



HAL
open science

Optimisation du système de 3DTV basé sur la technique d'imagerie intégrale

Shasha Shi

► **To cite this version:**

Shasha Shi. Optimisation du système de 3DTV basé sur la technique d'imagerie intégrale. Signal and Image Processing. Télécom Bretagne; Université de Rennes 1, 2014. English. NNT: . tel-01206260

HAL Id: tel-01206260

<https://hal.science/tel-01206260>

Submitted on 28 Sep 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sous le sceau de l'Université européenne de Bretagne

Télécom Bretagne

En habilitation conjointe avec l'Université de Rennes 1

Ecole Doctorale – MATISSE

Optimisation du système de 3DTV basé sur la technique d'imagerie intégrale

Thèse de Doctorat

Mention : Traitement du signal et télécommunications

Présentée par **Shasha Shi**

Département : Image et traitement de l'information

Directeur de thèse : Basel Solaiman
Supervisée par: Patrick Gioia, Gérard Madec

Soutenue le 31 Janvier 2014

Jury :

Mme Béatrice Pesquet-Popescu, Professeur, Télécom ParisTech (Rapporteur)
M. Laurent Lucas, Professeur, Université de Reims Champagne-Ardenne(Rapporteur)
M. Basel Solaiman, Professeur, Télécom Bretagne (Directeur de thèse)
M. Florent Dupont, Professeur, Université Lyon 1(Examineur)
Mme Luce Morin, Maître de Conférences, Université de Rennes 1(Examineur)
M. Gérard Madec, Enseignant Chercheur, Télécom Bretagne (Examineur)
M. Patrick Gioia, Ingénieur Recherche, Orange (Examineur)

Acknowledgement

This thesis arose out of three years of research work that has been done in co-supervising of Orange Labs and Image and Information Processing Department in Telecom Bretagne.

By that time, I have worked with a great number of people who contribute in various ways to my research and the making of the thesis. It is a pleasure to convey my most profound gratitude to them all.

First of all, I would like to give my great thanks to my supervisor, Patrick Gioia and Gérard Madec, who offered me the great opportunity to work in the 'compression video avancée (CVA)' group in Orange Labs and Image and Information Processing Department of Telecom Bretagne. Without their help, this thesis would not have been possible. Their invaluable support and guidance throughout the period of my research.

I gratefully acknowledge Professor Basel Solaiman, the director of this thesis, for his advice, supervision and constant support.

I am also indebted to the manager of the group Dider Gaubil for his warmly reception, insightful discussions and instructions on my works.

I would also like to thank the members of my committee for accepting the invitation and spent their precious time helping me to improve this thesis: Professor A, Professor B, Professor C and D. Their advice and patience are really appreciated.

Many thanks to all members in our research group: Youness Andam, Gordon Clare, Christophe Daguët, Félix Henry, Joel Houssais, BiHong Huang, Joel Jung, Elie Gabreil Mora, Stéphane Pateux, Pierrick Philippe, Jean-Marc Thiesse, Gilles Tenious and Philippe Vonwyl. The group has been a source of friendships as well as good advice and collaboration.

I am grateful for time spent with friends, Yaning Liu, Hao Lin, Wei Chen, Guozhi Wei, BiHong Huang, Han Yan, Qinglin Lu, Eric Villain, Nathalie Salmon, Jérôme Royan, Raphaële Balter, Nicolas Monfort and all my friends from Orange and Telecom Bretagne. You leave me many beautiful memories.

Where would I be without my family? Many thanks go in particular to Zhe LI for his attention and help for my research works and life. I am indebted to my parents Hongmin Shi and Yaling Zhang, who tried all their best in improving my living and studying conditions, gave me their full confidence and encouraged me during the most difficult times.

Finally, I would like to thank everybody who helped me for the successful realization of thesis.

Abstract

This thesis focuses on the challenging issues that prevent current integral imaging systems from displaying high quality and eyestrain-free 3DTV. We provide the problems formulation and corresponding solutions that ameliorate the viewers' 3D experience in integral imaging systems.

A new computational integral image generation (CIIG) method is proposed at first to expand the viewing range of the displayed 3D images. Concretely, the method simulates curve recording panel in computational system, which is realized by changing the viewing direction of each elemental image (EI) according to its own coordinate. The results of our experiments demonstrate that the proposed CIIG method enhances the viewing range of the traditional CIIG methods without increasing the source image size.

Then three different compression approaches are proposed and evaluated. The first compression method is based on multi-view video coding (MVC). It is applied on sub-images (SI). In this method, we arrange the group of SIs to suit the format of multi-view video (MVV) and then encode the generated MVV using the MVC standard.

The second compression method is inspired by Multi-View Video plus Depth (MVD) representation. It is also applied on sub-images. The principle of this method is to use the 3D information existing in depth maps to predict the SI for other directions and then remove the inter-view correlation between original SIs. Thanks to the 3D geometrical information in depth map, this method outperforms the MVC-based scheme mentioned before. Besides the MVC and MVD based compression approaches, we also considered applying the discrete wavelet transform (DWT) for integral image compression. In the proposed approach, the input IIs are filtered several times by determined high pass and low pass filters until the correlation within one EI and between adjacent EIs are completely eliminated. According to the simulation results, DWT works better than MVC-based method in low-debit situation.

Finally, we investigate the reconstruction process and develop two new computational integral imaging reconstruction (CIIR) techniques to improve the performance and reduce the complexity of the basic CIIR. The first method takes advantage of Parallel-Group Projection (PGP) techniques. Instead of projecting each point of integral image to the reconstructed plane pixel by pixel, the proposed method reconstructs the 3D image by mapping a series of sub-images onto the reconstructed plane successively. Because of the parallelization between the light rays in each SI, this approach reduces the pixel mapping complexity of the reconstruction process.

The second design aims at increasing the resolution of EI so as to improve the quality of the reconstructed 3D image. In this method, the high definition EI is composed of several synthesized SIs which are generated by depth image rendering. This approach provides various significant progresses for integral imaging system. It improves the reconstructed 3D image quality and offers scalability for EI resolution by simply augmenting the synthesized SIs number. Additionally, it requires only one color image and one depth map rather than transmitting multiple color images to reproduce the 3D scene, which makes the transmission in integral imaging systems more efficient.

Keywords: integral imaging system, computational integral image generation(CIIG), multi-view video coding, multi-view plus depth (MVD), computational integral image reconstruction (CIIR).

Contents

Acknowledgements	i
Abstract	iii
Contents	vii
List of Figures	xii
List of Tables	xiii
Résumé	xvii
1 Introduction	1
1.1 Development of 3DTV	1
1.2 Motivations and Objectives	3
1.3 Contributions	5
1.3.1 A new method for computational integral imaging generation	5
1.3.2 New techniques for Integral Image compression	5
1.4 Two novel method for computational integral imaging reconstruction	6
1.5 Thesis Organization	7
2 State of art of 3D technology	9
2.1 Introduction	9
2.2 Depth perception cues in Human Viewing System	10
2.3 (Auto)-stereoscopic 3DTV Systems	12
2.3.1 (Auto)-Stereoscopic 3D content creation	12
2.3.2 (Auto)-Stereoscopic content compression	17
2.3.3 (Auto)-Stereoscopic 3D content transmission	20
2.3.4 (Auto)-Stereoscopic 3D content display	21
2.3.5 Constrains of (Auto)-Stereoscopic 3DTV system	22
2.4 Alternative 3D display technologies	23
2.4.1 Integral imaging display	23
2.4.2 Volumetric display	24
2.4.3 Holographic display	25
2.5 Conclusions	27

3	Computational Integral image Generation	29
3.1	Introduction	29
3.2	Previous computational integral image generation technique	33
3.3	Proposed computational integral image generation method	35
3.3.1	Depth range controlling	35
3.3.2	Viewing range enhancement	35
3.4	Experimental results on proposed generation method	37
3.4.1	Evaluation of depth range controlling method	39
3.4.2	Evaluation of viewing range enhancement method	40
3.5	Conclusions	48
4	Integral image Compression	49
4.1	Introduction	49
4.2	Analysis of the correlation of integral image sequence	51
4.3	Previous integral image compression methods	52
4.3.1	MPEG-II-based encoding method	53
4.3.2	3D-DCT-based compression scheme	54
4.3.3	Hybrid compression method using DCT and DWT	55
4.3.4	Karhunen-Loeve transform (KLT)-based compression technique	56
4.4	Proposed integral image compression methods	57
4.4.1	Multi-view Video Coding (MVC)-based compression method	57
4.4.2	Multi-view Video plus Depth (MVD)-based compression method	59
4.4.3	Discrete Wavelet Transform (DWT)-based compression method	63
4.5	Evaluation and comparison of proposed methods	66
4.5.1	Evaluation for MVC-based encoding approach	69
4.5.2	Evaluation for MVD-based encoding approach	72
4.5.3	Evaluation for DWT-based encoding approach	75
4.6	Conclusions	79
5	Computational Integral image Reconstruction	81
5.1	Introduction	81
5.2	Two types of CIIR techniques	82
5.2.1	View-based CIIR	83
5.2.2	Plane-based CIIR	86
5.3	Previous plane-based CIIR Method	88
5.3.1	Traditional method	88
5.4	Pixel-to-pixel mapping method	88
5.5	Proposed plane-based CIIR methods	89
5.5.1	Parallel-group Projection (PGP)-based CIIR method	89
5.5.2	DIBR-based-CIIR	91
5.6	Experimental results	92
5.6.1	Evaluation of PGP-based CIIR	92
5.6.2	Evaluation of DIBR-based CIIR	100
5.7	Conclusions	103

6 Conclusion	105
6.1 Summary of Contributions	105
6.1.1 A new method for computational integral imaging generation	105
6.2 Three new techniques for Integral Image compression	106
6.2.1 Two novel methods for computational integral imaging reconstruction	106
6.3 Perspectives	107
6.3.1 Computational integral image generation	107
6.3.2 Integral image compression	107
6.3.3 Computational Integral image reconstruction	108
Bibliography	116
Publications	117
Glossary	120

List of Figures

1.1	The Wheatstone stereoscope, mirrors A' and A are set at an angle in order to reflect the left and right eye drawings (E' and E) towards the viewer's eyes	2
1.2	Anaglyph 3D technique	2
1.3	Linear polarization technique	3
1.4	Eclipse 3D technique	4
1.5	Flow chart of an integral imaging system (NHK image)	5
2.1	Depth cues in different depth distance (E.Cutting and M.Vishton 1995)	11
2.2	Structure of Complete 3DTV Processing Chain	12
2.3	Parallel cameras model	14
2.4	Toed-in camera setup	15
2.5	Example of this video plus depth representation format	15
2.6	Scheme of MVC	18
2.7	Original depth image and its decomposition results of piecewise-linear functions	18
2.8	Frame compatible transmission format	19
2.9	The principle of auto-stereoscopic display	20
2.10	The principle of a multi-view auto-stereoscopic display	22
2.11	Optical setup for generating (left) and displaying (right) 3D scene in integral imaging system	24
2.12	Hologram recording	26
2.13	Hologram reconstruction	26
3.1	Optical setup of integral imaging system	30
3.2	Real mode	30
3.3	Virtual mode	31
3.4	Focus mode	31
3.5	Viewing range in integral imaging system	32
3.6	Degradation of recording resolution for two separated objects	33
3.7	Concept of PRR method	34
3.8	Concept of MVR method	34
3.9	Concept of PGR method	35
3.10	Pickup and display process of traditional CIIG method	36
3.11	Pickup and display process of depth controlling CIIG method	36

3.12	An example of curve lens array and concept of the recording and display process	37
3.13	Flow chart of proposed CIIG method	37
3.14	CIIG system	38
3.15	Integral image generated by traditional CIIG method and its reconstructed 3D image	38
3.16	Integral image generated by depth controlling CIIG method and its reconstructed 3D image	39
3.17	Reconstructed results of further object by traditional CIIG method and depth controlling CIIG method	39
3.18	Integral image generated by traditional CIIG method with large resolution and its reconstructed 3D image for different viewing position for 3D object 'cube'	40
3.19	Integral image generated by traditional CIIG method with half recording density in each EI and its reconstructed 3D image for 3D object 'cube'	41
3.20	Integral image generated by traditional CIIG method with half recording range in each EI and its reconstructed 3D image in large viewing range (b, c, d) and halved viewing range (e, f, g) for 3D object 'cube'	42
3.21	Integral image generated by viewing angle enhanced CIIG method and its reconstructed 3D image for 3D object 'cube'	43
3.22	Integral image generated by traditional CIIG method with large resolution and its reconstructed 3D image in different viewing position for 3D object 'spaceship'	44
3.23	Integral image generated by traditional CIIG method with half recording density in each EI and its reconstructed 3D image for 3D object 'spaceship'	45
3.24	Integral image generated by traditional CIIG method with half recording range in each EI and its reconstructed 3D image in large viewing range (b, c, d) and halved viewing range (e, f, g) for 3D object 'spaceship'	46
3.25	Integral image generated by viewing angle enhanced CIIG method and its reconstructed 3D image for 3D object 'spaceship'	47
4.1	Integral image of 'cube'	51
4.2	Integral image of 'spaceship'	52
4.3	Spatial domain correlation in an integral image	52
4.4	Temporary domain correlation in an integral image sequence	53
4.5	Process of MPEG based integral image encoding method	53
4.6	Three different mode for scanning elemental images in an II	54
4.7	Scheme of 3D-DCT-based compression method	54
4.8	Scheme of hybrid compression method using DCT and DWT	55
4.9	Scheme of Packet partition process	55
4.10	Sub-image generation scheme	56
4.11	Block diagram of KLT-based encoding method	57
4.12	Viewing angle and object size are invariable for corresponding sub-images	58
4.13	Flow chart of the proposed integral images compression system	58

4.14	Converting SIs into MVV	59
4.15	Spiral scanning topology for SIs	59
4.16	Scheme of MVC	60
4.17	Concept of sub-image	60
4.18	Flow chart of proposed compression method	61
4.19	Mapping a 3D point to 2D clip plane by orthographic (left) and oblique (right) projection	62
4.20	Flow chart of the compression scheme	64
4.21	2D-DWT decomposition	65
4.22	DWT process for one II	66
4.23	Hybrid MPEG-4 based method	66
4.24	‘spaceship’ original integral image	67
4.25	‘spaceship’ generated Sub-images	68
4.26	Performance of MVC compression scheme for SIs, AVC scheme for SIs and baseline AVC scheme for integral images sequence	69
4.27	‘cube’ original integral image	70
4.28	‘spaceship’ generated Sub-images	71
4.29	Performance of MVC compression scheme for SIs, AVC scheme for SIs and baseline AVC scheme for integral images sequence	72
4.30	Original SIs and depth map	73
4.31	Synthesized SIs (up) and Errors (down) between original SIs and syn- thesized SIs	74
4.32	Performance of proposed compression method (labeled ‘SIs_MVD’), MVC-based method (labeled ‘SIs_MVC’) and baseline H.264/AVC method (labeled ‘II_AVC’)	75
4.33	‘spaceship’ DWT coefficients	76
4.34	Performance of proposed compression method (labeled ‘SIs_MVD’), MVC-based method (labeled ‘SIs_MVC’) and baseline H.264/AVC method (labeled ‘II_AVC’)	77
4.35	‘cube’ DWT coefficients	78
4.36	Performance of proposed compression method (labeled ‘SIs_MVD’), MVC-based method (labeled ‘SIs_MVC’) and baseline H.264/AVC method (labeled ‘II_AVC’)	79
5.1	Optical setup for displaying 3D scene in integral imaging system	82
5.2	Principle of view-based CIIR	83
5.3	view-based reconstructed images for 3D object ‘spaceship’	84
5.4	view-based reconstructed images for 3D object ‘cube’	85
5.5	Principle of plane-based CIIR	86
5.6	Example of reconstructed image by plane-based CIIR technique	87
5.7	Reconstruction principle of traditional CIIR method	88
5.8	Reconstruction principle of pixel-to-pixel mapping CIIR method	89
5.9	Concept of proposed CIIR method	90
5.10	Block diagram of proposed CIIR method	90
5.11	Computational reconstruction	91
5.12	Proposed computational integral imaging system	91

5.13	Generated SIs for 3D scene ‘spaceship’	92
5.14	II composed by the generated SIs for 3D scene ‘spaceship’	93
5.15	Comparison of reproduced ROI of three different methods for 3D scene ‘spaceship’	94
5.16	Generated SIs for 3D scene ‘cube’	95
5.17	II composed by the generated SIs for 3D scene ‘cube’	96
5.18	Comparison of reproduced ROI of three different methods for 3D scene ‘cube’	97
5.19	Computation cost of pixel-to-pixel and parallel-group projection CIIR for the 3D scene ‘spaceship’	99
5.20	Computation cost of pixel-to-pixel and parallel-group projection CIIR for the 3D scene ‘cube’	99
5.21	Reference SI from viewing angle equals to 0° and its corresponding depth map	100
5.22	Synthesized virtual SIs	100
5.23	Synthesized virtual SIs	101
5.24	Integral image composed by three sub-images shown in figure 5.22	102
5.25	Reconstructed images from -15° , 0° and 15°	102
5.26	Integral image composed by 11 sub-images shown in figure 5.23	103

List of Tables

2.1	Comparison of 3D display types	27
3.1	PSNR comparison of traditional and depth controlling CIIG method .	40
3.2	PSNR comparison of traditional and proposed CIIG method for ‘cube’	43
3.3	PSNR comparison of traditional and proposed CIIG method for ‘space- ship’	47
4.1	Comparison of proposed integral image compression methods	80
5.1	Comparison of the computation time of three CIIR methods for ‘space- ship’	98
5.2	Comparison of the computation time of three CIIR methods for ‘cube’	98
5.3	PSNR comparison of 3D image reconstructed by three sub images and eleven sub images	103

Résumé en Français

Résumé

Motivations de la Thèse

Cette thèse vise à lever les verrous existant dans l'application de l'imagerie intégrale aux systèmes de 3DTV. L'objectif de cette thèse est de trouver des solutions pour améliorer l'expérience 3D des téléspectateurs dans le système d'imagerie intégrale.

L'apparition de la 3D stéréoscopique remonte à 1838 lorsque Charles Wheatstone découvrit le principe de la vision binoculaire. Il a indiqué que l'homme est capable d'appréhender la profondeur du monde qui nous entoure grâce à différents indices ; la vision binoculaire étant un des indices importants de perception de la profondeur qui permet au cerveau d'interpréter l'information 3D. La rétine de chaque œil forme une image en deux dimensions de notre environnement. L'image d'un objet n'est pas projetée au même endroit sur chaque œil, cette différence de parallaxe horizontale est analysée par notre cerveau qui produit une représentation de son environnement en 3D.

La seconde phase de la technologie 3D a commencé avec l'apparition de la technique des lunettes anaglyphes. L'image anaglyphe superpose sur un écran deux images : une teintée en rouge pour l'image gauche et en vert pour l'image droite. un léger décalage horizontal appelé disparité est appliqué entre les deux images. Ensuite, des lunettes spéciales dites anaglyphes disposant de filtres colorés adaptés vont fournir les images indépendantes gauche et droite à chaque œil. Le cerveau fusionne les deux informations et perçoit grâce à la disparité une scène en trois dimensions.

Le soi-disant "engouement pour les films 3-D" dans les années 1952 à 1955 commence avec l'utilisation de lunettes polarisantes. Semblable à la technique anaglyphe, deux images sont superposées sur un même écran ; Elles sont projetées via des filtres polarisants différents. Une paire de lunettes 3D avec des filtres polarisants correspondants est utilisée pour séparer les images gauche et droite. Chaque filtre ne laisse passer que les rayons lumineux de même polarité et bloque la lumière polarisée différemment pour distribuer une image différente sur chaque œil. Comme elle est capable de reproduire un film de haute qualité stéréoscopique avec un assez faible coût, cette technique de polarisation ouvra une nouvelle ère à la technologie 3D.

A partir de 1985, avec la sortie de la technique IMAX 3D, le film en 3D a connu un énorme succès. IMAX, une norme de format de film pour très grand écran, a créé une technologie 3D mathématiquement exacte, de sorte qu'elle avait la capacité d'éliminer la fatigue de l'œil, un effet secondaire qui a été constaté par le public lors de la visualisation des films utilisant d'autres technologies. En fait, cette exactitude mathématique de la visualisation 3D a diminué de manière significative les géométries

approximatives et imprécises précédentes. En outre, le projecteur 3D IMAX offre une luminosité et un contraste inégalés par rapport aux autres technologies 3D disponibles aujourd'hui. Deux méthodes sont exploitées en IMAX pour créer l'illusion 3D pour le téléspectateur. La première est la technique de polarisation introduite précédemment. L'autre est le procédé Eclipse, qui utilise des lunettes actives qui obturent l'œil gauche lorsque l'image droite est affichée et vice versa.

2009 et 2010 constituent le point culminant pour les films 3D. Avec des blockbusters tels que "Monstres vs Aliens", "Up" et "Avatar" distribués au grand public, la stéréoscopie s'est attribuée une place sur le long terme. Au même moment, la technologie s'est étendue de l'industrie du cinéma à la production télévisuelle. Plusieurs chaînes de télévision ont annoncé leur intention de diffuser des programmes et des spectacles en 3D, ce qui a amené l'industrie à se poser les problèmes liés à la compression et à la transmission des contenus. Certains organismes de normalisation, tels que ISO/MPEG, ITU, DVB, HDM, ont développé des normes de transmission et d'interopérabilité des systèmes 3DTV. Tous les téléviseurs de nouvelle génération sont désormais compatibles 3D et ont atteint un niveau de prix compatible avec les exigences grand public. À terme, même s'il reste encore des réticences au niveau de l'utilisation de lunettes et de la qualité des contenus proposés, on peut penser que le téléviseur 3D deviendra un centre de divertissement omniprésent et irremplaçable à la maison.

Du fait qu'elle est beaucoup plus immersive que la technologie 2D, la 3D a eu un énorme succès sur le marché du cinéma grand public au cours des dernières années. Dans le même temps, la télévision 3D a également attiré une attention non négligeable. Cependant, en dépit de sa brillante perspective, la 3DTV a encore de nombreuses difficultés à franchir avant d'atteindre à la fois une qualité visuelle de premier plan et une expérience de visualisation confortable. Les limites de la 3DTV stéréoscopique actuelle sont les suivantes:

- Impossibilité de fournir des images de haute qualité 3D sans lunettes spéciales.
 - Conflit entre le point de convergence et le point d'accommodation de l'œil qui conduit à une gêne visuelle lors d'une observation prolongée. Comme nous l'avons mentionné précédemment, la création de sensation de profondeur est juste une simulation de la procédure de perception 3D humaine, elle n'est pas en mesure d'offrir un effet rigoureusement identique à ce qu'on perçoit dans la vie courante. La différence essentielle entre la perception de la profondeur par stéréoscopie et la perception de la profondeur dans le monde réel pourrait provenir de l'incohérence entre l'accommodation et la convergence. En effet, dans la vie réelle, les yeux convergent et focalisent dans le même plan, alors qu'en stéréoscopie, les yeux focalisent toujours dans le plan de l'écran, où l'image est générée alors qu'ils convergent dans le plan où semble apparaître l'objet, ce qui peut rendre spectateur mal à l'aise lorsque les valeurs de profondeur sont importantes.
 - Lorsque le spectateur se déplace, la scène le suit ; il est impossible de tourner autour des objets comme dans la vie réelle.
-

- Pas de parallaxe bidirectionnelle qui permettrait à l'observateur de voir différentes images correspondant à différentes positions verticale et horizontale. Une technologie 3DTV qui résout les problèmes ci-dessus permettrait aux observateurs de voir une image en 3D comme s'il s'agissait d'un objet réel. À cette fin, de nombreuses autres technologies d'affichage 3D basées sur la volumétrie, l'imagerie holographique intégrale ont été proposées et conçues ces dernières décennies. Parmi ces nouvelles technologies, l'imagerie intégrale est la plus attrayante en raison du bon compromis atteint entre la complexité et l'effet 3D. L'imagerie intégrale a été introduite par Gabriel Lippmann en 1908. Dans cette technologie, un réseau de microlentilles est utilisé pour acquérir une scène 3D, une même scène est enregistrée sous différents angles légèrement différents en horizontal et en vertical. Lorsque les images qui en résultent sont rectifiées et vues à travers un réseau similaire de lentilles, une image unique, composé de petites portions de l'ensemble des images, est vu par chaque œil. La partie observée de chaque microlentille est déterminée par la position de l'œil.

Système d'imagerie Intégrale

Différent des systèmes de vision binoculaire traditionnels qui simulent la perception 3D humaine en transmettant des images différentes dans les yeux gauche et droit, le système d'imagerie intégrale permet la génération de vraies images 3D en reproduisant les rayons lumineux de la scène 3D avec différentes couleurs et direction. Pour cette raison, il est théoriquement capable d'éliminer l'incohérence entre la convergence et l'accommodation dans le système de vision humain, de fournir une véritable expérience 3D pour les téléspectateurs sans sentiments désagréables.

Comme présenté auparavant, chaque image élémentaire est enregistrée derrière une lentille élémentaire. La coordonnée de cette lentille détermine la position de visionnement de cette image élémentaire. Chaque image élémentaire représente la scène 3D à partir d'un angle de vue différent. De même, dans le processus d'affichage, chaque sous-image est projetée à travers le réseau de lentille correspondant. L'image 3D reconstruite est observée aux points d'intersection des rayons de lumière projetés émanant du réseau de lentilles.

Construire un système d'imagerie optique intégrale est un projet très sophistiqué exigeant des dispositifs optiques de haute précision et un travail de calibrage précis. Afin de faciliter les activités de recherche, un système intégrale d'imagerie, qui vise à simuler le processus d'enregistrement et de reproduction d'un système d'imagerie optique intégrale (OII) de façon numérique, a été récemment développé et largement étudié. Comme le système optique, le système OII est également composé de deux processus fondamentaux, appelés la génération et la reconstruction. Le processus de génération est mis en œuvre par la technique de rendu par ordinateur pour créer une image intégrale (II) d'une scène 3D fournie, alors que le processus de reconstruction est obtenu en simulant l'effet optique de projection dans un procédé d'affichage d'une matrice de lentilles virtuelle par ordinateur.

L’affichage d’imagerie intégrale 1D est très similaire à l’affichage multi-vues. La principale différence réside dans la façon de projeter la distribution des faisceaux. En raison de projections de faisceaux parallèles, l’exposition de formation d’image intégrale peut fournir une parallaxe continue d’image en 3D avec une qualité constante. Néanmoins, la qualité de l’image perçue de l’affichage multi-vues va changer avec différents points d’observation en raison de la convergence du faisceau projeté.

Grâce à un grand nombre de microlentilles, l’imagerie intégrale est en mesure de fournir une vision 3D complète aux spectateurs sans dispositifs de visualisation de type lunettes, de façon à alléger les problèmes d’inconfort d’un système 3DTV traditionnel. Toutefois, certains inconvénients importants persistent. L’angle de vision est limité, la profondeur de l’image et le grand volume de données empêchent encore l’application pratique de système d’imagerie intégrale. L’objectif de cette thèse est d’améliorer les performances du système d’imagerie intégrale. Notre travail couvre les aspects majeurs de la technologie d’imagerie intégrale (1), la génération d’images intégrale, (2) la compression de l’image intégrale et (3) la reconstruction de l’image intégrale.

Cette thèse porte sur les questions difficiles qui empêchent le système d’imagerie intégrale d’afficher la 3DTV de haute qualité. Nous fournissons la formulation des problèmes et les solutions correspondantes qui améliorent l’expérience 3D du spectateur dans un système d’imagerie intégrale par trois aspects respectifs.

La première partie se concentre sur la génération d’image intégrale. Les problèmes de la plage de profondeur limitée et de l’étendue de l’affichage sont pris en compte et résolus dans cette partie.

Avec le développement des capacités du matériel informatique et des techniques graphiques sur ordinateur, la qualité de l’image générée par la technique CIIG (génération d’image intégrale) a été améliorée de manière significative. Cependant, certains problèmes inhérents tels que l’angle de vision limité, la profondeur de l’image et la résolution d’affichage faible sont encore en suspens.

Dans un système de base de l’imagerie intégrale, il existe toujours un compromis entre le champ visuel, l’image 3D en continue et la résolution d’image 3D. Ce compromis peut être expliqué de deux façons : D’une part, le nombre total de pixels sur la plaque de capture est le produit du nombre de lentilles élémentaires avec le nombre de pixels dans chaque image élémentaire, d’autre part, en cours d’affichage, le nombre de reproduction des vues 3D est le produit de la densité de l’angle de visualisation et de la portée de visualisation.

Un autre inconvénient du système d’imagerie intégrale réside dans la limitation de la gamme de profondeur enregistrée. Bien que le mode de mise au point soit capable de garder la même définition enregistrée sur l’ensemble de la scène 3D avec la direction de profondeur, la résolution d’un objet spécifique se dégrade si elle sa localisation est éloignée du réseau de lentilles. La résolution de l’enregistrement de l’objet lointain est inférieure à celle d’un objet proche en raison d’un nombre réduit de rayons d’échantillonnage recoupés dans sa surface. En conséquence, sur le côté de l’écran, l’impression de haute qualité est con-

trainte à la proximité de l'objet avec la matrice de lentilles. L'effet de distorsion avec un objet observé peut être observé dans certains angles de vue.

En conclusion, il existe un compromis entre ces trois critères importants de l'effet 3D: le champ visuel, l'angle de vision limité et la définition de l'image 3D affichée. L'amélioration de quelqu'un d'entre eux mènera à la dégradation des deux autres facteurs. La recherche d'une méthode appropriée pour briser cette restriction est un objectif essentiel pour de nombreuses équipes de recherche.

Nous avons proposé et démontré une nouvelle méthode de CIIG pour contrôler l'intervalle de profondeur et élargir le champ visuel du système CII(imagerie intégrale) traditionnel. Deux principales contributions ont été fournies dans notre système de génération. Premièrement, le déplacement des caméras virtuelles pour rendre chaque concentration de sous-image dans la même zone peut enregistrer un objet 3D lointain avec une résolution plus élevée. Deuxièmement, le changement de direction de chaque EI(image élément) améliore le champ visuel pour le spectateur. Les résultats de nos expériences montrent que la méthode de CIIG proposée permet de contrôler la portée de la profondeur des méthodes de CIIG traditionnelles, d'améliorer ainsi la qualité de l'image reconstruite de l'objet 3D lointain. En outre, il permet aussi d'améliorer la gamme visible de l'affichage d'une scène 3D sans effet de réduction des autres critères de l'effet 3D tel que la densité de l'angle de visualisation et de la résolution d'image reconstruite.

La seconde partie aborde les questions de compression de l'image intégrale. Comme nous le savons, la transmission d'importants volumes de données est un des plus gros problèmes pour l'application pratique du système d'imagerie intégrale. L'objectif de cette thèse est d'étudier la propriété de l'image intégrale, comprendre les limites des méthodes de compression et de proposer des nouvelles approches de codage afin d'améliorer l'efficacité de la compression.

