view $v \in [1, N]$ if $d_{l,v} < 0$ and $r_{l,v} < |d_{l,v}|$, then the octant O_l is called a W_l octant.
- (3) For all views $v \in [1, N]$ if $d_{l,v} > 0$ and $r_{l,v} - |d_{l,v}| < p$, then the octant O_l is called a GB_l octant; p is the specified projection upper bound and $r_{l,v} - |d_{l,v}|$ is called the white extent in the projected octant image. The representative view is one with $v^* = \underset{v}{\operatorname{argument}(\max\{r_{l,v} - |d_{l,v}|\})}$

(4) For those views $v \in [1, N]$ with $d_{l,v} < 0$ if $r_{l,v} - |d_{l,v}| < p$, then the octant O_l is called a GW_l octant; $r_{l,v} - |d_{l,v}|$ is called the lack extent in the projected octant image. The representative view is one with $v^* = \underset{v}{\text{argument}}(\max\{r_{l,v} + |d_{l,v}|\})$

(5) If an octant cannot be defined above, it is called a GG_l octant; a GG_l octant whose white and black extents exceed the prespecified p value.

The main differences between the octant types of method 1 and method 2 are:

(1) There is no GW_l octants in method 1. A GW_l^2 octant in the second method is categorized as either as a GG_l^1 octant if the white extent $r_{l,v} + |d_{l,v}|$ is greater than p or a GB_l^1 octant if the white extent $r_{l,v} + |d_{l,v}|$ is smaller than p .

(2) The definitions of a GB_l octant in the two methods are different:

(i) For a GB_l^2 octant, $d_{l,v} > 0$ and $r_{l,v} - |d_{l,v}| > p$ for all views $v \in [1, N]$.

(ii) For a GB_l^1 octant, there are two possibilities: $d_{l,v} > 0$ and $r_{l,v} - d_{l,v} < p$ or $d_{l,v} < 0$ and $r_{l,v} - d_{l,v} < p$ for all views $v \in [1, N]$.

3.3 The Octant Subdivision Algorithm of the Second Construction Method

With the new definitions of the octant types the octant subdivision algorithm of the second construction method is given below:

The Octant Subdivision Algorithm:

At each level $l = 0, 1, \dots, L$ only the GG_l octants need to be subdivided into eight child octants. The octant subdivision process is performed at all levels until there are no GG_l octants at the final level L .

3.4 The Relation between Object Spatial Resolution and Projection Error Upper Bound

In the first and second construction methods the projection error upper bound parameter p is needed to specify by the user. The proper value of p is determined by

the object spatial resolution. If the object detailed part or cavity has a width less than $2p$ then the octant covering the object detailed part will be classified as GB_i^1 or GB_i^2 octant by the two methods, respectively, so this detail disappears; similarly, the octant covering the object detailed cavity will be classified as GW_i^2 octant by the second method, so this detail disappears. Fig. 3.1 shows the depiction of the relation between the object spatial resolution and the projection error upper bound.

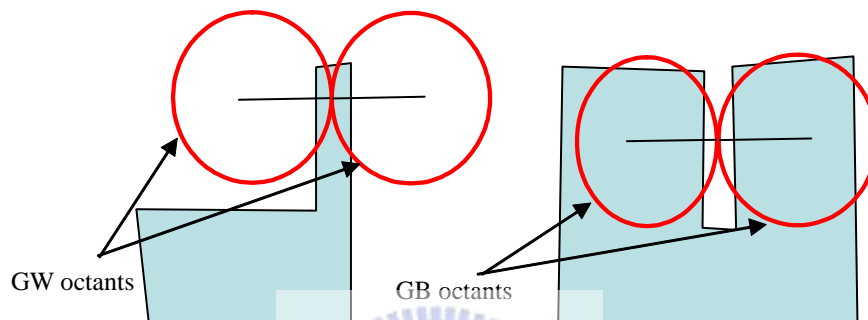


Fig. 3.1. The depiction of the relation between the object spatial resolution and the projection error upper bound.

3.5 Experimental Results

Experiment 1:

In this experiment, we feed the images of a tilted cube used in chapter 2 to method 1 and method 2, respectively. We then list all types of octants generated at each subdivision level, together with the 3D volume of the final constructed models in Tables 3.1 and 3.2 under different error bound and maximum subdivision level of our method and the conventional method. In addition, the graphical display of the constructed object models obtained by the conventional method and the new method is shown in Fig. 3.2.

Table 3.1. Number of the black, grey-black, grey-grey, grey-white and white octants of the constructed cube generated by method 2.

	Protrusion=15					Protrusion=7					Protrusion=2				
	Volume=11755.86					Volume=10333.71					Volume=9345.573				
	B	GB	GG	GW	W	B	GB	GG	GW	W	B	GB	GG	GW	W
0	0	0	1		0	0	0	1		0	0	0	1		0
1	0	0	8		0	0	0	8		0	0	0	8		0
2	0	0	48		16	0	0	48		16	0	0	48		16
3	12	42	125		205	12	16	151		205	12	2	165		205
4	32	177	264		527	130	115	436		527	234	28	531		527
5	0	891	2		1219	319	775	1139		1255	865	218	1910		1255
6	0	0	0		16	67	3473	300		5272	2098	1330	5527		5325
7						0	151	0		2249	9201	6559	9148		19308
8											18014	0	0		55170

Table 3.2. Number of black, grey-black, grey-grey and white octants of the cube generated by method 1.

	Protrusion=15					Protrusion=7					Protrusion=2				
	Volume=9515.625					Volume=9465.82					Volume=9372.925				
	B	GB	GG	GW	W	B	GB	GG	GW	W	B	GB	GG	GW	W
0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
1	0	0	8	0	0	0	0	8	0	0	0	0	8	0	0
2	0	0	34	14	16	0	0	44	4	16	0	0	46	2	16
3	12	42	52	54	112	12	16	122	27	175	12	2	156	9	189
4	32	145	0	125	114	130	115	288	129	314	234	28	498	33	455
5						319	775	1	789	420	865	218	1528	376	997
6						2	2	0	4	0	3098	1330	2270	2670	2856
7											7593	3315	0	4296	2956
8															

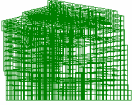
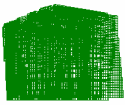

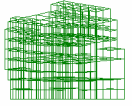
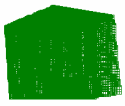
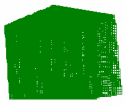
Method 1	Protrusion= 15	Protrusion= 7	Protrusion=2
Image of the constructed model			
Method 2	Protrusion= 15	Protrusion= 7	Protrusion=2
Image of the constructed model			

Fig. 3.2. The comparison between construction results of the synthetic cube obtained by method 1 and method 2.



Experiment 2:

In experiment 2, the same real image sets used in chapter 2 are used to compare the performance of our first method and second method.

Table 3.3 shows the size of a projection octant image at different level. We list the numbers of all types of octants generated at each subdivision level, together with the projection error upper bound value and the quality index of the final constructed models in Table 3.4, Table 3.7 and Table 3.10. Table 3.5, Table 3.8 and Table 3.11 are the octant numbers of the octree models constructed by our first method at $P= 25$, 15 and 7. In Table 3.6, Table 3.9 and Table 3.12, we show the results of applying the GB^2 and GW^2 classification criteria to the GB^1 octant shown in Table 3.5, Table 3.8 and Table 3.11, respectively. As mentioned in chapter two, it is sufficient to use two bits to represent the color of the octant of method 1. However, since we introduce three types of grey octant in this chapter, it requires three bits to represent the five colors of the octants for method 2. Let num_1 and num_2 be the total numbers of octants

generated in method 1 and method 2, respectively. From Lemma 3-2 introduced later in the next section, we obtain that $(num_1/8)$ is roughly equal to num_2 . That is, the total memory space in bits required for method 2 is roughly equal to $3*num_2$. The total memory space in bits required for method 1 is $2*num_1$. So the total number of bits required by method 2 is less, compared to method 1. In addition, the graphical display of the constructed object models and the XOR error images obtained by the conventional method and the new method are shown in Figures 3.3 to 3.5.

Besides feeding the image sets used in chapter 2 into method 2, we also apply the second method to construct the octree model of a flower pot and a dinosaur. Fig. 3.6 and Fig. 3.7 are the selected input images and the generated novel views of the constructed results.

Table. 3.3. The radius range of the circle containing the octant projection at different level.

Level \ Radius range	0	1	2	3	4	5	6	7	8
Min	327	150	72	35	18	9	4	2	1
Max	331	177	92	47	23	11	6	3	1

Table 3.4. Number of the black, grey-black, grey-grey, grey-white and white octants of the constructed cone generated by the second construction method.

	Protrusion=25					Protrusion=15					Protrusion=7				
	Xor=14213+9889					Xor=9010+5957					Xor=5849+3630				
	B	GB	GG	GW	W	B	GB	GG	GW	W	B	GB	GG	GW	W
0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
1	0	0	4	3	1	0	0	4	3	1	0	0	4	3	1
2	0	0	8	0	24	0	0	8	0	24	0	0	8	0	24
3	0	0	10	22	32	0	0	17	15	32	0	0	27	5	32
4	4	29	0	40	7	4	17	24	51	40	4	6	67	35	104
5						10	87	0	89	6	43	88	105	171	129
6											73	373	0	351	43
7															
8															

Table 3.5. Number of black, grey-black, grey-grey and white octants of the cone generated by the first construction method.

	Protrusion=25					Protrusion=15					Protrusion=7				
	Xor=56188+24					Xor=30737+125					Xor=10358+1068				
	B	GB	GG	GW	W	B	GB	GG	GW	W	B	GB	GG	GW	W
0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
1	0	0	7	0	1	0	0	7	0	1	0	0	7	0	1
2	0	0	8	0	48	0	0	8	0	48	0	0	8	0	48
3	0	0	32	0	32	0	0	32	0	32	0	0	32	0	32
4	4	44	64	0	144	4	17	91	0	144	4	6	102	0	144
5	0	137	0	0	375	10	191	140	0	387	43	88	298	0	387
6						0	202	0	0	918	73	582	595	0	1134
7											0	934	0	0	3826
8															

Table 3.6. Number of grey-black and grey-white leaf octants of the cone generated by applying the second method octant type evaluation criteria on grey-black octants in Table 3.3.

	Protrusion=25					Protrusion=15					Protrusion=7				
	B	GB	GG	GW	W	B	GB	GG	GW	W	B	GB	GG	GW	W
0		0		0			0		0			0		0	
1		0		0			0		0			0		0	
2		0		0			0		0			0		0	
3		0		0			0		0			0		0	
4		29		15			17		0			6		0	
5		9		128			99		92			88		0	
6							0		202			418		164	
7												170		764	
8															

Table 3.7. Number of the black, grey-black, grey-grey, grey-white and white octants of the constructed vase generated by the second construction method.

	Protrusion=25					Protrusion=15					Protrusion=7				
	Xor=58496+7317					Xor=33007+4825					Xor=22311+2777				
	B	GB	GG	GW	W	B	GB	GG	GW	W	B	GB	GG	GW	W
0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
1	0	0	8	0	0	0	0	8	0	0	0	0	8	0	0
2	0	0	25	16	23	0	0	30	11	23	0	0	34	7	23
3	1	11	62	53	73	1	6	82	43	108	1	0	107	26	138
4	30	165	0	245	56	52	114	132	233	125	88	55	348	125	240
5						47	483	0	483	43	297	524	577	798	588
6											530	1998	0	1828	260
7															
8															

Table 3.8. Number of black, grey-black, grey-grey and white octants of the vase generated by the first construction method.

	Protrusion=25					Protrusion=15					Protrusion=7				
	Xor=150755+4					Xor=88134+22					Xor=35498+498				
	B	GB	GG	GW	W	B	GB	GG	GW	W	B	GB	GG	GW	W
0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
1	0	0	8	0	0	0	0	8	0	0	0	0	8	0	0
2	0	0	41	0	23	0	0	41	0	23	0	0	41	0	23
3	1	11	122	0	194	1	6	127	0	194	1	0	133	0	194
4	30	233	269	0	444	52	114	406	0	444	88	55	477	0	444
5	0	716	0	0	1436	47	1046	658	0	1497	297	524	1498	0	1497
6						0	1043	0	0	4221	530	3153	3002	0	5299
7											0	6117	0	0	17899
8															

Table 3.9. Number of grey-black and grey-white leaf octants of the vase generated by applying the second method octant type evaluation criteria on grey-black octants in Table 3.6.

	Protrusion=25					Protrusion=15					Protrusion=7				
	B	GB	GG	GW	W	B	GB	GG	GW	W	B	GB	GG	GW	W
0		0		0			0		0			0		0	
1		0		0			0		0			0		0	
2		0		0			0		0			0		0	
3		11		0			6		0			0		0	
4		166		67			114		0			55		0	
5		78		638			577		469			524		0	
6							0		1043			2288		865	
7												1157		4960	
8															

Table 3.10. Number of the black, grey-black, grey-grey, grey-white and white octants of the constructed boy sculpture generated by the second construction method.

	Protrusion=25					Protrusion=15					Protrusion=7				
	Xor=94149+2783					Xor=52185+2410					Xor=30781+2402				
	B	GB	GG	GW	W	B	GB	GG	GW	W	B	GB	GG	GW	W
0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
1	0	0	8	0	0	0	0	8	0	0	0	0	8	0	0
2	0	0	38	13	13	0	0	44	7	13	0	0	47	4	13
3	7	32	96	91	78	7	12	148	71	114	7	6	192	34	137
4	49	309	0	347	63	121	263	238	366	196	160	100	630	218	428
5						94	864	0	884	62	747	1032	1046	1373	842
6											867	3677	0	3354	470
7															
8															

Table 3.11. Number of black, grey-black, grey-grey and white octants of the boy sculpture generated by the first construction method.

	Protrusion=25					Protrusion=15					Protrusion=7				
	Xor=192434+8					Xor=115143+8					Xor=48366+296				
	B	GB	GG	GW	W	B	GB	GG	GW	W	B	GB	GG	GW	W
0	0	0	1		0	0	0	1		0	0	0	1	0	0
1	0	0	8		0	0	0	8		0	0	0	8	0	0
2	0	0	51		13	0	0	51		13	0	0	51	0	13
3	7	32	200		169	7	12	220		169	7	6	226	0	169
4	49	414	444		693	121	263	683		693	160	100	855	0	693
5	0	1192	0		2360	94	1803	1132		2435	747	1032	2626	0	2435
6						0	1756	0		7300	867	5771	5210	0	9160
7											0	10985	0	0	30695
8															

Table 3.12. Number of grey-black and grey-white leaf octants of the boy sculpture generated by applying the second method octant type evaluation criteria on grey-black octants in Table 3.9.

	Protrusion=25					Protrusion=15					Protrusion=7				
	B	GB	GG	GW	W	B	GB	GG	GW	W	B	GB	GG	GW	W
0		0		0			0		0			0		0	
1		0		0			0		0			0		0	
2		0		0			0		0			0		0	
3		32		0			12		0			6		0	
4		311		103			263		0			100		0	
5		106		1086			974		829			1032		0	
6							9		1747			4110		1661	
7												1861		9124	
8															














First method	Protrusion= 25	Protrusion= 15	Protrusion= 7
Image of the constructed model			
Xor error image of the images			
			
Second method	Protrusion= 25	Protrusion= 15	Protrusion= 7
Image of the constructed model			
Xor error image of the images			

Fig. 3.3. The comparison between construction results of the cone obtained by the conventional method and the new method.













First method	Protrusion= 25	Protrusion= 15	Protrusion= 7
Image of the constructed model			
Xor error image of the image			
Second method	Protrusion= 25	Protrusion= 15	Protrusion= 7
Image of the constructed model			
Xor error image of the image			

Fig. 3.4. The comparison between construction results of the vase obtained by the conventional method and the new method.













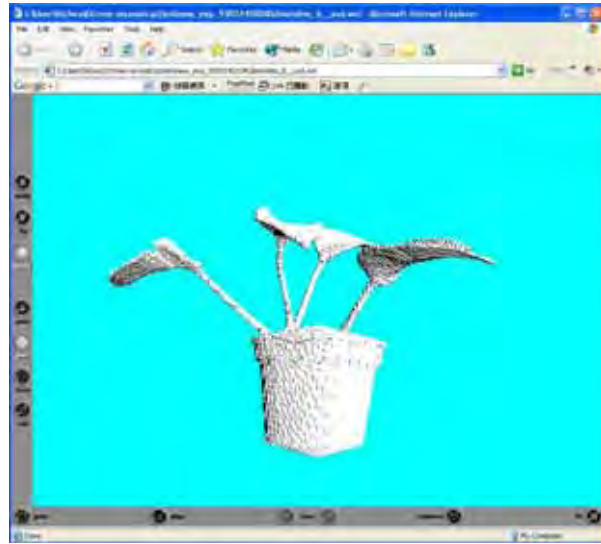
First method	Protrusion= 25	Protrusion= 15	Protrusion= 7
Image of the constructed model			
Xor error image of the image			
Second method	Protrusion= 25	Protrusion= 15	Protrusion= 7
Image of the constructed model			
Xor error image of the image			

Fig. 3.5. The comparison between construction results of the boy sculpture obtained by the conventional method and the new method.



(a)

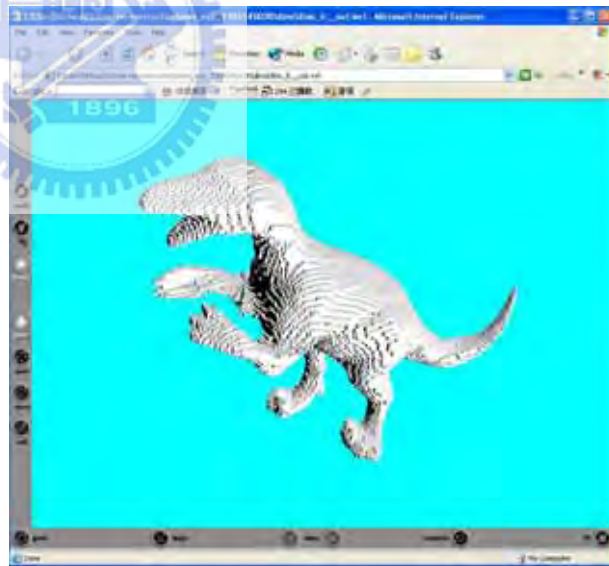


(b)

Fig. 3.6. (a) One of the input image to the second method. (b) The new view generated from the constructed octree model.



(a)



(b)

Fig. 3.7. (a) One of the input image to the second method. (b) The new view generated from the constructed octree model.

3.6 Analytical Analysis

Lemma 3-1:

For a given projection error bound p assume $2r_{L+1,v} < p < 2r_{L,v}, v \in [1, N]$, then the maximum level of subdivision of method 1, L_{\max}^1 , is equal to $L_{\max}^1 = L+1$ or $L_{\max}^1 = L$.

Proof:

Due to $2r_{L+1,v} < p < 2r_{L,v}$, any octant at subdivision level $l = L+1$ has a projection image whose bounding circle diameter is smaller than p . It implies the white extent is smaller than p , so it is a GB^1 octant; no further octant subdivision is needed. Sometimes, when $2r_{L,v}$ is closer to p than $2r_{L+1,v}$, if there is no octant O_L at level $l = L$ such that its white extent $> p$, then $L_{\max}^1 = L+1$.

Lemma 3-2:

The maximum subdivision level of method 2, L_{\max}^2 , is no greater than that of method 1, i.e., $L_{\max}^2 \leq L_{\max}^1$. In general, $L_{\max}^2 < L_{\max}^1$.

Proof:

Because the octants $GW_l^2, l = 1, \dots, L_{\max}^2$, in the second method are reclassified as:
 (1) GG_l^1 if both $r_{l,v} + |d_v| > p$ and $d_v < 0$ or (2) GB_l^1 if both $r_{l,v} - |d_v| < p$ and $d_v < 0$ in the first construction method. Therefore, the GG_l^1 octants will be further subdivided in the first construction method. Consequently, the first construction method has generally a greater number of subdivision levels, i.e., $L_{\max}^2 \leq L_{\max}^1$. Only when there are no GW_l^2 octants which are reclassified as GG_l^1 octants at the level $l = L_{\max}^2$, then $L_{\max}^2 = L_{\max}^1$.

Lemma 3-3a:

If $r_l > p$, then $GB_l^1 = GB_l^2$. If $r_l < p$, then $GB_l^1 \supseteq GB_l^2$ where $r_l = \max_v \{r_{l,v}\}$.

Proof:

Consider the case where $r_l > p$. It implies the second relation “ $d_{l,v} < 0$ and $r_{l,v} - d_{l,v} < p$ ” does not hold, since the fact that $r_{l,v} - d_{l,v} = r_{l,v} + |d_{l,v}| < p$ is contradictory to the assumption that $r_l > p$. So only the first relation holds. Furthermore, the first relation also implies the black extent in the GB_l^1 octant is equal to $r_{l,v} + d_{l,v} < p$. Therefore, the GB_l^1 octant is not contained in any GB_l^2 octant. Thus, $GB_l^1 = GB_l^2$. Next, consider the case where $r_l < p$. The first relation “ $d_{l,v} > 0$ and $r_{l,v} - d_{l,v} < p$ ” is satisfied by a GB_l^1 octant as well as a GB_l^2 octant. However, the second relation “ $d_{l,v} < 0$ and $r_{l,v} - d_{l,v} < p$ ” is satisfied by a GB_l^1 octant if it exists, but it is definitely not satisfied by a GB_l^2 octant since $d_{l,v} < 0$. Therefore, $GB_l^1 \supseteq GB_l^2$.

**Lemma 3-3b:**

If $\min_v \{2r_l\} > p$, then $B_l^1 = B_l^2$. If $\min_v \{2r_l\} < p$, then $B_l^1 \supseteq B_l^2$.

Proof:

If $\min_v \{2r_l\} > p$, then an octant B_l^2 contains the black extent whose length is $2r_{l,v}$, $v \in [1, N]$. Since $2r_{l,v}$ is greater than p , so this octant B_l^2 will not be contained in a larger B_l^1 with $l^- < l$. If $\min_v \{2r_l\} < p$, then for those views with $2r_{l,v} < p$, the black extent is smaller than p , so the octant may be contained in a larger B_l^1 with $l^- < l$, so this B_l^2 octant will not be produced. Thus, $B_l^1 \supseteq B_l^2$.

Lemma 3-3c:

For $l = 0, 1, 2, \dots, L_{max}^2$, $W_l^2 \subseteq W_l^1$ and $GG_l^2 \subseteq GG_l^1$.

Proof:

By definition a white octant in W_l^2 produced in the second method is also a white octant in W_l^1 and a grey-grey octant GG_l^2 is also a GG_l^1 octant. Due to the existence of GW_l^2 octants in the second method, a less number of octant subdivisions is performed in the second method, so the numbers of white, and grey-grey octants produced in the second method are less, compared to those produced in the first method.

Lemma 3-4a:

Let the GB difference set be defined as $DGB_m = GB_m^1 - GB_m^2$, $m = L, L+1$. There are more octants reclassified as GW_m^2 than those reclassified as GB_m^2 by the second method for the set of GB_l^1 octants belonging to DGB_m .

Proof:

It can be shown that if the octants in GW_l^2 , $l = 0, 1, 2, \dots, L_{max}^2$, are further processed in the first method, GW_l^2 will be reclassified as either GG_l^1 if the white extent $> p$ or GB_l^1 if the white extent $< p$. In the latter case the GB_l^1 octant is viewed a GW_l^2 octant in the second method. Then, the final octree model is the same as that obtained by the first method.

There are eight child octants in a subdivided parent octant. At level $m = L$ and $L+1$, the projected octant is sufficiently small so that the object silhouette boundary can be approximated by a straight line. The distance map value at the parent octant circle center indicates the shortest distance between the circle center and the object silhouette boundary, so draw a line perpendicular to the object silhouette boundary, as shown in Fig. 3.8. Suppose the object interior is on the left handed side of the boundary whose position is indicated by x_B .

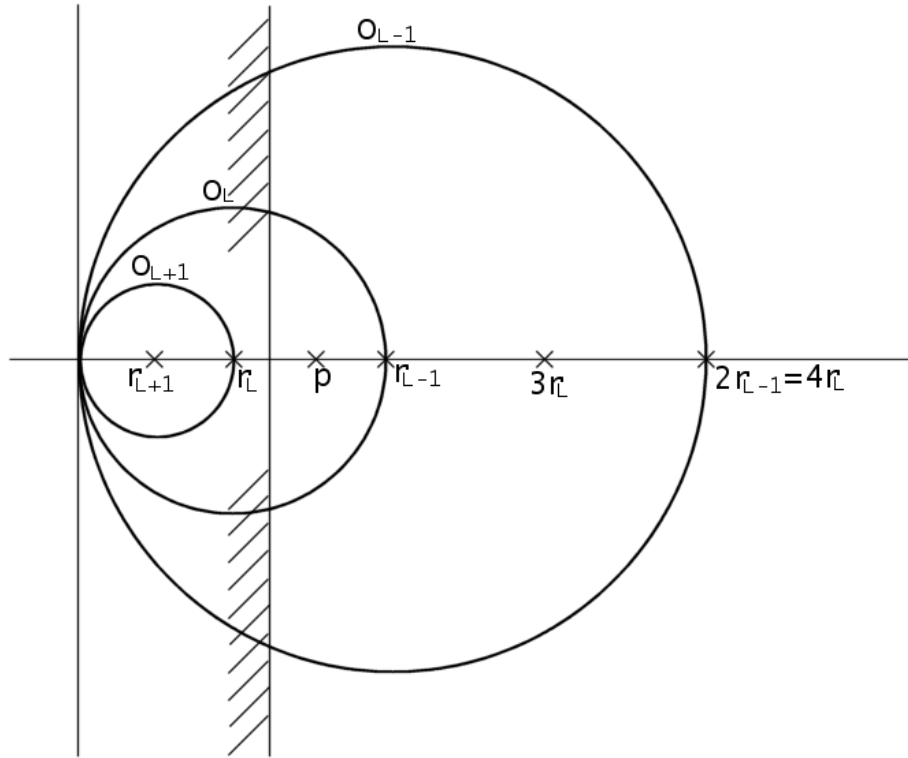


Fig. 3.8. The depiction of centroids of the bounding circles for an octant and its parent octant and child octant. (Assume $r_L < p < r_{L-1} = 2r_L$)

Proof of the lemma for the level L :

Now suppose O_{L-1} is viewed as a GG_{L-1}^1 octant in the first method and it will be subdivide into eight child octants. First, assume there exists a child octant O_L whose circle center is located at $x = r_L$. Later on, a child octant whose circle center is not located on the horizontal axis will be considered. Now suppose O_L is a GB_L^1 octant in the first method. Both octants O_{L-1} and O_L are offspring of a GW_L^2 ancestor at a lower level. Then

- (a) The parent GG_{L-1}^1 octant is with a white extent $> p$.
- (b) The child GB_L^1 octant is with a white extent whose value range is $0 < \text{white extent} < p$.
- (c) Both of the parent and child octants have a black extent less than p .

The above properties imply the following inequalities:

- (1) The white extent $= 2r_{L-1} - x_B > p$, so $x_B < 2r_{L-1} - p (= 2r_L + 2r_L - p)$.
- (2) The white extent is given by $0 < 2r_L - x_B < p$, so $2r_L - p < x_B < 2r_L$ or $r_L - (p -$

$$r_L) < x_B < 2r_L.$$

(3) the black extent is given by $0 < x_B < p - \Delta$.

Based on the above inequalities the allowed range of x_B is given by

$$r_L - (p - r_L) < x_B < p - \Delta.$$

The range of x_B can be further divided into two sub-ranges:

(a) If $r_L - (p - r_L) < x_B < r_L$, then the GB_L^1 is also as a GW_L^2 octant.

(b) If $r_L < x_B < p - \Delta$, then the GB_L^1 octant is also a GB_L^2 octant.

The ratio of the probabilities of these two cases is given by

$$\begin{aligned} \frac{\text{Probability of } GB_L^1 \text{ being a } GW_L^2 \text{ octant}}{\text{Probability of } GB_L^1 \text{ being a } GB_L^2 \text{ octant}} &= \frac{r_L - r_L + p - r_L}{p - \Delta - r_L} \\ &= \frac{p - r_L}{p - \Delta - r_L} \geq 1, \text{ since } \Delta \geq 0. \end{aligned}$$

Next, to prove the lemma for the case where the level is $L+1$:

Consider the child octant O_{L+1} , which is classified as a GB_{L+1}^1 by the first method, and its parent O_L , which is classified as a GG_L^1 by the first method, then

- (a) The parent GG_L^1 octant is with a white extent $> p$,
- (b) The child GB_{L+1}^1 octant is with a white extent satisfying $0 < \text{white extent} < p$,
- (c) Since both GB_{L+1}^1 and GG_L^1 are offspring of a GW_l^2 ancestor with $l \leq L-1$, so the black extent in either octant is less than p .

The above properties imply the following inequalities:

- (1) $2r_L - x_B > p$ or $x_B < 2r_L - p (= 2r_{L+1} - (p - 2r_{L+1}))$,
- (2) $0 < 2r_{L+1} - x_B < p$ or $2r_{L+1} - p < 0 < x_B < 2r_{L+1}$,
- (3) $0 < x_B < p - \Delta (\approx p)$

Combining these inequalities yields

$$0 < x_B < 2r_L - p, \text{ since } 2r_L - p < P \text{ or } r_L < p.$$

The range of x_B , $[0, 2r_L - p]$, can be written as $[0, r_{L+1}] \cup [r_{L+1}, 2r_L - p]$.

- (1) If $0 < x_B < r_{L+1}$, then GB_{L+1}^1 is a GW_{L+1}^2 ,
- (2) If $r_{L+1} < x_B < 2r_L - p$, then GB_{L+1}^1 is also a GB_{L+1}^2 .

$$\frac{\text{Probability of } GB_{L+1}^1 \text{ being a } GW_{L+1}^2 \text{ octant}}{\text{Probability of } GB_{L+1}^1 \text{ being a } GB_{L+1}^2 \text{ octant}} = \frac{r_{L+1} - 0}{2r_L - p - r_{L+1}}$$

$$= \frac{r_{L+1}}{3r_{L+1} - p} > 1, \text{ since } r_{L+1} > 3r_{L+1} - p, \text{ i.e., } p > 2r_{L+1}.$$

One more thing to consider is to remove the assumption that the child octant O_L whose circle center is located at $x = r_{L+1}$; that is, the child octant considered has its position shifted to the right, see Fig.3.9, then the object silhouette boundary becomes closer to its circle center, so the probability of the octant GB_L^1 being also a GW_L^2 becomes even larger than that the octant GB_L^1 being also a GB_L^2 , so the lemma still holds.

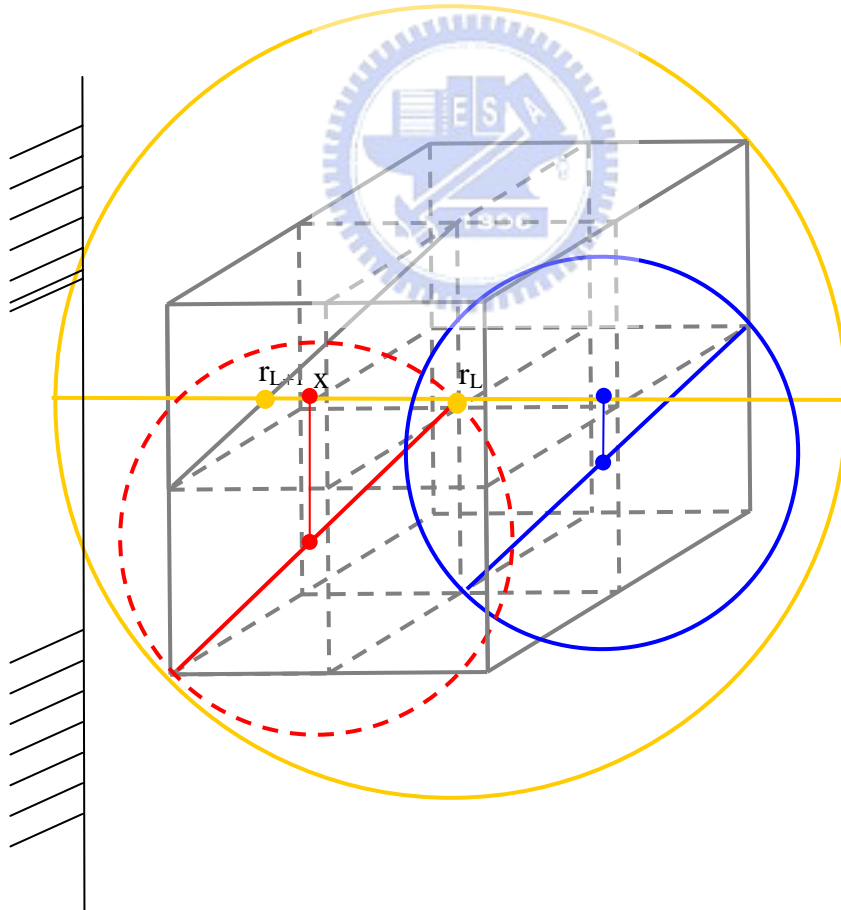


Fig. 3.9. The spatial relation depiction of bounding circles for an octant and its two sub-octants

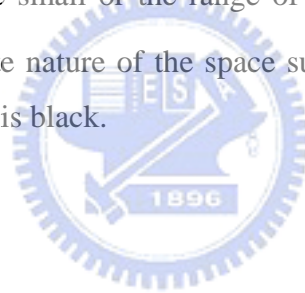
Lemma 3-4b

Consider the level $l = L$ or $L+1$. If $\min_v \{2r_{l,v}\} < p < \max_v \{2r_{l,v}\}$, then the GG_{l-1} octants may produce B_l child octants. However, the B_l child octants are much less in number compared to the GB_l child octants produced.

Proof:

Due to the equipment set-up, there is inevitably a small variation in the projected octant radius $r_{l,v}$ from view to view at each level l .

A GG_{l-1} parent octant O_{l-1} contains a white extent $> p$ for some view v^* , $v^* \in [1, N]$. In this view v^* , the octant O_{l-1} is half or even more occupied by the white (i.e. empty) space. Its eight child octants are more likely to be white or grey-black than black. The chance that the child octant is a black octant is rare. This happens only when p is quite small or the range of $\{r_{l,v}, v \in [1, N]\}$ is sufficiently large. Then due to the discrete nature of the space subdivision, the child octant of a half- or more-occupied octant is black.

**Lemma 3-5:**

The octree model constructed by the first method contains more terminal octants than the one constructed by the second method. This is also true for the non-terminal octants.

Proof:

The terminal octants of the two constructed octree models consist of the sets $\{B_1^1, W_1^1, GB_1^1, 1 = 0, 1, 2, \dots, L_{\max}^1\}$ and $\{B_1^2, W_1^2, GB_1^2, GW_1^2, 1 = 0, 1, 2, \dots, L_{\max}^2\}$, respectively. The non-terminal octants are the sets of $\{GG_1^1, 1 = 0, 1, 2, \dots, L_{\max}^1 - 1\}$ and $\{GG_1^2, 1 = 0, 1, 2, \dots, L_{\max}^2 - 1\}$. The lemma is a consequence of Lemmas 3a and 3b.

A remark is in order here. All the terminal and non-terminal octants are counted in the computer processing time for the octree model construction. However, only the black octants and grey black octants are required to store for displaying the

constructed object model. In these two respects, the second construction method outperforms the first one.

Lemma 3-6:

The octree model constructed by the first method has no less Exclusive-OR projection error than the one constructed by the second method.

Proof:

The lemma is a consequence of Lemmas 3-3a to 3-3c and 3-4a to 3-4b.

3.7 Conclusion

In this chapter the second construction method is presented. It is an extension to the first method. The octree model produced by the first method is one that covered the real object since the GB_i^1 octants are treated as the black octants, while the octree model produced by the second method is one that approximates the real object from both sides of the object surface since the GB_i^2 octants are treated as the black octants and the GW_i^2 octants are treated as the white octants.

The memory storage and the computer processing time required by the second method have been shown much less than those required by the first method.

CHAPTER 4

A PROGRESSIVE OCTREE CONSTRUCTION METHOD WITH A PROJECTION ERROR MEASURE

4.1 Introduction

In the previous two chapters, we recursively subdivide a partially occupied octree octant into smaller octants until all the generated octants are entirely inside, outside the object or within certain specified protrusion error. Since the octree is a hierarchical representation for the 3D space, there are two different traversal schemes can be used to traverse the grey octants in the octree structure: the depth-first traversal (DFS) and the breadth-first traversal (BFS). DFS is implemented with a stack data structure and the octants are processed on a basis of first-in and last-out (FILO). That is, descendent octants are processed before ascendant octants. On the contrary, BFS is implemented with the queue data structure and the octants are served in a first-in and first-out (FIFO) paradigm; that is, descendent octants are processed after ascendant octants. Since the ascendant octant is larger than the descendent octants in size in the octree construction method, the Exclusive-OR projection error decreases more rapidly by using the BFS scheme than by using the DFS scheme.

However, the octant selection in the above two implementations using BFS or DFS does not take the individual octant projection errors into account. The octant chosen to be processed is depended on the position of the octant in the structure of the octree. If we just have a limited amount of time to construct the object, we should spend much of the processing time on those octants which give greater improvements for the construction result. In this chapter, we propose a best-first tree traversal (BestFS) scheme for traversing the octree structure while constructing the octree model. Instead of using the stack or the queue to book keeps the grey octants to be processed, we propose to use a priority queue by sorting the projection errors of the octants to be processed. The advantages of using a priority queue to select an octant for processing are:

- (a) The maximum projection error reduction by processing the chosen octant

from the priority queue can be achieved.

- (b) The projection error reduction rate for per unit of time by using the best-first traversal scheme is better than by using the breadth-first traversal scheme or the depth-first traversal scheme.

This chapter is organized as follows. In Section 2, we briefly review the depth-first and breadth-first traversal for the octree structure. In Section 3, a best-first tree traversal algorithm is proposed by using the sorted octant queue based on the projection error. In Section 4, we provide a theoretical analysis on the performance of the progressive method of using a priority queue, a queue, a stack or a recursive call. Section 5 is the preliminary experimental results. Section 6 is the conclusion remarks.

4.2 Depth-First Tree Traversal Scheme vs. Breadth-First Tree Traversal Scheme

In this section we briefly introduce the depth first tree traversal and breadth-first tree traversal schemes using the stack and queue data structures, respectively. The traversal paradigms for depth-first and breadth-first traversal are:

- (A) Depth-first tree traversal (DFS) for traversing the octree structure

DFS is implemented with a stack data structure and the octants are processed on a basis of first-in and last-out (FILO). That is, descendent octants are processed before ascendant octants. Fig. 4.2 shows the visiting ordering for the depth-first tree traversal for the octree structure shown in Fig. 4.1. One can see that computation efforts are wasted on the small octants in the beginning of the construction process.

- (B) Breadth-first traversal (BFS).

BFS is implemented with the queue data structure and the octants are served in a first-in and first-out (FIFO) paradigm; parent octants are processed before the children octants. Sibling octants are processed according to their positions in the octree. Fig. 4.3 shows the visiting ordering for the breadth-first tree traversal.

The octant chosen to be processed in the depth-first traversal or breadth-first traversal is depended on the position of the octant in the structure of the octree. However, the intersection area of the octant image with the silhouette image is not taken into consideration. In the next section, we propose a best first traversal paradigm for the construction of the octree using the sorted octant queue based on the octant projection error.

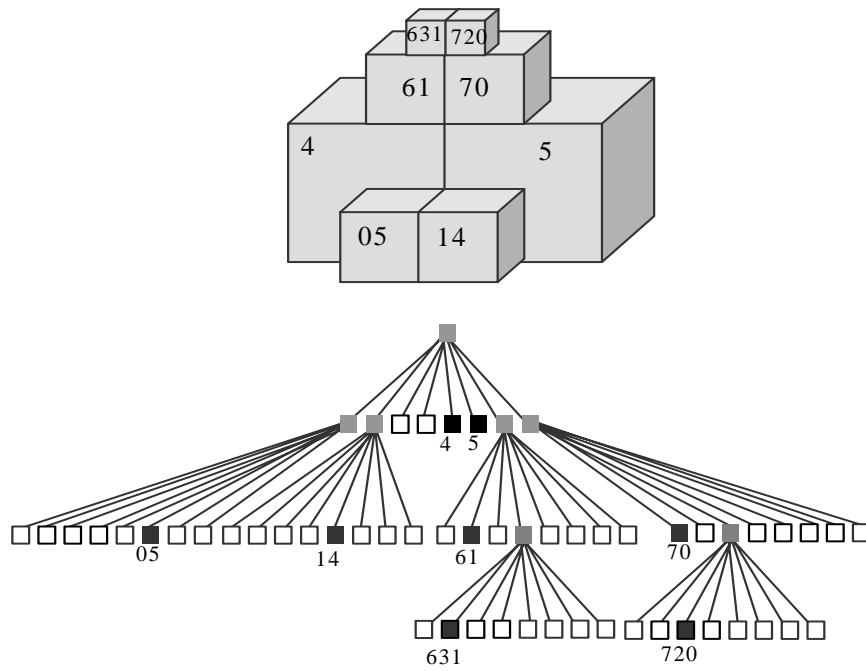


Fig. 4.1. A typical octree structure.