Comme expliqué précédemment, l'image intégrale est un cas particulière d'image 2D qui enregistre une scène 3D en un nombre d'images élémentaires des points de vue différents. Par conséquent, l'ensemble des EI peut décrire non seulement la couleur et l'intensité, mais aussi l'information de direction des rayons lumineux réfléchis par l'objet 3D. Dépendamment de l'information 3D complète, le système d'imagerie intégrale peut fournir une parallaxe complète et d'une parallaxe d'observation quasi-continue pour les téléspectateurs sans contrainte de visualisation indésirable dépendants des techniques stéréoscopiques 3D traditionnels.

En fonction de l'analyse, la définition de l'image 3D produite est égale au nombre des EI, tandis que la densité de l'angle de vision de l'image observée est en accord avec le nombre de pixels dans une EI. Pour cette raison, afin de fournir une image 3D de haute qualité pour le spectateur, nous devons fournir une image intégrale de très haute résolution dans le côté de l'écran. La question qui est posée est « comment compresser le volume important de II (image intégrale), qui est un problème important lors de la mise en œuvre pratique d'un système d'imagerie intégrale.

Afin de mieux comprendre les diverses techniques de codage pour l'image intégrée, on réalise une analyse générale de la distribution de corrélation dans un II. La corrélation d'une seule image intégrale peut être détectée non seulement entre les pixels voisins, mais aussi entre les EIs adjacentes. Ces deux types de corrélation sont composés des quatre dimensions spatiales redondantes au sein d'un II, plus la corrélation temporelle entre IIs successives, sachant qu'une séquence II contient cinq corrélations dimensionnelles au total.

B Diverses méthodes de compressions traditionnelles ont été testées récemment. Le principe et les limites de quatre méthodes de codage différentes sont examinés. Sur la base de notre étude, ces méthodes de compression sont parvenues à améliorer l'efficacité de l'encodage dans une certaine mesure, mais ils contiennent aussi certaines limites. Afin d'exploiter la redondance dans la séquence d'image intégrale de manière plus approfondie, trois nouvelles méthodes de compression sont proposées. Leurs structures, les avantages et les problèmes seront exposés en détail.

La première méthode est basée sur la technique du codage vidéo multi-vue (MVC). Cette méthode est appliquée sur les sous-images (SI), qui est une autre forme d'image 2D transformée à partir d'images élémentaires d'origine. Chaque SI représente la scène 3D à partir de directions d'observations parallèles et contient une capacité de compression supérieure par rapport à l'image élémentaire (EI) capturé à l'origine. Pour cette raison, nous considérons l'organisation du groupe de SIs comme le format de vidéo multi-vues (MVV) puis le codage du MVV généré selon la norme MVC. Cette approche améliore l'efficacité de la compression comparé à la méthode traditionnelle MPEG-4/AVC.

La deuxième méthode de compression est inspirée par la représentation Multi-View-plus-Depth (MVD). Elle est également appliquée aux sous-images. Le principe de cette méthode est d'utiliser les informations 3D dans la carte de profondeur pour prédire de la SI puis retirez la corrélation inter-vue entre les différentes sous-images. Grâce l'information 3D de la carte de profondeur, cette méthode est meilleure que MVC mentionnée précédemment.

Outre les approches de compression basées sur MVC et MVD, nous avons également considéré l'application de la transformée en ondelettes discrète (DWT) pour la compression de l'image intégrale. Contrairement à la DCT, la DWT est capable de transformer l'information 2D originale du domaine pixel dans le domaine pixel-fréquence tout en conservant les informations du domaine spatial des données de l'image. Dans l'approche proposée, les entrées IIs sont filtrées de multiple fois par un filtre passe-haut et un filtre passe-bas, jusqu'à ce que la corrélation au sein d'une EI et entre EI adjacentes soit totalement éliminée. Selon la simulation, les résultats montrent que DWT fonctionne mieux que la méthode MVC en situation de faible débit.

Enfin, nous étudierons le processus de reconstruction et le développement de deux nouvelles techniques de reconstruction d'imagerie intégrés qui améliorent les performances et réduisent la complexité de CIIR(reconstruction d'image intégrale).

Comme indiqué précédemment, dans le processus de reconstruction d'image intégré, la plaque d'affichage est placée là où était la plaque de capture, et est irradiée par une source de lumière incohérente. Les faisceaux lumineux passant à travers la plaque d'affichage et le réseau de lentilles retrace la route d'origine, puis convergent à l'endroit où l'objet était, ce qui forme une image 3D.

Dans un système optique, l'effet 3D affichée est affecté non seulement par la qualité de l'image intégrale, mais également par la diffraction et la limitation des dispositifs optiques. Afin d'éviter le facteur instable d'ouverture optique, le système CIIR a été développé et largement étudié récemment. L'objectif de CIIR est de simuler l'effet de projection de rayons optiques d'un EI dans OIIR (reconstruction d'image intégrale Optical) à l'aide d'un ordinateur simulant des lentilles virtuelles ou des trous microscopiques.

Grâce à l'amélioration de la capacité de calcul des ordinateurs, la performance du système CIIR a progressé de façon spectaculaire, mais il existe encore quelques limitations qui font obstacle à son utilisation pratique, comme une faible qualité de l'image en sortie reconstruite (ROI) et la complexité élevée du calcul. Dans cette thèse, deux nouvelles approches seront proposées et testées pour améliorer la performance de CIIR.

La première méthode s'appuie sur la technique de projection en groupes parallèles (PGP). Plutôt que de projeter chaque point de l'image intégrale par rapport au plan reconstruit pixel par pixel, la méthode proposée reconstruit l'image 3D en cartographiant une série de sous-images sur le plan reconstruit successivement. En raison de la parallélisations des rayons lumineux dans chaque SI, cette approche permet de simplifier la complexité du processus de reconstruction.

La seconde conception vise à augmenter la résolution de l'image élémentaire (EI) afin d'améliorer la qualité de l'image 3D reconstruite. Dans cette méthode, la définition de l'EI est constituée par le nombre de SIs synthétisés, qui sont générés en image de profondeur. En outre, cette méthode ne nécessite qu'une seule image couleur et une carte de profondeur plutôt que la transmission de multiples images de couleur pour reproduire la scène 3D, qui pourrait rendre la transmission dans le système d'imagerie intégrale plus efficace.

Organisation de la thèse

Cette thèse est composée de six chapitres, et elle est divisée en trois parties.

La première partie, contenant les chapitres 1 et 2, vise à donner un aperçu général et les informations de base qui sont requises pour la compréhension de cette thèse. Chapitre 1 fournit une description générale du contexte sur la base duquel la recherche est effectuée. Premièrement, nous identifions plusieurs grandes étapes de l'évolution de la technologie 3D. Deuxièmement, nous présentons les problèmes actuels de la télévision en 3D, et en particulier nous illustrons les défis et les solutions apportées par l'utilisation de la technologie "imagerie intégrale". Les contributions majeures de cette thèse sont également décrites

dans cette partie. Dans le chapitre 2 intitulé "Etat de l'art de la technologie 3D", nous détaillons les connaissances de base qui permettent de comprendre les étapes de développement de la partie technique et les problèmes associés à la technologie 3D actuelle. Nous examinons d'abord les études dans le système visuel humain, tout en décrivant les facteurs physiologiques et psychologiques pertinents à la sensation et à la visualisation en 3D, dans l'objectif de fournir les informations nécessaires pour comprendre les facteurs humains jouant un rôle dans la visualisation du contenu 3D. Ensuite, nous présentons le schéma d'une chaîne typique de diffusion Télé en 3D, y compris le processus d'acquisition du contenu 3D, la compression, la transmission et l'affichage. Chaque étape sera traitée et présentée à part. Par la suite, les avantages et les limites de plusieurs techniques alternatives de visualisation 3D (comme l'imagerie intégrale, volumétrique et holographique) sont analysées séparément. Nous terminons cette partie avec une étude approfondie du système de l'imagerie intégrale, où nous présentons une analyse de son principe de fonctionnement, et nous formulons les problèmes. Nous présentons également les motivations derrière nos travaux pour l'amélioration du système d'imagerie intégrale. Dans la deuxième partie de cette thèse (c.à.d. dans les chapitres 3, 4 et 5), le principe de fonctionnement et les problèmes dans différentes procédures de la technologie "image intégrale", comme la génération, la compression et la reconstruction sont présentées et étudiées respectivement. Après l'identification des problèmes et des limites dans chaque procédure du système d'imagerie intégrale, plusieurs méthodes sont proposées et évaluées comme solutions à ces problèmes.

Le chapitre 3 intitulé "Acquisition des images intégrales", traite le principe et propose des solutions pour la génération de l'image dans la technologie "imagerie intégrale". Tout d'abord, un résumé de la méthode de génération d'image intégrale est donné, les avantages et les inconvénients de la technique CIIG sont présentés en détail. Ensuite, nous illustrons les méthodes proposées pour résoudre les inconvénients des méthodes utilisées dans CIIG. Par la suite, une nouvelle approche basée CIIG est proposée pour surmonter les limitations des méthodes précédentes. Enfin, nous montrons la capacité de notre proposition à contrôler la profondeur d'enregistrement de la scène 3D, et à élargir la gamme de visualisation 3D pour les téléspectateurs sans augmenter le volume de données source.

Le chapitre 4 est intitulé : "Compression d'une image intégrale". Dans ce chapitre, nous analysons les différentes techniques de compression d'image utilisées pour réduire la consommation excessive de la bande passante lors de la transmission. Tout d'abord, l'importance et le principe de la compression d'image intégrale sont présentées. Ensuite, les méthodes de codage utilisées dans le système actuel d'imagerie intégrale sont présentées et analysées. Par la suite, trois techniques de compression d'image sont proposées et expliquées. Enfin, toutes les approches présentées sont évaluées pour mettre en évidence leurs avantages et leurs inconvénients.

Le chapitre 5 est intitulé : "La reconstruction de l'image intégrale". Ce chapitre traite la conception d'un système de reconstruction d'image intégrale. Tout

d'abord, deux techniques de reconstruction d'image intégrale sont présentées (une basée sur la vision et l'autre basée sur le plan). Ensuite, deux nouvelles méthodes de reconstruction (basées sur plan) sont proposées. L'une est basée sur la projection en groupes parallèles, et l'autre est réalisée par un algorithme d'analyse d'image en profondeur. Enfin, les avantages de la méthode proposée sont mis en évidence par les résultats de nos travaux expérimentaux.

Le chapitre 6, intitulé "Conclusion", résume toute la thèse. Les avantages et inconvénients des solutions proposées sont discutés. Les lignes directrices pour l'extension et les améliorations de nos travaux sont présentées.

Travail Future

Nos travaux de recherche permettent d'améliorer considérablement l'expérience des utilisateurs finaux dans le système d'imagerie intégrale utilisé pour le 3D. Les améliorations notables ont été accomplies dans divers aspects tels que la portée visuelle, la gamme de profondeur de l'image et la résolution d'affichage. Cependant, il y a des questions qui restent à résoudre dans de nombreuses directions. La proposition d'appliquer la technologie d'imagerie intégrale dans le système 3DTV doit encore progresser.

Dans cette thèse, nous avons proposé une méthode pour contrôler l'intervalle de la profondeur enregistrée de la scène 3D, qui est capable de sélectionner une région intéressante dans le sens de la profondeur et d'enregistrer cette région avec une résolution plus élevée. Cependant, cette approche ne peut pas agrandir la longueur totale de la profondeur. Elle améliore la qualité de certaines plages de profondeur en sacrifiant la qualité d'autres plages de profondeur. Par conséquent, dans nos travaux futurs, nous avons l'intention d'explorer une approche d'augmentation de la portée de la profondeur, qui sera en mesure d'étendre la plage d'enregistrement de l'image intégrale sans réduire la qualité d'affichage de l'autre intervalle de profondeur.

Un autre sujet de recherche intéressant est l'application de la technique CIIG pour générer hologramme, qui peut être utilisée pour fournir un alternatif processus d'acquisition de l'hologramme conventionnel.

Toutes les méthodes de compression proposées ont été appliquées sur des scènes en 3D très simple. Des évaluations devront être menées sur des scènes 3D plus complexes pour mettre en valeur le profit de nos contributions. Pour la méthode de compression à base de MVC, un algorithme d'estimation de disparité plus pertinente devra être trouvé afin de réduire l'erreur résiduelle de l'image prédite. Pour la méthode de compression basée sur MVD, nous pourrions envisager un nouvel algorithme de DIBR (image rendu basé sur image profondeur) avec une complexité de calcul inférieure de manière à fournir une approche de codage plus efficace. Pour la méthode de compression basée sur la décomposition en ondelette (Discrete Wavelet Transform - DWT), différentes méthodes de décomposition en ondelettes seront appliquées et testées pour comparer leur efficacité avec la méthode de Haar.

Pour le processus CIIR, outre les contributions présentées pour la reconstruction de l'image et qui sont basées sur le plan, nous prévoyons de proposer une nouvelle méthode pour améliorer l'efficacité de la reconstruction basée sur la vue.

En se basant sur nos travaux et les travaux de recherches actuelles, nous estimons que l'application de l'imagerie intégrale sera adoptée par les industrielles dans les années voir les mois à venir.

Chapter 1

Introduction

In this chapter, we first motivate our studies by providing a brief retrospection on the evolution of three-dimensional (3D) technology. Then, we present the brief of a particular 3DTV system that leverages the integral imaging technology. Thereafter, the objectives and contributions of this thesis are highlighted. Finally, we present the summary of the following chapters and the thesis organization.

1.1 Development of 3DTV

As a necessary technique to provide natural viewing experience, 3D technology is considered as the next major revolution in the development of Television. From rough black-and-white prototype to sophisticated volumetric or holographic instrument, 3D technology has gone through a long and tortuous course of development. In this chapter we try to summarize the main milestones along its evolution.

The appearance of 3D can be estimated to 1838 when Charles Wheatstone discovered the principle of binocular vision [WPM02]. He found that binocular vision is an important depth perception for the viewer to interpret the 3D information. As one looks around, the retina of each eye forms a two-dimensional image of the surrounding. The same objects are projected to two images in different horizontal positions, and the brain processes these images into a 3D visual experience. Based on this principle, he invented the world first 3D viewing device [WPM02]. As shown in Figure 1.1, this instrument produces separated drawing for left and right eyes by using several tilted mirrors; when both images are viewed at the same time, 3D effect is generated.

The first bloom of 3D technology began with the emergence of the anaglyph technique around 1915 [Lue11]. This technique superimposes a red and a green-tinted image upon each other to represent the 3D object, and the viewers use special anaglyph 3D glasses with corresponding to color filters so as to deliver the left and right images to each eye. In a red-green anaglyph, the eye viewing

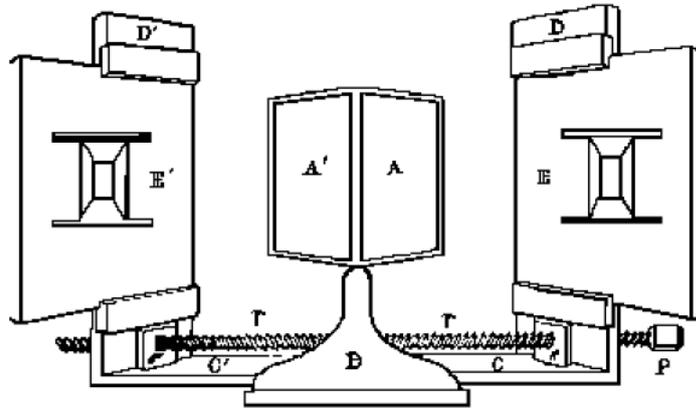


Figure 1.1: The Wheatstone stereoscope, mirrors A' and A are set at an angle in order to reflect the left and right eye drawings (E' and E) towards the viewer's eyes

through the red filter sees red within the anaglyph as "white", and the green within the anaglyph as "black", as illustrated by Figure 1.2. The eye viewing through the green filter perceives the opposite. The brain merges together the red and green channeled images into a perception of a three dimensional scene.

The so-called "3-D movie craze" in the years 1952 through 1955 started with the application of linear polarizing projection and glasses [Kaw02]. Similar to anaglyph technique, two images are superimposed on the same screen and projected by different polarizing filters. A pair of 3D glasses with corresponding polarizing filters is used to separate left and right images. Figure 1.3 shows an example of linear polarization 3D display. As illustrated, each filter passes only the light rays similarly polarized and blocks the light differently polarized to distribute different image to each eye. As it is able to produce the high quality 3D film with rather low cost the polarized technique opens a new era of 3D technology.

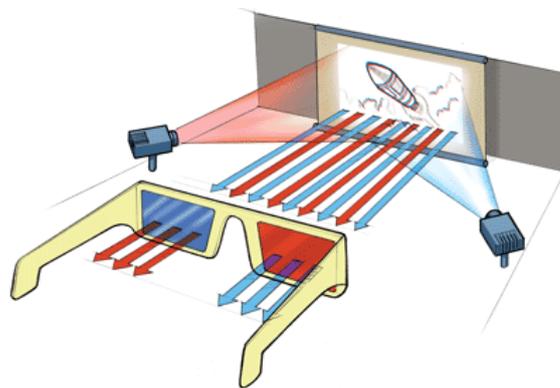


Figure 1.2: Anaglyph 3D technique

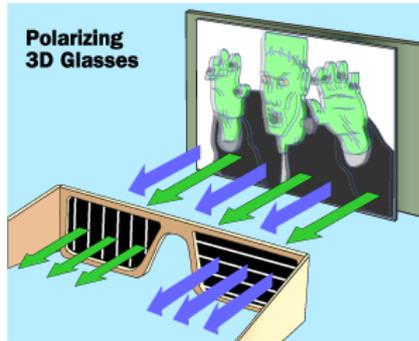


Figure 1.3: Linear polarization technique

From 1985, along with the release of IMAX 3D technique, 3D movies have experienced a huge success. IMAX, a projection standard and film format, created 3D technology that was mathematically accurate, so that it had the ability to eliminate fatigue of the eye, a side-effect that was experienced by audiences when viewing movies using other 3D technologies. Actually, it emphasized mathematical correctness of the 3D rendition and thus significantly diminished the imprecise approximate geometries of previous 3D experiences [Umb06]. Furthermore, IMAX 3D projectors deliver 3D images of unsurpassed brightness and clarity, better than any other 3D technology available at that time. Two methods are exploited in IMAX to create the 3D illusion for the viewers. One is the previous introduced polarization technique. The other is eclipse method, which applies a shutter to block light from each appropriate eye when the converse eye's image is projected on the screen [DN10]. As explained by Figure 1.4, two projectors display each frame alternatively at a high rate, and the glasses block or transmit light to each eye in synchronization with the projector, so that only the corresponding image is sent to each eye.

2009 and 2010 reached the pinnacle for 3D in movies. With blockbusters such as *Monsters v Aliens*, *Up*, and *Avatar* being served to audiences, 3D has set itself up for the long haul. Meanwhile, the bloom of 3D technology has extended from film industry to the television industry. Several TV broadcasters announced that they would be showing programs and shows in 3D, which yields transmission and compression issues. Some standards groups, such as MPEG and ISO, are seeking the norms of manufacturers and system interoperability for 3DTV. 3DTV is now becoming a household device gradually. It can be foreseen that 3DTV will turn to be a ubiquitous and irreplaceable entertainment facility at home. It suffices to say that 3DTV will widely replace current 2DTV as long as it is feasible and affordable.

1.2 Motivations and Objectives

Because of its great attraction of more natural perception and enhanced psychological impact than traditional 2D technology, 3D cinema has got a huge

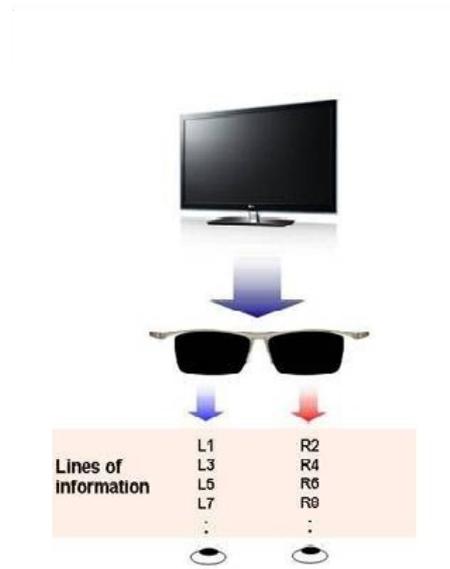


Figure 1.4: Eclipse 3D technique

success in the consumer market during the past few years. At the same time, 3DTV has also attracted a considerable attention. However, despite of its bright prospective, 3DTV still contains many difficulties to achieve both of the prominent visual quality and comfortable viewing experience. A brief summary of its defects are listed as follows:

- Impossibility to provide high quality 3D image without special glasses
- Mismatch between convergence point and the eye's accommodation point leads to visual discomfort in prolonged viewing
- Lack of full parallax which enables the observer to see different images corresponding to different positions horizontally and vertically

A 3DTV technology solving the above problems would enable observers to see a 3D image as though it were a real object. To this end, many other 3D display technologies such as volumetric, integral imaging and holography have been proposed and designed recently. Among these new technologies, integral imaging is often considered as the most attractive one due to its compromise in feasible complex and produced 3D effect.

Integral imaging was introduced by Gabriel Lippmann in 1908 [PHL09]. It uses an array of micro-lenses to record the 3D scene from many slightly different horizontal and vertical locations. When the resulting images are rectified and viewed through a similar array of lenses, a single integrated image, composed of small portions of every image, is seen by each eye. The observed part of each micro-lens is determined by the position of the viewer. As shown in Figure 1.5, the flow chart of a typical integral imaging system is almost the same as for traditional 2DTV system.

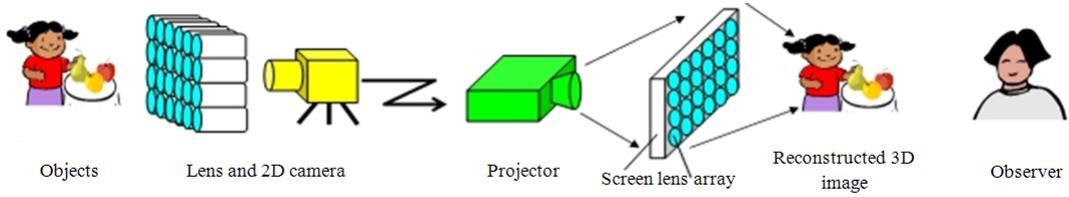


Figure 1.5: Flow chart of an integral imaging system (NHK image)

Thanks to the large number of micro-lenses, integral imaging is able to provide full and movement parallax to spectators without having to wear goggles, so as to alleviate the visual discomfort problem in traditional 3DTV systems. However, some severe drawbacks such as limited viewing angle, image depth and large volume of source data still limit the practical application of integral imaging system.

The objective of this thesis is to improve the performance of integral imaging systems. Our work covers the major aspects of integral imaging technology (1) computational integral image generation; (2) integral image compression and (3) computational integral image reconstruction.

1.3 Contributions

This thesis focuses on the challenging issues that prevent current integral imaging systems to display high quality 3DTV. We provide the problems formulation and corresponding solutions that improve the viewer's 3D experience in integral imaging systems.

1.3.1 A new method for computational integral imaging generation

The first part focuses on integral image generation. The problem of limited depth range and viewing range is addressed and solved in this part. The proposed method is composed of two parts. Firstly, a special multi-camera setup is applied to control the recorded depth range. Then, the curve lens array used in optical integral imaging (OII) system is simulated in computational approach so as to enlarge the reconstructed 3D image viewing range. The results of our experiments demonstrate that the proposed CIIG method is able to control the re-encoded depth range of the 3D scene, and enhance the viewing range of the traditional CIIG methods without increasing the source image size.

1.3.2 New techniques for Integral Image compression

Second part discusses the compression issues of integral image. Three different compression approaches are proposed and evaluated.

The first one is based on multi-view video coding (MVC) technique. This method is applied to sub-images (SI), which is an alternative form of 2D images transformed from original elemental images. Each SI represents the 3D scene from parallel viewing directions and contains superior compression capabilities than original captured EI. For this reason, we consider arranging the group of SIs as the format of multi-view video (MVV) and then encode the generated MVV by MVC standard. This approach improves the compression efficiency comparing to the traditional MPEG-4/AVC method.

The second compression method is inspired by Multi-View Video plus Depth (MVD) representation. It is also applied to sub-images. The principle of this method is to use the 3D information existing in depth maps to predict the SI for other direction and then remove the inter-view correlation between different sub-images. Thanks to the 3D geometrical information in depth map, this method outperforms the MVC-based scheme mentioned before.

Besides the MVC and MVD based compression approaches, we also considered applying the discrete wavelet transform (DWT) for integral image compression. As opposed to DCT, DWT is able to transform the original 2D information from pixel domain into pixel-frequency domain, which still preserves the spatial domain character of the source data. In the proposed approach, the input IIs are filtered multiple times by some given high pass and low pass filters until the correlation within one EI and between adjacent EIs are completely eliminated. According to the simulation results, DWT works better than MVC-based method in low-debit situation.

1.4 Two novel method for computational integral imaging reconstruction

Finally, we investigate the reconstruction process and developed two novel plane-based computational integral imaging reconstruction (CIIR) techniques to improve the performance and reduce the complexity of CIIR process.

The first method leverages the Parallel-Group Projection (PGP) technique. Instead of projecting each point of integral image to the reconstructed plane pixel by pixel, the proposed method reconstructs the 3D image by mapping a series of sub-images onto the reconstructed plane successively. Because of the parallelization between the light rays in each SI, this approach can simplify the pixel mapping complexity in the reconstruction process. Moreover, it supports convenient conversion between 2D and 3D perspective by switching between full SI and single SI projection.

The second design aims at increasing the resolution of EIs so as to improve the quality of reconstructed 3D image. In this method, the high definition EI is constituted by a large number of synthesized SIs that are generated by depth image rendering algorithms. This approach provides significant improvements for integral imaging systems. It improves the reconstructed 3D image quality

and offers scalability for EI resolutions by simply increasing the number of synthesized SIs. Additionally, it requires only one color image and one depth map rather than transmitting multiple color images to reproduce the 3D scene, which could make the transmission in integral imaging system more efficient.

1.5 Thesis Organization

This thesis consists of six chapters divided into three parts.

The first part, Chapter 1 and Chapter 2, aims to provide an overview and necessary background information of the thesis. Chapter 1, i.e. the current chapter, provides a broad description on the context based on which the research is carried out. We first identify several major evolutionary steps on the 3D technology. Second, we give a brief summary for the problems of current 3DTV systems, especially regarding the challenges and opportunities for integral imaging. Major contributions of this thesis are also described in the last section.

In chapter 2, we detail some background knowledge that help us to understand how the main technical part is developed and what are the problems in current 3D technology. We first review the studies in human visual system, describing the physiological and psychological factors relevant to depth sensation, thus providing background information for understanding the function human factors play in 3D content viewing. Then, we present the scheme of a typical (auto)-stereoscopic 3DTV broadcast chain, including the process of 3D content acquisition, compression, transmission and display. A short review of each step is given. Thereafter, the advantages and limitations of several alternative 3D visualization techniques such as integral imaging, volumetric and holography are analyzed individually. Finally, we conduct a substantial study of the current integral imaging systems, investigate the principle, develop the models and formulate the problems. The challenges for integral imaging systems that precisely motivate our work are highlighted in the end.

In the second part of this thesis, chapter 3, 4 and 5, the working principle, diagram, and problems of different procedures, i.e. integral image generation, compression and reconstruction are presented and studied respectively. Following the analysis of the limitations in each procedure of the integral imaging systems, several methods are proposed and evaluated.

Chapter 3, “Acquisition of integral image”, studies the principle and proposes the solutions for computational integral image generation. A short review of the optical integral image generation method and computational integral image generation method is given, and the advantages of CIIG technique are stressed. Then, we illustrate the previous CIIG methods, explain their structures and address the identified problems. Thereafter, a new CIIG approach is proposed to overcome the limitations of previous methods. Finally, we demonstrate that our proposal is able to control the recording depth range of the 3D scene, and

enlarge the 3D viewing range for the viewers without increasing the volume of source data.

Chapter 4, "Compression of integral image", this chapter conducts substantial studies on various integral image compression techniques in order to save valuable transmission bandwidth. First, the importance and principle of integral image compression is illustrated. Then, the encoding methods applied in current integral imaging system are presented and analyzed. Thereafter, three integral image compression techniques are proposed and explained. Finally, all the presented approaches are evaluated to highlight their benefits.

Chapter 5, "Reconstruction of integral image", this chapter addresses the design of a computational integral image reconstruction (CIIR) system. First, two basic types of computational integral image reconstruction methods are presented, known as view-based and plane based CIIR. Then, two novel plane-based reconstruction methods are proposed. One is based on parallel-group projection, and the other is realized by using depth image rendering algorithm. Finally, the advantages of the proposed method are emphasized by the results of our experiments.

Chapter 6, "Conclusion" summarizes the whole thesis. Pros and cons of the problem formulation, analysis, solutions and the verifications are discussed. Possible directions on which this work may be extended are given.

Chapter 2

State of art of 3D technology

This chapter presents the state-of-art of 3D technologies, with special emphasis on their applications, characteristics and problems. We start by the mechanism of the human viewing system and its role in human depth perception. Then, we provide an overview of an end-to-end (auto)-stereoscopic 3DTV broadcast chain. A thorough investigation is given about each component of the entire 3DTV model (i.e., generation, distribution, display) together with its challenges and design issues regarding the various techniques. Our analysis reveals that integral imaging system is a prospective technology for future 3DTV since it allows the production of satisfactory 3D effect with relatively low costs.

2.1 Introduction

Creating the illusion of the real world in its absence is the final target of visualization technologies. In the history, artists, engineers and scientists have made a great deal of effort to achieve this goal. Along with the creation of photography technology, invention of color and high definition television, current 2DTV provides a very high resolution near the limitation of human visual system [Pas91]. Therefore, the requirement of real natural visual effect with the third dimension becomes the most anticipated feature for next generation TV. Due to the high customer demand of 3D viewing experience, 3DTV is considered as the complement of HDTV in the near future, and will generate a huge potential market in the field of digital TV.

A successful 3DTV system should meet the following conditions [MIS04] [Red02]. First, it should be compatible to traditional 2DTV systems, which ensures a smooth transition from current broadcast systems to 3D broadcast systems. Next, it should be able to provide high quality 3D images, at least comparable to traditional 2DTV. Flexibility is also very important to make it possible to switch between 2DTV and 3DTV, and adapt the content to a wide range of 2D and 3D displays. Finally the system should not introduce any additional user-wearing devices, yield eye strain or other uncomfortable effect.