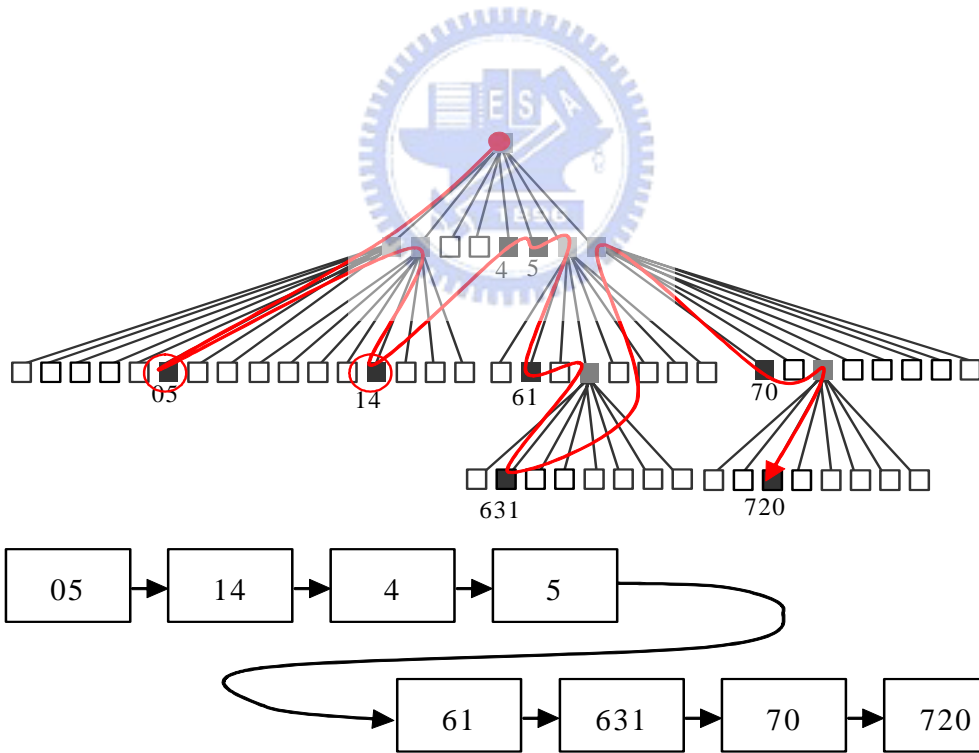


Fig. 4.2. The visiting ordering of the black nodes by the depth-first tree traversal scheme for the octree in Fig. 4.1.

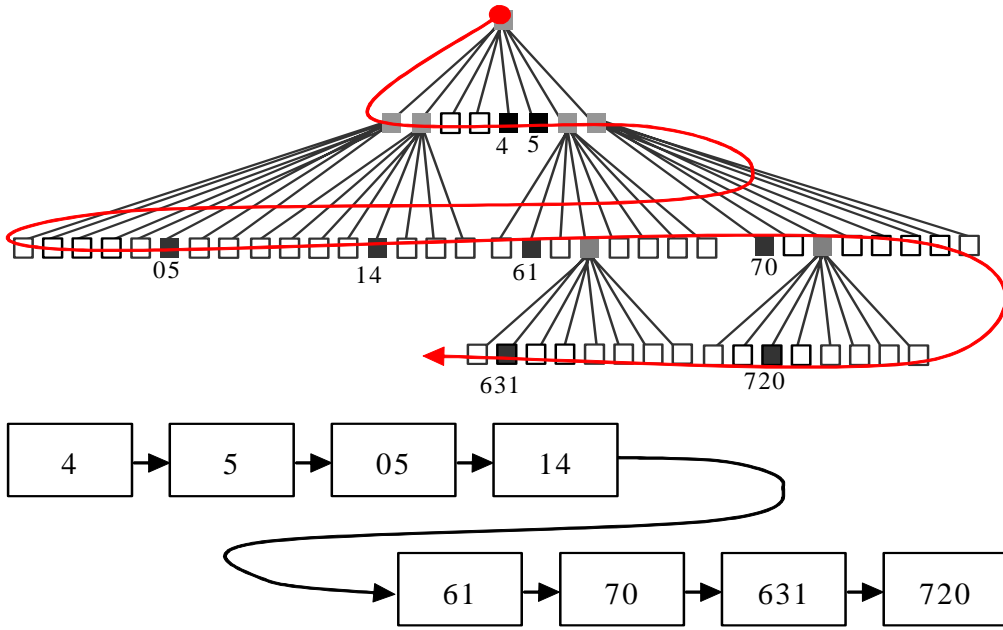


Fig. 4.3. The visiting ordering of the black nodes by the breadth-first tree traversal scheme for the octree in Fig. 4.1.

4.3 Best-First Tree Traversal Scheme with Octant Sorting Based on Projection Error

4.3.1 Octant Sorting Based on Projection Error

In the previous sections, an octant is further subdivided into 8 smaller sub-octants if its image intersects the silhouette images (it is a grey octant). The 8 smaller sub-octants are then pushed into the stack or appended to the tail of the queue for both breadth-first traversal and depth-first traversal schemes. In this section, we propose to use the priority queue to store the sub-octants to be processed. Once an octant is evaluated as a grey octant, it is subdivided into 8 smaller sub-octants. We first check the projection error of each sub-octant. The projection error of the sub-octant is defined as:

$$proj_error(O) = \max\{proj_error_i(O) \mid i = 0..N-1\}$$

$$\text{where } proj_error_i(O) = \begin{cases} r_i + |d_i| & \text{if } d_i \leq 0 \\ r_i - |d_i| & \text{if } d_i > 0 \end{cases}, \text{ which is the projection error}$$

of octant O in image I_i;

The sub-octant is then inserted into the priority queue, called *priority_queue*, based

on the projection error $proj_error(O)$.

The definitions of octant types described in chapter 2 and chapter 3 can be applied here to determine whether a grey octant is to be inserted into the priority queue or not. The advantage of using the extended grey octant definitions is that “GB¹”, “GB²” or “GW²” octants will not be inserted into the priority queue. Thus, both the memory storage and the number of insertion operations for maintaining the priority queue are reduced significantly. The modified insertion processes are shown in Fig. 4.4 and Fig. 4.5 for the first construction method (method 1) and the second construction method (method 2).

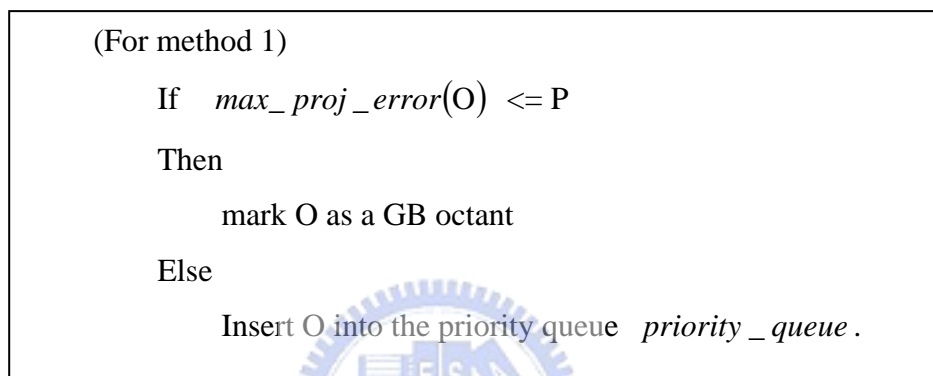


Fig. 4.4. The modified insertion process which avoids inserting GB¹ octant into the priority queue.

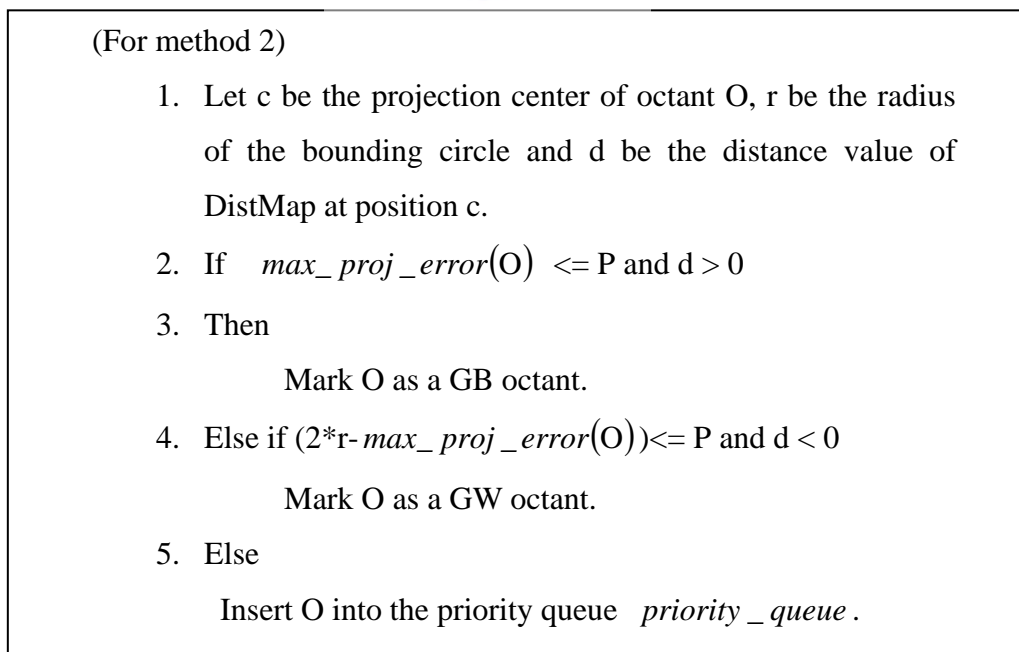


Fig. 4.5. The modified insertion process which avoids inserting GB² or GW² octant into the priority queue.

4.3.2 The Retrieval of Octant to Be Processed by the Best-First Tree Traversal Scheme

In this section, we will introduce the proposed octree construction method using the best first traversal Scheme with a priority queue. The algorithm is listed in Fig. 4.6.

1. Subdivide the root octant into 8 smaller octants, $SubO_i, i= 1, \dots, 8$.
2. Insert $SubO_i$ into the priority queue *priority_queue* if $SubO_i$ is classified as the grey octant, $i=1, \dots, 8$.
3. While(*priority_queue* is empty or program_termination is not issued by user)
4. {
 - Retrieve the first octant O_f in *priority_queue*.
 - Subdivide O_f into 8 smaller octants, $SubO_{fi}, i = 1, \dots, 8$.
 - Perform the intersection test of $SubO_{fi}$ against the silhouette images and compute the projection error of $SubO_{fi}, i = 1, \dots, 8$.
 - Insert $SubO_i$ into the priority queue *priority_queue* if $SubO_i$ is classified as the grey octant, $i = 1, \dots, 8$.}
5. Generate the construction octree.

4.4 Experimental Results

In this section, we implement the conventional method using the breadth-first traversal scheme, named BFS, the depth-first traversal scheme, named DFS, and the best-first traversal scheme, named BestFS, respectively. Three experiments are then conducted to compare the performance of the three methods, BestFS, BFS and DFS.

In the first experiment, we show the 2D projection results for the progressive octree construction method, the conventional octree construction method implemented using queue and stack data structure under different numbers of octants being processed. In the second experiment, we compute the projection XOR errors for the octree construction method using a priority queue, a queue or a stack for the first 5000 octants processed. In the third experiment, we show the priority queue length vs. the processed octant number of our progressive method with “GB¹”, “GB²” or “GW²” octant definitions associated with different protrusion values of p .

Experiment 1

In the first experiment, we apply BestFS, BFS and DFS method to construct a cone. The constructed octree model is generated when every grey octant is subdivided. We then project the constructed octree model on the image and show the XOR error map of the octree projection image and the silhouette image. Figs. 4.7, 4.8 and 4.9 are the XOR maps of the constructed octree obtained when subdividing the first 11 grey octants using the DFS, BFS and BestFS methods, respectively. One can see that the reduction of XOR error for DFS is little. This is due to the occlusion from the projection of the sibling octants at higher levels. The BFS and BestFS methods provide a fast refinement of the octree model visually. In the later experiment and section, we will compare these two methods in more detail.

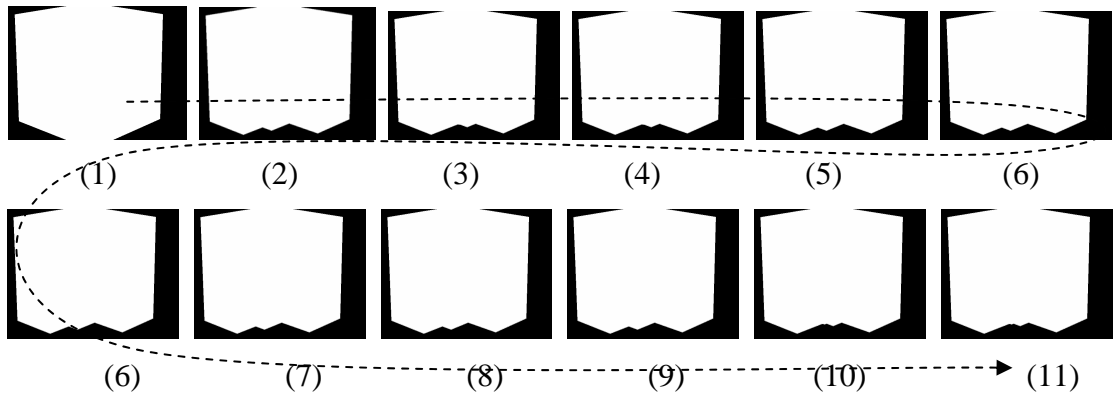


Fig. 4.7. The constructed octree model of the first 11 octants processed by the DFS method when viewed from the 1st viewing direction.

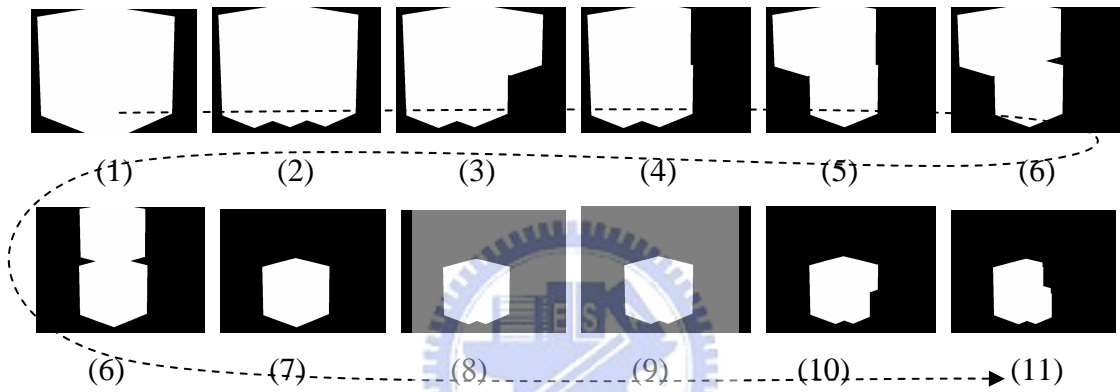


Fig. 4.8. The constructed octree model of the first 11 octants processed by the BFS method when viewed from the 1st viewing direction.

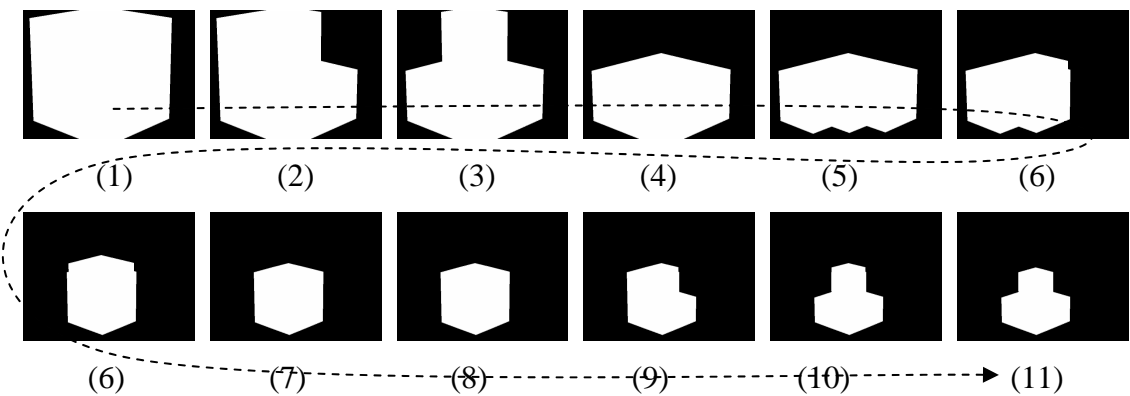


Fig. 4.9. The constructed octree model of the first 11 octants processed by the BestFS method when viewed from the 1st viewing direction.

Experiment 2

In this experiment, we use the previous methods to construct the models for two

objects: the vase and the boy sculpture. We then compare the performance of the three methods, BestFS, BFS and DFS from the results of the constructed model of the vase and boy sculpture under different octant number being processed (to be replaced see the previous wording) . Fig. 4.10 is the XOR projection error of the constructed vase vs. the subdivided octant number evaluated from three methods. Fig. 4.11 is the XOR projection error of the constructed vase vs. the subdivided octant number evaluated from three methods. In Fig. 4.10 and Fig. 4.11, we can see that the XOR projection errors decrease slightly for the octree model constructed from method DFS. The XOR projection errors for the octree model constructed from method BFS decrease much faster. However, the decreasing curve is not smooth. The XOR projection errors for the octree model constructed from the proposed method BestFS decrease much faster than the results from BFS. One can see that under any subdivided octant number, the XOR error of the octree model constructed from BestFS is smaller than the one constructed from BFS. In the later section, we will give an analytic analysis on this property.

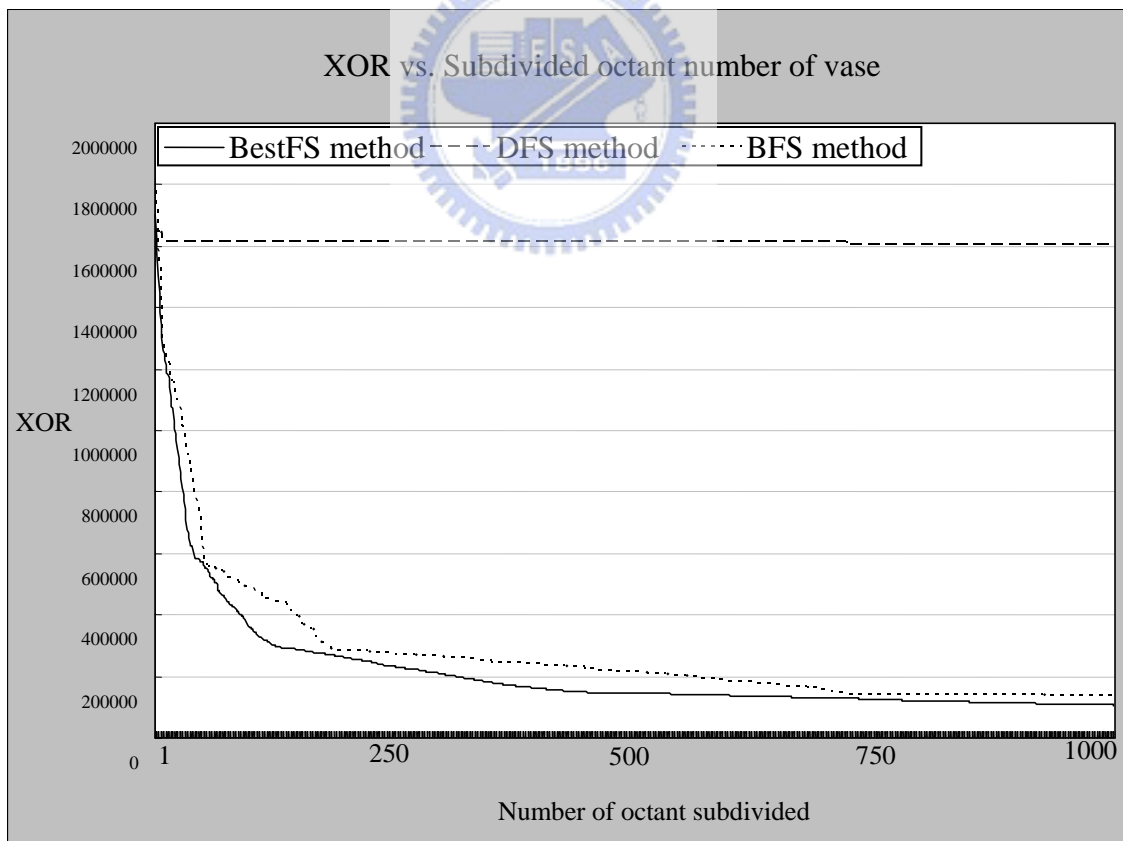


Fig. 4.10. The plot of XOR vs. subdivided octant number of the octree model for the vase constructed from the DFS, BFS and BestFS methods.