Compared to the big success of 3D cinema, 3DTV has not kept its promises of bringing high quality content without viewing strain 3D video to our living rooms. Some unsolved problems such as eye strain, glass wearing and unnatural 3D effect prevent the development of 3DTV to the massive consumer market [OSO⁺06].

In order to solve the problems raised by in current 3DTV system, it is important to first understand the depth perception principle underlying human visual system. The various depth cues and their meanings are reviewed in Section 2.2.

Once the depth interpretation principle explained, section 2.3 gives a thorough investigation of the current 3D broadcast chain. Most of the current 3DTV systems are constructed by use of stereoscopic technology, which provides the 3D effect by delivering two different views of the 3D scene to the respective eyes of the viewer. Similar to 2DTV, this 3DTV system is also composed by four basic blocks from content production, coding, transmission to display. Each component of them can be realized by diverse approaches, which are discussed individually in this section.

To overcome the limitations in the current binocular system, various alternative visualizing techniques are also presented and explored, e.g. integral imaging display, volumetric display and holographic display. Their performance and characters are illustrated and compared in section ???. According to the comparison, considering both complexity and quality, integral imaging system is selected as an appropriate technology for the future 3DTV.

2.2 Depth perception cues in Human Viewing System

The research related to human perception plays an important role in the 3DTV development. Relevant to perceptual and usability issues, the comprehension of human factors helps optimize the 3DTV technologies from an end-user perspective.

As illustrated in Figure 2.1, various depth cues help the human visual system to interpret the depth information [WPM02] [Pas91]. Some depth cues require both eyes to see the scene, known as binocular cues, such as convergence and binocular parallax. Others hold also when watching with only one eye; they are denoted as monocular cues, such as accommodation, occlusion, linear and aerial perspective, relative size, relative density, and motion parallax. That is why people with one eye vision could also evaluate depth in real world. Their relative potency at various distances is thoroughly discussed by E.Cutting and M.Vishton in 1995.

The meaning and effect of each depth cue are listed here[RHF10] [MMIW09]:

1. Accommodation and convergence: accommodation represents the tension of the muscle that impacts the focal length of our eyes, to make the eye focus on object and keep the retinal image sharp. Convergence represents
-

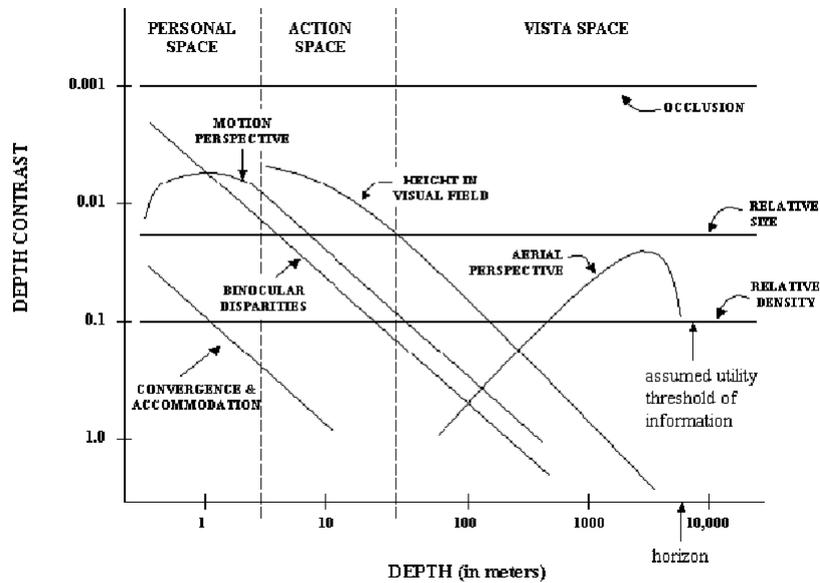


Figure 2.1: Depth cues in different depth distance (E.Cutting and M.Vishton 1995)

the difference of optical axes of two eyes to watching an object close. Both of them are only effective at short viewing distance.

2. Binocular parallax: the difference between two images perceived by left and right eyes. The HVS is very sensitive to this depth cue, and it is the most important depth cue for medium viewing distance.
3. Monocular movement parallax: the relative movements of objects caused by observer movement, to make us extract depth information in two successive images sensed by one eye.
4. Retinal image size: the difference between the real size of the object and its sensed size. The HVS can extract the depth information of the object by comparing this difference.
5. Linear perspective: the fact that two straight lines meet each other in depth direction.
6. Texture gradient: the fact that one can see more details of the same texture at shorter range.
7. Overlapping: when one object blocks another in our sight, the nearer objects can occlude the further one in the perceived image.
8. Aerial perspective: the dust and moisture in the air make the further object bluer and decreased in contrast comparing to foreground object.

All of the available cues listed above are used automatically by the human brain when watching a real scene. The objective of 3D display is to reproduce these depth cues for the end user. Diverse 3D display technologies are created

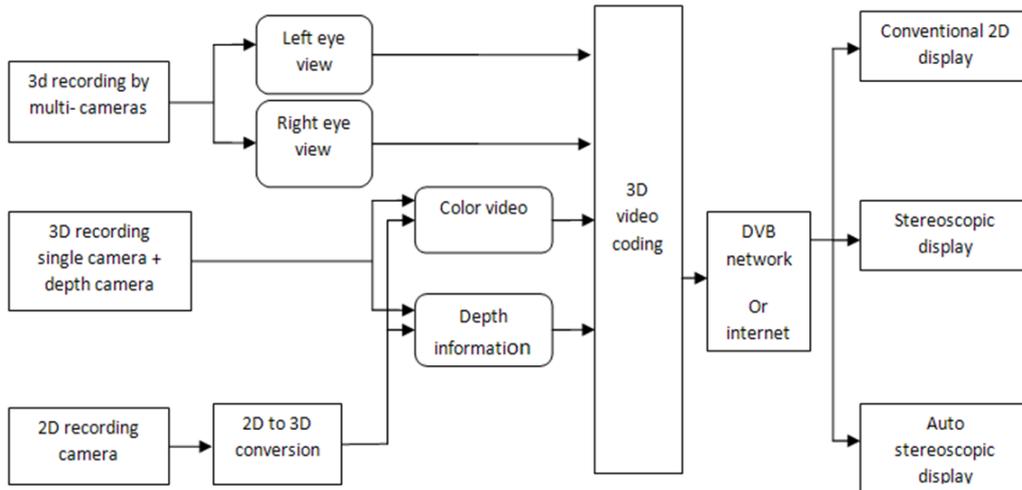


Figure 2.2: Structure of Complete 3DTV Processing Chain

with the use of one or several depth cues, such as (auto)-stereoscopic displays, head mounted displays, volumetric displays and holographic ones. Among all these different display devices, (auto)-stereoscopic ones are the most widely employed for 3DTV due to their simplicity of implementation and compatibility with existing 2DTV system.

2.3 (Auto)-stereoscopic 3DTV Systems

As reviewed in the precedent section, for an (auto)-stereoscopic 3DTV, the human visual system interprets depth from images by using the binocular parallax cues. Based on this principle, (auto)-stereoscopic approaches distribute two different views through a flat screen for left and right eyes respectively [SS99] [DMY88]. This way, the 3D effect can be perceived by the viewer together with binocular parallax. Thanks to its relatively simple design, (auto)-stereoscopic systems have been favored by 3D service providers and content producers ever since its emersion. Figure 2.2 shows an example of an (auto)-stereoscopic 3DTV system. A precise coupling of various technical components is required to build such a system. In the same way as conventional 2DTV system, it is composed of four main blocks including content creation, video compression, stream transmission and display. In this section, we give a detailed description of the (auto)-stereoscopic 3DTV system and discuss the alternative methods applied in each technical component.

2.3.1 (Auto)-Stereoscopic 3D content creation

Content acquisition is the beginning of a 3DTV system. Before the investigation of content generation technology, we should at first understand how the 3D

scene is represented in current (auto)-stereoscopic system. Two basic types of formats are widely used. The first representation format is standardized in MVC. It is developed by MPEG and VCEG, which attempt to represent the 3D scene through several different views. The second one is the texture plus depth representation, that depict the 3D scene by several 2D images and their corresponding depth maps. This kind of 3D format is standardized in MPEG-c part 3 [Red02] and multiview video plus depth (3DV) [KMW11].

Current 3D content generation techniques can be classified into two basic types: camera shooting and digital rendering. Furthermore, two basic camera shooting techniques are applied to record different 3D formats. The first one is using several (at least two) traditional 2D cameras to record the views for left and right eye separately. The other one consists in applying a “2D plus Z” camera to generate a monoscopic color video and the associated per-pixel depth information simultaneously to represent the 3D scene.

Besides camera shooting, computational creation is another important method for 3D content generation. The existing digital creation approaches can be classified into two main categories. The first one is to model the mathematical representation of the 3D scene via some specialized software. Arbitrary perspective of the generated 3D model can be displayed through a 3D rendering process. This method is widely used in 3D animation production. The other digitally 3D content generation approach is 2D to 3D conversion approach. This method is realized by firstly computing a depth map from the original video and then synthesis the virtual views through 3D warping algorithm. An overview of the existing 3D content creation approaches in the following:

Stereoscopic two camera system

Today, most of the 3D video materials used for broadcast are generated by a dual monoscopic cameras configuration. Left and right eye views are recorded from slightly different perspectives to simulate the binocular parallax of human visual system. To avoid potentially conflicting information between the two views and eliminate the visual discomfort, the stereoscopic shooting parameters have to be carefully designed and implemented, and the two cameras need to be calibrated in terms of position, luminance and space transformation in order to confirm that the disparity they get is exactly the same as the viewer would get with a real scene.

Two types of camera configuration are possible for 3D content acquisition [FP08]. One is the parallel camera configuration, where the optical axes of both cameras are parallel to each other. The other one is the converging camera configuration, where the optical axes of the two cameras intersect.

In parallel cameras model (see figure 2.3), the lenses (cameras) are parallel and the optical axes of the two cameras meet at infinity. In the simplest situation, without any post shifting of the images to correct the zero parallax distance (ZPD), objects at infinity will be cast at the surface of the display and all

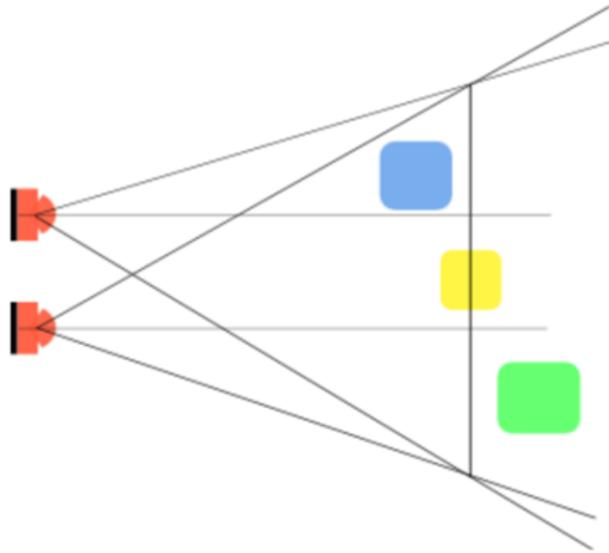


Figure 2.3: Parallel cameras model

other images will be cast in front of the display. Nothing will appear behind the screen surface (no positive screen parallax). The zero parallax distance (sometimes called the convergence distance) can be changed by shifting the two imaging sensors (e.g. CCD) behind the lenses. In this model, the ZPD is usually set mid way across the scene, meaning that some objects will appear behind the screen and some objects will appear in front of it.

In toed-in camera setup, two cameras (and lenses) are rotated inwards. Their optical axes intersect at a point usually mid-way through the scene. So that the objects will appear behind the screen and some objects will appear in front of the screen, as shown in Figure 2.4. For the converging dual camera acquisition system, the projection of a real world point is more complex than in the parallel set-up, it also contains vertical disparity component between the two viewing images, which causes the keystone distortion and depth plane curvature [WPM02]. For this reason, parallel camera configuration is used more widely in 3D content generation.

Depth camera

Besides stereoscopic image representation, 3D content can also be represented by texture image and depth map. This kind of 3D content is created by simultaneously capturing one traditional RGB image and a depth image, which records the depth information of each pixel in the color image.

In order to get the required depth image, several new types of camera have been introduced. ZcamTM is one of them [CKH11], whose idea comes from light reflection property. It is built by integrating a high-speed emitted infrared light

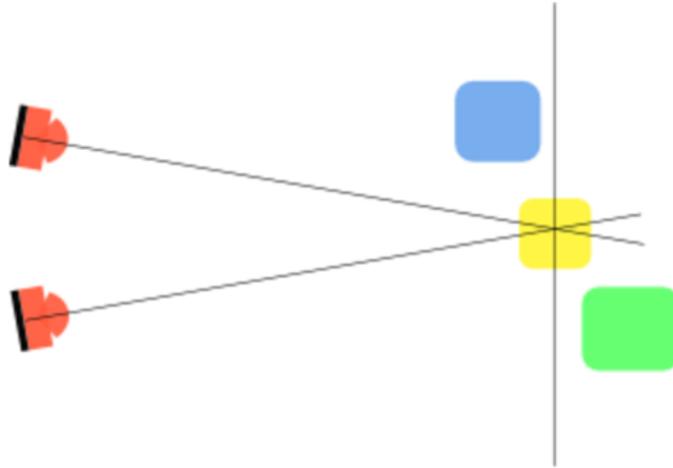
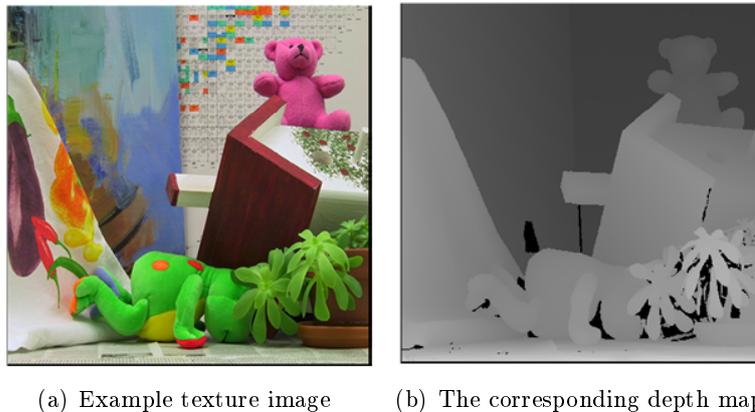


Figure 2.4: Toed-in camera setup

source into a traditional video capture camera. To record depth information, the Zcam™ camera pulses a light wall to the object; when the light hits the object and is reflected back to the camera, one can measure the time of flight of the emitted light and derive the distance from the camera to the object; this distance represents the depth of the scene which can be used to construct the depth map. The output of this camera includes two video streams: the traditional color image and corresponding depth image with the same spatial resolution. An example of this video plus depth representation format is given by Figure 2.5.

Texture image plus depth map representation separates the process of capture and display, and offers remarkable flexibility and adaptability at the display side. But it will miss some information of the scene since the two views presented to the viewers are reconstructed from a single recorded view and a corresponding



(a) Example texture image

(b) The corresponding depth map

Figure 2.5: Example of this video plus depth representation format

depth map. The missing information of occlusion regions in original texture image and depth map causes some gaps in reconstructed stereoscopic images. Recently, various “hole-filling” algorithms have been developed to solve this problem [FW10].

3D modeling

3D modeling is the process of creating a 3D model to represent a moving or still scene. A 3D model is represented by a set of points in 3D space, which are connected by various geometric data such as lines, and curved surfaces. Three popular methods are applied to create 3D models:

The first method is polygonal modeling; this method creates a 3D model by connecting line segments through points in 3D space. These points in space are also called vertices. Polygonal models are very flexible and can be rendered by a computer very quickly. However, it cannot create an exact curved surface of a 3D object.

The second method is known as primitive modeling. It creates the complex models from some geometric primitives such as cylinders, cones, cubes and balls. This is the simplest way of 3D modeling as the forms are mathematically defined and precise.

Another one is parametric surfaces, such as NURBS (Non-uniform rational B-spline) modeling. This representation is used in some popular software such as Maya. Unlike polygonal modeling techniques that can only approximate curved surfaces using numerous polygons, this method allows the user to create smooth curved surfaces directly.

With a given 3D model, arbitrary perspective of the scene can be generated through a 3D rendering process.

2D-3D conversion

2D to 3D conversion is another relevant technique for generating 3D content. The goal of 2D to 3D conversion is to generate the stereo video via a 2D video, where the generated stereo video should have similar quality to the original stereo video, which is shot in stereo camera. Two kinds of conversion approaches are defined: first one is low-quality automatic conversion method which outputs the stereo video from one 2D input without any user participation [HFFD02], the other one is high-quality semi-automatic conversion method with human-computer interactive operations [PRA11].

Automatic conversion is widely used in low quality 3DTV because of its simplicity for implementation and its real time processing. Technically the automatic 2D to 3D conversion method is carried out in two steps: 1) depth estimation and 2) depth image based rendering (DIBR).

Depth map generation is the critical process of automatic 2D to 3D conversion. Three criteria can be applied to estimate the depth information of a given

2D video. 1) Blur amount of the object, 2) Position of vanishing point: the vanishing point is always the farthest point of the whole scene. 3) Motion parallax: this is an important human depth cue, which have been introduced in the first section of this chapter. It represents the fact that objects at different depths always have different apparent speed, e.g. near objects move faster than far objects. Therefore, depth value can be estimated by relative motion.

Then, the resulting depth map and original 2D video are used to generate the stereo video through a depth-image-based-rendering algorithm [Feh04]. DIBR is a technique for view synthesis that uses given texture images and their associated per-pixel depth information to generate a new 2D image from another viewing direction. Conceptually, DIBR method is performed in two steps: 1) 3D image warping: using known camera parameters and depth map to back-project pixel samples from a given texture images to the proper 3D locations and re-project them onto the new synthesized image space. 2) reconstruction and re-sampling: determining pixel sample values in the synthesized image.

Automatic conversion methods are very simple to implement in hardware and supply a real-time application. Most of the 3D ready TVs are embedded with real-time automatic 2D-to-3D converters.

In semi-automatic 2D to 3D conversion, the human intervention is required at first for annotating 3D information (e.g. depth map) to certain frames (key frames). The depth maps of other frames are then generated automatically. Like automatic 2D to 3D techniques, the resulting depth map and original 2D video are used to generate the stereo video through a depth-image-based-rendering algorithm. Usually, the original 2D image is applied as the center view image. Two virtual views for left and right side are generated from this center image and its depth map. A painting process to fulfill the uncovered holes of stereo images is executed in the last step.

It is obvious that there is a trade-off between algorithm complexity and 3D conversion effect. More natural 3D effect require more time spent for annotating depth value to key frames. Many efforts have been deployed to find more efficient conversion methods facilitating the user operation and generating high quality stereo videos.

2.3.2 (Auto)-Stereoscopic content compression

After the acquisition of 3D content, a compression algorithm is applied to encode the source data into a much smaller volume, which allows the video sequence to be transformed into a stream with less bit-rate and to be stored with less space.

Multiview Video Coding (MVC)

In a (auto)-stereoscopic configuration, the video stream for 3D content can be regarded as two or more sequences of 2D image. It is clear that the image information in distinct perspective views is highly correlated. Therefore, high coding

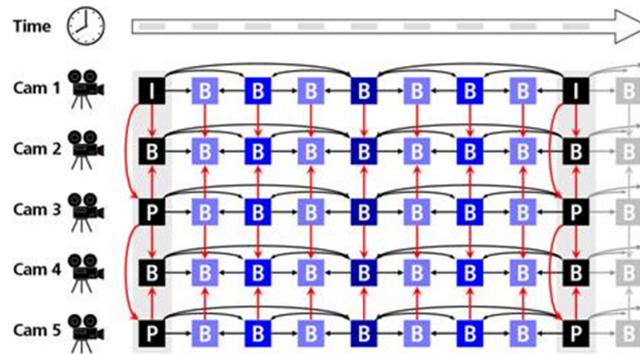


Figure 2.6: Scheme of MVC

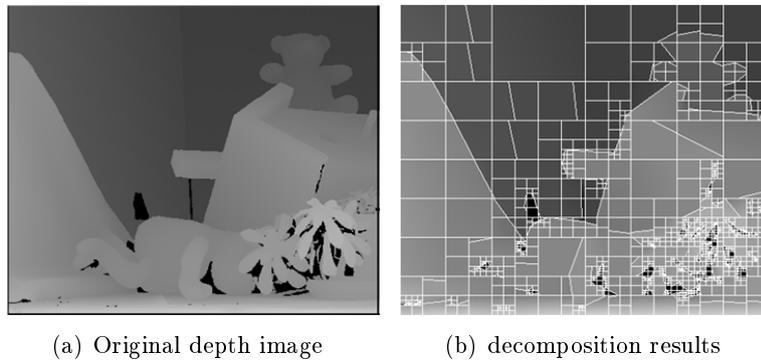


Figure 2.7: Original depth image and its decomposition results of piecewise-linear functions

gain can be reached by removing the dependencies between views. Multiview Video Coding (MVC) approach has been proposed to achieve this objective.

MVC intends to remove the redundancy of multi-view video sequence in two ways. In one hand, the correlation among successive 2D image from one view is eliminated by a traditional H264/AVC approach. On the other hand, the additional correlation between different views is exploited by disparity estimation [CP12]. In a similar way to motion estimation in regular 2D video compression, disparity estimation applies the reference view image to predict the target view image, so as to represent it by the computed disparity vectors and residual error. Figure 2.6 shows an example of MVC process. In this graph the hierarchical B prediction in temporal domain is combined with the inter-view disparity estimation; one then applies the inter-view prediction for all the pictures in the sequence.

Depth image compression

As described previously, in the depth image based 3D content representation the 3D video is represented by a monoscopic color video and an associated

depth map. In this case, depth map can be regarded as a particular color image with the YUV format 4:0:0. Therefore, the sequence of depth image can be compressed by traditional 2D coding method too. In this way, we ignore the different property between depth image and texture image. As a result, some disturbing artifacts distortions such as ringing or blurring effect will arise in the decoded video.

In fact, since the depth map represents the scene geometry and thereby covers the 3-D perception of the scene, it is usually made of smooth regions divided by sharp edges. Here the sharp edges representing object boundaries should be preserved during the compression process [RBBHS01]. Therefore a new algorithm using piecewise-linear functions is proposed. It decomposes the depth map into platelets of variable size with various depth values. In order to preserve sharp object boundaries, the support area of each platelet is adapted to the object boundary. Thanks to the careful handling of edges, this algorithm is able to achieve good compression performances [Md06]. Figure 2.7 shows an original depth image and its decomposition results of piecewise-linear functions. Besides the specific encoding algorithm, the bit-rate allocation between texture and depth images is another challenge for multi-view video plus depth (MVD) data compression, as discussed in [FS11] and [EB13] the bit-rate allocation strategy should be optimized based on the image quality of synthesized virtual view. In [CG11] and [BE12], the author proposed an asymmetric bit-rate allocation method based on a view synthesis distortion model.

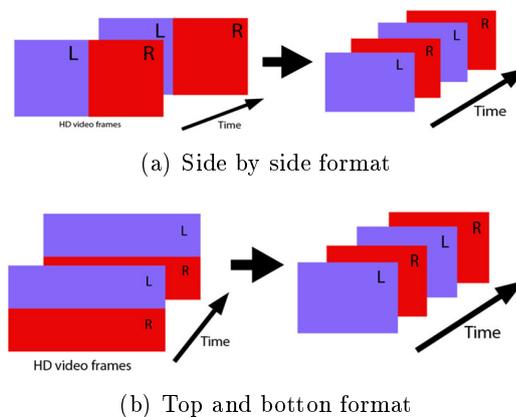
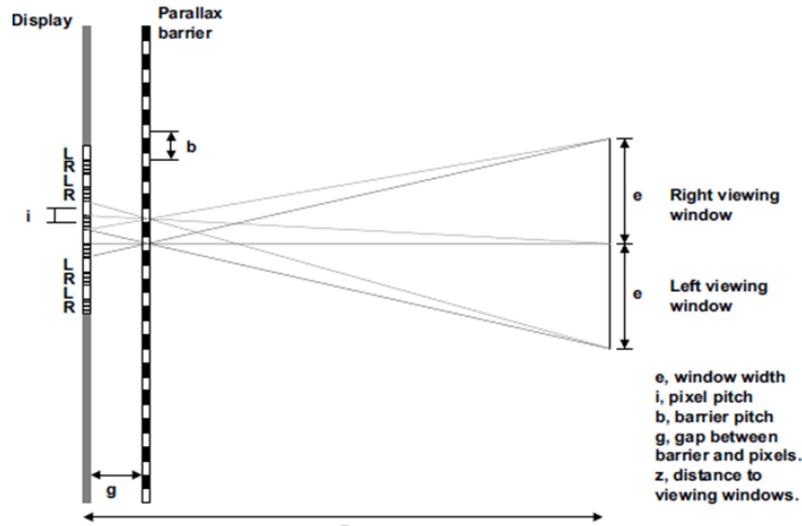
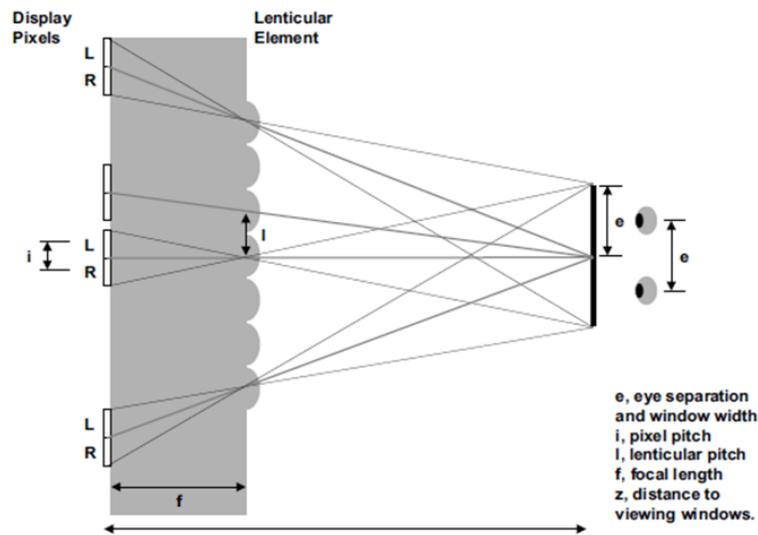


Figure 2.8: Frame compatible transmission format



(a) Barrier adopted



(b) Lenticular adopted

Figure 2.9: The principle of auto-stereoscopic display

2.3.3 (Auto)-Stereoscopic 3D content transmission

Transmission is the process of sending and receiving data over a point-to-point or point-to-multipoint network, either wired or wireless. The transmission should be fast enough to provide comfortable visualization at the client side, without noticeable delays.

With the current stereoscopic video format, two views should be transmitted in conventional 2D broadcast band. The amount of data captured by the stereo-camera system is however twice larger than traditional single-view videos. In

order to attain this higher demand of transmission capacity, new standards are under definition to set the techniques and procedures for sending 3D video signals. The Digital Video Broadcast (DVB)-3DTV is such a standard that came out at the end of 2012 [YHLJ11][ATGA08].

Currently, stereoscopic videos are transmitted in frame compatible format, which represents the stereo view signal through a single frame or sequence of frames [MP04]. Usually, the two views are spatially sub-sampled and interleaved into a single sequence of frames that can be delivered by traditional 2D broadcast technologies. Two types of sub-sampling and interleaving methods are universally used in 3D video transmission. The first one is the left-right multiplexing approach, which is realized by simply interleaving the left and right views as two separate parts in horizontal axis, as shown in Figure 2.8(a). Another multiplexing method illustrated in Figure 2.8(b) is top-bottom mode, in which left and right views are deployed one over another. In both methods, the frame resolution of each view is reduced to half of the original video.

2.3.4 (Auto)-Stereoscopic 3D content display

When it comes to the viewing conditions, 3D display can be divided into two types, either they are stereoscopic or auto-stereoscopic [SS99].

Stereoscopic display

In stereoscopic displays, spectators perceive two different images for left and right eyes by wearing a pair of specific goggles. Different kinds of glasses are applied to assign the correct view to each eye.

Anaglyph glasses separate the image according to the wavelength. The color filters are applied to distribute the correct view to each eye. LCD-shutter glasses separate the image in function of time-division. With the shutter method, the light for one eye is blocked when the image of the opposite eye is projected on the screen. The alternate display of right and left images produces the binocular parallax for the viewer. Polarized glasses separate the image in function of light polarization direction. The glasses use polarizing filters to transmit the light with similar polarization and block the rest.

Auto-stereoscopic display

Auto-stereoscopic displays use the conventional stereo principle to create depth illusion without having to rely on any extra device. The binocular parallax for the viewer is realized by generating a fixed viewing zone for each eye with the utilization of parallax-barrier or lenticular array. The parallax-barrier is placed in front of a classical high resolution screen. It allows each eye to see a different set of pixels, thus creating a sensation of depth through the parallax effect. The typical setup of auto-stereoscopic display adopting lenticular and parallax-barrier are illustrated in Figure 2.9 [Lue11].

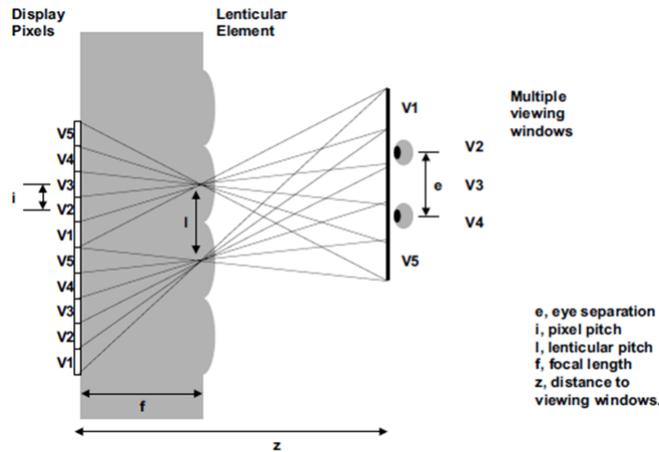


Figure 2.10: The principle of a multi-view auto-stereoscopic display

Auto-stereoscopic display emancipates viewer from the restriction of wearing 3D glasses, but it also presents drawbacks. First, it suffers from a reduced horizontal resolution. An auto-stereoscopic reduces the resolution of native display, since each eye only sees every other pixel. The second important disadvantage is the size of the defined stereo zone; if the viewer moves outside of the area, the image becomes inverted: hills become valleys, and vice versa. Indeed the viewer must stand within a very narrow angle that is nearly perpendicular to the screen to benefit the best 3D experience. What else, the perceived brightness of the screen is also reduced by the barrier.