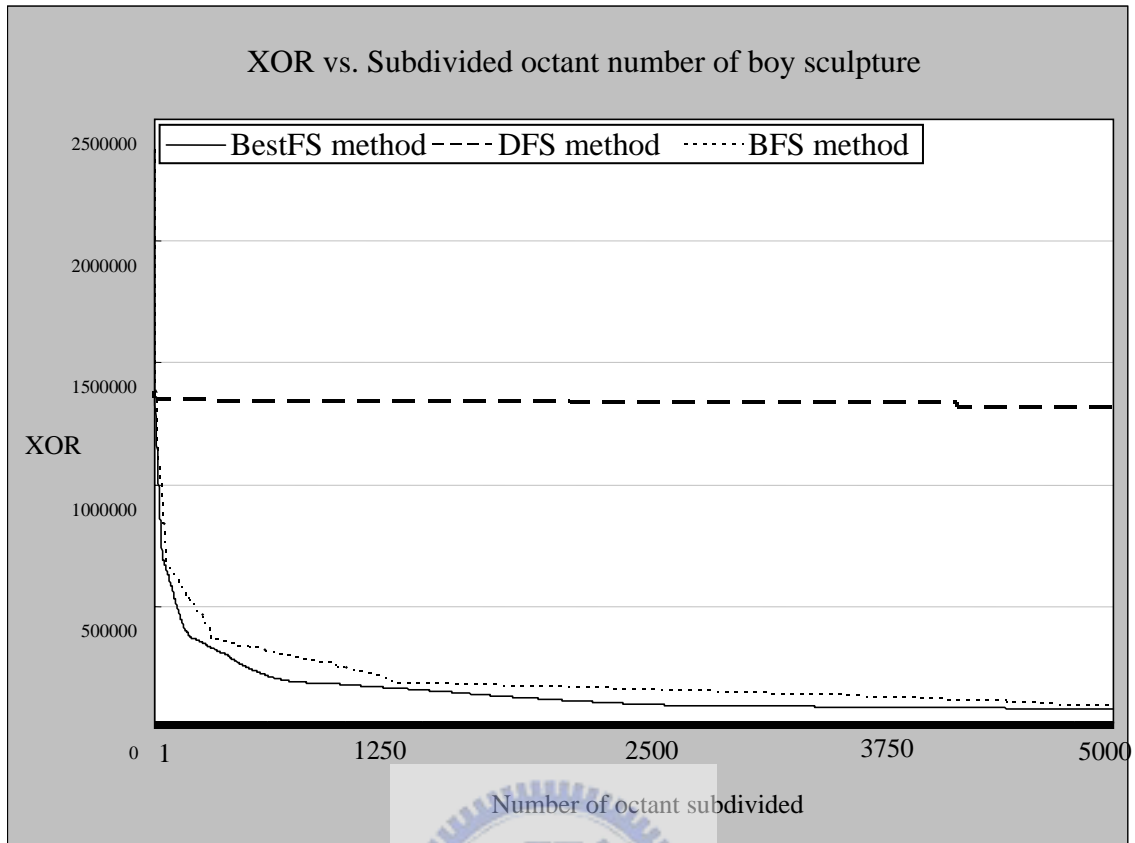


Fig. 4.11. The plot of XOR vs. Subdivided octant number of the octree model for the boy sculpture constructed from the DFS, BFS and BestFS methods.

Experiment 3

In this experiment, we incorporate GB^1 , GB^2 and GW^2 classification criteria proposed in chapter 2 and chapter 3 before inserting the grey octants into the priority queue. The proposed BestFS method implemented with method 1 and method 2 are named BestFS_m1 and BestFS_m2, respectively. We then compare the queue length vs. subdivided octant number for BestFS, BestFS_m1 and BestFS_m2. Fig. 4.12 and Fig. 4.14 are the queue length vs. subdivided octant number using BestFS_m1 which avoid inserting GB^1 octant into the priority. Fig. 4.13 and Fig. 4.15 are the queue length vs. subdivided octant number using BestFS_m2. The queue length grows at the beginning of the construction process. When it decays and reaches 0, the construction process is complete and the octree model is generated. The black dash lines in Fig. 4.13 to Fig. 4.15 are the plots of the queue length under different subdivided octants number for BestFS. The rest of 5 color lines are the plots of the queue length under different subdivided octants number for protrusion value $P=2, 5, 17, 15$ and 25 , respectively. One can find that the queue length of BestFS_m1 and

BestFS_m2 is greatly reduced when P is set to be large. While P is approaching 0, the queue length of BestFS_m1 and BestFS_m2 is getting closer to the queue length of BestFS.

Besides, one can also find that the queue length in BestFS_ms2 is also less than the queue length in BestFS_m1. This is because that some of the GG^1 octants in the first method will be classified as the GW^2 octants in the second method which has been shown in chapter 3. In the later section, we will provide an analysis on the queue length of BestFS_m1 and BestFS_m2, in more detail.

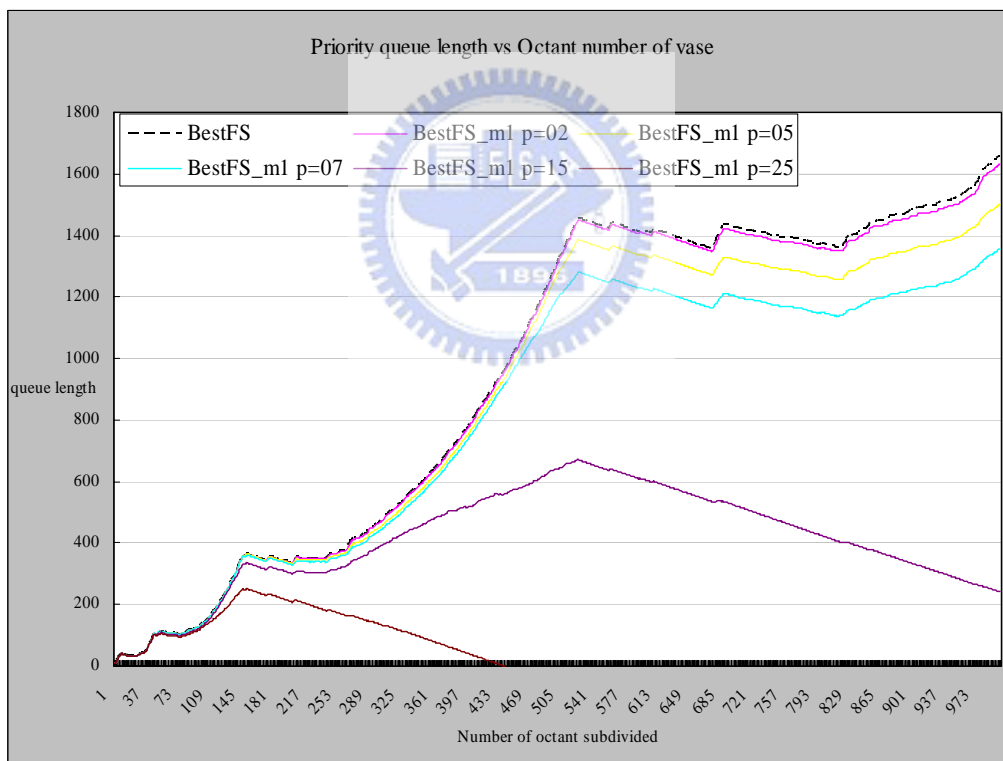


Fig. 4.12. The priority queue length when constructing the model of the vase using the BestFS_m1 method.

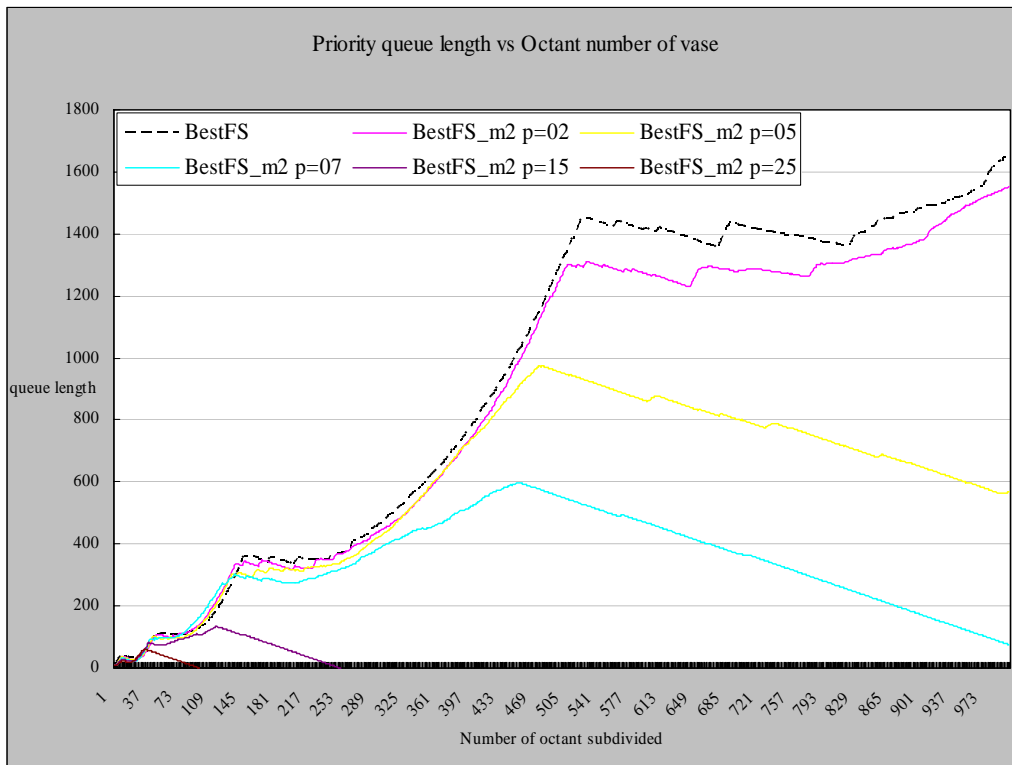


Fig. 4.13. The priority queue length when constructing the model of the vase using the BestFS_m2 method.

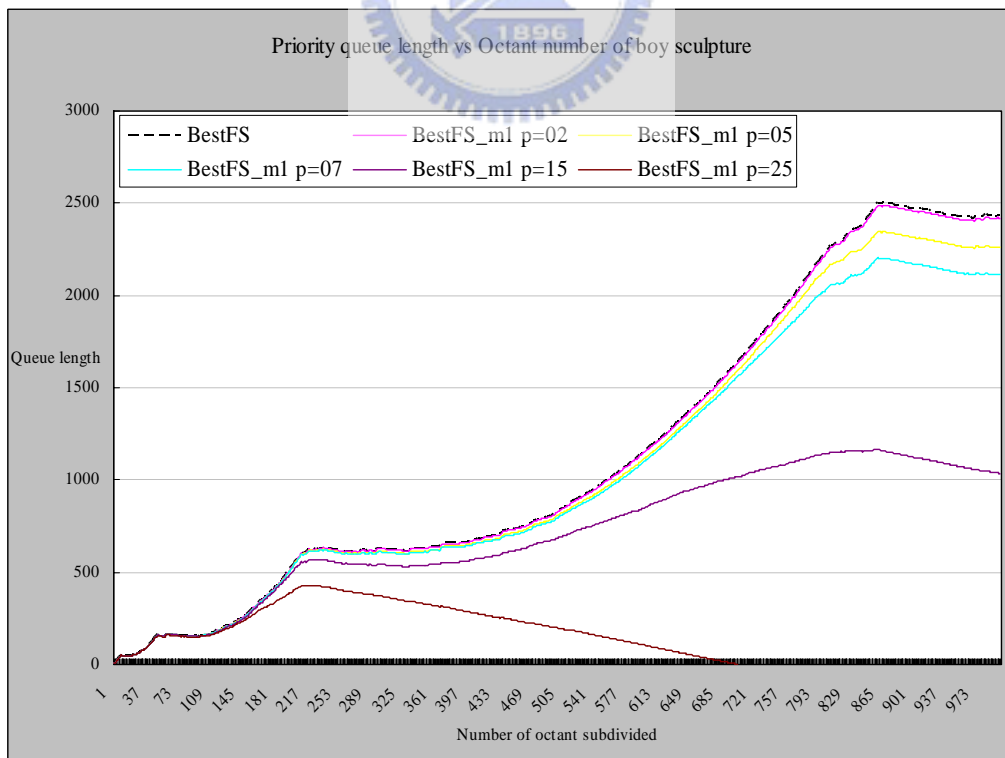


Fig. 4.14. The priority queue length when constructing the model of the boy sculpture using the BestFS_m1 method.

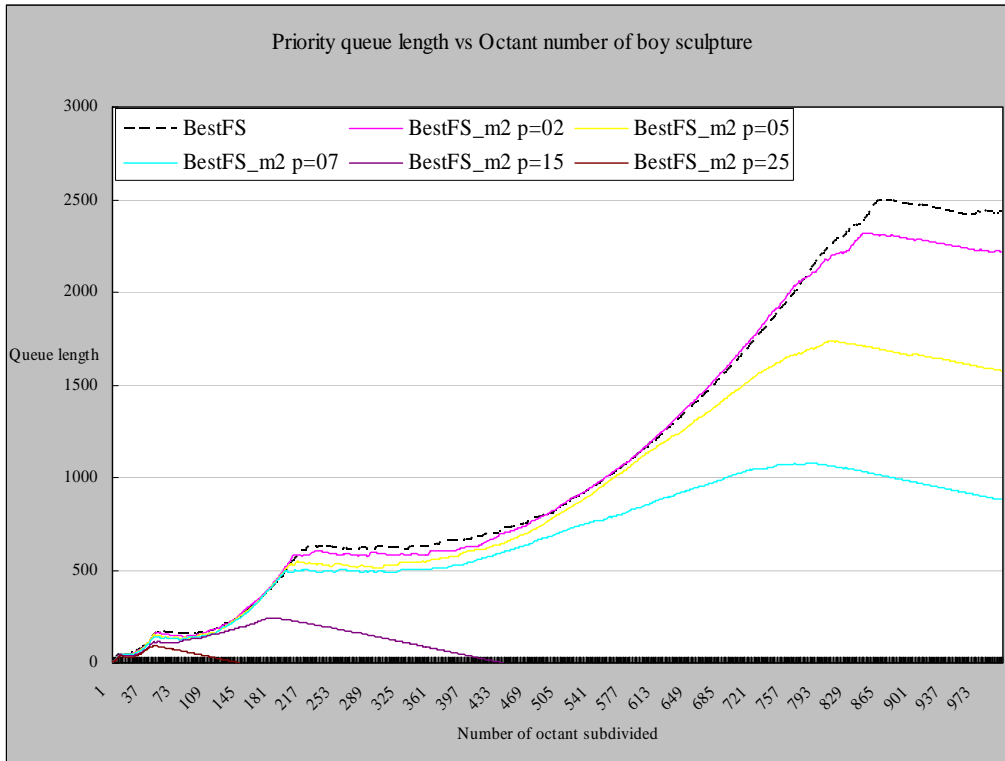


Fig. 4.15. The priority queue length when constructing the model of the boy sculpture using the BestFS_m2 method.



4.5 Analytical Analysis on the Performance of the Proposed Method

In this section we analyze the performance of the proposed progressive method theoretically. Before introducing the lemmas, we define the definitions and notations first:

Definition:

Assume N silhouette images, $I_0 \dots I_{N-1}$, are used to construct the octree model for the object.

Let

- (i) O_i be the i -th octant projected onto the images and applied test intersection against the silhouette image in the octree construction method;
- (ii) c_i be the center (or centroid) of the 8 projected octant vertices of O_i ;
- (iii) r_i be the radius of the bounding circle of the projected octant image for O_i ;
- (iv) d_i be the distance value at the circle center c_i in the distance map $DistMap$ defined in Section 2;

Also let

(i) $\max_proj_error(O)$ be the maximum of the projection errors, $proj_error_i$, $i = 0, 1, 2, \dots, N-1$, of octant O on the silhouette images I_0 to I_{N-1} , where

$$proj_error_i = \begin{cases} r_i + |d_i| & \text{if } d_i \leq 0 \\ r_i - |d_i| & \text{if } d_i > 0 \end{cases} \text{ is the projection error of octant } O \text{ on image } I_i;$$

$priority_queue(i)$ be the i -th octant in the priority queue which has the below property:

if $i < j$ then $O_i = priority_queue(i)$, $O_j = priority_queue(j)$ and

$$\max_proj_error(O_i) < \max_proj_error(O_j);$$

(ii) V_m be the constructed octree model after examining m octants in the octree structure against the silhouette images I_0 to I_{N-1} and $V_m^{priority_queue}$, V_m^{queue} , V_m^{stack} be the constructed octree model V_m from the octree construction methods using the priority queue, queue and stack data structure to store the octants to be processed, respectively;

(iii) S_m be the set of m octants examined in the octree structure against to the silhouette images and S_m^{stack} , S_m^{queue} , $S_m^{priority_queue}$ be the set S_m with octants retrieved from the stack, queue and priority queue data structure respectively;

Lemma 4.1: If V_m^{queue} , $V_m^{priority_queue}$ are the volumes generated by the octree construction methods implemented with the queue and the priority queue, respectively, when m octants have been processed, then $XOR(V_m^{queue}) \geq XOR(V_m^{priority_queue})$.

Proof:

Let V_m^{queue} and $V_m^{priority_queue}$ be the volumes generated by performing octant projection and intersection test on the set of octants S_m^{queue} and $S_m^{priority_queue}$, respectively.

(1) If V_m^{queue} and $V_m^{priority_queue}$ are identical then $XOR(V_m^{queue}) = XOR(V_m^{priority_queue})$ holds.

(2) If V_m^{queue} and $V_m^{priority_queue}$ are not identical then there exists at least one octant different between S_m^{queue} and $S_m^{priority_queue}$.

Assume $O_i \in S_m^{queue}$ and $O_i \notin S_m^{priority_queue}$, $O_j \in S_m^{priority_queue}$ and $O_j \notin S_m^{queue}$, r_i and r_j are

the radii of the bounding circles of the octant projection images for O_i and O_j . The rest of the octants in S_m^{queue} and $S_m^{\text{priority_queue}}$ are the same. Then the subdivision of O_i and O_j will affect the difference between $\text{XOR}(V_m^{\text{queue}})$ and $\text{XOR}(V_m^{\text{priority_queue}})$. In other words, let $V_{m-1}^{\text{priority_queue}}$, V_{m-1}^{queue} be the octree models generated by examining the octants in $S_m^{\text{priority_queue}} - \{O_j\}$ and $S_m^{\text{queue}} - \{O_i\}$, respectively. Since $S_m^{\text{priority_queue}} - \{O_j\} = S_m^{\text{queue}} - \{O_i\}$, we have $\text{XOR}(V_{m-1}^{\text{priority_queue}}) = \text{XOR}(V_{m-1}^{\text{queue}})$. Now we only need to analyze the influence of the subdivision of octant O_i and O_j to $V_{m-1}^{\text{priority_queue}}$, V_{m-1}^{queue} . Fig. 4.15 is the depiction of the relation among the projection images O_i and O_j and the silhouette images.

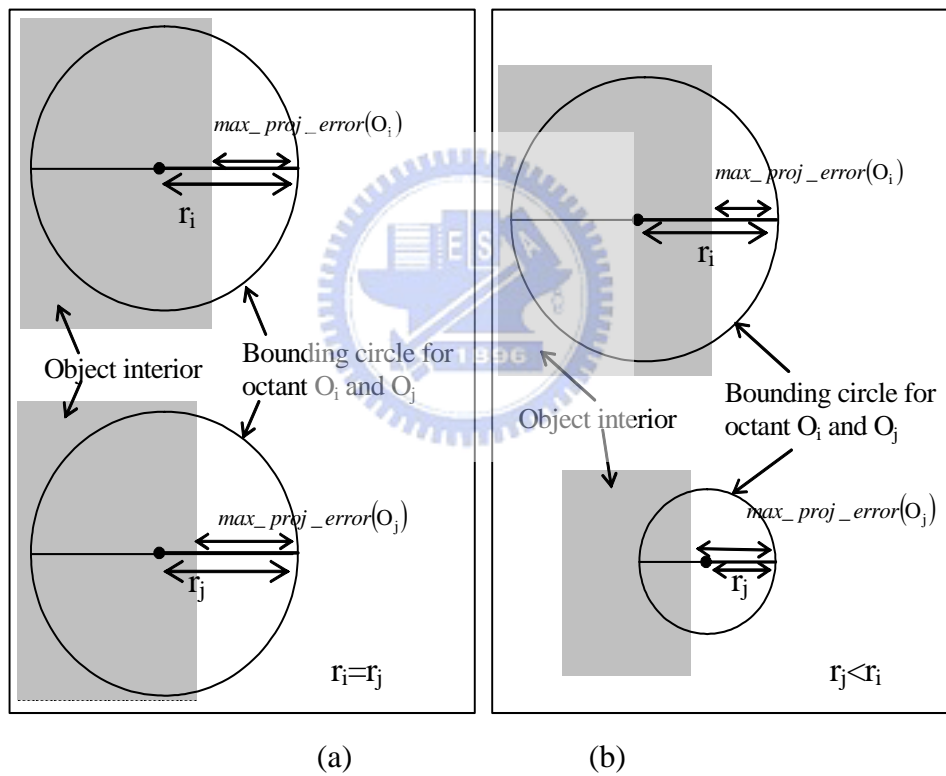


Fig. 4.16. The relation of radii of the bounding circles for octant O_i and O_j and their $\text{max_proj_error}(O_i)$ and $\text{max_proj_error}(O_j)$.

Since Octants in S_m^{queue} are the breadth-first traversal of the octree, if O_j is not the same octant as the O_i , then the level of O_i should be smaller than or equal to the level of O_j in $S_m^{\text{priority_queue}}$.

(a) If level of O_i and level of O_j are equal, then $r_i = r_j$.

Since O_i is not in $S_m^{\text{priority_queue}}$, it implies that $\max_proj_error(O_i) < \max_proj_error(O_j)$.

The relations among $\max_proj_error(O_i)$, $\max_proj_error(O_j)$ and r_i could be

$$(i) \max_proj_error(O_i) < \max_proj_error(O_j) < r_i (=r_j)$$

The subdivisions of O_i and O_j will produce black or grey sub-octants which will not affect $XOR(V_{m-1}^{\text{queue}})$ and $XOR(V_{m-1}^{\text{priority_queue}})$.

$$(ii) \max_proj_error(O_i) < r_i (=r_j) < \max_proj_error(O_j)$$

The subdivisions of O_j will produce white octants and cut off half of the error. Thus the $XOR(V_{m-1}^{\text{priority_queue}})$ will be less than the $XOR(V_{m-1}^{\text{queue}})$.

$$(iii) r_i (=r_j) < \max_proj_error(O_i) < \max_proj_error(O_j)$$

Both the subdivision of O_i and O_j will produce white octants and cut off half of the error induced by the octant O_i and O_j . Thus $XOR(V_{m-1}^{\text{queue}})$ and $XOR(V_{m-1}^{\text{priority_queue}})$ reduce equally.

(b) If level of O_i is less than the level of O_j , then $r_i \geq 2*r_j$

The relations among $\max_proj_error(O_i)$, $\max_proj_error(O_j)$, r_i and r_j could be

$$(i) \max_proj_error(O_i) < \max_proj_error(O_j) < r_j < r_i$$

The subdivision of O_i and O_j will produce black or grey sub-octants which will not affect $XOR(V_{m-1}^{\text{queue}})$ and $XOR(V_{m-1}^{\text{priority_queue}})$.

$$(ii) \max_proj_error(O_i) < r_j < \max_proj_error(O_j) < r_i$$

The subdivisions of O_j will produce white octants and cut off half of the error. Thus the $XOR(V_{m-1}^{\text{priority_queue}})$ will be less than the $XOR(V_{m-1}^{\text{queue}})$.

$$(iii) r_j < \max_proj_error(O_i) < \max_proj_error(O_j) < r_i$$

The subdivisions of O_j will produce white octants and cut off half of the error. Thus the $XOR(V_{m-1}^{\text{priority_queue}})$ will be less than the $XOR(V_{m-1}^{\text{queue}})$.

To summarize, $XOR(V_m^{\text{queue}}) \geq XOR(V_m^{\text{priority_queue}})$.

Lemma 4.1: If V_m^{stack} , $V_m^{\text{priority_queue}}$ are the volumes generated by the octree construction methods implemented with the queue and the priority queue, respectively, when m octants have been processed, , then $XOR(V_m^{\text{stack}}) \geq XOR(V_m^{\text{priority_queue}})$.

Proof:

Lemma 4.2 can be proved in a similar way as in the proof of Lemma 4.1.

Lemma 4-3: Let $\text{priority_queue}(m)$, $\text{priority_queue}_p^1(m)$, and $\text{priority_queue}_p^2(m)$ be the queues after processing m grey octants based on the projection error definitions given in the conventional and the two new construction methods introduced previously. Also let $\text{priority_queue_length}(m)$, $\text{priority_queue_length}_p^1(m)$ and $\text{priority_queue_length}_p^2(m)$ be the queue lengths of $\text{priority_queue}(m)$, $\text{priority_queue}_p^1(m)$ and $\text{priority_queue}_p^2(m)$, respectively.

Then for any given projection error bound p

$$\begin{aligned} \text{priority_queue_length}_p^2(m) &\leq \text{priority_queue_length}_p^1(m) \\ &\leq \text{priority_queue_length}(m). \end{aligned}$$

Proof:

(1) We first prove that $\text{priority_queue_length}_p^1(m) \leq \text{priority_queue_length}(m)$ by induction.

(a) While $m=1$, the first octant, the root octant, is projected onto the images and examined by the intersection test. If the root octant is partially intersected with the object silhouette images, then it is subdivided into 8 smaller sub-octants.

Let n be the number of sub-octants of the root octant which are grey and d be the number of sub-octants which are grey and their $\text{max_proj_error}(O)$ is smaller than P . Then

$$\text{priority_queue_length}(1) = n \text{ and } \text{priority_queue_length}_p^1(1) = n-d.$$

It follows that

$$\text{priority_queue_length}_p^1(1) = n-d \leq n = \text{priority_queue_length}(1).$$

That is,

$$\text{priority_queue_length}_p^1(1) \leq \text{priority_queue_length}(1).$$

(b) Suppose $\text{priority_queue_length}_p^1(k) \leq \text{priority_queue_length}(k)$ sustains for $m = k$. We now prove that $\text{priority_queue_length}_p^1(k+1) \leq \text{priority_queue_length}(k+1)$ also sustains for $m=k+1$.

Let O_i and O_j be the octants in the head of $\text{priority_queue}(k)$, $\text{priority_queue}_p^1(k)$, respectively.

(i) Assuming $O_i = O_j$:

Let n be the number of sub-octants of O_i which are grey and d be the number of sub-octants which are grey and their $\text{max_proj_error}(O)$ is smaller than P then the length of

$\text{priority_queue}(m)$, $\text{priority_queue}_p^1(m)$ after subdividing O_i are:

$\text{priority_queue_length}(k+1) = \text{priority_queue_length}(k) - 1 + n$ and

$\text{priority_queue_length}_p^1(k+1) = \text{priority_queue_length}_p^1(k) - 1 + n - d$.

Since $\text{priority_queue_length}_p^1(k) \leq \text{priority_queue_length}(k)$, it follows that $\text{priority_queue_length}_p^1(k+1) \leq \text{priority_queue_length}(k+1)$.

(ii) Assuming $O_i \neq O_j$:

Since O_i is not in $\text{priority_queue}_p^1(k)$, it implies $\text{max_proj_error}(O_i)$ is smaller than P .

Also, O_i is the head of $\text{priority_queue}(k)$, that is, $\text{max_proj_error}(O_i)$ is the maximum value in the queue. This implies that $\text{priority_queue}_p^1(k)$ is empty.

Thus, $\text{priority_queue_length}_p^1(k+1) = 0 \leq \text{priority_queue_length}(k+1)$.

From (a) and (b), we can conclude that

$\text{priority_queue_length}_p^1(m) \leq \text{priority_queue_length}(m)$ for any projection error bound p .

(2) We can also prove that $\text{priority_queue_length}_p^2(m) \leq \text{priority_queue_length}_p^1(m)$ by induction. It is similar to the proof for $\text{priority_queue_length}_p^1(m) \leq \text{priority_queue_length}(m)$.

4.6 Conclusion

In this chapter the octree construction method using the best first traversal scheme with a priority queue is presented. It is an extension to the first method and the second method which construct the octree model of an object progressively. The octants to be processed are arranged in the order of maximum projection error. The octant having maximum projection error is always the one first to be processed.

The reconstruction quality in terms of XOR error of the octree projection image with the silhouette image has been shown better than the octree construction method using the breadth first traversal of depth first traversal. Besides, we also show that the memory requirement for the priority queue becomes less if we can avoid inserting octants with tolerable projection errors into the priority queue by incorporating the first method and the second octant color classification criteria into the proposed progressive method.



CHAPTER 5

A NOVEL 3D PLANAR OBJECT RECONSTRUCTION FROM MULTIPLE UNCALIBRATED IMAGES USING THE PLANE-INDUCED HOMOGRAPHIES

5.1 Introduction

In the physical world (especially the man-made world) planar surfaces such as walls, windows, table, roof, road, and terrace can be found in the indoor as well as the outdoor scenes. Our task is to reconstruct the 3D planar surfaces in a scene from multiple uncalibrated images taken by a camera placed at different viewpoints. In general, the methods for 3D projective or uncalibrated reconstruction [33][36][37][40][41][50] are point-based. They estimate the fundamental matrix from a sufficient number of corresponding point pairs first, and then derive the epipole and the canonical geometric representation for projective views using the fundamental matrix. Then, for each pair of corresponding points, they use a triangulation technique or bundle adjustment technique to compute the 3D point coordinates in the projective space. Finally, for the determination of the uncalibrated planar scene structure [34][39][46][47][49][54][56][60], the 3D points found are fitted by planes. However, it is desirable to derive the 3D planar scene structure in terms of plane features in the images directly, for these features are more reliable than the point or line features [49]. The estimation of the 3D projective planar structure based on the projected plane feature information exclusively has not yet received much attention, although it is known that the corresponding projected plane regions in a pair of stereo images induce a homography. It is also known that homographies are useful to many other practical applications including:

- (a) Fundamental matrix estimation or canonical projective geometry representation [48][49].
- (a) 2D image mosaicing or view synthesis [43].
- (a) Plane + parallax analysis [34][36][56].
- (a) Planar motion estimation and ego-motion [45][54][60].

Recently, two methods have been proposed for the 3D projective reconstruction of planes and cameras. The first method assumes all planes are visible in all images and the second method assumes a reference plane is visible in all images [51][52]. In practice, it is not realistic to have all planes or even one plane visible in all images unless a very large ground plane is available. When there is no reference plane visible in all images, the reconstruction problem cannot be formulated within a common projective space and the reconstruction results will be inevitably obtained in different projective spaces.

We shall recover the 3D scene planar structure from the uncalibrated images using the plane-induced homographies without assuming that all planes or one plane must be seen in all images. To obtain the homographies, we must locate the projected regions of planar surfaces in the images. There are methods for detecting regions corresponding to planar surfaces in the image [55][44][59]. After the image regions of planar surfaces have been extracted, we use the Gabor filtering technique [57] to identify at least four point correspondences for every plane in the stereo images in order to obtain the initial value of the homography. Then we iteratively refine the homography based on a nonlinear minimization method given in [53]. Next, we use two homographies to compute the epipole and to find the compatible projection equations in terms of the estimated homography and an assigned plane coefficient vector of a reference plane, together with the estimated epipole. With the projection equations thus derived we then prove that the 3D equation of any other plane visible in the stereo images can be computed with respect to the reference plane equation as long as its homography is determined. Finally, we merge or integrate all reconstructed plane equations found in individual projective spaces within a common space through the coordinate (or space) transformations. Again, each required coordinate transformation matrix is expressed by the homography and plane coefficient vector information of two planes visible in the involved image pairs. Fig. 5.1 shows the flow diagram of our method.

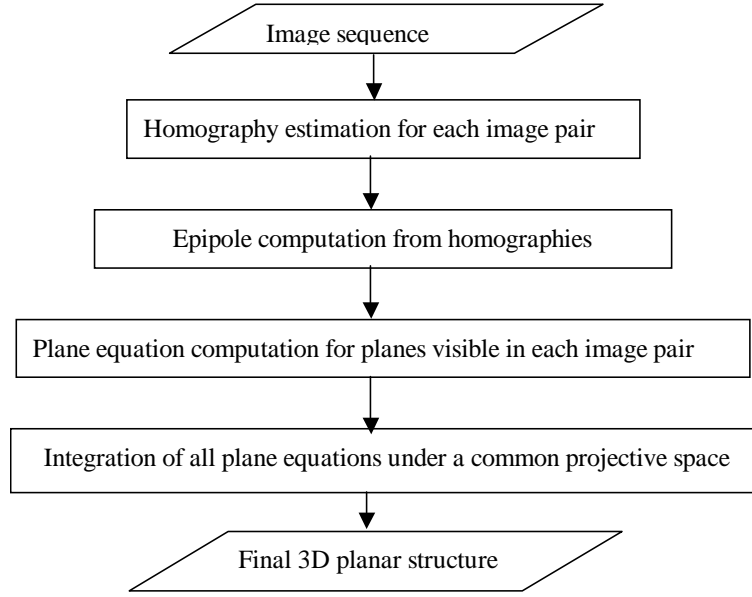


Fig. 5.1. The flow diagram of our reconstruction method.

The remaining sections of the chapter are organized as follows. Section 2 is the preliminaries and mathematical notations for the projective reconstruction. Section 3 shows how the 3D equations of all planar surfaces visible in the stereo images can be determined from their homographies. Section 4 presents the integration of the reconstruction results obtained in different projective spaces through the coordinate transformations. Section 5 shows the estimation of the plane-induced homographies and the related epipole. Section 6 reports the experimental results on both the synthetic and real images. Section 7 is the concluding remarks.

5.2 Preliminaries and Mathematical Notations for Projective Reconstruction

Consider any two consecutive images (I_i, I_j) in an image sequence for reconstructing the visible planar surfaces. Let $\mathbf{R}_i, \vec{\mathbf{t}}_i$ be the extrinsic parameters and \mathbf{M}_i be the 3x3 upper triangular intrinsic camera matrix of the i -th camera. Then the coordinates of a 3D point $\vec{\mathbf{p}}_E = [x_E \ y_E \ z_E]^T$ and its 2D projection point $\vec{\mathbf{u}}_i = [u_i \ v_i]^T$ in image I_i are related by a pinhole camera model [40][37][43]:

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} \cong \mathbf{M}_i \left[\mathbf{R}_i | \vec{t}_i \right] \begin{bmatrix} x_E \\ y_E \\ z_E \\ 1 \end{bmatrix}.$$

To represent the point in the projective space or the homogeneous coordinate system, we use the vectors with a tilde to denote the homogeneous coordinates of the 3D points and its image projection point such that $\tilde{\mathbf{u}}_i = [\vec{\mathbf{u}}_i^T \ 1]^T$ and $\tilde{\mathbf{p}}_E = [\vec{\mathbf{p}}_E^T \ 1]^T$. The symbol \cong indicates an equality up to a non-zero scale in the homogeneous coordinate system.

Assume the world coordinate system is chosen to be the i-th camera coordinate system; namely, $\mathbf{R}_i = \mathbf{I}$ and $\vec{t}_i = \vec{\mathbf{0}}$.

$$\tilde{\mathbf{u}}_i \cong \mathbf{M}_i \cdot [\mathbf{I} | \vec{\mathbf{0}}] \cdot \tilde{\mathbf{p}}_E \triangleq \mathbf{M}_i \cdot \tilde{\mathbf{p}}_E. \quad (5-1)$$

Similarly, let $\mathbf{M}_j, \mathbf{R}_j, \vec{t}_j$ be the camera parameters of the j-th camera. For image I_j , we have

$$\tilde{\mathbf{u}}_j \cong \mathbf{M}_j \cdot [\mathbf{R}_j | \vec{t}_j] \cdot \tilde{\mathbf{p}}_E. \quad (5-2)$$

Since the epipole on image I_j is given by $\tilde{\mathbf{e}}_j \cong \mathbf{M}_j [\mathbf{R}_j | \vec{t}_j] [0 \ 0 \ 0 \ 1]^T \cong \mathbf{M}_j \cdot \vec{t}_j$ or $\lambda_e \tilde{\mathbf{e}}_j = \mathbf{M}_j \vec{t}_j$ (λ_e is the lens depth parameter), we rewrite Eq. (2) as

$$\tilde{\mathbf{u}}_j \cong [\mathbf{M}_j \mathbf{R}_j | \lambda_e \tilde{\mathbf{e}}_j] \cdot \tilde{\mathbf{p}}_E. \quad (5-3)$$

Consider a plane Π_A , which does not pass through the optical center of the i-th camera (otherwise, its image will be degenerated into a line). Let its plane equation be $\vec{\mathbf{a}}_E^T \vec{\mathbf{p}}_E + 1 = 0$ with $\vec{\mathbf{a}}_E^T = [a_{E1} \ a_{E2} \ a_{E3}]^T$. After eliminating the variable $\vec{\mathbf{p}}_E$ in the two projection equations (5-1) and (5-3), we can obtain a homography \mathbf{A}_{ij} as follows [58][48][53]:

$$\tilde{\mathbf{u}}_j \cong \mathbf{A}_{ij} \tilde{\mathbf{u}}_i \quad \text{with} \quad \mathbf{A}_{ij} \cong \left\{ \mathbf{M}_j \mathbf{R}_j \mathbf{M}_i^{-1} - \lambda_e \tilde{\mathbf{e}}_j \vec{\mathbf{a}}_E^T \mathbf{M}_i^{-1} \right\}.$$

The homography \mathbf{A}_{ij} from image I_i to image I_j is said to be induced by plane Π_A . In Section 5 we shall show how to compute the homography \mathbf{A}_{ij} from image pair (I_i, I_j) .