Multi-view display

Multi-view display is based on the same principle as binocular display. As shown in Figure 2.10, in multi-view display the 3D illusion is also created through the different images for left and right eyes, e.g. binocular parallax. However, instead of distributing only two views, the multi-view displays create a discrete set of perspective views and deliver them across the viewing field [Lue11] [JTLC92]. This way, viewers can observe the motion parallax of the scene when they move in the viewing zone. Furthermore, multi-view displays are capable of providing different perspectives to several viewers [HM93]. However, it still suffers from discontinuous motion parallax and reduced image resolution.

2.3.5 Constrains of (Auto)-Stereoscopic 3DTV system

According to the long-term studies and experiences, constrains of current (Auto)-Stereoscopic 3DTV system could be considered in the following aspects [TIY95]:

1. The visual discomfort for long time watching comes from the mismatch between convergence and accommodation operation. As we mentioned be-

fore, the creation of depth sensation from pair of planar image is just a method to simulate the human 3D perception procedure, it is not able to offer the identical effect as our natural scene. The critical differences between the depth perception by dual 2D images and the depth perception in real world could be found in the inconsistency between the accommodation and convergence. For constructing the depth perception by the planar stereo image pairs, the convergence of the eyes need to verge off the stereoscopic image plane to fixate points in depth while their accommodation must always keep the image plane itself in focus [BAB95]. This requirement breaks the normal link between convergence and accommodation, makes the viewer feel uncomfortable for looking at the stereoscopic image with large depth value.

2. Restricted visual area to perceive natural 3D effect in auto-stereoscopic 3DTV.
3. Necessity of wearing special glasses in stereoscopic 3DTV.

2.4 Alternative 3D display technologies

In order to solve the problems in (Auto)-Stereoscopic 3DTV, several alternative 3D display technique are presented recently.

2.4.1 Integral imaging display

Integral imaging was first invented by G. Lippmann in 1908. Due to its extraordinary advantages in terms of free-viewpoint visualization, it is regarded as a promising technique for future 3DTV and has become a vigorous research topic for the past two decades. Better than conventional stereoscopic technologies, it is able to provide high quality 3D images with full parallax and continuous viewing angle without any need of additional devices such as dedicated glasses.

Different from traditional binocular vision systems that simulate the human 3D perception procedure by transmitting correspondence images into left and right eyes separately, the integral imaging system allows the generation of true 3D images by reproducing the light rays of the 3D scene with different colors and intensities in different direction [VA08] [PHL09]. For this reason, it is theoretically capable to eliminate the inconsistency between convergence and accommodation in human viewing system, supply a real 3D experience for the viewers without any uncomfortable feelings such as eye strain or headache.

Optical integral imaging (OII) system is implemented by use of a lenticular lens array or a parallax barrier. As illustrated in figure 2.11, each elemental image is recorded behind one elemental lens. The coordinate of this lens determinates the viewing position of this elemental image [N.94]. Numbers of elemental images represent the 3D scene from different viewing position. Similarly, in the display process, each EI is projected through its corresponding lenticular. The

observed 3D image is reconstructed by the intersection of the projected light rays emanated from every lens.

Building an optical integral imaging system is a very sophisticated project requiring highly precise optical devices and careful calibration [SCA⁺08]. To facilitate research activity, a computational integral imaging (CII) system, aiming at simulating the recording and reproducing process of an optical integral imaging (OII) system in the digital manner, has been developed and studied recently. Like the optical system, the CII system is also composed of two basic processes, known as generation and reconstruction. The generation process is implemented by the computer rendering technique to create an integral image (II) of a given 3D scene, whereas the reconstruction process is achieved by simulating the optical projection effect in a display process by a computer-built virtual lens array [HJJ04].

1D integral imaging display is very similar to multiview display. The main difference between them relies in the mode of projecting beams distribution [FTS⁺09]. Because of the parallel beam projection, integral imaging display can provide a continuous parallax of 3D image with constant quality [BFD95]. Nevertheless, the perceived image quality of multiview display will change along the different viewing points because of converging beam projection.

2.4.2 Volumetric display

Instead of generating a 3D illusion to the viewer using conventional plane display, the objective of volumetric display is to create true 3D perception by illuminating each points of a scene at its actual position in space. Two types of volumetric displays have been developed to achieve this target: swept-volume display and static-volume display [FE05].

The principle of swept-volume display is based on the human persistence of vision (the phenomenon that an image persists on the retina for approximately one twenty-fifth of a second). The 3D scene is computationally decomposed

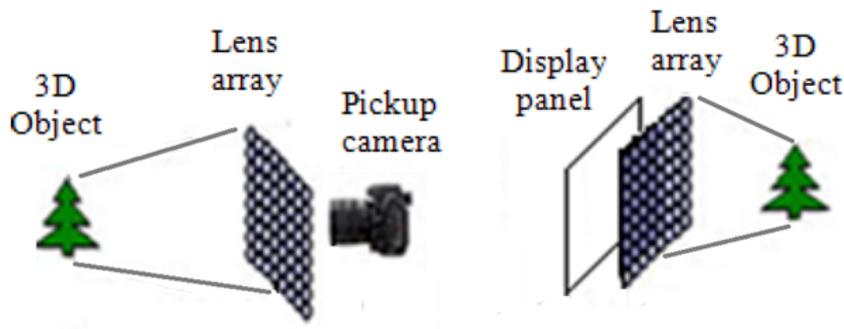


Figure 2.11: Optical setup for generating (left) and displaying (right) 3D scene in integral imaging system

into a series of 2D images. Then, the 2D images are swept through a 3D spatial volume at a higher frequency than human can resolve. By this way, the viewer can perceive a 3D volumetric image.

In static-volume display, the 3D image is created without any mechanical motion [FDH⁺01]. It is regarded as the most straightforward form of 3D display. In a nutshell, static-volume displays create an addressable physical 3D space filled with active elements (called voxels), where the 3D scene is generated by changing the state of each voxel. Voxels in 'off' state are set as transparent, while the voxels in the 'on' state are either opaque or luminous. The collection of all the voxels constructs a real 3D image to the viewer.

Due to their particular setup, volumetric displays can provide a correct 3D image very close to the real scene. However, limited 3D volume and constrains in image transparency prevent its industrial application. Sophisticated acquisition and display mechanism are demanded to satisfy the correspondent represented 3D videos. The hardware devices should be carefully implemented to achieve the best viewing performance. All these requirements prevent its application in 3DTV.

2.4.3 Holographic display

Holographic display is regarded as the ideal 3D display, which can provide exactly the same depth cues as real world to the viewer. It was invented by the British physicist Dennis Gabor in 1948. This is the only visual recording and playback process that can record our three dimensional world on a two dimensional recording medium and playback this world as a true three dimensional image to naked eyes.

In a holographic system a photosensitive plate is used to record the interference between the light from the reference beam and the object reflected light, and illuminate the recording plate in the receiver side to reproduce the object with naturally 3D perception.

The holography technique used to define a 3D image produced from the interference of light-beams scattered from photographed 3D object and reference beam. This interference information is recorded by a two dimensional photosensitive plate; the recorded image is called fringe pattern. Figure 2.12 illustrates the recording process of a hologram [PMFN10].

In the reconstruction process shown in Figure 2.13, the wave front of the original object is reproduced by illuminating the recorded fringe pattern with the reconstruction beam, which is generated by a laser identical to the one used for the reference beam. This way, a viewer located in the reconstruction beam is able to see the exactly same scene as he would see in the captured real scene.

Because to its promises to produce true 3D image without any viewing assistance, holography is highly appropriate for television applications. But the practical holographic system still has many difficulties to overcome. Some drastic conditions are required to record and display holographic image, such as

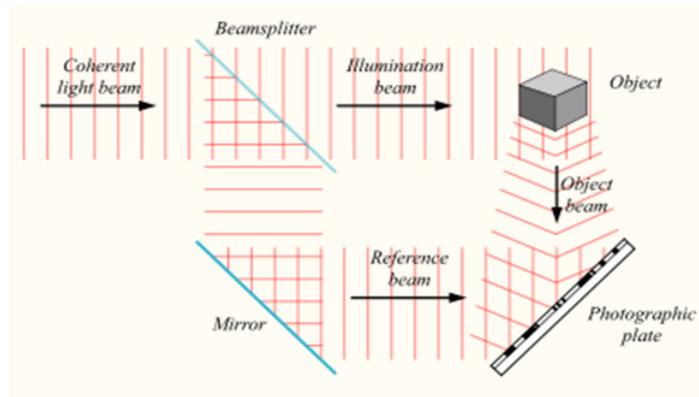


Figure 2.12: Hologram recording

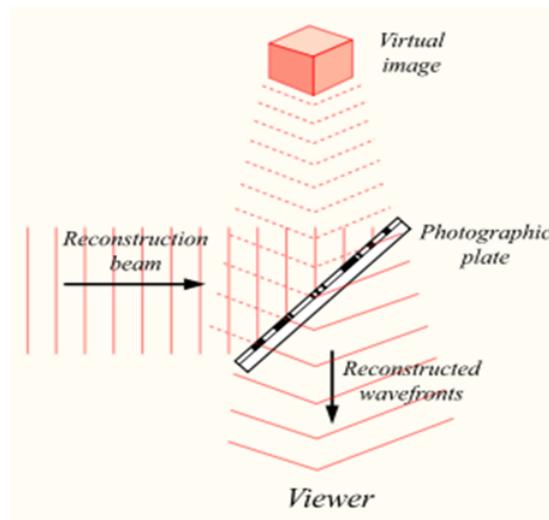


Figure 2.13: Hologram reconstruction

coherent light source, dark room environment and stable mechanical aperture. In addition, a holographic system is complex, requires very powerful processing systems as well as transmission and storage capacity [PMFN10]. Moreover, on the display side, an extremely high resolution display screen with a pixel pitch of about 0.5m (about one wavelength of visible light) is required to reproduce the picture [MFN⁺96], which is unreachable for currently HD screens. All these considerations decrease the feasibility of the holographic technique for 3DTV application. As the most visible procedure of 3DTV system, the success of 3DTV is judged by the performance of display process. Therefore, choosing a suitable display technique is the fundamental for constructing a successful 3DTV system. In Table 2.1, we provide a summary and comparison of all the five types of 3D display introduced above:

3D displays	Provided depth cues			Limitations
	Disparity	Motion parallax	Accommodation and convergence	
Binocular	✓	×	×	Glasses dependant Uncomfortable experience
Multi-view	✓	--	×	Image flipping limited resolution Uncomfortable experience
Integral imaging	✓	✓	✓	Limited resolution Huge data volume
Volumetric	✓	✓	✓	Too complicate for 3DTV
Holographic	✓	✓	✓	Too complicate for 3DTV

Table 2.1: Comparison of 3D display types

2.5 Conclusions

Depending on the character of various display technologies and requirement of 3DTV, integral imaging is considered a wise choice for meeting the requirements of an ideal 3DTV setup. Actually, it can be regarded as a trade-off between stereoscopic and holographic techniques, which is easier to be implemented than holographic systems and provides better 3D effect than stereoscopic ones. It eliminates the conflict between convergence and accommodation, solves the eye strain and fatigue problem, allows the viewer to see natural 3D image without wearing dedicated glasses. Apart from these advantages, it still holds some limitations for practical application, such as restricted viewing angle, large volume of source data and limited image resolution. In the next chapter, we will thoroughly study the problems raised above, and derive corresponding solutions.

Chapter 3

Computational Integral image Generation

In this chapter, we study the process of integral image acquisition. A thorough description of integral image generation in computational and optical operation is given first. Then, the scheme and limitation of existing computational integral image generation (CIIR) techniques are interpreted and analyzed in Section 3.2. Thereafter, a new CIIR method is proposed to overcome the mentioned problems. Finally, the simulation results are demonstrated and conclusions are drawn.

3.1 Introduction

Figure 3.1 illustrates the basic acquisition and display process of the integral imaging system. As shown in the left Figure, a lens array composed of many convex micro lenses is positioned in front of the capture plate to record an integral image. The captured II is composed of several elemental images (EIs) recorded through different lenses, each EI representing a slightly different view of the 3D scene.

On the display side (shown in figure 3.1), the recorded II will be displayed behind an identical lens array to reproduce the light rays in the opposite direction to the captured ones so as to reconstruct the 3D scene.

Depending on the relationship between the focal length of micro-lens and the gap between the lens array and the display panel, integral imaging systems can be classified into three different modes: real, virtual and focused[MHKL05].

In the case of real mode, the gap between the lens array and the display panel is larger than the focal length of the micro-lenses. The integrated image is located in front of the lens array (cf. Figure 3.2). The plane of reconstructed image is called integration plane, and the focal plane of the lens array is called the central depth plane (CDP). The location of CDP from the lens array (d) can be computed by the

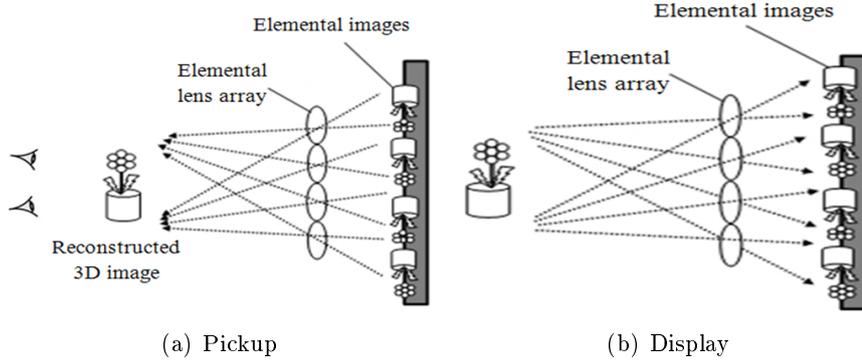


Figure 3.1: Optical setup of integral imaging system

well-known Gaussian lens law [SCCKES11], which is detailed in Equation 3.1. According to this law, the quality of $\frac{1}{d} + \frac{1}{g}$ (3.1) where f represents the focal length of micro lens, d is the location of CDP from the lens array and g represents the gap between the lens array and the display panel.

As opposed to the real mode, when the gap between the display device and the lens array is smaller than the focal length of the elemental lenses, the central depth plane and the integration plane will be sited behind the lens array (see figure 3.3). In this case, the integrated image is displayed in the virtual mode. Same as real mode, the location of the reconstructed image is restricted to a tight range around the CDP.

As mentioned above, both real and virtual modes contain a focused image in central depth plane. The quality of the integrated image will decline when the distance between integration plane and the CDP exceeds a certain value.

As shown in Figure 3.4, the focused mode occurs when the gap between the display device and the lens array is equal to the micro-lens focal length. In this mode, the lens array samples one pixel in each elemental image to construct the corresponding perspective image for different viewing direction. Since only one pixel of each elemental image can be perceived from a certain direction, this mode limits the resolution of integrated 3D image.

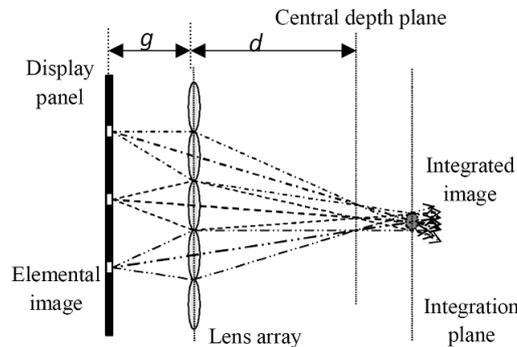


Figure 3.2: Real mode

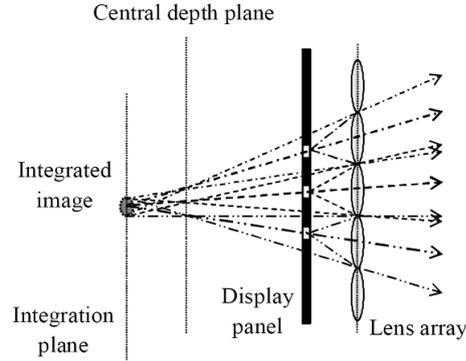


Figure 3.3: Virtual mode

Despite the limitation in image resolution, integral imaging system in focused mode produces 3D images with dramatically expanded depth range than the other two modes. The quality of the reconstructed image is stable when its distance from the lens array varies. Because of this important advantage, all the research works in this thesis are realized in the focus mode.

Besides the traditional optical methods described above, integral imaging systems can also be realized in digitally. The computational integral image generation process is implemented by using 3D rendering to generate integral images of a given 3D scene, in which the 3D intensity distribution of the object has been stored [McMCS⁺08] [NYH01]. While the computational integral image reconstruction (CIIR) process is achieved by simulating the optical projection effect in displaying process by a computer-built virtual lens array [CCJW13] [YM12]. Due to its simplicity of implementation, computational integral imaging (CII) provides obvious advantages for integral imaging research. In this thesis, all of the proposed methods are implemented and evaluated in CII system.

With the development of computer hardware capacity and computer graphic technique, the quality of generated images has been improved significantly. However, some inherent problems such as limited viewing angle, image depth and low viewing resolution are still unsettled [MWH97].

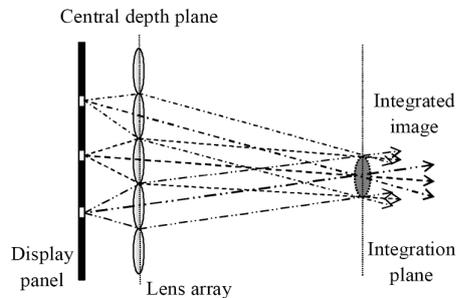


Figure 3.4: Focus mode

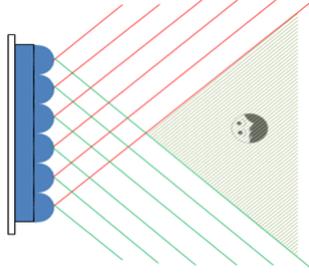


Figure 3.5: Viewing range in integral imaging system

In a basic integral imaging system, there is always a trade-off between viewing range, 3D image continuum and 3D image resolution [OAMO06] [KCKY12a]. This trade-off can be explained in two folds. On the one hand, the total number of pixels on the capture plate N_t , is the product of the number of elemental lenses N_m , and the number of pixels in each elemental image N_e :

$$N_t = N_m \times N_e \quad (3.2)$$

As we mentioned before, in the focused mode, the resolution of displayed 3D image is equal to the elemental lenses number (N_m), and the number of reproduced views (V_n) is highly dependent on the number of pixels in each EI (N_e). Therefore, for a created integral image with a determined resolution, there is a reciprocal relation between reconstructed image resolution and number of visual directions [Kak08].

On the other hand, in display process, the number of reproduced 3D views (V_n) is the product of viewing angle density (V_d) and the viewing range (V_r):

$$V_n = V_r \times V_d \quad (3.3)$$

Another tradeoff is raised between viewing angle density and viewing range [MHKL05]. Here the viewing angle density (V_d) impacts the continuity of motion parallax. With a high viewing angle density, one can perceive an integrated 3D image with more fluid motion parallax. The viewing range (V_r) symbolizes the zone where each EI can be correctly projected, which is constrained in the overlap area of light rays emitted from each elemental lens. A viewer moving inside this area can observe the correct 3D scene with continuous parallax, whereas a viewer located outside this zone will observe a distorted or flipped image. In Figure 3.5, the viewing range is restricted in the shaded area.

In conclusion, a trade-off has to be found between viewing range, viewing angle density and definition of displayed 3D image. Improving anyone of them will lead to the degradation of the other two. Searching a proper method to break this restriction is an essential objective for many research teams.

Another drawback of integral imaging system lies in the limitation of recorded depth range. Although the focus mode is able to keep the same recorded definition of the entire 3D scene for different depth location, the resolution of one specific object is degraded if its distance from the lens array increases [KKH⁺12]. As shown in Figure

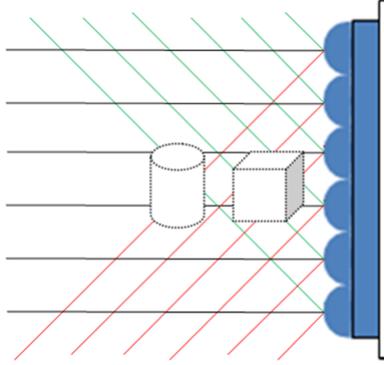


Figure 3.6: Degradation of recording resolution for two separated objects

3.6 the recording resolution for the distant object (cylinder) is smaller than the near one (cube) since less sampling rays intersect in its surface. Consequently, on the display side, the object near the lens array can be perceived at higher quality than the distant object.

In this chapter, a new CIIG method will be provided to solve both problems reviewed above.

3.2 Previous computational integral image generation technique

Along with the improvement of computer graphic technology, the CIIG technique has been widely used to generate integral image. In this section, we will review previous CIIG methods, such as Point Retracing Rendering (PRR), Multiple Viewpoint Rendering (MVR) and Parallel group rendering (PGR) [LCP⁺06, PJKL08]. Their concepts, designs and characteristics are exposed in the following paragraphs.

Point Retracing Rendering (PRR) is the first and most straightforward approach for integral image generation. In this method, the integral image is created by drawing every points of the 3D object pixel by pixel to the array of elemental images following a geometric principle, as illustrated in Figure 3.7.

Since it is easy to implement, the PRR approach is usually used for integral image creation for simple 3D object. As shown in Figure 3.8, Multiple Viewpoint Rendering (MVR) CIIG methods can be regarded as an inverse operation of PRR approaches. They apply an image mapping to simulate a virtual camera, and then use this camera to generate the elemental images of the 3D object. The virtual camera is moved by the pitch of one elemental lens of the optical integral imaging system.

MVR can create the correct elemental images through a simple pickup algorithm. Compared to the PRR approach, the complexity of MVR is not affected by the size of the 3D objects. However, since it needs to construct the EIs for each lens of the lens array, its processing time will become longer when the lens number increases.

To overcome the time consuming problem of precedent two CIIG methods, another image mapping method named Parallel Group Rendering (PGR) has been derived. In

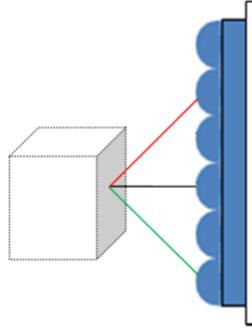


Figure 3.7: Concept of PRR method

PGR, elemental images are not obtained directly from the recording process but transformed from the first generated directional image. The directional image represents the 3D object from a certain viewing direction, which will be named as sub-image (SI) in the rest of this thesis.

The creation of sub-images is similar to elemental image capturing process, except that the applied virtual camera is no longer a perspective camera as common 2D cameras but an orthographic one that records the parallel group of light rays from the 3D object [PBK⁺08]. Moreover, the settled virtual camera no longer moves in horizontal or vertical axis but stays in a determined position. The number of sub-images with different viewing angle is obtained by changing the orientation of the virtual camera. As shown in Figure 3.9, three virtual cameras are applied to capture the 3D object from distinct viewing direction, then all the pixels from the same position of each SI are extracted and composed together to form an elemental image. The number of pixel in one EI is identical to the number of generated sub-image and the resolution of one SI represents the lens density in the array.

In the majority of cases, the number of pixels in one EI is several orders less than the number of lenses, which allows PGR to simplify the CIIG process significantly [MCND⁺11]. However, it still cannot solve the two main drawbacks of integral image systems, i.e. limited depth range and trade-off between reproduced 3D image quality,

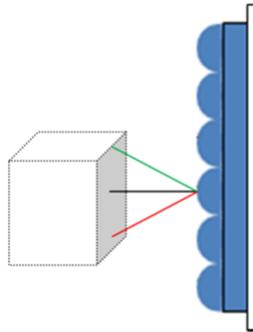


Figure 3.8: Concept of MVR method

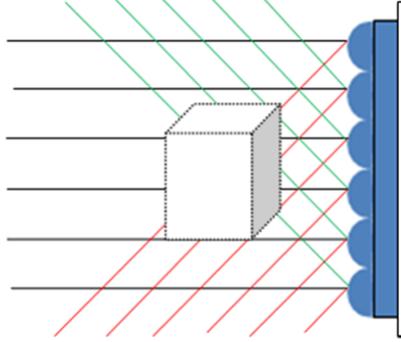


Figure 3.9: Concept of PGR method

viewing angle density and viewing range.

3.3 Proposed computational integral image generation method

In this section, a new PGR-based CIIG method is proposed to solve the mentioned problems in traditional integral imaging systems. This method is composed by two parts. First, a special multi-camera setup is applied to control the recorded depth range of the 3D scene. Then, a pixel selection process is executed to simulate the curve recording lens array used in OII system so as to enlarge the reconstructed 3D image viewing range.

3.3.1 Depth range controlling

As shown in Figure 3-10, in the traditional PGR technique, a series of virtual cameras with the same extrinsic parameter, e.g. identical location and orientation, are applied to record the sub images. The only variation between these cameras is the optical axis direction; different SIs are captured by the orthographic or oblique projecting virtual cameras with rotated optical axis. This kind of CIIG technique works well for the objects near the lens array. However, for the object far away from it, the recording resolution will decrease because of the spreading effect of the recorded beams. This kind of reduction may lead to distortion effects in the reconstructed pictures [Hir09].

A new CIIG method is proposed to increase the recording resolution of the distant objects. As described in 3.11(a) and 3.11(b), its idea is to select an interesting plane which we want to perceive in high quality, then make the recording area of every virtual camera overlap on this plane. In this way, all the beams would intersect on the plane of interest, and the object nearby could be recorded with relatively high resolution.

3.3.2 Viewing range enhancement

In articles [KPC⁺04] and [PLL11], an optical integral image capturing method is proposed to enlarge the viewing range of reconstructed 3D scene. The idea of this

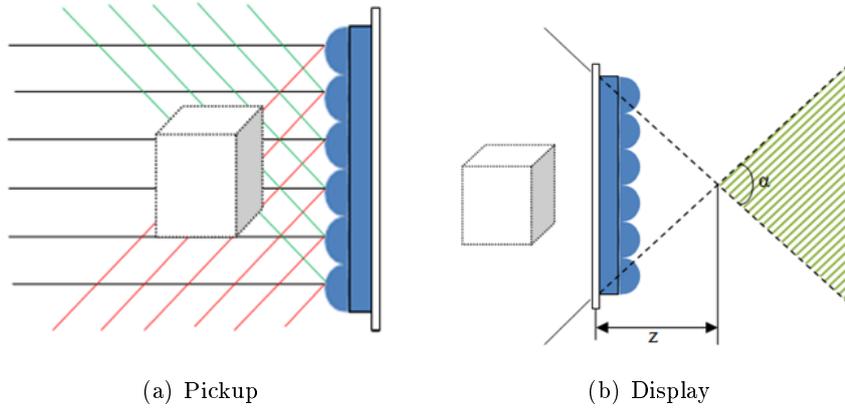


Figure 3.10: Pickup and display process of traditional CIIG method

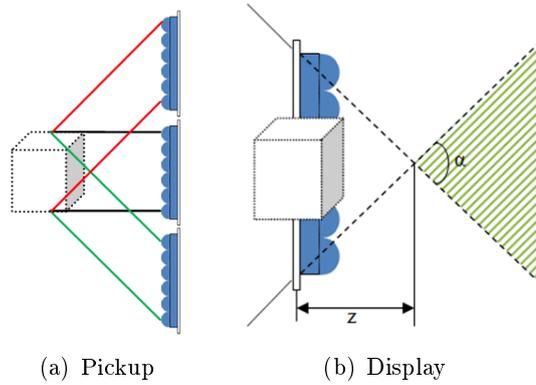


Figure 3.11: Pickup and display process of depth controlling CIIG method

method is to apply a curved lens array to record the 3D scene. As shown in Figure 3.12, thanks to the special disposition, the recording direction of each EI changes gradually along the curved lens array, so that their overlapping region is expanded due to the concentrative beam directions [ETKS13]. Meanwhile, the viewing continuous and perceived image resolution is the same as traditional planar system since the disparity between the neighboring views and number of lenslets is not changed.

Theoretically, this is a possible method to improve the 3D effect of integral imaging system. However, since a curved lens array and CCD device is difficult to build, it is hard to realize this method in practical optical systems.

A new CIIR method is proposed to simulate the curved lens array digitally. As illustrated in Figure 3.13, this approach doesn't decrease the viewing angle resolution; a larger numbers of SIs are generated in this step. Then the generated SIs are combined together to create an initial array of elemental images which contains more pixels than expected. As shown in Figure 3.14, size of II in Figure 3.14(b) is larger than the size of II in Figure 3.14(a). In order to avoid increasing the generated II volume, the number of pixels in this EI array should be reduced in the next step. Finally, according to the position of the EI, the pixels projected outside the view-

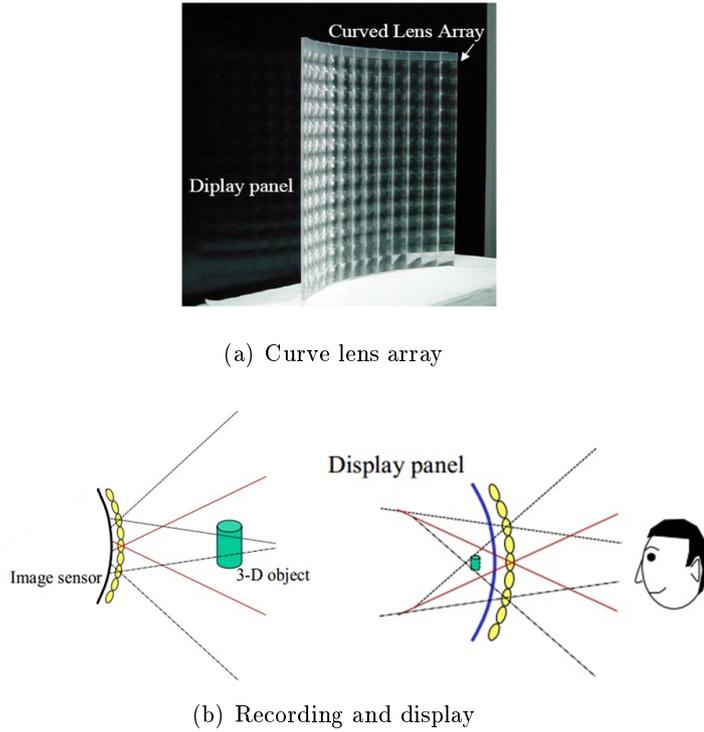


Figure 3.12: An example of curve lens array and concept of the recording and display process



Figure 3.13: Flow chart of proposed CIIG method

ing range will be eliminated so as to decrease the size of II to our expected volume. For example, as shown in Figure 3.14(b) the pixels which are projected out of the shade area will be eliminated in the new integral image. Once the removing process is accomplished, a new II with larger viewing range is created (cf. Figure 3.14(c)).

As illustrated in 3.14(c), due to the different recording direction of every EI, the viewing angle is largely enhanced compared to previous CII systems.

3.4 Experimental results on proposed generation method

The view-based CIIR method proposed in [SLsK06] is applied to demonstrate the performance of the proposed CIIR method. This method simulates the observed image of the viewer, which provides a direct sight to evaluate the quality of displayed 3D image.