For an uncalibrated camera the intrinsic and extrinsic camera parameters in Eqs. (5-1) and (5-3) cannot be estimated. We need to replace these two equations by some new parameters that can be estimated. This is done as follows:

Let \mathbf{A}_{ij} be rewritten as $\mathbf{A}_{ij} = \lambda_A \{ \mathbf{M}_j \mathbf{R}_j \mathbf{M}_i^{-1} - \lambda_e \tilde{\mathbf{e}}_j \tilde{\mathbf{a}}_E^T \mathbf{M}_i^{-1} \}$, then

$$\mathbf{M}_j \mathbf{R}_j = \frac{1}{\lambda_A} \mathbf{A}_{ij} \mathbf{M}_i + \lambda_e \tilde{\mathbf{e}}_j \tilde{\mathbf{a}}_E^T.$$

Eq. (5-3) can be rewritten as

$$\begin{aligned} \tilde{\mathbf{u}}_j &\cong [\mathbf{M}_j \mathbf{R}_j \mid \lambda_e \tilde{\mathbf{e}}_j] \begin{bmatrix} \tilde{\mathbf{p}}_E \\ 1 \end{bmatrix} \cong \frac{1}{\lambda_A} [\mathbf{A}_{ij} \mid \tilde{\mathbf{e}}_j] \begin{bmatrix} \mathbf{M}_i & \mathbf{0} \\ \lambda_A \lambda_e \tilde{\mathbf{a}}_E^T & \lambda_A \lambda_e \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{p}}_E \\ 1 \end{bmatrix} \\ &\cong \frac{1}{\lambda_A} [\mathbf{A}_{ij} \mid \tilde{\mathbf{e}}_j] \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \tilde{\mathbf{a}}_{ij}^T & a_{ij4} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \tilde{\mathbf{a}}_{ij}^T & a_{ij4} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{M}_i & \mathbf{0} \\ \lambda_A \lambda_e \tilde{\mathbf{a}}_E^T & \lambda_A \lambda_e \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{p}}_E \\ 1 \end{bmatrix} \end{aligned}$$

where $\tilde{\mathbf{a}}_{ij}$ (with $a_{ij4} \neq 0$) is assumed to be a non-null column vector. Then

$$\begin{aligned} \tilde{\mathbf{u}}_j &\cong \frac{1}{\lambda_A} [\mathbf{A}_{ij} \mid \tilde{\mathbf{e}}_j] \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \tilde{\mathbf{a}}_{ij}^T & a_{ij4} \end{bmatrix} \begin{bmatrix} \mathbf{M}_i & \tilde{\mathbf{0}} \\ -\tilde{\mathbf{a}}_{ij}^T \mathbf{M}_i + \frac{\lambda_A \lambda_e}{a_{ij4}} \tilde{\mathbf{a}}_E^T & \frac{\lambda_A \lambda_e}{a_{ij4}} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{p}}_E \\ 1 \end{bmatrix} \\ &\cong \frac{1}{\lambda_A} [\mathbf{A}_{ij} \mid \tilde{\mathbf{e}}_j] \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \tilde{\mathbf{a}}_{ij}^T & a_{ij4} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{p}}_{ij} \\ p_{ij4} \end{bmatrix}. \end{aligned} \quad (5-4)$$

A similar formulation of Eq. (5-4) has been derived in [48]. In this way, the original Euclidean point $\tilde{\mathbf{p}}_E$ becomes point $\tilde{\mathbf{p}}_{ij} = [\tilde{\mathbf{p}}_{ij}^T \quad p_{ij4}]^T$ in the new projective space, denoted by $\{\tilde{\mathbf{p}}_{ij}\}$, which describes the projective geometry associated with images i and j .

The coordinate transformation from the Euclidean space $\{[\tilde{\mathbf{p}}_E^T \quad 1]^T\}$ to the projective space $\{[\tilde{\mathbf{p}}_{ij}^T \quad p_{ij4}]^T\}$ is given by

$$\tilde{\mathbf{p}}_{ij} = \mathbf{M}_i \tilde{\mathbf{p}}_E, \quad p_{ij4} = \left(\frac{\lambda_A \lambda_e}{a_{ij4}} \right) (\tilde{\mathbf{a}}_E^T \tilde{\mathbf{p}}_E + 1) - \frac{\tilde{\mathbf{a}}_{ij}^T}{a_{ij4}} \mathbf{M}_i \tilde{\mathbf{p}}_E.$$

Also,

$$\tilde{\mathbf{a}}_{ij}^T \tilde{\mathbf{p}}_{ij} + a_{ij4} p_{ij4} = \tilde{\mathbf{a}}_{ij}^T \tilde{\mathbf{p}}_{ij} + (\lambda_A \lambda_e) (\tilde{\mathbf{a}}_E^T \tilde{\mathbf{p}}_E + 1) - \tilde{\mathbf{a}}_{ij}^T \tilde{\mathbf{p}}_{ij} = \lambda_A \lambda_e (\tilde{\mathbf{a}}_E^T \tilde{\mathbf{p}}_E + 1) = 0$$

It implies that $\tilde{\mathbf{a}}_{ij}^T \tilde{\mathbf{p}}_{ij} + a_{ij4} p_{ij4} = 0$ is the new 3D equation of the plane in the

projective space $\{\tilde{\mathbf{p}}_{ij}\}$. Since the projective structure can only be determined up to a 4x4 non-singular projective matrix [43], the new plane coefficient vector $\tilde{\mathbf{a}}_{ij} = [\tilde{\mathbf{a}}_{ij}^T \ a_{ij4}]^T$ can take on some general value, say, $[1 \ 1 \ 1 \ 1]^T$ (more discussion on the values of $\tilde{\mathbf{a}}_{ij}$ is given in Section 6). In this new space the parameters including homography \mathbf{A}_{ij} , epipole $\tilde{\mathbf{e}}_j$ and plane coefficients $\tilde{\mathbf{a}}_{ij} = [\tilde{\mathbf{a}}_{ij}^T \ a_{ij4}]^T$ involved in Eq. (5-4) are now all known.

Next, we shall describe how to obtain the projective reconstruction for the other planes visible in the image pair (I_i, I_j) in the newly defined projective space $\{\tilde{\mathbf{p}}_{ij}\}$.

5.3 Reconstruction of All Visible Planes from a Given Image Pair

In the new projective space the projection equations become

$$\begin{aligned} \tilde{\mathbf{u}}_i &\cong [\mathbf{I} | \tilde{\mathbf{0}}] \begin{bmatrix} \tilde{\mathbf{p}}_{ij} \\ \mathbf{p}_{ij4} \end{bmatrix}, \\ \tilde{\mathbf{u}}_j &\cong \frac{1}{\lambda_A} [\mathbf{A}_{ij} \ | \ \tilde{\mathbf{e}}_j] \begin{bmatrix} \mathbf{I} & \tilde{\mathbf{0}} \\ \tilde{\mathbf{a}}_{ij}^T & a_{ij4} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{p}}_{ij} \\ \mathbf{p}_{ij4} \end{bmatrix}. \end{aligned} \quad (5-5)$$

Similarly, for any other plane Π_B visible in (I_i, I_j) the induced homography between the plane regions in image pair (I_i, I_j) is expressed by

$$\mathbf{B}_{ij} = \lambda_B \left\{ \mathbf{M}_j \mathbf{R}_j \mathbf{M}_i^{-1} - \lambda_e \tilde{\mathbf{e}}_j \tilde{\mathbf{b}}_E^T \mathbf{M}_i^{-1} \right\} \quad (5-6)$$

with the plane equation of Π_B being $\tilde{\mathbf{b}}_E^T \tilde{\mathbf{p}}_E + 1 = 0$.

Next, we shall prove the fact that the relation between plane coefficient vectors of planes Π_B and Π_A is determined once their homographies \mathbf{A}_{ij} and \mathbf{B}_{ij} are found.

From above we have

$$\mathbf{M}_j \mathbf{R}_j \mathbf{M}_i^{-1} = \frac{1}{\lambda_A} \mathbf{A}_{ij} + \lambda_e \tilde{\mathbf{e}}_j \tilde{\mathbf{a}}_E^T \mathbf{M}_i^{-1} = \frac{1}{\lambda_B} \mathbf{B}_{ij} + \lambda_e \tilde{\mathbf{e}}_j \tilde{\mathbf{b}}_E^T \mathbf{M}_i^{-1}$$

or

$$\mathbf{A}_{ij} = \frac{\lambda_A}{\lambda_B} \mathbf{B}_{ij} + \lambda_A \lambda_e \tilde{\mathbf{e}}_j [\tilde{\mathbf{b}}_E^T - \tilde{\mathbf{a}}_E^T] \mathbf{M}_i^{-1} \triangleq \frac{\lambda_A}{\lambda_B} \mathbf{B}_{ij} + \tilde{\mathbf{e}}_j \tilde{\mathbf{n}}_{ij}^T \quad (5-7)$$

where $\vec{\eta}_{ij} = \lambda_A \lambda_e [\vec{b}_E - \vec{a}_E] \mathbf{M}_i^{-1}$. We can apply the least-squares method to estimate the unknowns $\vec{\eta}_{ij}$ and $\frac{\lambda_A}{\lambda_B}$ in a system of nine linear equations; here the epipole \vec{e}_j can be determined in advance from the two homographies \mathbf{A}_{ij} and \mathbf{B}_{ij} based on the fact that $\mathbf{B}^T [\vec{e}_j]_{\times} \mathbf{A} (\cong \mathbf{B}^T [\vec{e}_j]_{\times} \mathbf{B})$ is skew symmetric.

Therefore,

$$\tilde{\mathbf{u}}_j \cong \frac{1}{\lambda_A} [\mathbf{A}_{ij} \mid \vec{e}_j] \left[\begin{array}{c|c} \mathbf{I} & \vec{0} \\ \hline \vec{a}_{ij}^T & a_{ij4} \end{array} \right] \left[\begin{array}{c} \vec{p}_{ij} \\ p_{ij4} \end{array} \right] \cong \frac{1}{\lambda_B} [\mathbf{B}_{ij} \mid \vec{e}_j] \left[\begin{array}{c|c} \mathbf{I} & \vec{0} \\ \hline \vec{b}_{ij}^T & b_{ij4} \end{array} \right] \left[\begin{array}{c} \vec{p}_{ij} \\ p_{ij4} \end{array} \right].$$

Substituting $\mathbf{A}_{ij} = \frac{\lambda_A}{\lambda_B} \mathbf{B}_{ij} + \vec{e}_j \vec{\eta}_{ij}^T$ into the above equation, we have

$$\frac{1}{\lambda_A} \left[\frac{\lambda_A}{\lambda_B} \mathbf{B}_{ij} + \vec{e}_j \vec{\eta}_{ij}^T \mid \vec{e}_j \right] \left[\begin{array}{c|c} \mathbf{I} & \vec{0} \\ \hline \vec{a}_{ij}^T & a_{ij4} \end{array} \right] \cong \frac{1}{\lambda_B} [\mathbf{B}_{ij} \mid \vec{e}_j] \left[\begin{array}{c|c} \mathbf{I} & \vec{0} \\ \hline \vec{b}_{ij}^T & b_{ij4} \end{array} \right].$$

Then

$$[\mathbf{B}_{ij} \mid \vec{e}_j] \left[\begin{array}{c|c} \mathbf{I} & \vec{0} \\ \hline \frac{\lambda_B}{\lambda_A} \vec{\eta}_{ij}^T & \frac{\lambda_B}{\lambda_A} \end{array} \right] \left[\begin{array}{c|c} \mathbf{I} & \vec{0} \\ \hline \vec{a}_{ij}^T & a_{ij4} \end{array} \right] \cong [\mathbf{B}_{ij} \mid \vec{e}_j] \left[\begin{array}{c|c} \mathbf{I} & \vec{0} \\ \hline \vec{b}_{ij}^T & b_{ij4} \end{array} \right].$$

Since $a_{ij4} \neq 0$ and $b_{ij4} \neq 0$ (i.e., planes Π_A and Π_B do not contain the lens center), we obtain,

$$\left[\begin{array}{c|c} \mathbf{I} & \vec{0} \\ \hline \frac{\lambda_B}{\lambda_A} \vec{\eta}_{ij}^T & \frac{\lambda_B}{\lambda_A} \end{array} \right] \left[\begin{array}{c|c} \mathbf{I} & \vec{0} \\ \hline \vec{a}_{ij}^T & a_{ij4} \end{array} \right] \cong \left[\begin{array}{c|c} \mathbf{I} & \vec{0} \\ \hline \vec{b}_{ij}^T & b_{ij4} \end{array} \right].$$

Thus

$$\left[\frac{\lambda_B}{\lambda_A} \vec{\eta}_{ij}^T + \frac{\lambda_B}{\lambda_A} \vec{a}_{ij}^T \mid \frac{\lambda_B}{\lambda_A} a_{ij4} \right] \cong \left[\vec{b}_{ij}^T \mid b_{ij4} \right].$$

In other words, the relationship between two plane coefficient vectors is given by:

$$\tilde{\mathbf{b}}_{ij} \cong \left(\tilde{\mathbf{a}}_{ij} + \begin{bmatrix} \vec{\eta}_{ij} \\ 0 \end{bmatrix} \right). \quad (5-8)$$

Thus, $\tilde{\mathbf{b}}_{ij}$ can be determined with respect to $\tilde{\mathbf{a}}_{ij}$ once the planar homography \mathbf{B}_{ij} is known.

5.4 Integration of Planes Reconstructed from Different Image Pairs

Next, we consider the integration of reconstructed planes obtained from different image pairs (I_i, I_j) and (I_j, I_k) , which contain the projections of two commonly visible planes. We shall use the plane-based coordinate transformation method for integrating the reconstruction results defined in different spaces.

Let the 4x4 coordinate transformation matrix \mathbf{H}_{ijk} , mapping the points in the projective space $\{\tilde{\mathbf{p}}_{ij}\}$ to the points in the projective space $\{\tilde{\mathbf{p}}_{jk}\}$, be defined by

$$\tilde{\mathbf{p}}_{jk} = \mathbf{H}_{ijk} \tilde{\mathbf{p}}_{ij}.$$

Then, the plane coefficient vectors $\tilde{\mathbf{c}}_{ij}$, $\tilde{\mathbf{c}}_{jk}$ of a common plane, which are respectively defined in the two different projective spaces $\{\tilde{\mathbf{p}}_{ij}\}$ and $\{\tilde{\mathbf{p}}_{jk}\}$, will be related by:

$$\tilde{\mathbf{c}}_{jk} \cong \mathbf{H}_{ijk}^{-T} \tilde{\mathbf{c}}_{ij}. \quad (5-9)$$

Thus, it requires the information of 5 common planes in the two different projective spaces in order to solve for the transformation matrix \mathbf{H}_{ijk} . It is usually not very practical to find five common planes in the image pairs.

On the other hand, the two respective 3x4 projection matrices associated with image I_j defined in the two projective spaces $\{\tilde{\mathbf{p}}_{ij}\}$ and $\{\tilde{\mathbf{p}}_{jk}\}$ are related directly by the matrix \mathbf{H}_{ijk} [38]. This relationship provides 11 linear equations in the 15 matrix elements in \mathbf{H}_{ijk} . Then, it is reduced to a need of two plane information to provide 6 additional linear equations to solve for the 15 unknowns. In the following we shall give a system of 24 linear equations using the information of two planes for solving for the 15 unknowns; the result will be more reliable.

From Eq. (4), we have

$$l_j \tilde{\mathbf{u}}_j = \frac{1}{\lambda_B} [\mathbf{B}_{ij} \mid \tilde{\mathbf{e}}_j] \left[\begin{array}{c|c} \mathbf{I} & \tilde{\mathbf{0}} \\ \hline \tilde{\mathbf{b}}_{ij}^T & b_{ij4} \end{array} \right] \tilde{\mathbf{p}}_{ij}.$$

Combining this equation with a plane equation $\tilde{\mathbf{b}}_{ij}^T \tilde{\mathbf{p}}_{ij} = 0$, we have

$$\begin{bmatrix} l_j \tilde{\mathbf{u}}_j \\ 0 \end{bmatrix} = \frac{1}{\lambda_B} \begin{bmatrix} \mathbf{B}_{ij} & | & \tilde{\mathbf{e}}_j \\ \bar{\mathbf{0}}^T & | & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I} & | & \bar{\mathbf{0}} \\ \bar{\mathbf{b}}_{ij}^T & | & b_{ij4} \end{bmatrix} \tilde{\mathbf{p}}_{ij}.$$

Thus,

$$\tilde{\mathbf{p}}_{ij} = \lambda_B \left(\begin{bmatrix} \mathbf{B}_{ij} & | & \tilde{\mathbf{e}}_j \\ \bar{\mathbf{0}}^T & | & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I} & | & \bar{\mathbf{0}} \\ \bar{\mathbf{b}}_{ij}^T & | & b_{ij4} \end{bmatrix} \right)^{-1} \begin{bmatrix} l_j \tilde{\mathbf{u}}_j \\ 0 \end{bmatrix}.$$

Similarly, for the same point on plane Π_B , but represented as $\tilde{\mathbf{p}}_{jk}$ in the different projective space $\{ \tilde{\mathbf{p}}_{jk} \}$, we can relate it to the same 4x1 vector $\begin{bmatrix} l_j \tilde{\mathbf{u}}_j^T & 0 \end{bmatrix}^T$ by

$$\tilde{\mathbf{p}}_{jk} = \left(\begin{bmatrix} \mathbf{I} & | & \bar{\mathbf{0}} \\ \bar{\mathbf{0}}^T & | & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I} & | & \bar{\mathbf{0}} \\ \bar{\mathbf{b}}_{jk}^T & | & b_{jk4} \end{bmatrix} \right)^{-1} \begin{bmatrix} l_j \tilde{\mathbf{u}}_j \\ 0 \end{bmatrix}.$$

Then

$$\begin{aligned} & \left(\begin{bmatrix} \mathbf{I} & | & \bar{\mathbf{0}} \\ \bar{\mathbf{0}}^T & | & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I} & | & \bar{\mathbf{0}} \\ \bar{\mathbf{b}}_{jk}^T & | & b_{jk4} \end{bmatrix} \right)^{-1} \begin{bmatrix} l_j \tilde{\mathbf{u}}_j \\ 0 \end{bmatrix} = \mathbf{H}_{ijk} \\ & \lambda_B \left(\begin{bmatrix} \mathbf{B}_{ij} & | & \tilde{\mathbf{e}}_j \\ \bar{\mathbf{0}}^T & | & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I} & | & \bar{\mathbf{0}} \\ \bar{\mathbf{b}}_{ij}^T & | & b_{ij4} \end{bmatrix} \right)^{-1} \begin{bmatrix} l_j \tilde{\mathbf{u}}_j \\ 0 \end{bmatrix} \end{aligned}$$

After some algebraic manipulation, this can be reduced to

$$\begin{bmatrix} \mathbf{I} \\ -\bar{\mathbf{b}}_{jk}^T / b_{jk4} \end{bmatrix} \tilde{\mathbf{u}}_j = \lambda_B \mathbf{H}_{ijk} \begin{bmatrix} \mathbf{B}_{ij}^{-1} \\ -\bar{\mathbf{b}}_{ij}^T \mathbf{B}_{ij}^{-1} / b_{ij4} \end{bmatrix} \tilde{\mathbf{u}}_j.$$

Since this equality holds for all image points on plane Π_B , it further implies

$$\begin{bmatrix} \mathbf{I} \\ -\bar{\mathbf{b}}_{jk}^T / b_{jk4} \end{bmatrix} \mathbf{B}_{ij} = \lambda_B \mathbf{H}_{ijk} \begin{bmatrix} \mathbf{I} \\ -\bar{\mathbf{b}}_{ij}^T / b_{ij4} \end{bmatrix}.$$

This leads to a system of 12 linear equations in 16 unknowns: 15 from the matrix \mathbf{H}_{ijk} plus one from λ_B . Therefore, we need another system of equations provided by a second visible plane, say, Π_G :

$$\begin{bmatrix} \mathbf{I} \\ -\bar{\mathbf{g}}_{jk}^T / g_{jk4} \end{bmatrix} \mathbf{G}_{ij} = \lambda_G \mathbf{H}_{ijk} \begin{bmatrix} \mathbf{I} \\ -\bar{\mathbf{g}}_{ij}^T / g_{ij4} \end{bmatrix}.$$

Combining the above two systems of equations, we have a total of 24 linear equations in 17 unknowns. Here we give a least squared solution by placing the two systems of equations in the following form

$$\mathbf{H}_{ijk} \left[\begin{bmatrix} \mathbf{I} \\ -\bar{\mathbf{b}}_{ij}^T / b_{ij4} \end{bmatrix} \mid \begin{bmatrix} \mathbf{I} \\ -\bar{\mathbf{g}}_{ij}^T / g_{ij4} \end{bmatrix} \right] \cong \left[\begin{bmatrix} \mathbf{I} \\ -\bar{\mathbf{b}}_{jk}^T / b_{jk4} \end{bmatrix} \mathbf{B}_{ij} \mid \frac{\lambda_B}{\lambda_G} \begin{bmatrix} \mathbf{I} \\ -\bar{\mathbf{g}}_{jk}^T / g_{jk4} \end{bmatrix} \mathbf{G}_{ij} \right] \quad (5-10)$$

where the ratio of λ_b/λ_c has been estimated during the plane reconstruction phase (see Eq. (5-7)). We can find the matrix \mathbf{H}_{ijk} using the pseudo-inverse matrix of the 4x6 matrix on the left-hand side of the above equation.