To simplify the implementation, our simulation is only carried out for the horizontal direction. Two experiments are executed to evaluate the presented approach. First

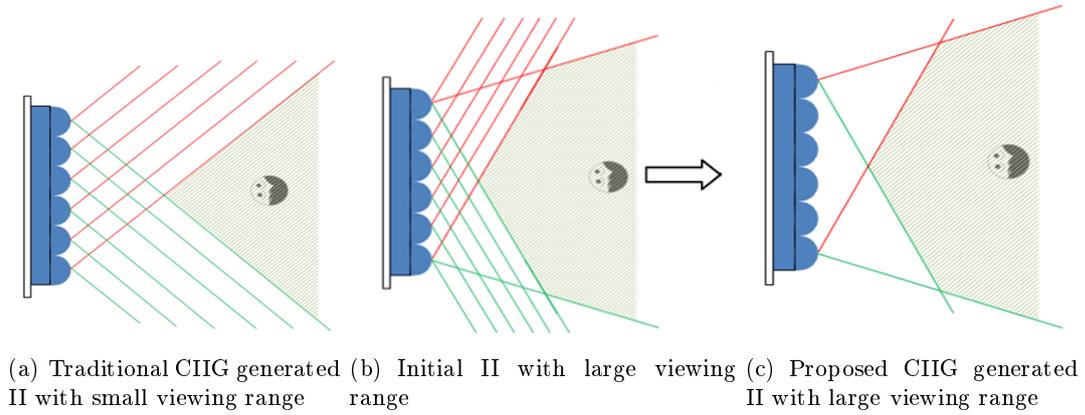
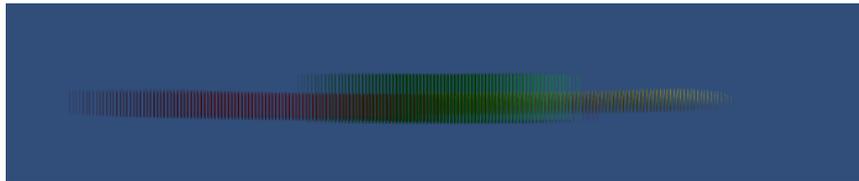
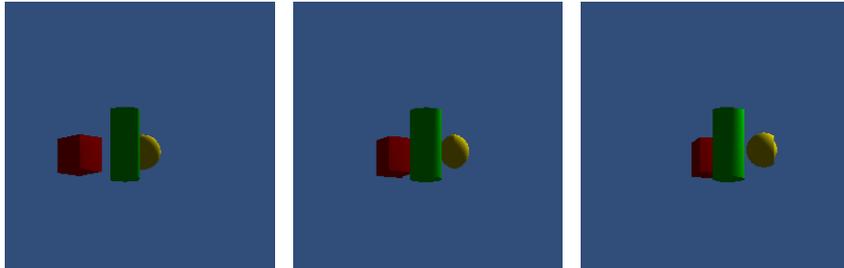


Figure 3.14: CIIG system

experiments prove the depth range controlling effect of the proposed CIIG method. The second experiment demonstrates the viewing range enhancement effect without degrading the 3D image resolution and visual fluency.



(a) Integral image generated by traditional CIIG method



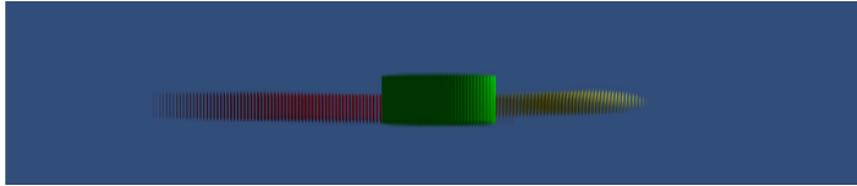
(b) Reconstructed image (c) Reconstructed image (d) Reconstructed image

Figure 3.15: Integral image generated by traditional CIIG method and its reconstructed 3D image

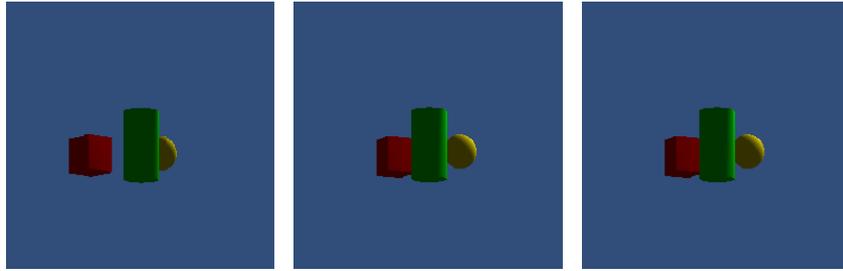
3.4.1 Evaluation of depth range controlling method

Figure 3.15(a) shows the II generated by traditional CIIG method. By use of this integral image, three perceived images from different viewing position are reconstructed and illustrated in figure (3.15(b), 3.15(c), 3.15(d)).

Then, another integral image generated by proposed depth controlling CIIG method is exposed in Figure 3.16(a). Using the same contest as the first test, three observed images from the applied viewing position in Figure 3.15(b), 3.15(c), 3.15(d) are reconstructed and illuminated in Figure 3.16(b), 3.16(c), 3.16(d).



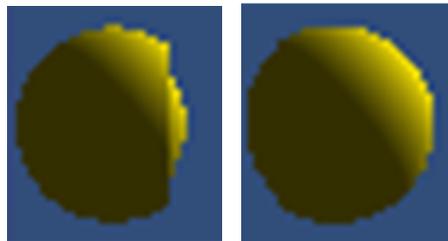
(a) Integral image generated by depth controlling CIIG method



(b) Reconstructed image (c) Reconstructed image (d) Reconstructed image

Figure 3.16: Integral image generated by depth controlling CIIG method and its reconstructed 3D image

Figure 3.17 and Table 3.1 show the detailed comparison of two reconstructed images of 3.15(d) and 3.16(d). According to this comparison, the quality of the furthest object (yellow sphere) is obviously improved with the depth controlling CIIG method.



(a) Result of traditional CIIG (b) Result of depth controlling CIIG

Figure 3.17: Reconstructed results of further object by traditional CIIG method and depth controlling CIIG method

	Traditional CIIG method	Depth controlling CIIG mehtod
PSNR(dB)	31.50	33.14

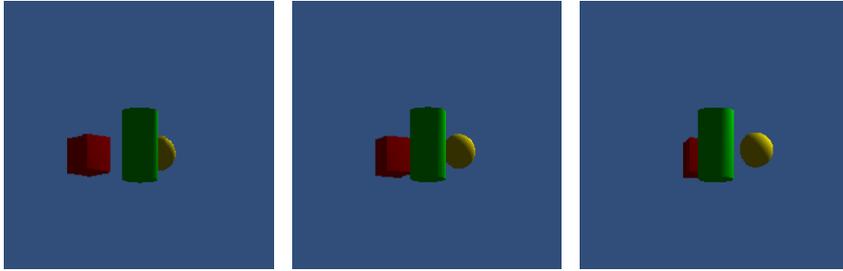
Table 3.1: PSNR comparison of traditional and depth controlling CIIG method

3.4.2 Evaluation of viewing range enhancement method

The second experiment is implemented to demonstrate the viewing range enhancement effect of the proposed approach. It is composed of four tests.



(a) Integral image generated by traditional CIIG method with large resolution



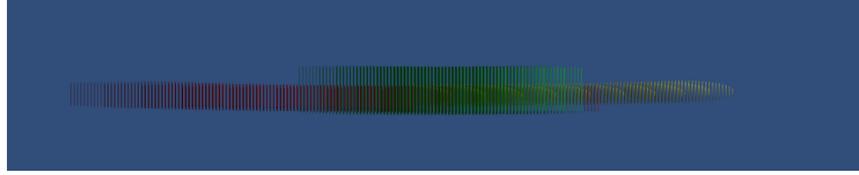
(b) Reconstructed image (c) Reconstructed image (d) Reconstructed image

Figure 3.18: Integral image generated by traditional CIIG method with large resolution and its reconstructed 3D image for different viewing position for 3D object ‘cube’

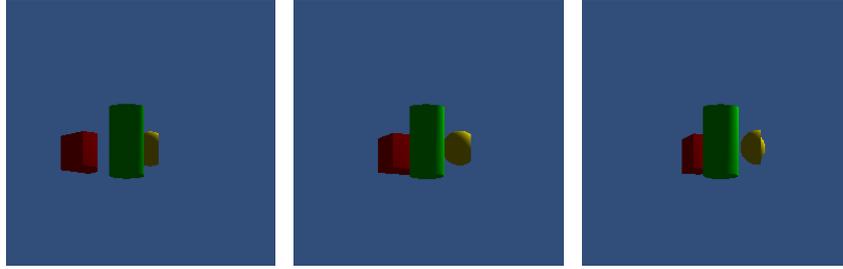
The first test gives a reference result to evaluate the other tests. In this test, an original II with high enough resolution is generated to produce a 3D scene with continuous motion parallax and large viewing range. The other three tests decrease the resolution of generated II by half the number of pixels in each EI. In the second and third tests, the reduction of the number of pixels causes the degradation of the viewing range and viewing angle density respectively. The fourth test provides the high quality reconstructed images as the reference results. The results of these four tests are illustrated as follows.

Figure 3.18 illustrates the results of test one. In the beginning of this test, an original II with a larger viewing range is generated by traditional PGR-based CIIG method. In order to achieve a satisfactory viewing experience, this II is generated with both large recording range and high recording angle density. The generated II is composed by 256×256 elemental images, where each EI contains 20×1 pixels. The recording range of each EI is set to 60° .

Then, by use of the classical view-based reconstruction method introduced in [SLsK06], the observed images from three different viewing positions are calculated and shown in Figure (3.18(b), 3.18(c), 3.18(d)). Thanks to the extremely high resolution of the II, the quality of the perceived image is rather high. These reconstructed images are applied as the reference result to evaluate the performance of other tests.



(a) Integral image generated by traditional CIIG method with half recording density in each EI



(b) Reconstructed image (c) Reconstructed image (d) Reconstructed image

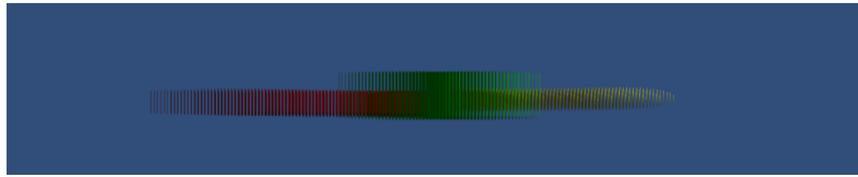
Figure 3.19: Integral image generated by traditional CIIG method with half recording density in each EI and its reconstructed 3D image for 3D object ‘cube’

In the second test, the number of pixels in each EI is decreased to half as in the first test but the recording range of each EI is kept the same as the reference test, so that the recording angle density in each EI is reduced. Figure 3.19 illustrates the generated II and the reconstructed 3D image of this test. As we already mentioned, the number of pixels in each EI is reduced to half the one in of test 1. Only 10 pixels are generated for each EI. However, the range of viewing angle remains the same as in test 1 which is still configured as 60° . Compared to the reference results shown in 3.18(b), 3.18(c), 3.18(d), the quality of the reconstructed image and the fluency of the viewing image in horizontal direction is severely degraded.

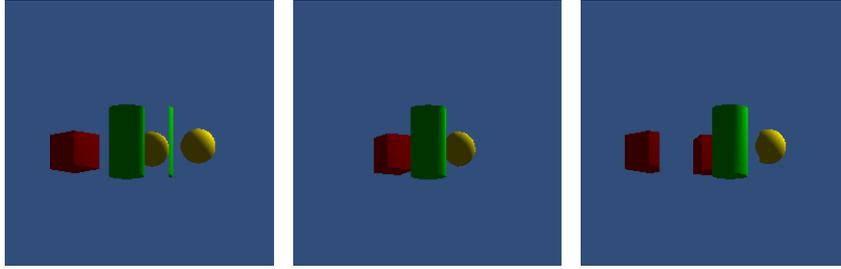
As opposed to the second test, the third test keeps the recording angle density of test one but halves recording range of each EI. In this case, the number of pixels for each EI is also reduced to 10. Contrary to the second test, we attempt to reduce the viewing range of each EI from 60° to 30° instead of diminishing the viewing angle density. As a result, the reconstructed viewing range of the 3D image is largely reduced. As shown in Figure 3.20(b), 3.20(c), 3.20(d), severe distortion effect can be observed in the left and right side reconstructed image.

If the viewing range is reduced, the 3D image can be perceived correctly. As shown in Figure 3.20(e), 3.20(f), 3.20(g), the double image distortion is completely eliminated but the parallax effect in horizontal direction is significantly decreased.

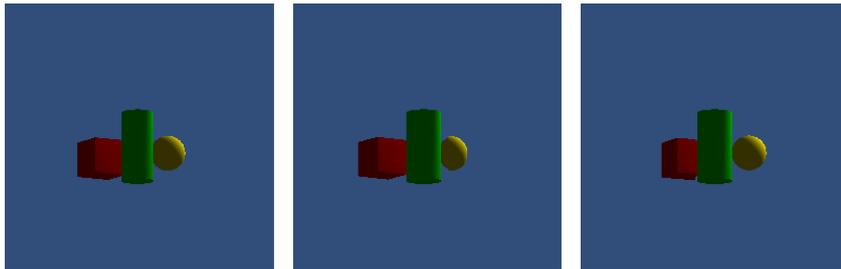
From the previous three experiments, we observe that reducing the II size by decreasing the recording angle resolution or narrowing the recording angle range will cause certain kinds of viewed image quality degradation. Concretely speaking, less viewing angle resolution leads to lower image quality and degraded parallax fluency, while smaller recording angle range induces minor viewing range and limited parallax effect.



(a) Integral image generated by traditional CIIG method with half recording range in each EI



(b) Reconstructed image in large viewing range (c) Reconstructed image in large viewing range (d) Reconstructed image in large viewing range

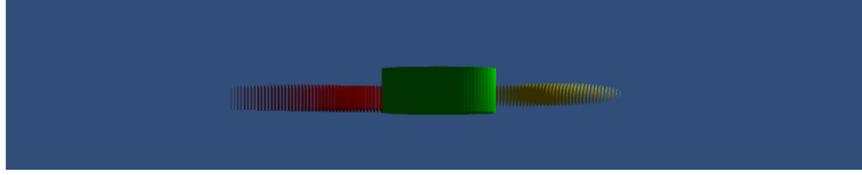


(e) Reconstructed image in reduced viewing range (f) Reconstructed image in reduced viewing range (g) Reconstructed image in reduced viewing range

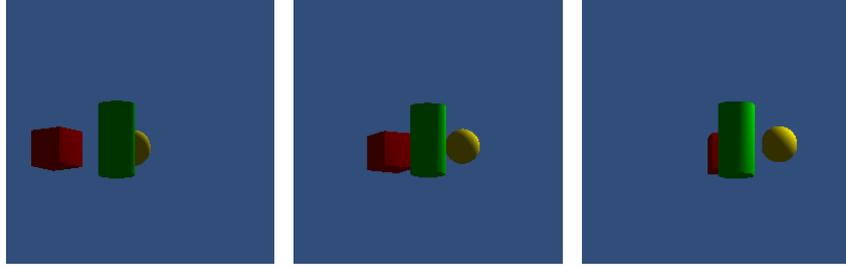
Figure 3.20: Integral image generated by traditional CIIG method with half recording range in each EI and its reconstructed 3D image in large viewing range (b, c, d) and halved viewing range (e, f, g) for 3D object ‘cube’

The last test applies our proposed CIIG method, which aims to expand the viewing range without decreasing the motion parallax smoothness and displayed image quality. This method is implemented by eliminating the useless pixels in each EI and regenerating a novel EI array with less pixels. Based on this array of elemental image, we are able to construct a 3D image with the same viewing angle density and viewing range as the reference results.

Starting from the original II shown in Figure 3.18(a), the proposed CIIG method is applied to eliminate the useless pixels in each EI. Then, we can obtain a new array of EI of half size of original EI. The regenerated II is shown in Figure 3.21(a), where each EI possesses 10×1 pixels. Finally, the same reconstruction technique is adopted to produce the observed 3D image from different points of view. The reconstructed images are illustrated in Figure 3.21(b), 3.21(c), 3.20(d).



(a) Integral image generated by viewing angle enhanced CIIG method



(b) Reconstructed image (c) Reconstructed image (d) Reconstructed image

Figure 3.21: Integral image generated by viewing angle enhanced CIIG method and its reconstructed 3D image for 3D object ‘cube’

According to the reconstructed results shown in Figure 3.21(b), 3.21(c), 3.20(d), we can detect that, by use of the proposed CIIG method the viewer can perceive a high quality reconstructed image from the same viewing range as in Figure 3.18. Actually, the size of the II in Figure 3.21(a) is identical to the II in Figure 3.19(a) and 3.20(a), but its provided viewing effect is almost as good as the result of Figure 3.18 which derived from an II with double size. In sum, our proposed CIIG method is able to enlarge the viewing range to the viewer without impact the source data size and viewing angle density. In order to precise the evaluation, we also calculated the PSNR values of each reconstructed image, the comparison results are shown in Table 3.2, which also demonstrates the higher reconstructed image quality of the proposed CIIG method.

	Traditional CIIG method			Proposed CIIG method
	Half recording density	Half recording range		
		Small viewing range	Double effect	
PSNR(dB)	33.20	30.14	31.12	35.40

Table 3.2: PSNR comparison of traditional and proposed CIIG method for ‘cube’



(a) Integral image generated by traditional CIIG method with large resolution



(b) Reconstructed image (c) Reconstructed image (d) Reconstructed image

Figure 3.22: Integral image generated by traditional CIIG method with large resolution and its reconstructed 3D image in different viewing position for 3D object ‘spaceship’

Another example implemented on ‘spaceship’ is also illustrated. Similarly as the simulation for object ‘cube’, this experiment is also implemented by four tests.

Firstly, a high resolution II is generated by traditional CIIG method. Like in the first experiment, the virtual camera only moves in horizontal direction. The generation parameters are configured as follows:

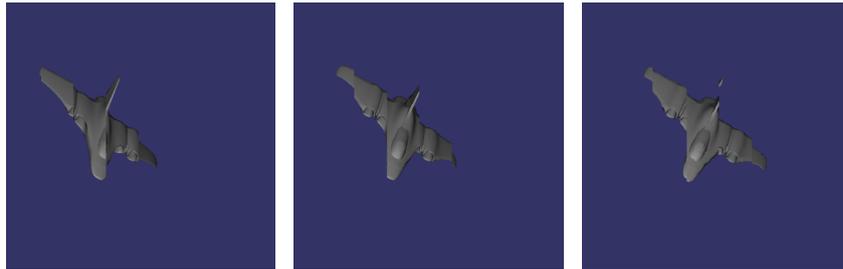
- Viewing angle for each EI is settled to 30°
- Numbers of EI is 512×512
- Size of EI is $20pixels \times 1pixel$

The resolution of generated II is 10240×512 , image shown in Figure 3.22(a)

Then, applying the classical view-based CIIR, we can obtain the observed images from various viewing positions. The viewing images for -15° , 0° and 15° are illustrated in Figure 3.22(b), 3.22(c), 3.20(d). From the reconstructed images, we can find that an II with high enough definition can produce a fluent motion parallax to the viewer within its viewing range.



(a) Integral image generated by traditional CIIG method with half recording density in each EI



(b) Reconstructed image (c) Reconstructed image (d) Reconstructed image

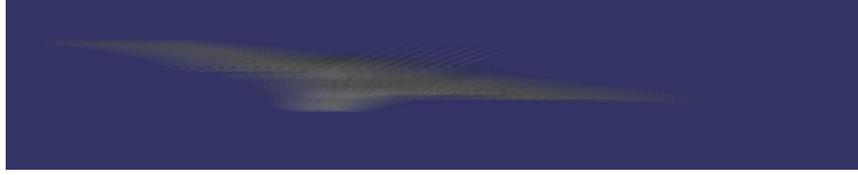
Figure 3.23: Integral image generated by traditional CIIG method with half recording density in each EI and its reconstructed 3D image for 3D object ‘spaceship’

In the second test, a new input II with half resolution is generated by traditional CIIG methods (cf. Figure 3.23(a)). The generation parameters are configured identically as in the precedent experiment; only the size of EI is divided by two.

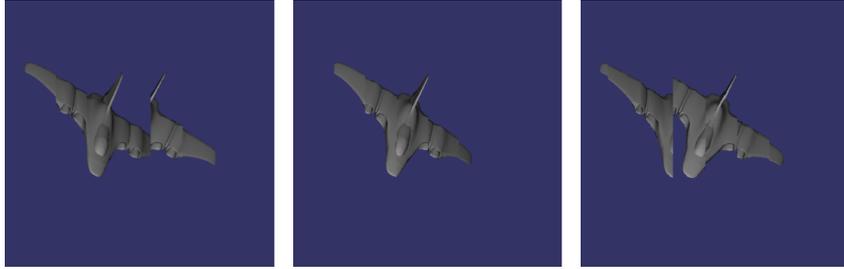
- Viewing angle for each EI is settled to 30°
- Numbers of EI is 512×512
- Size of EI is $10pixels \times 1pixels$

Due to the reduction of the EI size, the resolution of the generated II is also divided by two, from 10240×512 to 5120×512 .

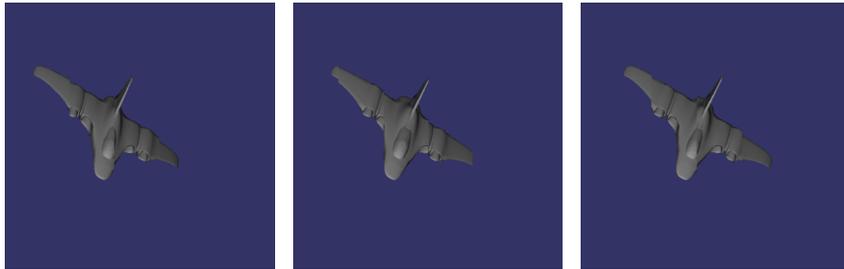
Using the same view-based CIIR method, three images from -15° , 0° , 15° are generated and shown in Figures 3.23(b), 3.23(c), 3.23(d). Compared to previous reconstructed images in Figures 3.22(b), 3.23(c), 3.23(d), a severe degradation in image quality is observed by the viewer.



(a) Integral image generated by traditional CIIG with half recording range in each EI



(b) Reconstructed image (c) Reconstructed image (d) Reconstructed image



(e) Reconstructed image (f) Reconstructed image (g) Reconstructed image

Figure 3.24: Integral image generated by traditional CIIG method with half recording range in each EI and its reconstructed 3D image in large viewing range (b, c, d) and halved viewing range (e, f, g) for 3D object ‘spaceship’

The third test tries to decrease the size of II by abasing the viewing range but maintaining the viewing angle density. Parameters are set as:

- Viewing angle for each EI is settled to 15°
- Numbers of EI is 512×512
- Size of EI is $10pixels \times 1pixels$

The generated II is shown in Figure 3.24(a) with resolution of 5120×5120 .

Figure 3.24(b), 3.24(c), 3.24(d) show the reconstructed image from different viewing positions. Viewing angles for the three reconstructed images are still -15° , 0° and 15° . However, because of the narrowed recording angle range, a severe double effect appears in the reconstructed images out of the viewing range (Figures 3.24(b) and 3.24(d)).

To eliminate the double effect, the end-user can only change his position within the viewing range which is reduced to half size of the previous two experiments. Viewed images inside the limited viewing range are shown in Figures 3.24(e), 3.24(f), 3.24(g).

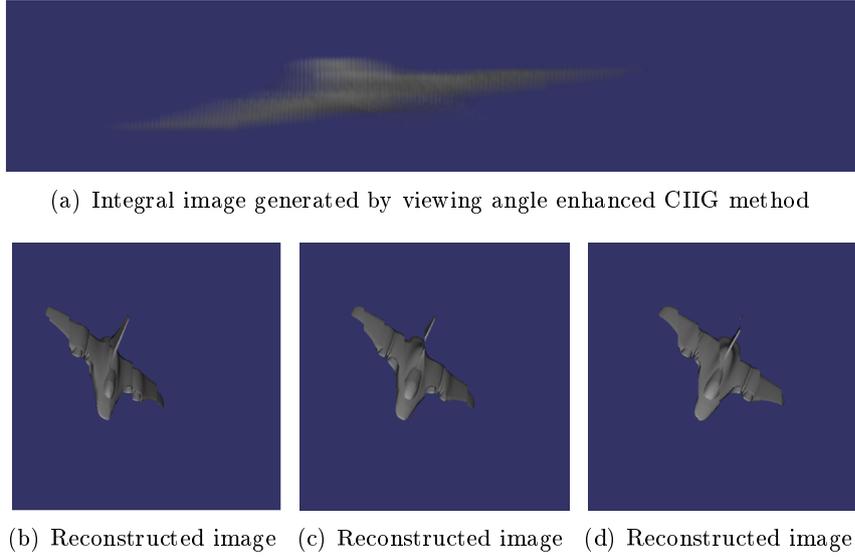


Figure 3.25: Integral image generated by viewing angle enhanced CIIG method and its reconstructed 3D image for 3D object ‘spaceship’

Contrasting with the observed images in 3.24(b), 3.24(c), 3.24(d), the double effects in marginal images are removed. However, the viewing range is reduced dramatically, which leads to a certain constraint to viewing effect.

The fourth experiment is intended to apply the proposed CIIG method. In this simulation, we first generate an II with high enough resolution which is able to produce the 3D image with continuous parallax and high-grade quality, as shown in 3.22(a).

Thereafter, according to the position of each EI, some unnecessary pixels in every EI are eliminated correspondingly to generate a novel array of EIs, shown in Figure 3.25(a). Dividing by two the number of pixels in every EI, the size of recomposed II is reduced by the same amount.

Observed images from -15° , 0° and 15° are listed in Figure 3.25(b), 3.25(c), 3.25(d), compared to the second and third experiments, it is easy to note a remarkable improvement in image quality or viewing range in the proposed CIIG method. The comparison of PSNR values listed in Table 3.3 demonstrates the improvement of reconstructed image quality in an objective manner.

	Traditional CIIG method			Proposed CIIG method
	Half recording density	Half recording range		
		Small viewing range	Double effect	
PSNR(dB)	32.50	31.13	28.38	34.35

Table 3.3: PSNR comparison of traditional and proposed CIIG method for ‘spaceship’

3.5 Conclusions

In this chapter, we proposed and demonstrated a new CIIG method to control the depth range and expand the viewing range of traditional CII system. Two main contributions were provided in our generation scheme. In the first place, moving the virtual cameras to make every sub image focus in the same region helps record the 3D object in far distance with higher resolution. Secondly, changing the viewing direction of the EIs in different position enhances the viewing range for the viewer. The results of our experiments demonstrate that the proposed CIIG method is able to control the depth range of the traditional CIIG methods, thus improving the reconstructed image quality of the distant 3D objects. Moreover, it also succeeds to enhance the visible range of the displayed 3D scene without reducing the other 3D effects criteria such as viewing angle density and reconstructed image resolution.

Chapter 4

Integral image Compression

This chapter tackles the issues of integral image compression techniques. A thorough analysis is conducted to stress the correlations exploitable within one given integral image. Two deductions are derived from this analysis; in one hand, an efficient encoding method is necessary to build a feasible integral imaging system since the volume of integral image sequences is too large to be transmitted and stored in the current broadcast chain. On the other hand, this compression technique is possible to be implemented because a sequence of integral images contains more correlation than traditional 2D videos. Some precedent methods for integral image compression, exploiting the particular characteristics of integral images, are presented. Their principles and limitations are explored. According to these limitations three different approaches based on multi-view video coding (MVC), multi-view plus depth (MVD) coding and discrete-wavelet transform (DWT) are proposed and demonstrated. The results of simulations show that all these three methods lead to substantial improvements on the compression efficiency of integral images.

4.1 Introduction

To our knowledge, transmitting the huge volume of source data is one of the most important problems for the practical application of integral imaging system. The objective of this chapter is to study the property of integral image, understand the limitations of current compression methods, propose new encoding approaches to improve the compression efficiency.

As explained previously, an integral image is a special sort of 2D image that records the 3D scene by several elemental images from different points of view. Therefore, the whole set of EIs can describe not only the color and intensity but also the direction information of the reflected light rays from the 3D scene [KHL10] [KUO10] and [ZW12]. Depends on the complete 3D information, integral imaging system can provide a full parallax and quasi-continuous viewing parallax for the viewers without undesirable viewing strain in traditional stereoscopic 3D techniques.

As analyzed in Section 3.1, the definition of the produced 3D image is equal to the number of EI, while the viewing angle density of the observed image is in accordance with the number of pixels in one EI. For this reason, in order to provide a

high quality 3D image to the viewer, we need to supply an extremely high resolution integral image in the display side. As a consequence, compressing the large volume of II is a considerable problem for the practical implementation of integral imaging system.

In order to better understand the diverse encoding techniques for integral image, Section 4.2 gives a general analysis of the correlation distribution in a given II. The correlation of a single integral image can be detected not only within one EI but also among adjacent EIs. These two types of correlation consist in the four spatial dimensional redundancies within one II, plus the temporal correlation between successive IIs.

According to their particular characteristics, plenty of traditional compression methods are provided recently. The principle and limitations of four different encoding methods are reviewed in Section 4.3. Based on our study, these compression methods are able to somehow improve the encoding efficiency, but they also contain some limitations. In order to exploit the redundancy in the sequence of integral image more thoroughly, three new compression methods are proposed in section 4.4. Their structure, advantages and problems are expounded then.

The first approach is implemented by use of multi-view video coding (MVC). This method is applied on sub-images. As previously mentioned, SI is an alternative form of 2D image transformed from original II. Each SI represents the 3D scene from parallel viewing directions and is prone to superior compression compared to the original captured elemental images (EI). For this reason, we arrange the group of SIs as the format of multi-view video (MVV) and then encode the generated MVV by MVC.

Another encoding method using multi-view video plus depth (MVD) representation is presented next. This method is also applied on sub-images. The principle of this approach is to use geometrical information in the depth map to predict side SIs from the reference SI so as to eliminate the two-dimensional redundancies between different SIs. Then the remaining two-dimensional correlations within one SI and the temporal redundancy between consecutive SIs are eliminated respectively by the discrete cosine transform (DCT) and inter-frame prediction techniques in traditional H.264/AVC technology. Therefore, this compression method is able to fully remove the five dimensional correlations in an II sequence described in Section 4.1.

The last proposed method is based on Discrete Wavelet Transform (DWT). As opposed to previous two methods, this approach is applied on original integral image. It leverages the DWT technique to remove the spatial correlation in one II, and then applying the inter prediction function of MPEG to encode the transformed coefficients of successive IIs.

Experimental results show that the compression approach we propose improves the compression efficiency when compared to the traditional MPEG-4/AVC compression method for II.