5.5 Computation of Homographies

We need to estimate \mathbf{A}_{ij} from the image data associated with the planar surface Π_A . We shall use the region-based matching, instead of point-based matching, to find the homography. First of all, we use the Gabor filtering technique [57] to identify at least four point correspondences in order to obtain the initial solution of the homography. We then use the Levenberg-Marquardt iterative nonlinear minimization algorithm [53] to minimize the sum of the squared intensity differences of the transformed and original image points due to the plane Π_A in the image pair

$$E = \sum_k \left\{ I_j \left((u_k^i)', (v_k^i)' \right) - I_i \left(u_k^i, v_k^i \right) \right\}^2.$$

Here the transformed location $[(u_k^i)', (v_k^i)', 1]^T$ is obtained from the image point (u_k^i, v_k^i) using an estimated \mathbf{A}_{ij} , and $I_j \left((u_k^i)', (v_k^i)' \right)$ is the intensity obtained by a bilinear interpolation from the original image I_j . The intensity values of the image points in the common region of the two images are normalized to remove the possible illumination difference. The above minimization method converges in a few iterations.

After finding two homographies, recall that we can compute the epipole $\tilde{\mathbf{e}}_j$ from the skew symmetry property of $\mathbf{B}^T [\tilde{\mathbf{e}}_j]_{\times} \mathbf{A}$. Also, in turn, we can compute the fundamental matrix \mathbf{F} using the epipole $\tilde{\mathbf{e}}_j$ as follows:

$$\begin{aligned} [\tilde{\mathbf{e}}_j]_{\times} \mathbf{A} &\cong [\tilde{\mathbf{e}}_j]_{\times} \left(\left\{ \mathbf{M}_j \mathbf{R}_j \mathbf{M}_i^{-1} - \lambda_e \tilde{\mathbf{e}}_j \tilde{\mathbf{a}}_E^T \mathbf{M}_i^{-1} \right\} \right) \\ &\cong [\tilde{\mathbf{e}}_j]_{\times} \mathbf{M}_j \mathbf{R}_j \mathbf{M}_i^{-1} = \mathbf{F}. \end{aligned}$$

5.6 Experimental Results

Experiment 1:

In the first experiment we use a synthetic tower whose feature points and schematic diagram are given in Table 5.1 and Fig. 5.2. We take a sequence of six pictures to cover all aspects of the tower using a virtual camera looking down from the upper positions. The image resolution is 640x480 in pixel. Three consecutive images of the sequence, I_1 , I_2 , and I_3 , are shown in Fig. 5.3. We apply the reconstruction process to this data set. We employ a linear least-squares method based on eight corresponding image point pairs available in the synthetic data to get the true homography for each of the five planes, Π_{Gr} , Π_A , Π_B , Π_E , Π_F visible in the pair (I_1, I_2) . In addition, to handle the possible problems caused by data translation and scaling change, we also use the normalization transform proposed by Hartley [42] to compute the homographies. We choose Π_{Gr} as the reference plane. During the reconstruction process, we find the plane coefficient vectors with respect to the reference plane Π_{Gr} vector designated as $[1 \ 1 \ 1 \ 1]^T$. To check the correctness of the final 3D projective reconstruction result, we convert the 3D camera centered projective space back to the 3D object centered Euclidean space using the 3D Euclidean data of the tower available in Table 5.1 to measure the reconstruction errors in the metric space. The computation times for estimating the plane coefficient vectors and the coordinate transformation matrix for space integration are within a second.

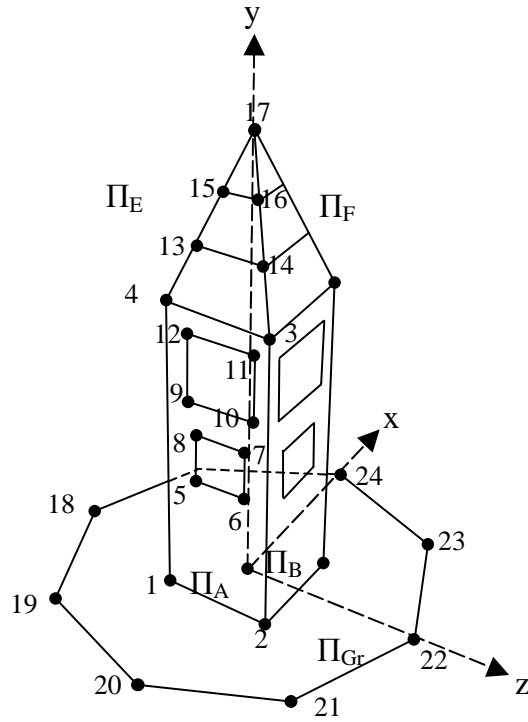


Fig. 5.2. The schematic diagram of the tower. The dimensions of the tower are 40 inches in depth (the x-direction), 40 inches in width (the z direction) and 180 inches in height (the y direction).

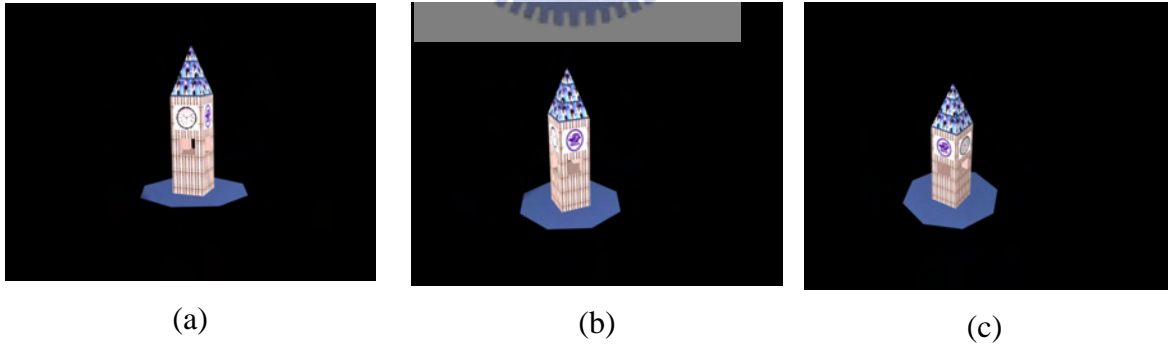


Fig. 5.3. Three distinct images I_1 , I_2 and I_3 taken at a distance of about 500 inches. The visible planes in the three images are Π_{Gr} , Π_A , Π_B , Π_E , Π_F in I_1 and I_2 , and Π_{Gr} , Π_B , Π_F , Π_C , Π_G in I_3 .

Table 5.1. 3D object centered coordinates of the tower feature points.

Point		1	2	3	4	5	6	7	8	9	10	11	12
	X	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20
	Y	0	0	120	120	50	50	70	70	80	80	110	110
	Z	-20	20	20	-20	-10	10	10	-10	-15	15	15	-15
Point		13	14	15	16	17	18	19	20	21	22	23	24
	X	-13.333	-13.333	-6.667	-6.667	0	0	-50	-70	-50	0	50	70
	Y	140	140	160	160	180	0	0	0	0	0	0	0
	Z	-13.333	13.333	-6.667	6.667	0	-70	-50	0	50	70	50	0

In what follows we assume the noise is uniformly distributed over the interval $[-R, R]$, where R indicates the noise strength or level. We generate 500 copies of noisy image data using the given noise model with $R = 0.5, 1.0, 1.5, 2.0$ pixels, respectively. Then we find the 500 reconstruction results and compute the mean and standard deviation of the differences between the true and the estimated values of the 3D coordinates of the tower feature points. Table 5.4 lists the statistics of the relative distance errors into the $x, y,$ and z components. The results indicate our reconstruction method is quite stable in the presence of the noise.

Table 5.2. The statistics of the distance errors of the reconstruction results.

Error value			Noise level R (in pixels)				
			0	0.5	1	1.5	2
Error type	Mean distance error	x	6.624e-5	0.0547	0.1096	0.2108	0.2505
		y	2.492e-5	0.0896	0.1811	0.2944	0.3923
		z	7.341e-5	0.0761	0.1514	0.3379	0.3798
	Standard deviation	x	2.027e-4	0.0855	0.1707	0.3655	0.3850
		y	3.188e-5	0.1323	0.2673	0.4416	0.5745
		z	1.192e-4	0.1165	0.2320	0.6085	0.9093

Since the rest of the planes visible in the images are estimated relative to the reference plane, we shall examine the effect of the assigned value of the reference

plane coefficient vector on the reconstruction. Five hundreds of the reference plane coefficient vectors are uniformly generated from the range $[10^{-3}, 10^2]$; we also randomly select the positive or negative sign for the coefficients. Fig. 5.4 depicts the reconstruction results under the effects of the random selection of the reference plane coefficient vector and the noise at different levels. The horizontal axis indicates the trial number of the reconstruction process and the vertical axis indicates the resulting distance errors. The various marks ".", "o", "x", "+" and "*" stand for mean errors of the computed relative distances associated with the uniform noise levels of $R=0, 0.5, 1, 1.5, 2$ pixels, respectively. The figure indicates the reconstruction results are virtually not affected by the random selection of the reference plane coefficient vector under the given specified noise. A remark is in order here. That is, we must avoid using $(1,1,1,0)^T$ for the reference plane coefficient vector, since the camera origin $(0,0,0,1)$ is supposed not to lie on the plane.

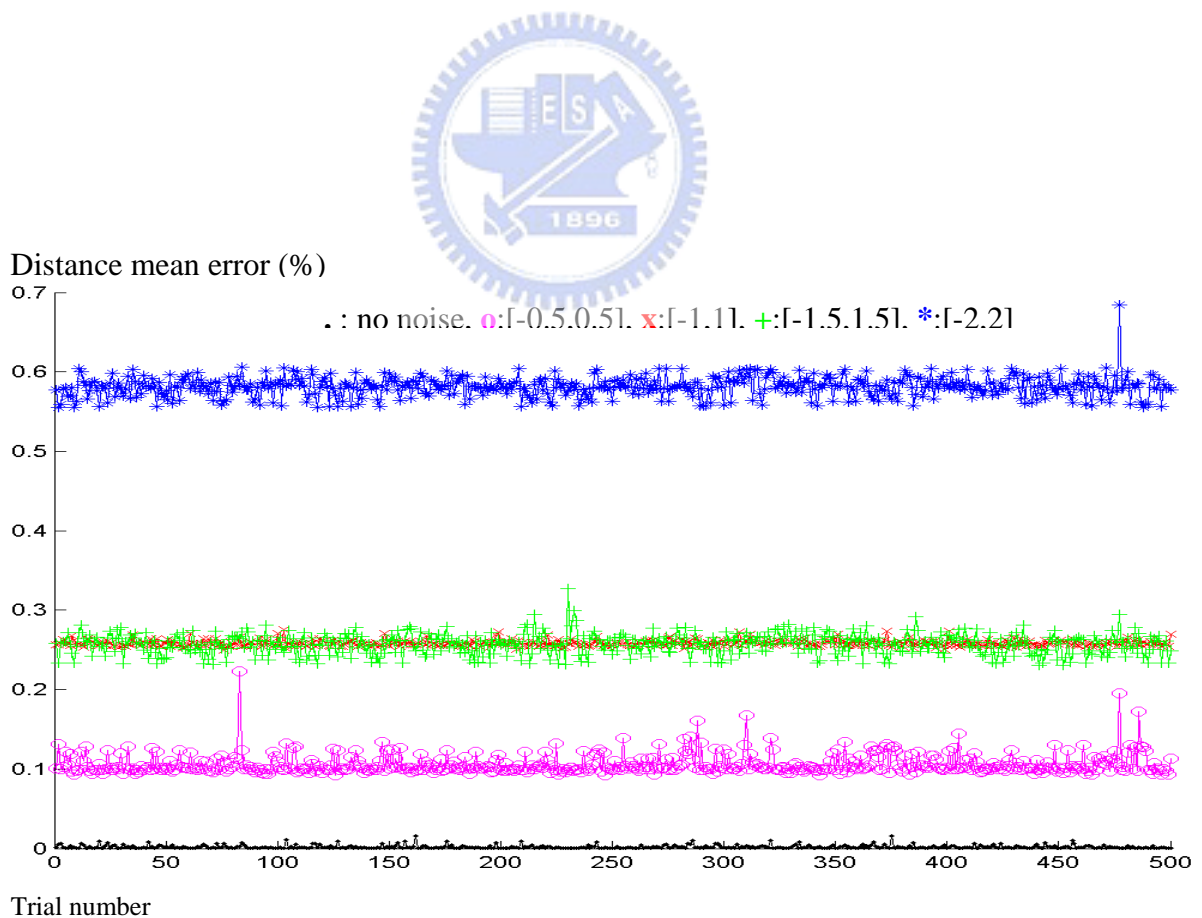


Fig. 5.4. The effects of the uniformly generated reference plane coefficient vectors and the noise at different levels on the reconstruction result.

Experiment 2:

For a comparison between our method and the point-based method employing the fundamental matrix estimated from two arbitrary planes with the aid of hallucinated points [54], we use the same setup as in the previous experiment and run the experiment 500 times with a uniform distribution at different noise levels. The reconstructed Euclidean position errors are computed and tabulated in Table 5.3. The position errors are in the unit of inch.

Table 5.3. The mean errors of the reconstructed Euclidean point positions for different setups

Noise level (in pixel) \ Method	0	0.2	0.5	1.0
Ours	0.0000301259	0.2030772297	0.4996064197	1.2803144642
(4,2,2)*	0.0000781737	1.6271384793	14.5382797879	47.6350445089
(4,4,1)*	0.0000125658	0.2113535985	0.5662224535	1.4202651215
(4,4,2)*	0.0000134805	0.2128984139	0.5343441024	1.4456471793

* (n, m, p): n, m are the respective numbers of points on the two planes, p is the number of points hallucinated per plane.

From this table, we observe that in the noiseless cases where the noise level is 0, all the reconstruction results obtained by the two methods are almost equally good and very small; the errors are due to the rounding/ truncation errors arising from numerical computations. As the noise level increases from 0 to 1.0 for all the (n, m, p) cases, we notice that the homography estimation using the four noisy data points varies dramatically, and, thus, the fundamental matrix computation with the resulting noisy data points is bad. These lead to the final reconstruction results with large errors. For a more interesting comparison, we compute the homography using the four outmost data points and then use the estimated homography to generate the hallucinated points located inside the area surrounded by the four outmost points; we denote these hallucinated points as the p (p = 1 or 2) points. In our method since the homographies are iteratively estimated in a region-based way, so our reconstruction results are good even in the presence of image noise. In these simulations, the reconstruction results of the two methods are nearly equally good. Even so, our method is better in the sense that we can efficiently find each 3D plane without the

need of computing 3D points, while the authors in the other method alternated a plane estimation stage with the point reconstruction stage. Thus, their method conducted two kinds of estimations: plane and points.

Experiment 3:

In this experiment the real images of a polyhedral, depicted in Fig. 5.5, are used to reconstruct the model of seven major planar surfaces. We go through the whole reconstruction process as we did in Experiment 1. The line parallelism and perpendicularity properties of the scene are used to compute the projective-to-Euclidean coordinate transformation matrix. First, the line parallelism is used to compute the plane at infinity which is then used to transform the reconstruction results from projective space to affine space. Secondly, the line perpendicularity is used to transform the reconstruction results from affine space to metric space. Further details can be found in [35][61]. Fig. 5.6 shows the four new views of the reconstructed model, which look like the real ones. Besides, the metric angles between individual object plane and the ground plane are shown in Table 5.4. The reconstructed object is found to be rather close to the true one.

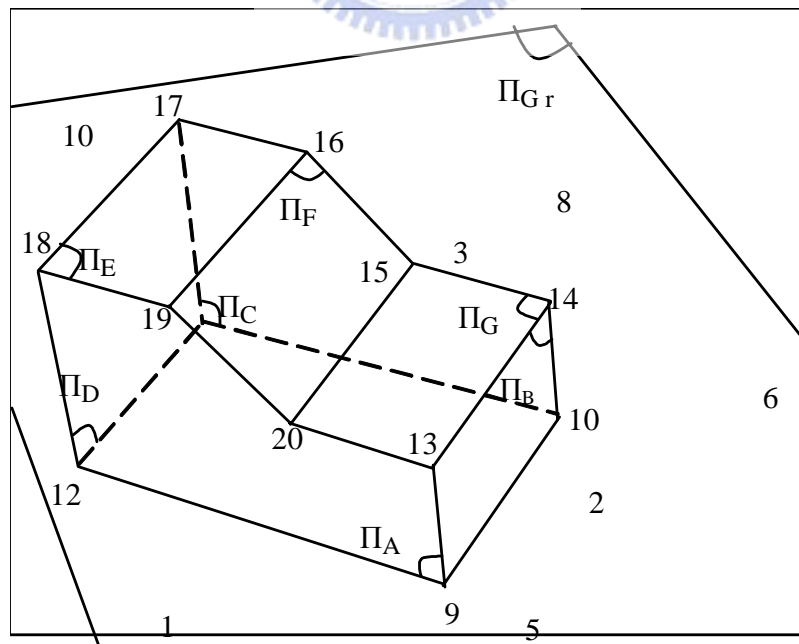


Fig. 5.5. The indices of the vertices and planes of the object.

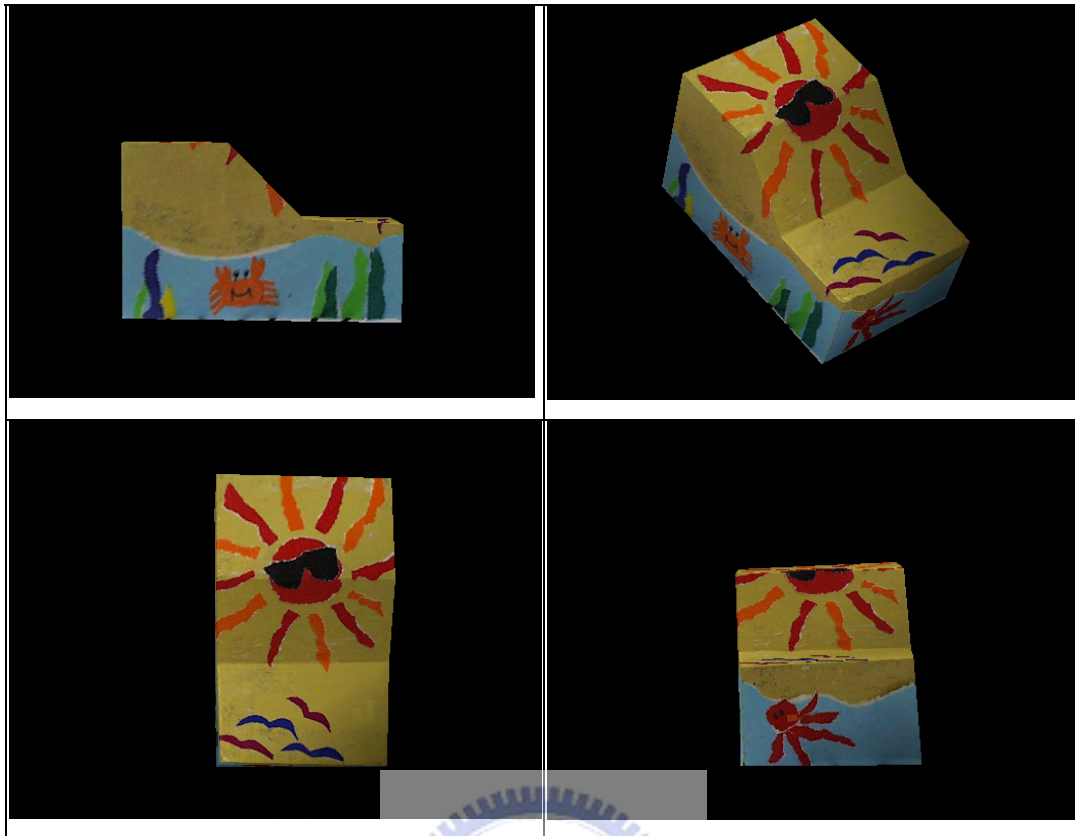


Fig. 5.6. New views of the reconstructed object with texture mapping.

Table 5.4. The estimated angles between the object planes and the ground plane

Angle Planes form the angle	Estimated value	Actual value
(Π_{Gr}, Π_A)	92.21	90
(Π_{Gr}, Π_B)	92.69	90
(Π_{Gr}, Π_C)	97.09	90
(Π_{Gr}, Π_D)	91.67	90
(Π_{Gr}, Π_E)	1.69	0
(Π_{Gr}, Π_F)	45.31	45
(Π_{Gr}, Π_G)	1.66	0

5.7 Conclusions

An uncalibrated planar object reconstruction method has been described in which we rely on the plane information. We first estimate the homography for all planar surfaces using the region features of planar surfaces, and then we use the homographies induced by two planes to compute the epipole. We represent explicitly the compatible projection equations for the stereo images using the information of planar homographies and an assigned reference plane coefficient vector. We continue to derive the 3D equations of the planes visible in the stereo images with respect to the assigned reference plane once the planar homography is determined. Finally, to integrate the reconstruction results obtained from different image pairs under a unified projective space, we use the homography and the plane coefficient vector information of two planes to derive the coordinate transformation matrix. We then compute the new plane equation for the planes in the unified projective space. In the experiments we conduct the sensitivity analysis on our method by introducing image noise. We also consider the effect of assigning the different values of the reference plane coefficient vector on the reconstruction results. Experimental results on the synthetic and real images indicate the reconstruction method works quite successfully. In the future, we shall consider combining this plane based reconstruction method with other methods to determine the 3D structure of more complex objects.

CHAPTER 6

SUMMARY AND FUTURE RESEARCH

6.1 Summary and Conclusions

In this dissertation, four different methods for constructing the 3D model of a real object from multiple images are proposed. In the first method, an error bound controlled subdivision scheme is presented to determine whether an octant is to be subdivided or not. Besides, the computation of the projection of 3D octants onto the 2D image planes is reduced by using the invariant property of cross ratio. The experiments are conducted on three real objects to demonstrate the performance of the new method. The results show the improvement of the new method over the conventional method in terms of memory space, computation time, and quality of the construction result. Theoretical analysis on the new method is presented.

The second construction method is an extension to the first method. The octree model produced by the first method is one that covered the real object since the GB_i^1 octants are treated as the black octants, while the octree model produced by the second method is one that approximates the real object from both sides of the object surface since the GB_i^2 octants are treated as the black octants and the GW_i^2 octants are treated as the white octants. The memory storage and the computer processing time required by the second method have been shown much less than those required by the first method.

In the third method, the octree construction method using the best first traversal scheme implemented by a priority queue is presented. The octants to be processed are arranged in the order of maximum projection error. The octant having maximum projection error is always the one first to be processed. The reconstruction quality in terms of XOR error of the octree projection image with the silhouette image has been shown better than the octree construction method using the breadth first traversal or depth first traversal. Besides, we also show that the memory requirement for the priority queue becomes less if the new subdivision scheme for grey nodes defined in

chapter 2 and chapter 3 is used to eliminate the necessity of the sorting and inserting of the “grey-black” and “grey-white” nodes.

The fourth construction method is to fit the object surface with planes. An uncalibrated planar object reconstruction method has been described in which we rely on the plane information. We first estimate the homography for all planar surfaces using the region features of planar surfaces, and then we use the homographies induced by two planes to compute the epipole. We represent explicitly the compatible projection equations for the stereo images using the information of planar homographies and an assigned reference plane coefficient vector. We continue to derive the 3D equations of the planes visible in the stereo images with respect to the assigned reference plane once the planar homography is determined. Finally, to integrate the reconstruction results obtained from different image pairs under a unified projective space, we use the homography and the plane coefficient vector information of two planes to derive the coordinate transformation matrix. We then compute the new plane equation for the planes in the unified projective space. In the experiments we conduct the sensitivity analysis on our method by introducing image noise. We also consider the effect of assigning the different values of the reference plane coefficient vector on the reconstruction results. Experimental results on the synthetic and real images indicate the reconstruction method works quite successfully.

6.2 Topics for Future Research

The possible extensions to the methods presented in the dissertation are:

- (a) Converting the octree model into polygonal representation.

In the future, we plan to convert the reconstructed octree model to a polygonal representation. One of the simplest way is to project the reconstructed octree model in the orthogonal directions of the 6 six faces of the octant and apply the z-buffer technique to obtain the outmost surface surrounded by the reconstructed octree model. Then, we can extract textures from the real object images and map them onto the model to produce realistic computer generated images.

- (b) Embedding image based rendering techniques to the construction method to provide photo-realistic rendering of the constructed object.

Instead of mapping single texture onto 3D polygon, image based rendering

technique like surface light field technique [62] proposed recently can be combined with the construction method to provide a photo-realistic rendering effect.

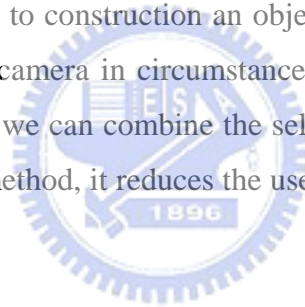
- (c) Combining several methods to determine the 3D structure of complex objects.

The octree construction method fails to construct the cavity surface of the object. Voxel coloring [27][65] or active vision techniques [63] which use the color consistency check or the deformation information of the projection pattern can be combined with the proposed method to construct the cavity surface.

Besides, we also plan to combine the plane based reconstruction method with other methods to determine the 3D structure of complex objects

- (d) Constructing the octree model without calibrating the camera in a priori.

For now, we have to calibrate the camera before using the proposed octree construction methods to construction an object. However, it is not easy for a user to calibrate the camera in circumstances, in particular, in the case of a hand-held camera. If we can combine the self-calibration techniques into the octree construction method, it reduces the user's effort to construct the model outdoors.



REFERENCES

1. Y. F. Wang, A. Mitiche, and J. K. Aggarwal, "Computation of surface orientation and structure of objects using grid coding," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 9, no. 1, pp. 129-137, 1987.
2. M. Asada, H. Ichikawa, and S. Jsuji, "Determining surface orientation by projecting a stripe pattern," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, no. 5, pp. 749-752, 1988.
3. G. Hu and G. Stockman, "3-D surface solution using structure light and constraint propagation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 4, pp. 390-402, 1989.
4. N. Shrikande and G. Stockman, "Surface orientation from a projected grid," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 6, pp. 650-655, 1989.
5. Z. Chen, T. L. Chia, and S. Y. Ho, "Measuring 3-D location and shape parameters of cylinders by a spatial encoding technique," *IEEE Trans. Rob. and Automat.*, vol. 10, no. 5, pp. 632-647, 1994.
6. Z. Chen, S. Y. Ho, and D. C. Tseng, "Polyhedral face reconstruction and modeling from a single image with structured light," *IEEE Trans. Syst. Man Cybernet.*, vol.23, no. 3, pp. 846-872, 1993.
7. T. L. Chia, Z. Chen and C. J. Yueh, "A method for rectifying grid junctions in grid-coded images using cross ratio," *IEEE Trans. Image Processing*, vol. 5, no. 8, pp. 1276-1281, 1996.
8. T. L. Chia, Z. Chen and C. J. Yueh, "Curved surface reconstruction using a simple structured light method," *Proc. ICPR, Vienna, Austria*, pp. 844-848, 1996.
9. C. S. Chen, C. C. Chiang, Y. P. Hung and J. L. Wu, "Range data acquisition using color structured lighting and stereo vision," *Proc. CVGIP, Taiwan, ROC*, pp. 88-95, 1995.
10. K. L. Boyer and A. C. Kak, "Color-encoded structured light for rapid active ranging," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 9, no. 1, pp. 14-28, 1987.
11. R. Vaillant and O. D. Faugeras, "Using Extremal Boundaries for 3-D Object Modeling," *IEEE Trans. on PAMI*, vol. 14, no. 2, 1992, pp. 157-173.

12. W. B. Seales and O. D. Faugeras, "Building three-dimensional object models from image sequences," *Computer Vision and Image Understanding*, vol. 61, no. 3, pp. 308-324, 1995.
13. C. H. Chien and J. K. Aggarwal, "Volume/surface octrees for the representation of three-dimensional objects," *CVGIP*, 36, pp. 100-113, 1986.
14. C. H. Chien and J. K. Aggarwal, "Identification of 3D objects from multiple silhouettes using quadtrees/octrees," *CVGIP*, 36, pp. 256-273, 1986.
15. H. Noborio, S. Fukuda, and S. Arimoto, "Construction of the octree approximation three-dimensional objects by using multiple views," *IEEE Tran. PAMI.*, vol. 10, no. 6, pp. 769-782, 1988.
16. Narendra Ahuja and Jack Veenstra, "Generating octrees from object silhouettes in orthographics Views," *IEEE Tran. PAMI*, vol. 2, no. 2, pp. 137-149, 1989.
17. H. H. Chen and T. S. Huang, "A survey of construction and manipulation of octrees," *CVGIP*, 43, pp. 409-431, 1988.
18. M. Potmesil, "Generating octrees models of 3D objects from their silhouettes in a sequence of images," *CVGIP*, 40, pp. 1-29, 1987.
19. S. K. Srivastava and N. Ahuja, "Octree generation from object silhouettes in perspective views," *CVGIP*, 49, pp. 68-84, 1990.
20. R. Szeliski, "Rapid octree construction from image sequences," *CVGIP: Image Understanding*, vol. 58, no. 1, pp. 23-32, 1993.
21. A. Laurentini, "How many 2d silhouettes does it take to reconstruct a 3D object?" *Computer Vision and Image Understanding*, vol. 67, no. 1, pp. 81-87, 1997.
22. Q. Delamarre and O. Faugeras, "3D articulated models and multi-view tracking with silhouettes," *Proceedings of International Conference on Computer Vision, Corfu, Greece, September*, pp. 716-721, 1999.
23. L. Bottino and A. Laurentini, "Interactive reconstruction of 3D objects from silhouette," *Proceedings of the 9th International Conference on Computer Graphics, Visualization and Computer Vision*, pp. 5-9, 2001.
24. Y. H. Cheng, Y. Y. Hsu, H. L. Chou and Z. Chen, "Reconstruction of 3D Objects from Multiple Silhouette Images: Shape Estimation and Texture Mapping," *Proc. CVGIP*, 2000.
25. Y. S. Fang, H. L. Chou and Z. Chen, "3D Shape Recovery of Complex Objects Multiple Silhouette Images," *Proc. CVGIP*, 2001.

26. W. Niem, "Automatic reconstruction of 3D objects using a mobile camera," *Image and Vision Computing*, 17, pp. 125-134, 1999.
27. S. M. Seitz and C. R. Dyer, "Photorealistic Scene Reconstruction by Voxel Coloring," *International Journal of Computer Vision*, vol. 35, no. 2, pp. 151-173, 1999.
28. G. K. M. Cheung, T. Kanade, J.-Y. Bouguet and M. Holler, "A real time system for robust 3D octant reconstruction of human motions," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp.714-720, 2000.
29. W. Matusik, C. Buehler, and L. McMillan, "Polyhedral visual hulls for real time rendering," *Proceedings of the 12th Eurographics Workshop on Rendering*, pp. 115-125, 2001.
30. G. Cheung, S. Baker, and T. Kanade, "Visual hull alignment and refinement across time: a 3D reconstruction algorithm combining shape-frame-silhouette with stereo," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 375-382, 2003.
31. R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision, Volume II*, Addison-Wesley, New York, 1993.
32. A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, New York, 1988.
33. P. A. Beardsley, A. Zisserman, D. Murray, "Sequential updating of projective and affine structure from motion," *Int. J. of Computer Vision*, vol. 23, no.3, pp. 235-259, 1997.
34. R. Criminisi, A. Zisserman, "Duality, rigidity and planar parallax," *Proc. European Conf. on Computer Vision*, vol. 2, pp. 846-861, 1998.
35. K. Daniilidis, J. Ernst, "Active intrinsic calibration using vanishing points," *Pattern Recognition Letters*, 17, pp. 1179-1189, 1996.
36. O. D. Faugeras, "What can be seen in three dimensions with an uncalibrated stereo rig?" *Proc. European Conf. on Computer Vision*, pp. 563-578, 1992.
37. O. D. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, Cambridge, MA, 1993.
38. A. W. Fitzgibbon, A. Zisserman, "Automatic Camera Recovery for Closed or Open Image Sequences," *Proc. European Conf. on Computer Vision*, pp. 311-326, 1998.
39. M. Fradkin, M. Roux, H. Maitre, U. M. Lelogu, "Surface reconstruction from

- multiple aerial images in dense urban areas,” Proc. Int. Conf. on Computer Vision, pp.262-267, 1999.
40. R. I. Hartley, R. Gupta, T. Chang, “Stereo from uncalibrated cameras,” Proc. Conf. on Computer Vision and Pattern Recognition, pp. 761-764, 1992.
 41. R. I. Hartley, “Projective reconstruction and invariants from multiple images,” IEEE Trans. Pattern Analy. Mach. Intell., vol. 16, no. 10, pp. 1036-1041, 1994.
 42. R. I. Hartley, “In defense of the eight-point algorithm,” IEEE Trans. Pattern Analy. Mach. Intell., vol. 19, no. 6, pp. 580-593, 1997.
 43. R. I. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision (Chapter 9). Cambridge University Press, Cambridge, UK, 2000.
 44. N. H. G. Hamid, R. Cipolla, “Identifying planar regions in a scene using uncalibrated stereo vision,” Proc. British Machine Vision Conference, 1997.
 45. M. Irani, B. Rousso, S. Peleg, “Recovery of ego-motion using region alignment,” IEEE Trans. Pattern Analy. Mach. Intell., vol. 19, no. 3, pp.268-272, 1997.
 46. M. Irani, P. Anandan, D. Weinshall, “From reference frames to reference planes: multi-view parallax geometry and applications,” Proc. European Conf. Computer Vision, vol. 2, pp. 829-845, 1998.
 47. B. Johansson, “View synthesis and 3D reconstruction of piecewise planar scenes using intersection lines between the planes,” Proc. Int. Conf. on Computer Vision, pp.54-59, 1999.
 48. Q. T. Luong, T. Vieville, “Canonic representations for the geometries of multiple projective views,” Computer Vision and Image Understanding, vol. 64, no. 2, pp.193-229, 1999.
 49. Q. T. Luong, O. D. Faugeras, “Determining the fundamental matrix with planes: instability and new algorithms,” Proc. Conf. on Computer Vision and Pattern Recognition, pp.489-494, 1993.
 50. R. Mohr, E. Arbogast, “It can be done without camera calibration,” Pattern Recognition Letters, vol. 12, no. 1, pp. 39-43, 1991.
 51. C. Rother, S. Carlsson, and D. Tell, “Projective Factorization of Planes and Cameras in Multiple Views,” Proc. Int. Conference on Pattern Recognition, pp. 737-740, 2002.
 52. C. Rother, Multi-View Reconstruction and Camera Recovery using a Real or Virtual Reference Plane. Doctoral Dissertation, ISBN 91-7283-422-6, KTH, Stockholm, Sweden, 2003.

53. R. Szeliski, "Video mosaics for virtual environments," *IEEE Computer Graphics and Applications*, vol. 6, pp. 22-30, 1996.
54. R. Szeliski, P.H.S. Torr, "Geometrically constrained structure from motion: points on planes," *Proc. European Workshop on 3D Structure from Multiple Images of Large-Scale Environments*, pp. 171-186, 1998.
55. D. Sinclair, A. Blake, "Quantitative planar region detection," *Int. J. of Computer Vision*, vol. 18, no. 1, pp.77-91, 1996.
56. H. S. Sawhney, "3D geometry from planar parallax," *Proc. Conference on Computer Vision and Pattern Recognition*, pp.929-934, 1994.
57. S. K. Sun, Z. Chen, T. L. Chia, "Invariant Feature Extraction and Object Shape Matching Using Gabor Filtering," *Proc. Int. Conf. on Visual Information System*, Hsinchu, Taiwan, pp. 95-104, 2000.
58. R. Y. Tsai, T. S. Huang, "Estimating three-dimensional motion parameters of a rigid planar patch, II: singular value decomposition," *IEEE Trans. Acoustic, Speech and Signal Processing*, vol. 30, no. 8, pp. 525-533, 1982.
59. L. Theiler., H. Chabbi, "Projective decomposition of planar facets," *Proc. Int. Conf. on Image Analysis and Processing*, pp.556-561, 1999.
60. L. Zelnik-Manor, M. Irani, "Multi-frame estimate of planar motion," *IEEE Trans. Pattern Analy. Mach. Intell.*, vol. 22, no. 10, pp. 1105-1116, 2000.
61. Z. Zhang, K. Isono, S. Akamatsu., "Euclidean structure from uncalibrated images using fuzzy domain knowledge: application to facial images synthesis," *Proc. Int. Conf. on Computer Vision*, pp. 784-789, 1998.
62. D. N. Wood, D. I. Azuma, K. Aldinger, B. Curless, T. Duchamp, D. H. Salesin, W. Stuetzle. "Surface Light Fields for 3D Photography," *Proc. SIGGRAPH*, 2000.
63. M. Kampel, S. Tosovic, and R. Sablatnig, "Octree-based fusion of shape from silhouette and shape from structured light," *Proc. IEEE 3DPVT*, 2002.
64. S. Avidan and A. Shashua, "Novel view synthesis by cascading trilinear tensors," *IEEE Trans. on Visualization and Computer Graphics*, vol. 4, no.4, pp.293-306, 1998.
65. K. N. Kutulakos and S. M. Seitz, "A Theory of Shape by Space Carving" *International Journal of Computer Vision*, vol. 38, no. 3, pp. 199-218, 2000.
66. W. Matusik, C. Buehler, R. Raskar, L. McMillan, and S. J. Gortler, "Image-Based Visual Hulls", *Proc. SIGGRAPH*, pp. 369-374, 2000.

VITA

NAME: Hong-Long Chou

BIRTH: Oct. 4th, 1971, Taipei, R.O.C.

EDUCATION:

B.Sc., Department of Computer Science and Engineering, Yuan Ze University, Sep. 1990 – Jun. 1994.

M.Sc., Department of Computer Science and Information Engineering, National Chiao Tung University, Sep. 1994 – Jun. 1996.

Ph.D., Department of Computer Science and Information Engineering, National Chiao Tung University, Sep. 1996 – present.

AWARDS:

Best paper of 1999 IPPR Conference on Computer Vision, Graphics and Image processing with K. H. Yu and Prof. Z. Chen.

Best paper of 2001 IPPR Conference on Computer Vision, Graphics and Image processing with Y. S. Fang and Prof. Z. Chen.

RESEARCH INTERESTS:

Computer Vision	Computer Graphics
Image processing	Virtual Reality

PUBLICATIONS

(A) Journals:

1. Y. S. Fang, H. L. Chou, Z. Chen, "3D Shape Recovery of Complex Objects Multiple Silhouette Images," Pattern Recognition Letters, pp.1279-1293, 2003.
2. H. L. Chou and Z. Chen, "Fast Octree Construction from Multiple Object Silhouettes with Quality Control," To appear in Journal of Information Science and Engineering.
3. H. L. Chou and Z. Chen, "A Novel 3D Planar Object Reconstruction from Multiple Uncalibrated Images Using the Plane-induced Homographies," To appear in Pattern Recognition Letters.
4. Z. Chen, H. L. Chou and Y. S. Fang, "A Quality Controllable Octree Construction Method with Time and Space Complexity Reduction," to be submitted.
5. H.L. Chou, Y. S. Fang, and Z. Chen, "Progressive Octree Construction with an Error Rate Measure," to be submitted.

(B) Conferences:

1. H. L. Chou, Z. Chen and T. L. Chia, "A Virtual Person Imitating the Body Motion of a Real Person via Computer Vision Technique," Proc. CSIA, 1997.
2. C. Y. Chen, Y. P. Fang, Z. Chen and H. L. Chou, Hand Gesture Commands for Slide View Control in A PC Based Presentation," Proc. of ICS, 1998.
3. K. H. Yu, H. L. Chou and Z. Chen, "Calibration-Free Augmented Reality For Embedding Virtual Objects In A Real Image Sequence," Proc. CVGIP, pp. 532-541, 1999.
4. Y. H. Cheng, Y. Y. Hsu, H. L. Chou and Z. Chen, "Reconstruction of 3D Objects from Multiple Silhouette Images: Shape Estimation and Texture Mapping," Proc. CVGIP, 2000.
5. Y. S. Fang, H. L. Chou and Z. Chen, "3D Shape Recovery of Complex Objects Multiple Silhouette Images," Proc. CVGIP, 2001.
6. H. L. Chou, Z. Chen and Y. C. Chen, "3D Plane-Based Reconstruction from Uncalibrated Images," Proc. CVGIP, pp. 399-405, 2002.
7. H. L. Chou and Z. Chen, "Homography-based Planar Object Reconstruction from Uncalibrated Images," Proc. ACCV, pp.312-317, 2004.