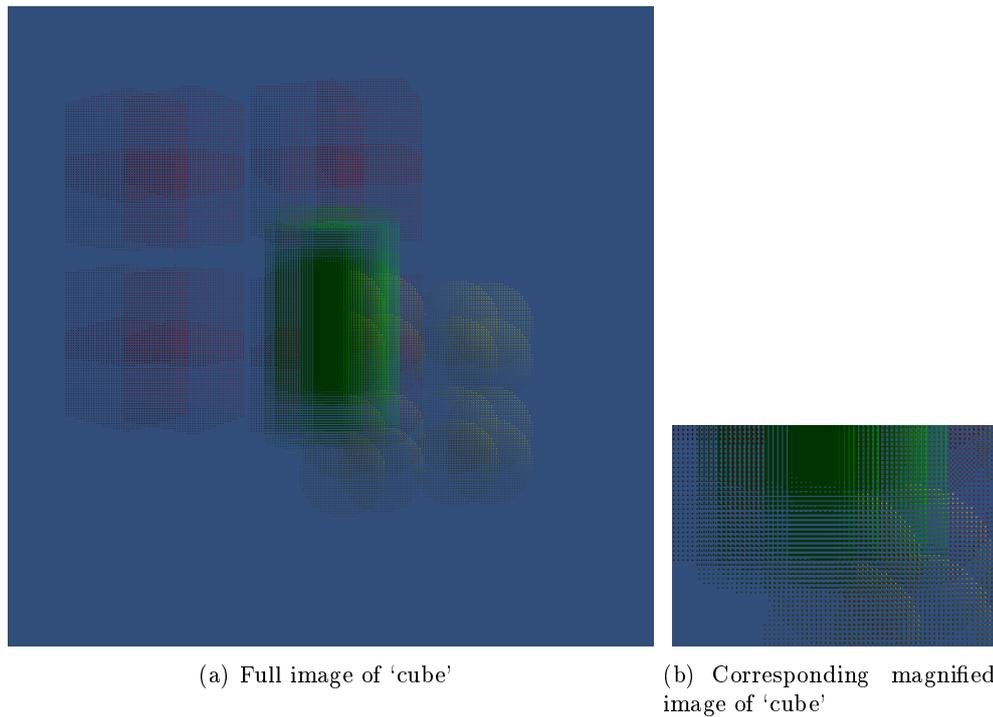


Figure 4.1: Integral image of 'cube'

4.2 Analysis of the correlation of integral image sequence

To design an efficient compression scheme for integral imaging system, we should firstly understand the nature of the recorded intensity distribution data. This section gives a general view of the conformation of the acquired source data.

Figures 4.1(a) and 4.2(a) show two examples of computational generated integral image, which are generated by PGR-based CIIG method introduced in Section 3.2. The magnified parts of the recorded source data shown in Figures 4.1(b) and 4.2(b) expose the detail of elemental image. As shown in Figure 4.1(b), 4.2(b), the regular blocked structure in the intensity distribution represents an elemental image, and one integral image is composed by a huge number of elemental images.

As highlighted by Figure 4.3, the correlation in one integral image can be explored in two parts as 'inter-EI correlation' and 'intra-EI correlation'. Inter-EI correlation represents the cross-correlation between neighboring EIs. Intra-EI correlation means correlation between adjacent pixels in each EI.

Besides the spatial correlation present in one single integral image, the sequence of integral images possesses inter-frame redundancies (cf. Figure 4.4). This kind of temporary correlation is denoted as 'inter-II correlation'.

In summary, the features of integral image can be exposed in four points:

- The generated intensity distribution has a block structure conforming to the form of the lenticular unit;
- An important correlation exists between blocks in the intensity distribution, and

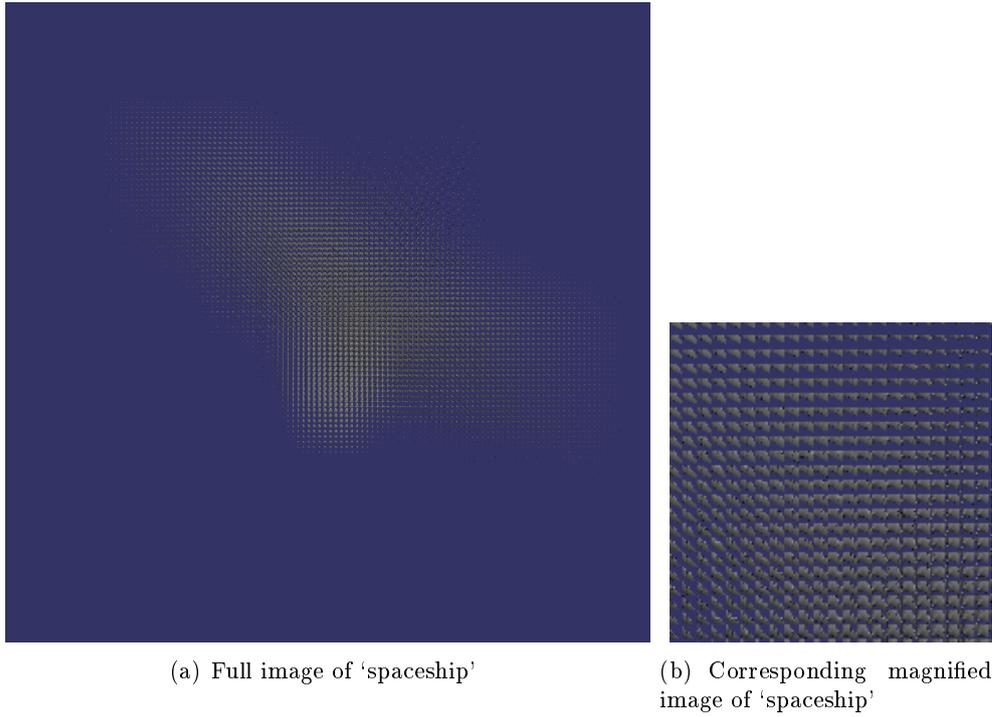


Figure 4.2: Integral image of 'spaceship'

this property should be taken into account in compression scheme;

- Inter-pixel correlation can be observed;
- Inter-frame correlation is present between consecutive integral images.

4.3 Previous integral image compression methods

Following the understanding of integral image characteristics, several different compression schemes are proposed and developed to encode the still and moving integral images.

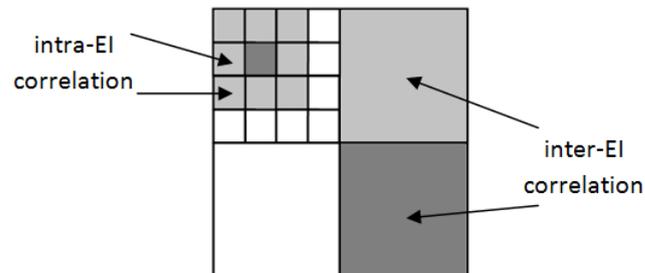


Figure 4.3: Spatial domain correlation in an integral image

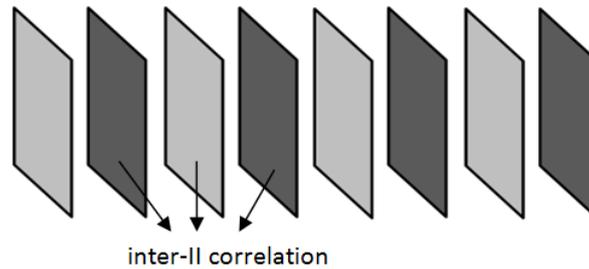


Figure 4.4: Temporary domain correlation in an integral image sequence

The most straightforward approach one can think of is using the classical MPEG-4 method directly to encode a sequence of integral images. Here, the integral image sequence is regarded as an ordinary 2D video. This approach is very simple to implement but its performances are limited since it omits the inter-EI correlation.

Besides the direct MPEG-4 technique, a set of solutions has been specially designed and implemented in the past decades. Several of them are exposed in this section, including MPEG-II based approach in [YSJ04], 3D-DCT-based compression scheme in [ZAM02], motion compensated technique for SI [KSK10] and [HHKK12], hybrid compression method using discrete wavelet transform and discrete cosine transform [ESHJ07] and [Agg11a].

4.3.1 MPEG-II-based encoding method

Since integral images obviously exhibit high correlation between different EIs, it is not surprising that previous compression methods sought to remove these redundancies using MPEG-4 for EIs. Proposed by Sekwon Yeom, Adrian Stern and Bahram Javidi in 2004 [YSJ04], the process of this compression method is relatively simple. It applies the MPEG standard directly to remove the high cross-correlations between elemental images.

As shown in Figure 4.5, this encoding scheme is implemented in two steps. Firstly, the elemental images from one integral image are scanned in a special topological mode to form a 2D sequence. Three different scanning modes in parallel, perpendicular and spiral directions are illustrated in Figure 4.6. Thereafter, the traditional MPEG technique is adopted to compress the obtained elemental image sequence.

Considering each elemental image as one frame in a moving picture leads to exploiting the inter-EI correlations, which improves the compression efficiency. However, the de-correlation effect is restrained to the consecutive EIs along the determined scanning direction (illustrated in Figure 4.6). While the correlation in the other direction



Figure 4.5: Process of MPEG based integral image encoding method

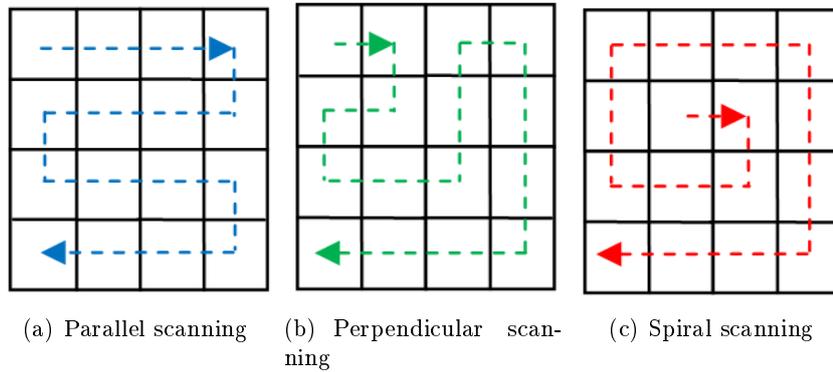


Figure 4.6: Three different mode for scanning elemental images in an II

and the temporal correlation are not taken into account, which limits the performance of this MPEG-based compression technique.

4.3.2 3D-DCT-based compression scheme

Instead of using conventional MPEG-based compression technique, the authors in [ZAM02] propose another 3D-DCT based method to compress the significant quantity of data of integral image. The block diagram of this algorithm can be shown in Figure 4.7.

As illustrated in Figure 4.7, the 3D-DCT-based method is carried out via four steps. Firstly, the mean value of every elemental image is calculated to generate a planar mean image. Then, this planar mean image is segmented into several homogeneous blocks by use of slit-merge algorithm. In this step, the elemental images are divided into different groups with respect to the planar mean image segmentation. Thereafter, groups of elemental images formed in the last step are encoded by 3D-DCT technique so as to de-correlate the redundancy in both inter and intra-elemental image dimension simultaneously. Finally, the classical quantization and entropy coding processes are executed to complete the compression scheme.

This method applies an adaptive algorithm to segment the array of elemental images, and uses the 3D-DCT algorithm to eliminate the correlation intra-EI and inter-EI. In this way, it succeeds to improve the compression efficiency of MPEG-based method. However, as in the case of the MPEG-based technique, the temporal correlation between successive IIs is still not considered. Also, it is unable to explore the correlation among adjacent EIs in the other direction apart from the scanning direction.

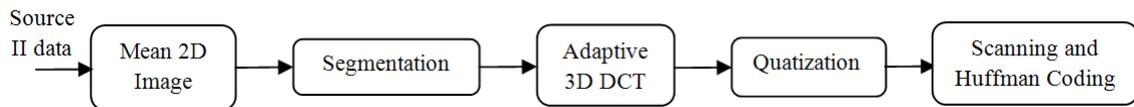


Figure 4.7: Scheme of 3D-DCT-based compression method

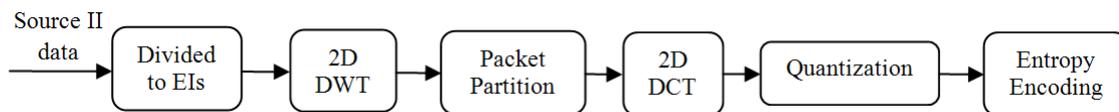


Figure 4.8: Scheme of hybrid compression method using DCT and DWT

4.3.3 Hybrid compression method using DCT and DWT

Discrete-cosine transformation (DCT) is applied in both last compression approaches to convert the spatial correlation into frequency domain. In 2007, a novel integral image compression method applying both the DCT and DWT was proposed [ESHJ07].

The process of this approach is illustrated in Figure 4.8. In this method, the four dimensional correlation within one integral image are exploited by a two steps pixel transformation process. Firstly, a 2D-DWT transformation is applied in each EI so as to remove the redundancy within one EI. Then, the wavelet coefficients in the same energy band from every EI are collected together to form several data blocks through a ‘packet partition’ process; the scheme of this procedure is specified in Figure 4.9. These composed data blocks, consisting of similar groups of DWT coefficients represent the redundancy of adjacent elemental images. Finally, the grouped wavelet coefficients are de-correlated by a classical 2D- DCT process.

According to the simulation results, this hybrid DCT and DWT method achieves a better rate-distortion performance because of the thorough elimination of the four dimensional correlation within one II. However, it is worth mentioning that this evaluation is derived from encoding an II with small numbers of EI but a large number of pixels of each EI. In this case, it is logical to implement 2D DWT on elemental images with high resolution. However, in practical applications, this kind of integral images produce a very coarse 3D image on the display side [Agg11b]. For 3D visualization applications, a large number of EIs with pixel number are required for an integral image. In this situation, it is not suitable to apply the 2D DWT on each elemental image. As a result, the hybrid DWT and DCT approach is not a good choice for encoding integral image. Moreover, a 2D DWT can be applied only on two dimensional EIs, which excludes its application for mono-dimensional EI arrays.

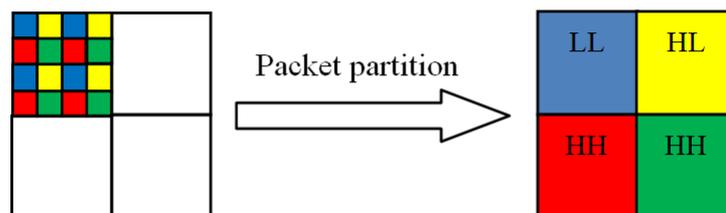


Figure 4.9: Scheme of Packet partition process

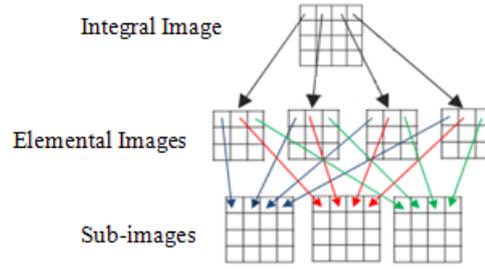


Figure 4.10: Sub-image generation scheme

4.3.4 Karhunen-Loeve transform (KLT)-based compression technique

All three methods presented previously are applied on elemental image. In these approaches, a single EI is considered as a unit of compressing process. Therefore, the performance of these methods is impacted by the number of EI. The calculation time will increase relatively for the II sequence with large number of EI.

Due to the high compression cost, we envision another format of image to represent the original II. Sub-images are considered for this purpose. A sub-image is composed of the pixels from the same position of each EI. The resolution of a SI is equal to the number of lenses while the number of SI is the same as the resolution of one EI. Since in a typical 3D display the total number of lens is a few orders higher than the resolution of each EI, so that SI encoding brings remarkable cost reduction in compression process.

Sub-image characteristic

Figure 4.10 illustrates the concept of generating sub-images from integral image. The pixels at the same position in the EIs are extracted to compose the corresponding sub-image. Since there are 16 elemental images in Figure 4.10, each EI consists of 3×3 pixels so 9 different sub-images will be generated with size 4×4 , repeat this work for every II and finally generate a sequence of SIs for the entire II series. For the focused mode integral imaging system, pixels from the same location of the EIs record the intensity and color of parallel light rays. For this reason, all the pixels in a given sub-image represent the perception of a 3D scene from a specific viewing angle and different SIs possess different viewing directions.

Among the generated SIs, the SI located in center position is denoted as reference image while the rest images are side SIs. For example, as illustrated in Figure 4.10, the SI in red is applied as the reference image, which is composed by the centric pixels of each EI. While the other two SIs represented in blue and green are side SIs.

Process of Karhunen-Loeve transform (KLT)-based method

The generated SI records the parallel light ray information of the 3D scene. Based on its character, the authors in [KSK10] propose a KLT-based method for encoding motion compensated SIs.

Figure 4.11 depicts the overall flow chart of the proposed approach, which is realized in four steps. Firstly, the original array of elemental images is rearranged to from the set of sub images by use of the transforming rule reviewed in Section 4.3.4. Then, a motion vector extraction process is applied to the rest SIs for extracting their motion value from reference SI. After that, with the calculated motion vectors, the displacement between the reference SI and the rest SIs are compensated via motion compensation process. Finally, the Karhunen-Loeve transform (KLT) algorithm is applied to compress these motion-compensated sub-images.

The advantages of this method are of two kinds. Firstly, since the compression process is carried out on sub image instead of elemental image, the inefficiency problem from enormous number of EIs is solved. Secondly, thanks to the higher similarity among the sub-images than traditional elemental image, this method succeeds in improving the compression efficiency of integral images. In spite of its out-performance compared to previous encoding methods, the KLT-based approach is still constrained to the single integral image with a static scene, and does not involve instances of moving pictures. Moreover, the high computational complexity of the KLT algorithm blocks the practical employment of this method.

4.4 Proposed integral image compression methods

Three integral image encoding methods are proposed and discussed in this section to improve the compression efficiency and overcome the limitation of existing encoding methods.

4.4.1 Multi-view Video Coding (MVC)-based compression method

According to the particular structure of SIs introduced in Section 4.3.4, we can detect two helpful features of SIs for efficient compression compare to EI [PKL05], which are explained by figure 4.12. The first one is the fact that the perception angle of each sub-image is a constant value that is independent from the 3D position of the object, and the corresponding SIs from successive integral images always possess the same viewing angle. For example, the sub-images collecting the information of slope light rays in Figure 4.12 record the 3D scene from the viewing angle as follows:

$$\theta = \tan^{-1} \frac{y_i}{f} \quad (4.1)$$

where f is the focal length of the convex lens, y_i is the distance between the i -th pixel and the optical axis in the EI.

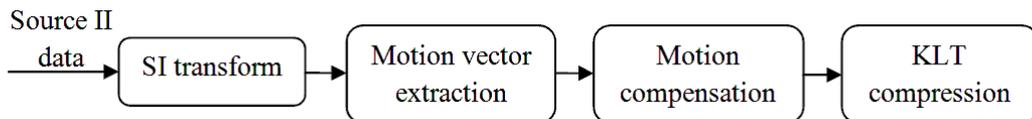


Figure 4.11: Block diagram of KLT-based encoding method

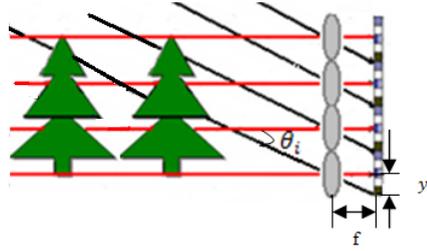


Figure 4.12: Viewing angle and object size are invariable for corresponding sub-images

The second useful characteristic of SIs is that the object sizes from corresponding SIs with same viewing angle are always invariant and also independent from the object 3D location. For example, the size of the 3D object represented by the horizontal light rays in Figure 4.12 is independent of its actual location. The valuable properties presented above can be exploited to compress the series of SIs using MVC [SSG11].

Figure 4.13 illustrates the scheme of the proposed integral images compression approach. The whole procedure of this method can be divided into three steps. In the first one, if the integral image is rendered directly from the 3D scene, it is necessary to generate the sub-image through created II by the regular transformation introduced in Section 4.3.4. Otherwise if the initial rendered image is already in format of sub image, we can skip over the transform step, and go directly to the encoding procedure. Secondly, the generated SIs are properly arranged as a multi-view video (MVV). Finally, the organized MVV is encoded by multi-view video coding (MVC). Correspondingly, in the reconstruction part, we need to rearrange the decoded SIs into II and then reproduce the 3D scene with the reconstructed II.

Figure 4.14 illustrates the process of arranging a sequence of SIs into a multi-view video sequence. The SIs from different viewing angles are regarded as the multi-view images in spatial domain and the successive SIs with the same viewing angle constitute the video sequence in temporal domain. Sub-images from different viewing angles are scanned in spiral order in order to minimize the motion compensation between them. For example, in Figure 4.15 the SI made up of the center pixels from each elemental image will serve as the reference view denoted as V_0 . The remaining SIs, denoted as V_1, V_2, \dots, V_8 , are composed of the other pixels of EIs. These nine different sub-images are scanned successively in spiral order, and form the multi-view

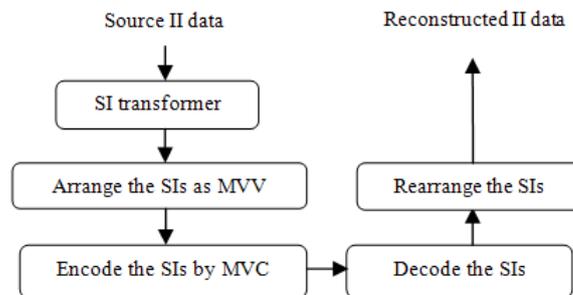


Figure 4.13: Flow chart of the proposed integral images compression system

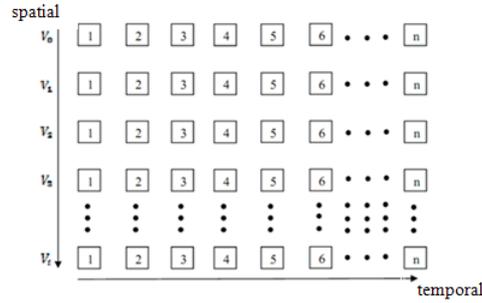


Figure 4.14: Converting SIs into MVV

images in the spatial domain.

Once the sub-images are rearranged as MVV, the multi-view video coding standard is used to encode this video sequence. We employed the JMVC_2_1 software from the Multi-view Video Coding (MVC) project of the Joint Video Team (JVT) and the ITU-T Video Coding Experts Group (VCEG) to encode our SIs sequence. The structure of the employed MVC scheme is shown in Figure 4.16. In this coding algorithm we combine the hierarchical B prediction in the temporal domain with the inter-view disparity estimation in the spatial domain and apply the inter-view prediction for all the pictures in the video.

This method can completely eliminate the redundancies in one II sequence and significantly improves encoding efficiency, but its application is limited by the computational complexity and the limited number of input views of the MVC standard.

4.4.2 Multi-view Video plus Depth (MVD)-based compression method

Another integral image compression method based on Multi-view Video plus Depth (MVD) representation is presented in this section. Like the MVC-based encoding approach, this method is applied on sub-image. The principle of this method is to use the 3D information existing in depth maps to predict the SIs for various viewing direction and then remove the inter-view correlation between original SIs.

As mentioned before, the source data of the proposed system is MVD video. This video format is composed of several sequences of color images and their associated depth images. Conventionally, these color images and depth images are recorded by ordinary cameras with perspective projection mode. However, the SIs applied in our

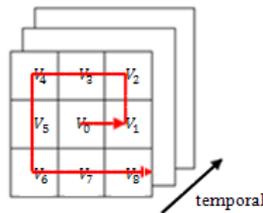


Figure 4.15: Spiral scanning topology for SIs

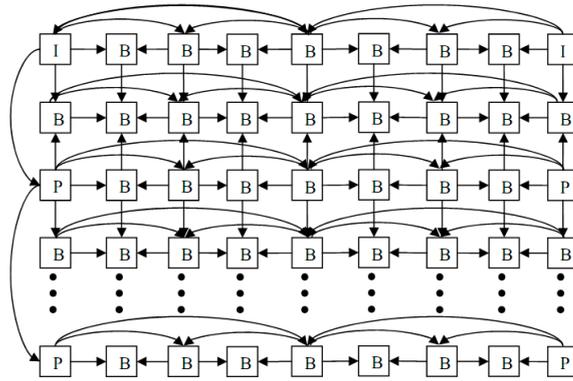


Figure 4.16: Scheme of MVC

system are generated by particular cameras with orthographic or oblique projection modes, which aim to record the parallel light rays reflected from the 3D scene. Similar to the MVC-based encoding approach, here the center SI is regarded as the reference SI, and the other SIs are denoted as side SIs. The source MVD video is composed of all the recorded SIs and a single serial of depth maps associated to reference SI.

For example, as illustrated in Figure 4.17, the light rays in red are recorded by an orthographic projection camera. Thus, the SI from this viewing direction is applied as the reference image, while the other two SIs representing light rays in blue and green are recorded by oblique projection cameras, employed as side SIs. In this example, the source MVD data is made of these three sequences of SIs and the depth maps of the reference view.

As shown in Figure 4.18, the compression process of this method is carried out in three steps. First, the original side SIs is predicted from the reference SI and its associated depth map via the DIBR technique. Then, the residual errors are calculated by subtracting the predicted SIs from their corresponding original SIs. Finally, the reference SI, its associated depth image and the residual error are encoded by classic H.264/AVC techniques.

The compression efficiency of this method largely depends on the performance of side SI prediction process. The detailed study of the prediction algorithm is given in

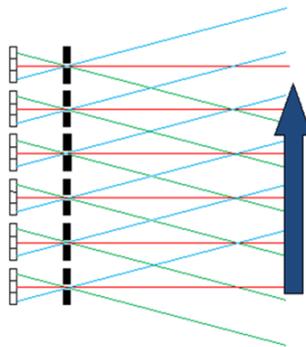


Figure 4.17: Concept of sub-image

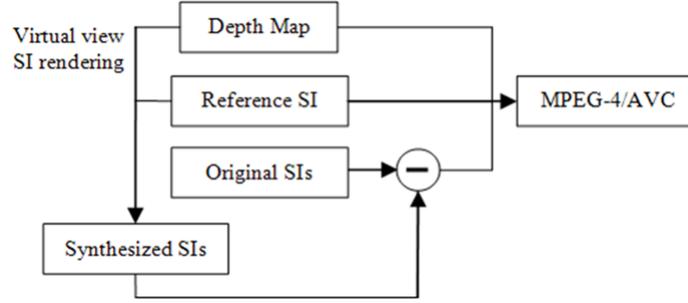


Figure 4.18: Flow chart of proposed compression method

the following paragraphs.

Conceptually, the side SI can be predicted by depth image based rendering algorithm, which can be achieved in two procedures. In the first place, a 2D pixel in one image is transformed into a 3D vertex via proper projection, then the 3D vertex can be re-projected into another 2D image through the inverse projection operation. However, the traditional DIBR algorithms introduced in [SGwHS98], [CW93] and [MB95] are provided to deal with the perspective projection mode, which is not suited to our proposed CII system. For this reason, a special DIBR technique adapted to the properties of sub-image is required in the proposed compression technique.

With the analysis of camera projection mode addressed in Figure 4.17, we note that for a given sub-image generation setup, there is no physical movement between any two cameras. All the virtual cameras are configured with the same extrinsic parameters during the SI generation process. Therefore, the diversification between these cameras lies uniquely in their intrinsic parameters, which are represented by the projection matrix.

In the same way as the conventional DIBR algorithm, the synthesis of side SIs is also achieved in two steps. Firstly, the 2D points in the reference image are projected inversely into the 3D world, using the respective depth data and its own projection matrix. Then these 3D points are mapped onto the clipping plane of a side camera with its own projection matrix. The mapping effects of orthographic and oblique projection modes are reviewed in Figure 4.19.

As shown in Figure 4.19, through the orthographic projection, a 3D point (x_v, y_v, z_v) in camera coordinate system (X_c, Y_c, Z_c) is linearly mapped to a pixel (x_p, y_p) in the clipping plane of the reference camera with the following projection matrix:

$$M_p = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.2)$$

According to the camera arrangement illustrated in Figure 4.17, a side camera could be regarded as the modification of the reference camera with rotated angles in horizontal and vertical directions. From the projection process shown in Figure 4.19, the relationship between the 2D pixel coordinate in the clipping plane of a side

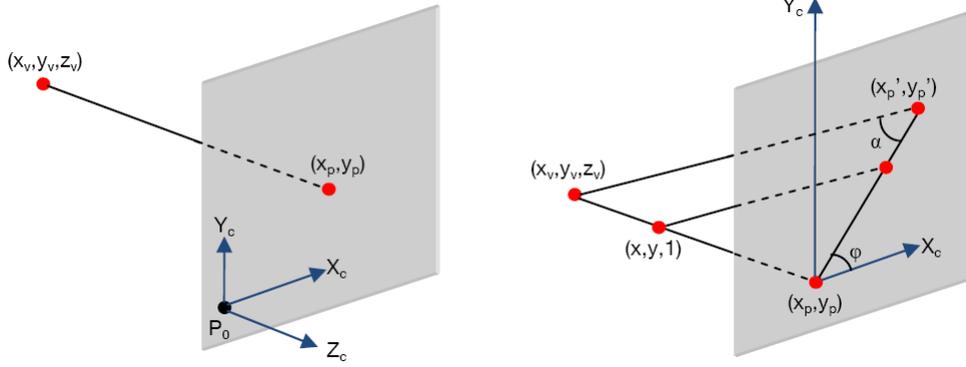


Figure 4.19: Mapping a 3D point to 2D clip plane by orthographic (left) and oblique (right) projection

camera and the 3D point in the camera coordinate can be expressed as:

$$x'_p = z \times \frac{1}{\tan(\alpha)} \times \cos(\phi) \quad (4.3)$$

$$y'_p = z \times \frac{1}{\tan(\alpha)} \times \sin(\phi) \quad (4.4)$$

where ϕ and α denote the optical axis rotation angles between the side camera and the reference camera for the vertical and horizontal axes respectively.

As a consequence, the oblique projection matrix can be written as:

$$M'_p = \begin{bmatrix} 1 & 0 & \cos(\phi)/\tan(\alpha) & 0 \\ 0 & 1 & \sin(\phi)/\tan(\alpha) & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.5)$$

To simplify the computation, only horizontal rotation is discussed in this paper, so the vertical rotation angle ϕ is set to 0, and the projection matrix of the side cameras could be rewritten as:

$$M'_p = \begin{bmatrix} 1 & 0 & 1/\tan(\alpha) & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.6)$$

Using the projection matrix of the reference camera expressed in Equation 4.2, we can project the 2D points in reference SI onto the 3D space as follows:

$$\begin{bmatrix} x_p \\ y_p \\ 0 \\ 1 \end{bmatrix} = M_p \times \begin{bmatrix} x_v \\ y_v \\ z_v \\ 1 \end{bmatrix} \quad (4.7)$$

where (x_v, y_v, z_v) represents the 3D space point in the reference camera coordinate system and (x_p, y_p) symbolizes the 2D pixel coordinate in the reference SI.

Substituting Equation 4.2 into 4.7 leads to the following formulas for calculating the 3D vertexes from 2D pixels in the reference SI, where z_v is derived from the depth image associated with the reference SI.

$$x_v = x_p \quad (4.8)$$

$$y_v = y_p \quad (4.9)$$

Then, we project the generated 3D vertex into the side SI with the known oblique projection matrix M'_p figured in Equation 4.6:

$$\begin{bmatrix} x'_p \\ y'_p \\ 0 \\ 1 \end{bmatrix} = M_p \times D \times \begin{bmatrix} x_v \\ y_v \\ z_v \\ 1 \end{bmatrix} \quad (4.10)$$

where (x'_p, y'_p) is the 2D image point coordinate in the side SI. D is the extrinsic transformation matrix between the reference camera and the side camera.

Since no extrinsic transformation exists between diverse virtual cameras, the transformation matrix D is the identity matrix. Substituting Equations 4.6, 4.8 and 4.9 into Equation ?? then leads to the following formulas which are used to generate an arbitrary SI with a known oblique angle:

$$x'_p = x_p + z_v / \tan(\alpha) \quad (4.11)$$

$$y'_p = y_p \quad (4.12)$$

Equations 4.11 and 4.12 define the 3D warping equations that can be used to effectively generate an arbitrary side sub-image in a horizontal direction from a known reference sub-image and its associated depth map.

This view synthesis is used as the inter-view prediction procedure in the proposed compression scheme. After the prediction process, the residual errors between original side SIs and their predictions are calculated and encoded by classic H.264/AVC schemes. While on the display side, the corresponding decoding method will be applied to reconstruct the original MVD video.

As opposed to MVC-based compression method, this approach applies the 3D information in depth map to predict the sub image. Since DIBR is outperform than direct disparity estimation in sub image prediction, this method is more efficient than MVC-based approach.

4.4.3 Discrete Wavelet Transform (DWT)-based compression method

The two methods introduced previously are based on discrete cosine transform (DCT) and applied on sub-images. Using DCT, the original signal in spatial domain can be transformed into frequency domain, and the energy of the signal is compacted

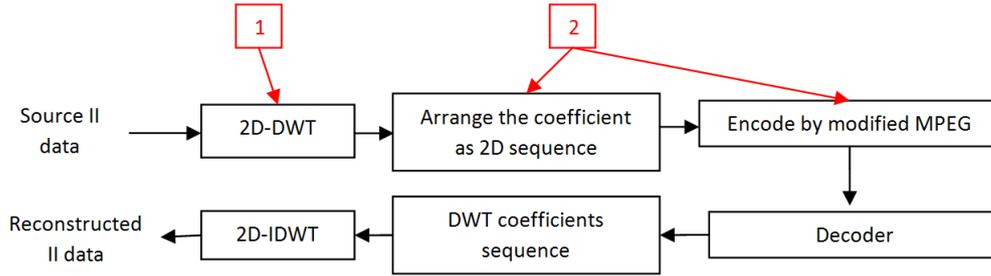


Figure 4.20: Flow chart of the compression scheme

in the low frequency part. The resulting coefficients of DCT can be compressed by quantization and entropy coding process. Due to its simplicity and high performance, DCT is a widely used and robust method for image compression. However, it also contains some undesirable defects. As we know, the input image needed to be divided into several ‘blocks’ before DCT, so the correlation across the block boundaries is not eliminated and the detail information of each block is discarded [rKG⁺12] [LK00]. For this reason, the ‘block effect’ artifact is detected in the reconstructed image, and the quality of decoded image is rapidly degraded with the decrease in bit rate.

To overcome the limitations of previous compression methods, we considered applying the discrete wavelet transform (DWT) instead of DCT for integral image compression. As opposed to DCT, the DWT process is no longer applied on sub-images but directly applied on original integral image.

As shown in Figure 4.20, the proposed method is carried out by two general steps. In the first step, several times of 2D-DWT is applied to one II. Through certain times of decomposition in both of the horizontal and vertical direction, the correlation of pixels within one EI can be fully exploited. Thereafter, one more step of DWT will be applied to remove the correlation between adjacent EIs. Once the four dimensional correlations in one II are completely decorrelated, a slightly modified MPEG-4 technique will be used to eliminate the temporal redundancy between consecutive IIs.

The objective of the first step is to remove the four dimensional correlation of one II in spatial domain. Choosing the optimal wavelets is an important procedure for DWT implementation. Haar wavelet is used in the proposed approach due to its simplicity and good property for evaluating the physical contrast of the original II [LK00]. The 2D Haar DWT is implemented in a separable fashion by successive one dimensional transformation along the x, y dimensions of the data. One $1 - D$ decomposition aims at splitting a signal of length N into two subsequences of length $N/2$: the lower resolution approximation (low-pass component) and the wavelet (high-pass) component. The splitting process is repeated on the lower resolution version of the signal until a predefined decomposition level is reached. The decomposition process is illustrated in Figure 4.21.

During the computation, we may transform first the columns of X , by multiplying by T , and then the rows of the result by multiplying by T^T . Hence the $2D$ Haar

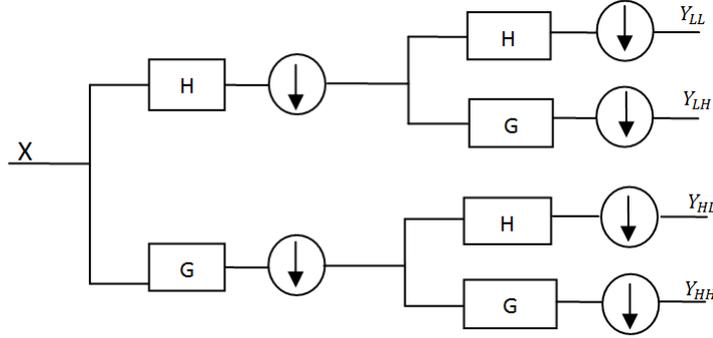


Figure 4.21: 2D-DWT decomposition

transform can be expressed mathematically in the following equation:

$$Y = T \times X \times T^T \quad (4.13)$$

$$T = \frac{1}{\sqrt{2}} \times \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (4.14)$$

Where T is the Haar transform matrix, X is the input image and Y is the decomposed coefficients. Performing this transform to the entire II, we can decompose the original image into four sorts of coefficients, including LL coefficients to represent the approximation of input image, HL coefficients contains the near-vertical edges, LH coefficients contains the near-horizontal edges and the HH coefficients contains the diagonal edges. According to the property of wavelet transform, most of the energy will be compacted in the LL coefficients, which need be further decomposed by another DWT.

In our implementation, the level of decomposition is determined by several pixels in one EI. Suppose that each EI possesses pixels, after N times $2D$ wavelet decomposition, we can remove the $2D$ correlation between adjacent pixels in one EI. Thereafter, the low pass coefficients obtained possess the redundancy between neighboring EIs, for removing these correlations one more step of DWT will be performed to LL coefficients, the final decomposition level is $N + 1$. The entire DWT process for one II is in illustrated in figure 4.22.

After performing $N + 1$ times Haar decomposition, the four dimensional correlations in spatial domain of one integral image is thoroughly eliminated. Thereafter, the inter-frame previous technique in MPEG will be applied to exploit the temporal correlations between consecutive IIs.

According to the particular conformation of the obtained coefficients, an adjusted MPEG-based scheme is applied to remove the inter-II redundancy. In this method, the wavelet coefficients are divided into two sets, as LL and the rest. Since LL coefficients contain the average information of the original image, which possess the majority of the energy, it will be encoded by a classical high efficiency MPEG-4 method. Otherwise, the remaining high pass coefficients containing the detail information will be

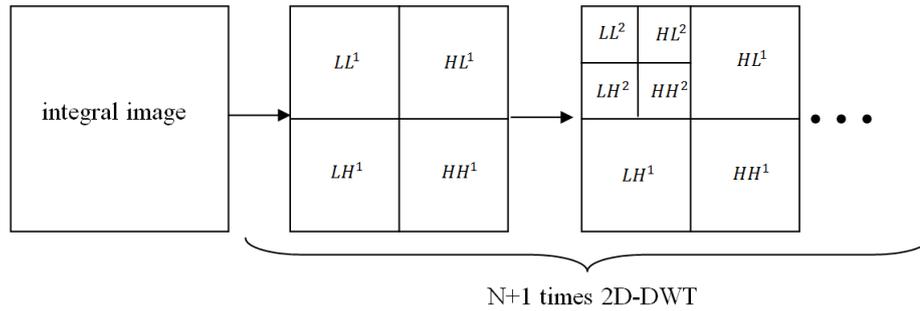


Figure 4.22: DWT process for one II

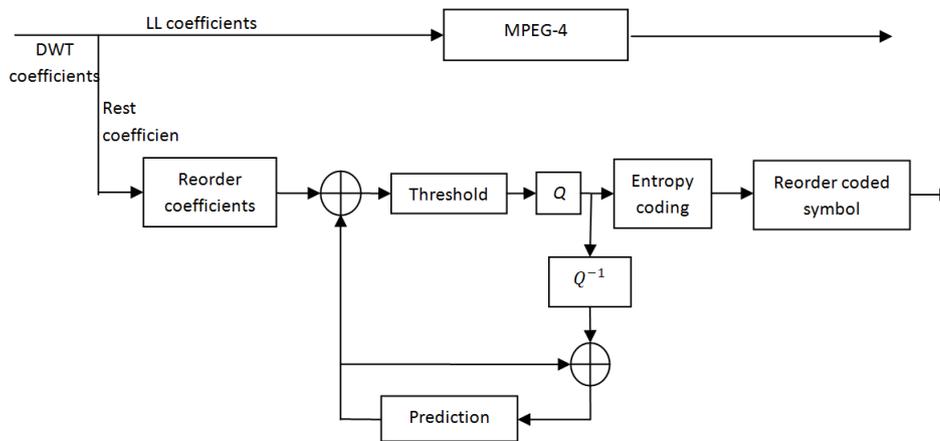


Figure 4.23: Hybrid MPEG-4 based method

encoded by a slightly modified MPEG-based method without DCT. The architecture of this hybrid encoding method is illustrated in Figure 4.23.

This method provides three important improvements. First of all, since the spatial and temporal correlations of one IIs sequence can be eliminated by using of DWT and MPEG respectively, it is able to completely remove the redundancy in an integral image sequence. Secondly, it solves the ‘block effect’ problem in the DCT-based method. Finally, since each II is treated entirely in this method, it is no longer constrained by the view number limitation in MVC technique.

4.5 Evaluation and comparison of proposed methods

After the detailed description of three proposed encoding methods, the evaluation results of these approaches will be illustrated in this section with the test of two video sequences named as ‘spaceship.yuv’ and ‘cube.yuv’. In this thesis, the performances of proposed compression methods are evaluated by PSNR value in function of the bit-rate.

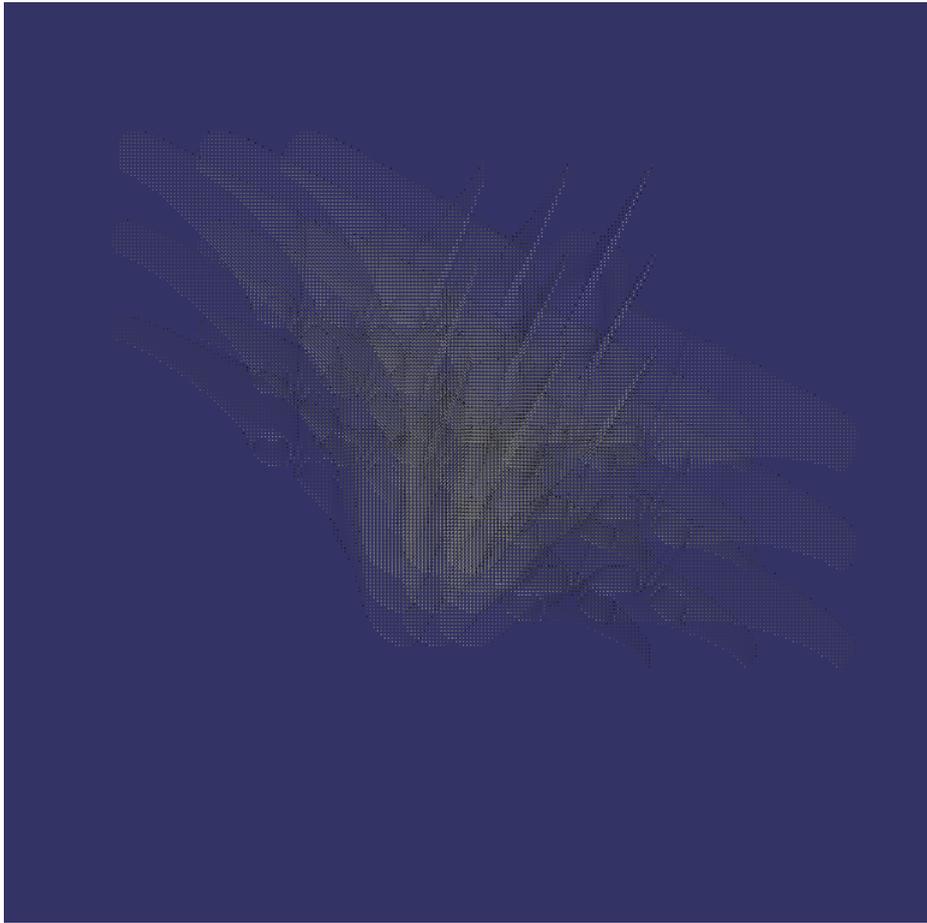


Figure 4.24: 'spaceship' original integral image

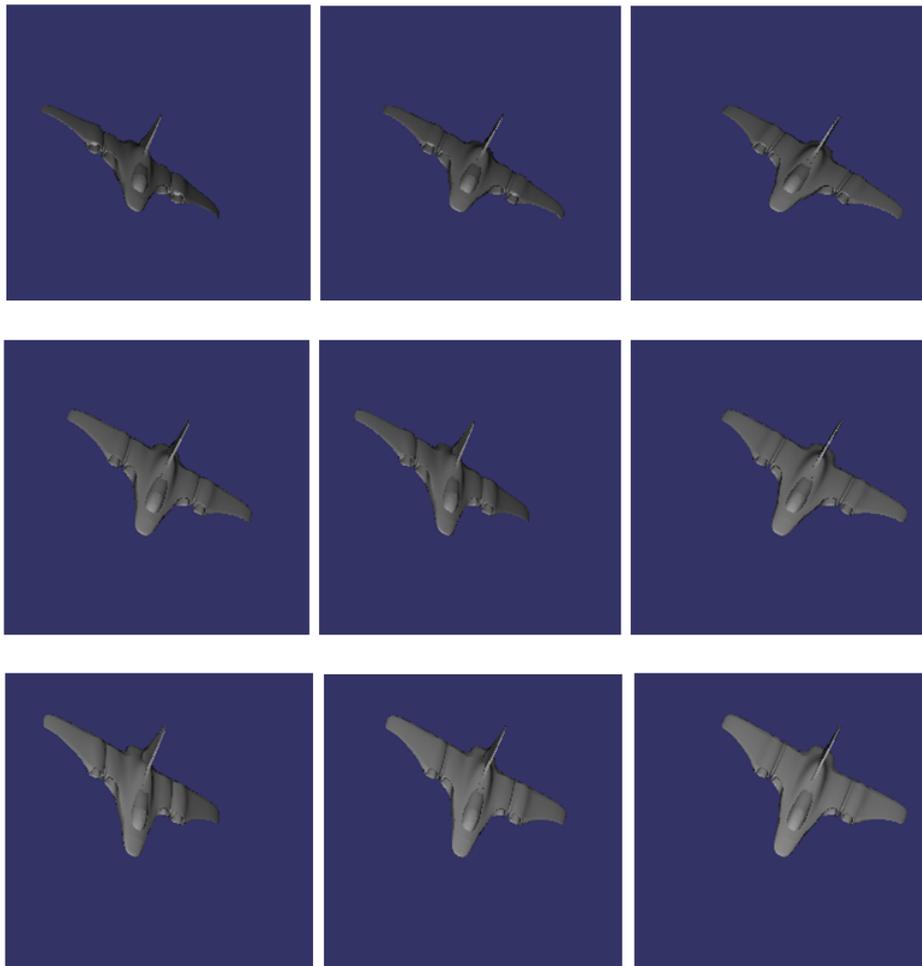


Figure 4.25: 'spaceship' generated Sub-images

4.5.1 Evaluation for MVC-based encoding approach

The ‘spaceship.yuv’ used in this experiment is a sequence of 30 frames integral images, where the resolution of II is 1536×1536 . Along with the 30 frames, the object moves in horizontal direction. The number of elemental images in each II is 512×512 where each EI contains 3×3 pixels. Through the transformation process explained in 4.3.4, the set of original elemental images can be transformed into 3×3 sub-images with distinct viewing angles. Each SI contains 512×512 pixels. Therefore, the final constructed multi-view video is composed by 9 sequences, each one possess 30 frames.

Figure 4.24 shows the image of original integral image sequence.

Figure 4.25 shows the generated sub-images from the integral image in figure 4.24. As we mentioned before, 9 SIs with different viewing angles were constructed in our simulation.

Figure 4.26 shows the comparison results between various algorithms. Three algorithms were tested and compared. The first one is the traditional method that encodes the sequence of integral images using MPEG-4/AVC. The second approach is also based on MPEG-4/AVC but specifically implemented for sub-images. In this method, we arrange the transformed sub-images as a sequence of moving pictures and then use MPEG-4/AVC to encode it. The last compression method is our proposed MVC algorithm. In this approach, the transformed sub-images are arranged as multi-view video and encoded by standard MVC.

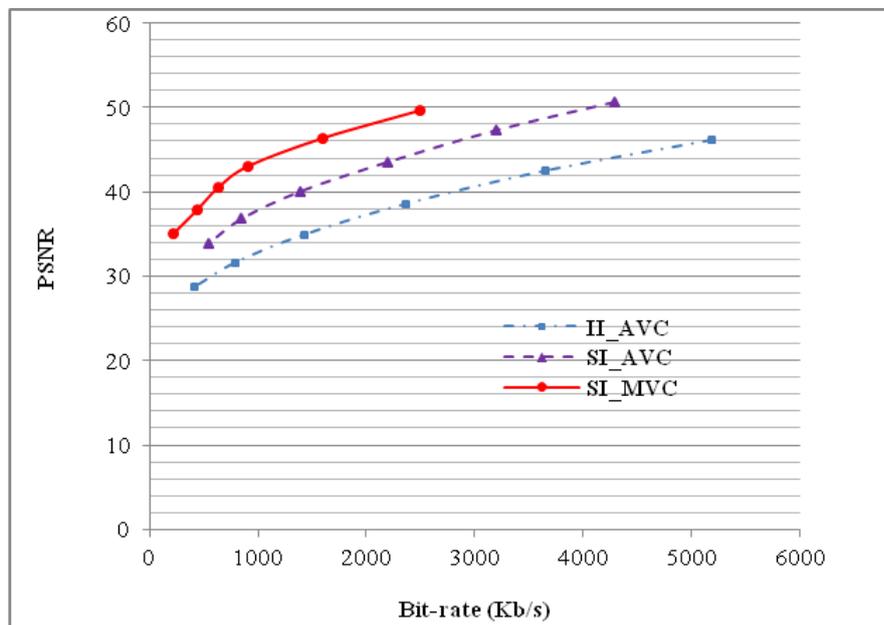


Figure 4.26: Performance of MVC compression scheme for SIs, AVC scheme for SIs and baseline AVC scheme for integral images sequence

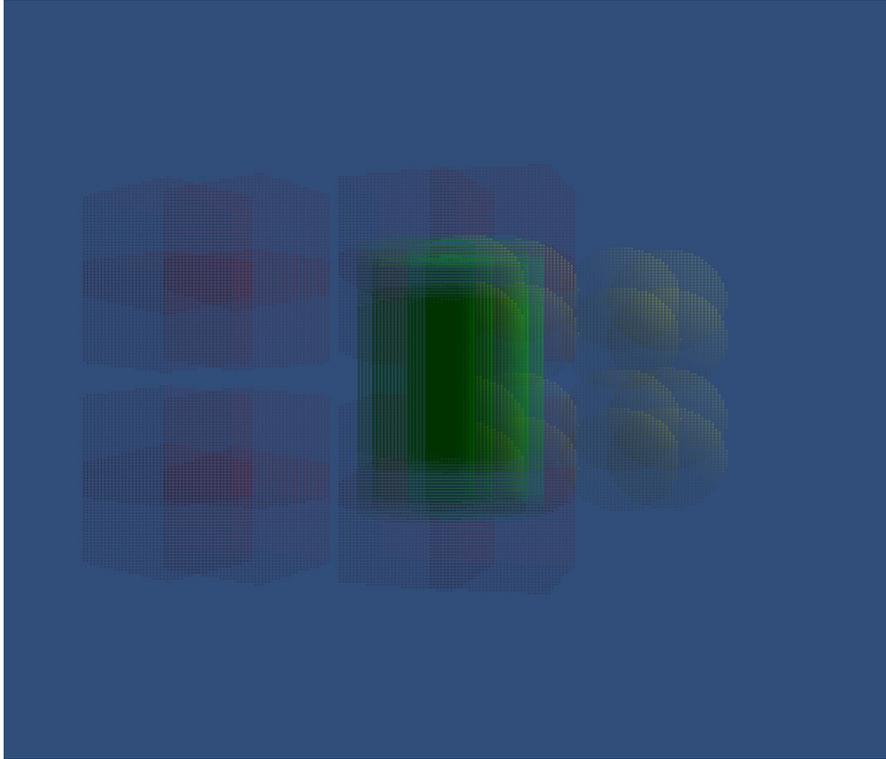


Figure 4.27: ‘cube’ original integral image

The compression efficiency of the algorithms is evaluated by the curve representing the PSNR in function of the bit-rate. According to Figure 4.26, we can see that the MPEG-4 based algorithm for sub-images outperforms the regular MPEG-4 approach for original integral images because the transformed SIs contain more redundancy than initial integral images. Furthermore, our proposed MVC algorithm performs better than the other two schemes. This demonstrates the advantage offered by the correlation removing process among inter-view sub-images.

Another example to evaluate the performance of MVC-based algorithm is performed in the second video content, (*e.g.*, ‘cube.yuv’), which processes more several sub-images than the first tested video sequence. As shown in Figure 4.27, the originally generated II contains 4×4 sub-images; the size of each SI is 256×256 . So the original II can be transformed into 16 sub-images with different viewing angle (cf. Figure 4.28).

Figure 4.29 highlights the evaluation results of the MVC based algorithm implemented in the video ‘cube.yuv’. Like the previous experiment executed in ‘space-ship.yuv’, the MVC-based method largely outperforms the other two. Furthermore, because of the larger number of sub-images, the PSNR gain of MVC-based algorithm in this experiment is more significant than the first test.

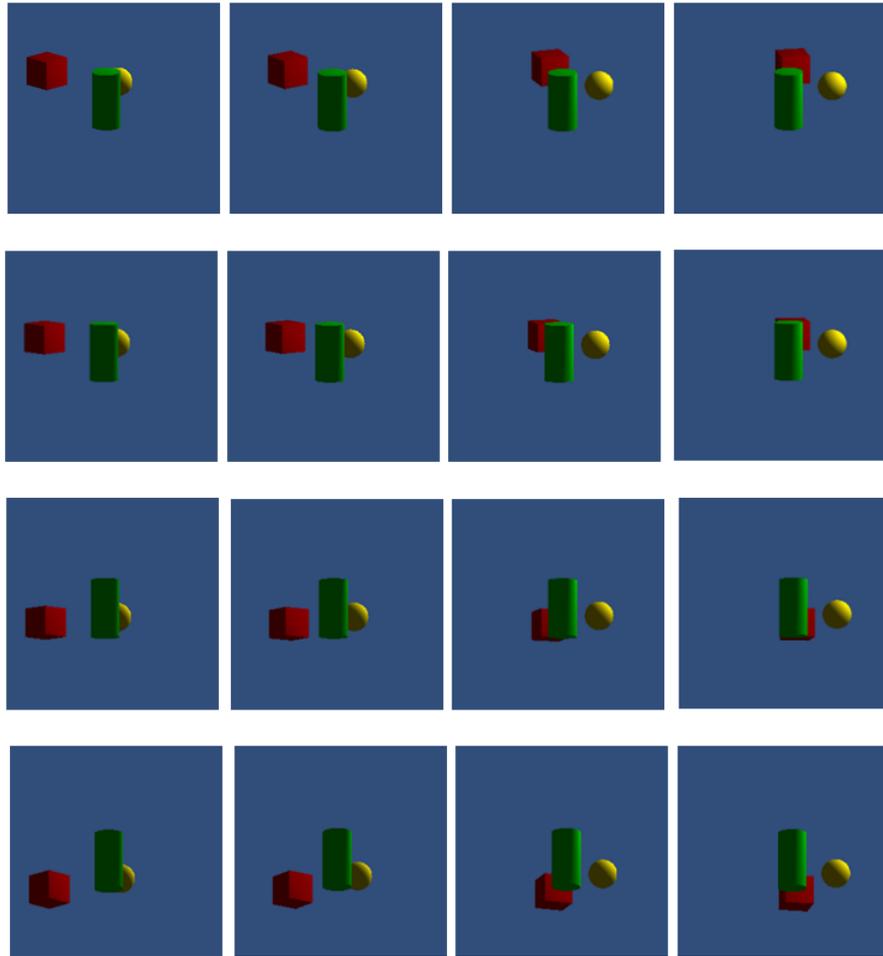


Figure 4.28: 'spaceship' generated Sub-images

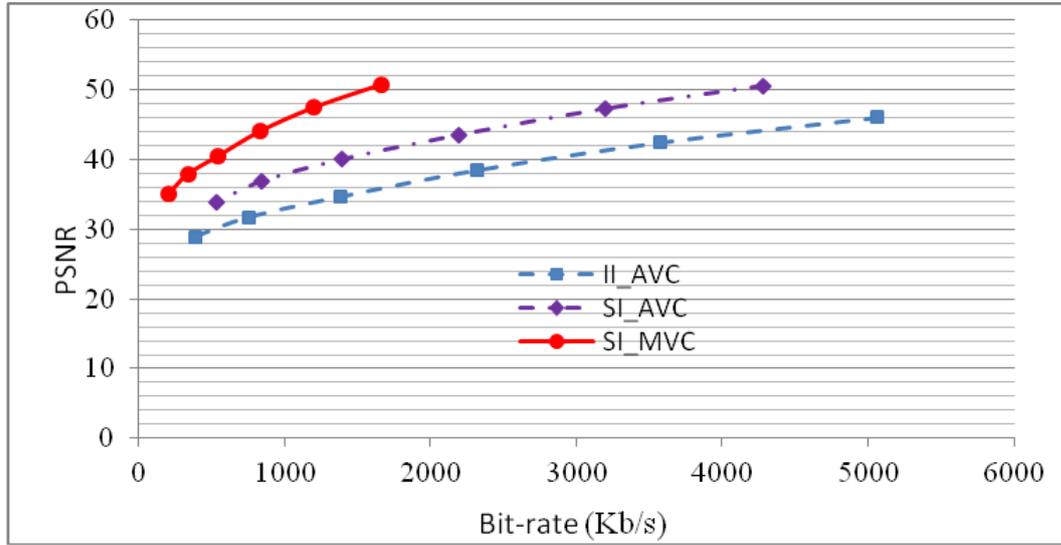


Figure 4.29: Performance of MVC compression scheme for SIs, AVC scheme for SIs and baseline AVC scheme for integral images sequence

4.5.2 Evaluation for MVD-based encoding approach

As in the precedent experiments, we have compared the results of three algorithms in this test. The first one, denoted as ‘II_AVC’ in Figure 4.30 describes the traditional method that encodes the sequence of integral images by H.264/AVC. The curve denoted as ‘SIs_MVC’ in Figure 4.30 represents the MVC-based compression method, which arranged the original sub-image sequences as multi-view video and encoded the sequence of SIs by MVC. Joint Multi-view Video Coding (JMVC) reference software was applied in the second test with the same temporal configuration and an ‘IBP’ interview prediction structure for the multi-view SI sequence. The last one is our proposed method based on MVD representation, which is labeled as ‘SIs_MVD’. H.264 JM 17.3 software is applied to encode the depth maps, reference SIs and generated residual errors.

To reduce computational complexity, only horizontal parallax is implemented in our simulation. The 3D object used in our simulation is represented by an existing 3D file that stores all the 3D information, including depth value of the object. Therefore, no depth estimation is needed in the experiment to generate the depth image from sub-images. The depth map is derived directly from the depth value stored in the 3D file. As shown in Figure 4.30, three SI sequences (figures 4.30(b), 4.30(c) and 4.30(d)) were generated in our experiment. The reference SI shown in Figure 4.30(c) is recorded by an orthographic projection camera. While the side SIs captured by two oblique projection cameras are shown in Figures 4.30(b) and 4.30(d). The depth image (Figure 4.30(e)) is also generated by an orthographic projection virtual camera in order to record the depth value of the corresponding points in the reference SI. Here the OpenSceneGraphic (OSG) graphic library is applied to simulate the virtual cameras used in our experiment. All of the generated video sequences are composed of 30 frames with 512×512 resolution.

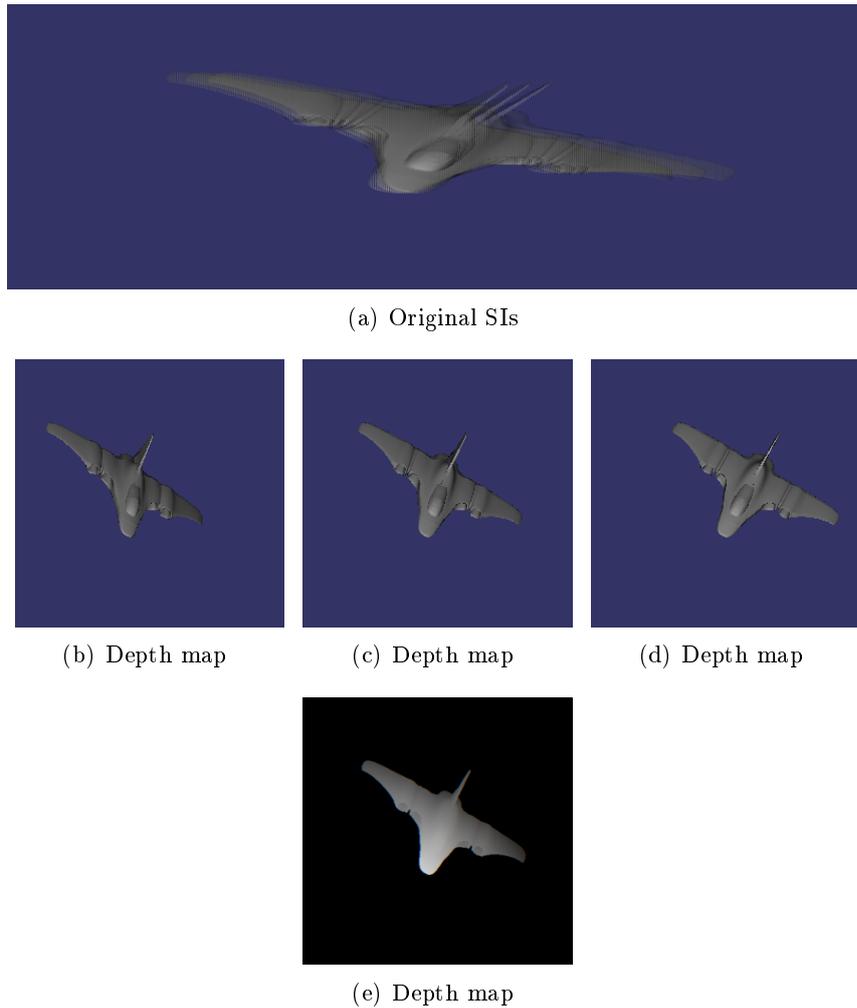


Figure 4.30: Original SIs and depth map

Thereafter, the left and right viewing SI sequences were predicted successively by the DIBR algorithm introduced in Section 4.4.2 (shown in Figures 4.31(a) and 4.31(b)). Then the residual data for each SI were calculated by subtracting the predicted SI from its corresponding original image (cf. Figures 4.31(c) and 4.31(d)). After the de-correlation process of inter-view SIs, the source data can be divided in three groups, i.e. reference SIs, depth images and generated residual errors. Then the H.264 reference software JM 17.3 is applied here to encode all the three groups of video sequences respectively. Thanks to the high correlation between successive SIs, only a small quantity of bits is needed to represent reference SI sequence and its associated depth sequence. In our experiment, the bit-rate of encoded depth sequence ranges from 39 Kbit/s to 156Kbit/s for different quantization parameters (QP).

Figure 4.32 illustrates the comparison results between the tested algorithms. According to this Figure, we can see that the MVC-based algorithm used for SIs outperforms the regular MPEG-4 approach used for integral images because it can fully remove the five dimensional correlations in the II sequence. At the same time, our

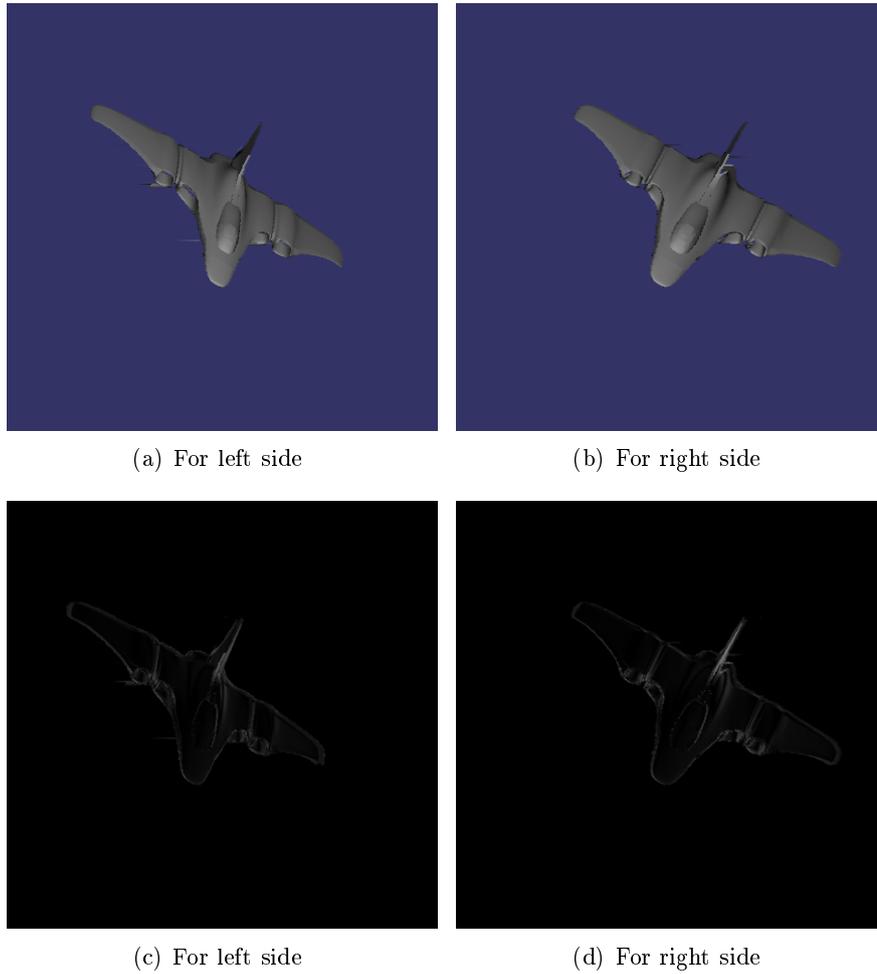


Figure 4.31: Synthesized SIs (up) and Errors (down) between original SIs and synthesized SIs

proposed algorithm based on MVD representation performs better than the two other schemes. This demonstrates that the geometry-based prediction can remove the inter-view redundancy between SIs more thoroughly.

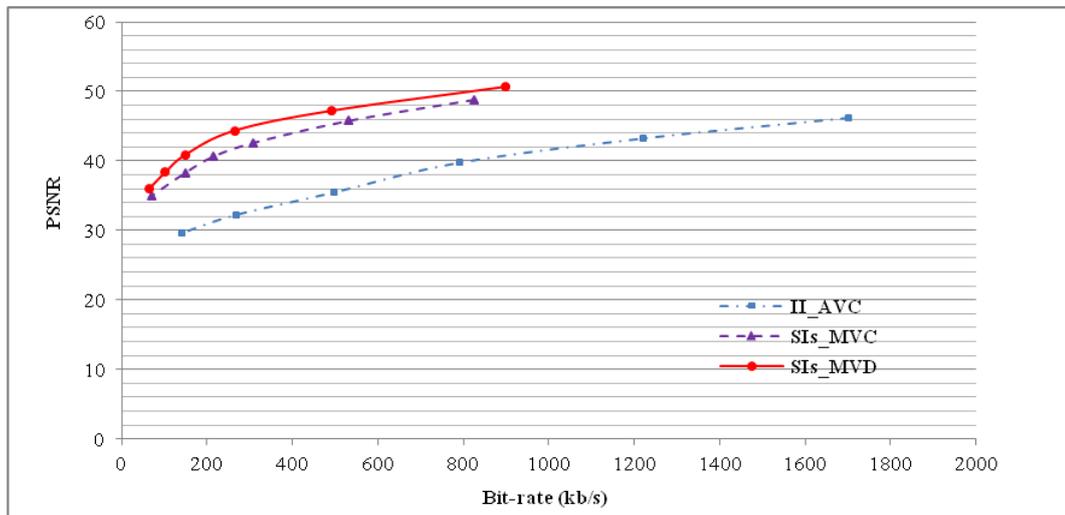


Figure 4.32: Performance of proposed compression method (labeled ‘SIs_MVD’), MVC-based method (labeled ‘SIs_MVC’) and baseline H.264/AVC method (labeled ‘II_AVC’)

4.5.3 Evaluation for DWT-based encoding approach



Figure 4.33: ‘spaceship’ DWT coefficients

DWT-based algorithm is evaluated by two video sequences: ‘spaceship.yuv’ and ‘cube.yuv’. In the simulation for ‘spaceship.yuv’, the source sequence comprises 10 frames of integral images, within which the 3D object moves in horizontal direction. Each original II is composed by 128×128 elemental images, where each EI contains 8×8 pixels.

In this simulation, the original II is decomposed by Haar wavelet for four levels. Then we can obtain a transformed image with 56×56 LL coefficients, shown in Figure 4.33.

Then the modified MPEG-based encoding scheme is applied to compress the transformed coefficient sequence. In the modified encoding method, DCT process is only realized for LL coefficient while the rest coefficients are skipped from this process and directly quantized.

Just like for the previous tests, the traditional AVC method and the first proposed MVC-based method are applied as the reference to evaluate the performance of this compression approach. According to the results shown in Figure 4.34, DWT-based algorithm is able to improve the PSNR value in the low bit-rate than proposed MVC-based method.

Another experiment is performed for ‘cube.yuv’; the same original sequence applied in the MVC-based method is reused here. As shown in Figure 4.27, this II sequence is composed of 30 frames, where each II is composed by 256×256 elemental images and each EI possesses 4×4 pixels.

In this case, three stages of DWT decomposition are needed to exploit the cor-

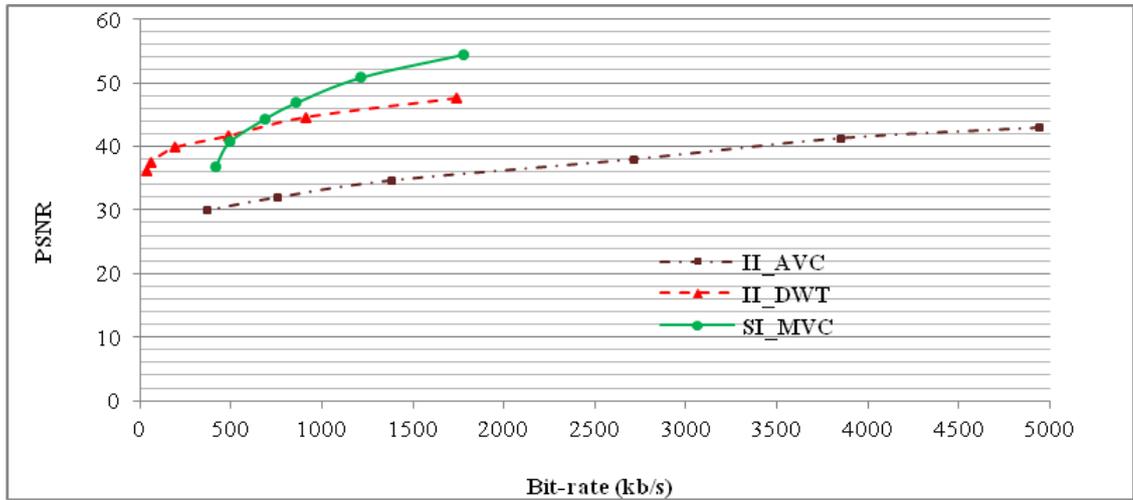


Figure 4.34: Performance of proposed compression method (labeled ‘SIs_MVD’), MVC-based method (labeled ‘SIs_MVC’) and baseline H.264/AVC method (labeled ‘II_AVC’)

relation in neighboring EIs. After three times of two dimensional transformations, a coefficient sequence with 128×128 LL is generated (cf. Figure 4.35). Then, this coefficient sequence is encoded by the modified H.264 encoder. As in the first simulation, DWT works better than MVC-based method in low-bandwidth situation, but in high-bandwidth compression, MVC still outperforms DWT.

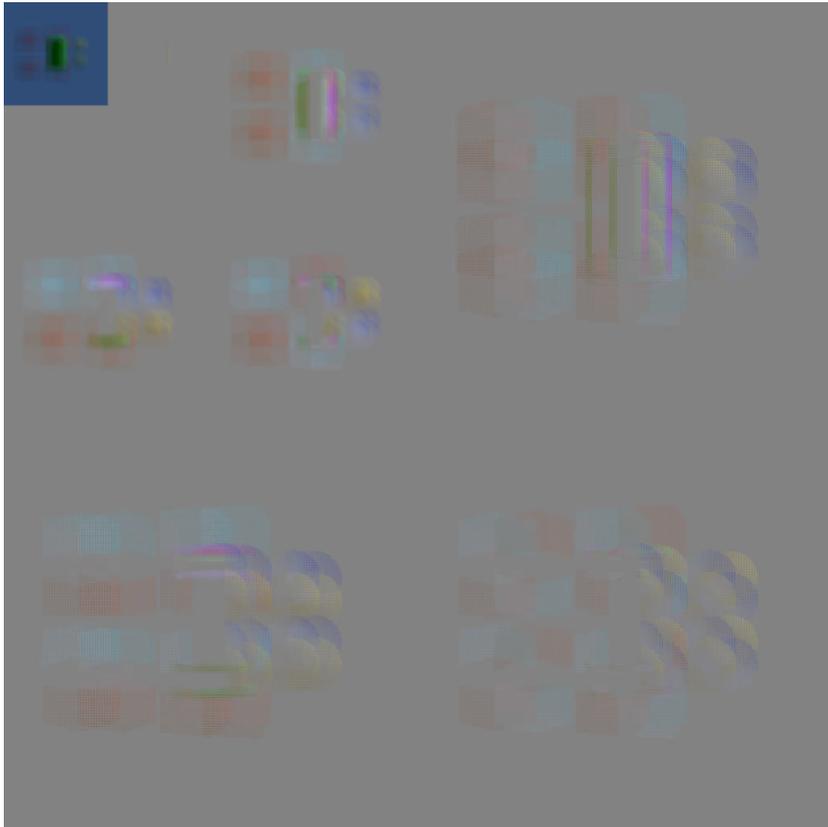


Figure 4.35: 'cube' DWT coefficients

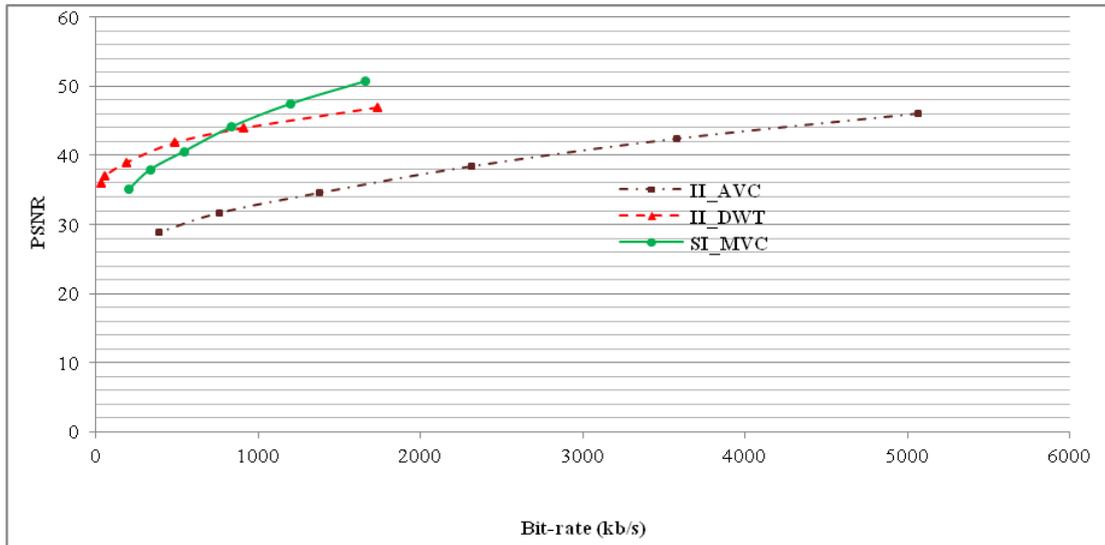


Figure 4.36: Performance of proposed compression method (labeled ‘SIs_MVD’), MVC-based method (labeled ‘SIs_MVC’) and baseline H.264/AVC method (labeled ‘II_AVC’)

4.6 Conclusions

In this chapter, we have reviewed the existing compression techniques for computation integral imaging system. Their schemes and problems have been exhibited and discussed. Then, to overcome the limitations of previous approaches, new encoding methods have been developed and implemented.

In the MVC-based encoding method, for exploiting the five dimensional correlation in an integral image sequence and resisting the calculation complexity, sub-images were studied and employed. According to previous research, sub-image contains several advantages for compression, on original integral image. In order to adopt these useful features for video compression, we proposed to arrange the sub-images as multi-view video and encode the generated MVV using MVC standard. The experimental results show that our proposed algorithm is more efficient than the other two compression approaches base on MPEG-4/AVC.

Thereafter, another compression method applied to sub image based on multi-view video plus depth representation was proposed and demonstrated. The invention of this system is using the three-dimensional information existing in depth map to remove the correlation between SIs. In the compression procedure, a dedicated depth-based image rendering method is provided to predict side SIs from a known reference SI and its associated depth map. Due to more thorough correlation elimination, the novel approach is more efficient than the other two compression approaches based on MPEG/AVC and MVC.

Besides the two approaches based on DCT algorithm, we proposed a new encoding scheme using DWT technique. This method directly transforms the original integral image into sets of coefficients at different frequencies. It provides substantial improvements in picture quality at high compression rates due to better energy compaction property of the wavelet transform.

The following table gives a general summary of the three proposed encoding methods, highlight their qualities and lacks in several various aspects:

	Efficiency	Cost	Stability	Constraint	Manageable Size
MVC	Medium	Low	High	Input number	Medium
DWT	Medium	Medium	High	None	High
DIBR	High	High	Low	Occlusion	Medium

Table 4.1: Comparison of proposed integral image compression methods

Chapter 5

Computational Integral image Reconstruction

This chapter addresses the design of computational integral image reconstruction (CIIR). We first highlight the existing challenges in an optical integral image reconstruction (OIIR) system, and describe the advantages of the CIIR system. Then, two basic CIIR modes known as view-based CIIR and plane-based CIIR are introduced and interpreted. Thereafter, the principles and limitations of the previous CIIR methods are reviewed followed by an explanation and evaluation of two new CIIR approaches.

The first method is realized by use of the parallel-group projection (PGP) technique. Similar to the parallel-group rendering (PGR) technique provided in section 3.2, it applies the parallel property of sub images to improve the efficiency of 3D image reconstruction. The second method is achieved by applying the sub image predicting technique used in the MVD-based compression method to synthesize the intermediate SIs, and the increase of number of SIs leads to an improvement of reconstructed 3D image quality.

5.1 Introduction

As highlighted in figure 5.1, in the optical integral image reconstruction process, the light beams pass through the display plate in front of the integral image and then converge at the point where the object had been, forming a 3D image.

In an OIIR system, the displayed 3D effect is not only impacted by the quality of integral image but also influenced by the diffraction and limitations of optical devices. In order to avoid the unstable factors of optical aperture, the CIIR system is produced and widely studied recently. The objective of CIIR is to simulate the projection effect of optical rays of EIs in OIIR using a computer built virtual lens or pinhole array.

Thanks to the improvement of computer calculation capacity, the performance of CIIR systems has progressed dramatically, but there still exist some limitations which obstruct its practical utilization, such as low reconstructed output image (ROI) quality and high calculation complexity [SLsK06] [KCKY12b]. In this chapter, two new approaches are proposed and tested to improve the performance of the CIIR

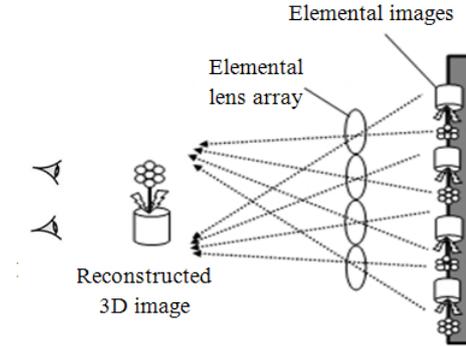


Figure 5.1: Optical setup for displaying 3D scene in integral imaging system

process.

The first approach is applying the parallel group projection technique in order to reduce the complexity of the CIIR process. Differing from previous techniques which project each point of integral image to the reconstructed plane pixel by pixel, the proposed method reconstructs the 3D image by mapping a series of sub images (SIs) onto the reconstructed plane successively, where each SI records pixels from parallel light rays with identical viewing angles. According to experimental results, this approach is able to significantly simplify the calculation of the reconstruction process.

Besides the PGP reconstruction algorithm, we also provide another depth image based rendering (DIBR)-based approach to improve the reconstructed image quality. The principle of this method is applying the DIBR algorithm proposed in chapter 4 to synthesize virtual SIs from a reference SI and its associated depth map. By use of the synthesized sub-images, a new array of elemental images with higher numbers of pixels is generated. Reconstructing the final 3D image with the new series of EIs produces a better 3D effect to the viewers.

This chapter is organized as follows. Section 5.2 investigates two types of CIIR techniques, view-based and plane-based CIIR, and describes their processes and applications. Section 5.3 presents some related work on plane-based computational integral imaging reconstruction, introduces the process of two previous CIIR methods, and illuminates their advantages and disadvantages. Then, two new techniques are proposed to improve efficiency and output image quality. Simulation results and detailed analyses are reported in section 5.4. Finally, conclusions and discussions are drawn in section 5.6.1.

5.2 Two types of CIIR techniques

As we remarked before, the objective of a computational integral image reconstruction approach is to simulate the output effect of an optical display system. These methods can be categorized in two types, denoted as view-based CIIR and plane-based CIIR [HJ01] [HJJ04]. As the name suggests, view-based CIIR generates the viewing image of the 3D scene for different viewing positions [GL12]. It reflects the real viewed effect

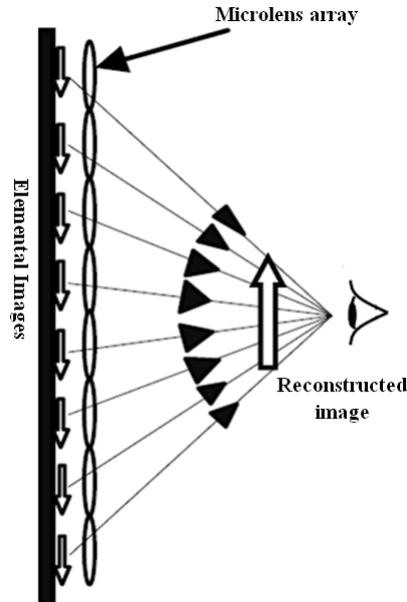


Figure 5.2: Principle of view-based CIIR

of the end-user. Plane-based CIIR reconstructs the 3D image for diverse output planes at different distances. It does not reproduce the observed image directly, but gives an integrative output image for arbitrary viewing distances.

5.2.1 View-based CIIR

As we mentioned, in the focused mode integral imaging system, each EI is regarded as one pixel for the viewer at the display side. Depending on this principle, view-based CIIR generates the 3D image for different viewing positions by extracting a single pixel from corresponding element images through a virtual pinhole or lens model so as to imitate the viewing effect of human eyes. Along with the movement of the viewing position, the observer can experience a natural perception of the 3D object with continuous parallax.

The principle of view-based CIIR is illuminated in figure 5.2. The end-user observes the series of elemental images through a computer built virtual pinhole array or lens array. One single pixel is viewed through each micro-lens according to the position of the observer. As a result, the resolution of viewed image is equal to the number of elemental images. Thanks to its direct representation of the 3D scene, the view-based CIIR method gives a direct evaluation of the 3D effect.

Two concrete examples of view-based reconstructed images for 3D objects ‘spaceship’ and ‘cube’ are illuminated in figure 5.3 and figure 5.4. In both of the two examples, the view-based reconstruction process is carried out for ‘x, y, z’ directions individually.



(a) Perceived image in x direction

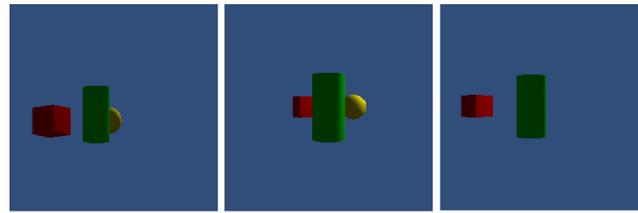


(b) Perceived image in x direction

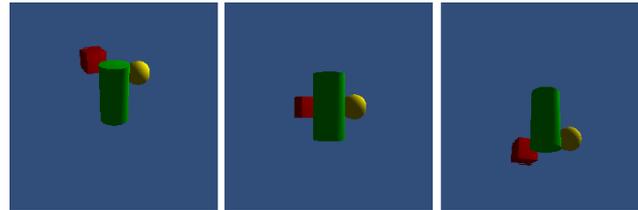


(c) Perceived image in z direction

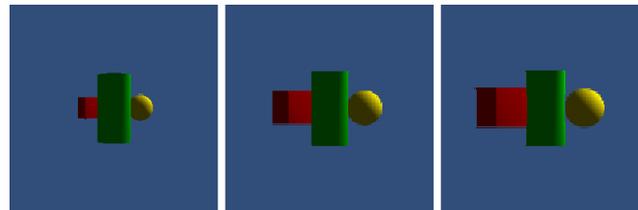
Figure 5.3: view-based reconstructed images for 3D object ‘spaceship’



(a) Perceived image in x direction



(b) Perceived image in x direction



(c) Perceived image in z direction

Figure 5.4: view-based reconstructed images for 3D object ‘cube’

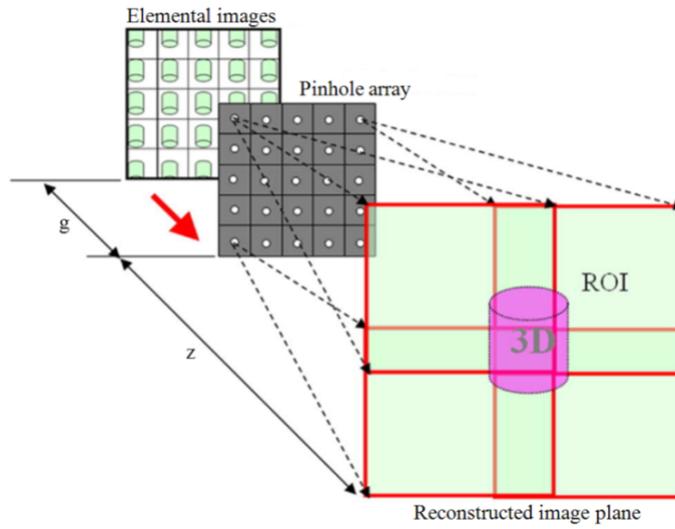


Figure 5.5: Principle of plane-based CIIR

5.2.2 Plane-based CIIR

Figure 5.5 describes the process of plane-based CIIR, which is realized in two basic steps. Firstly, each elemental image is projected individually through its corresponding micro-lens. Then, the projected elemental images are overlapped together to form a reconstructed output image at the given distance.

In this method, the reconstructed output image could be generated at any arbitrary distance from the virtual pinhole array or micro-lens array [SMsK05]. In this thesis, our research work is concentrated on the plane-based CIIR technique.

In order to demonstrate the result of plane-based CIIR, two concrete examples of reconstructed images are shown in figure 5.6. Here, the reconstruction is calculated for three different output image planes.

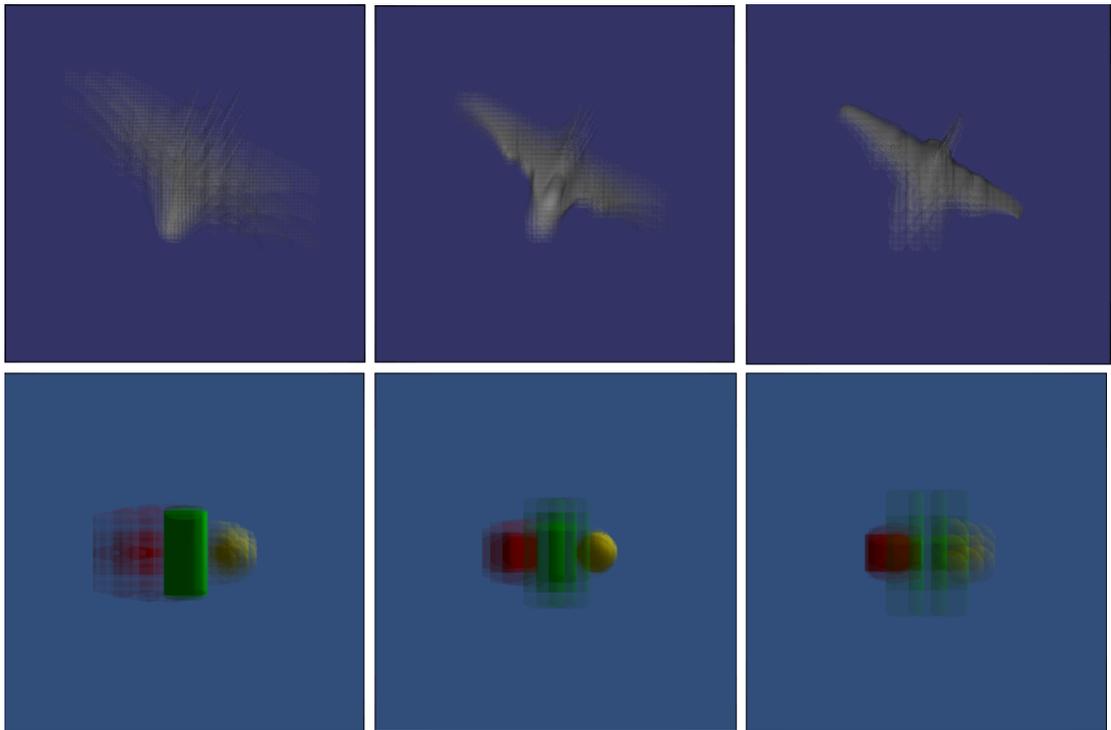


Figure 5.6: Example of reconstructed image by plane-based CIIR technique

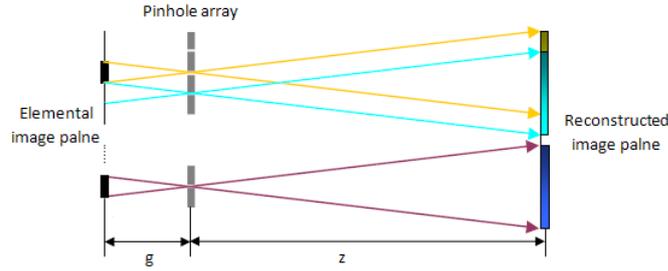


Figure 5.7: Reconstruction principle of traditional CIIR method

5.3 Previous plane-based CIIR Method

Two previous CIIR methods are reviewed in this section. Their concept, design and character are shown in the following paragraphs.

5.3.1 Traditional method

Figure 5.7 illustrates the process of the conventional plane-based CIIR method. In this method, each elemental image is inversely projected through its corresponding pinhole, magnified by the magnification factor M , which is the quotient between the distance from the virtual pinhole array to the reconstructed image plane (denoted by z in figure 5.7) and the distance from elemental image plane to the pinhole array (denoted by g in figure 5.7). The magnified pixels are overlapped with each other in the reconstructed image plane to reproduce the ROI at the determined distance. Series of reconstructed images with different z values are generated continuously to compose a volumetric reconstruction of the 3D scene. This CIIR approach is straightforward but it contains a crucial problem related to image quality degradation. This quality degradation comes from the magnification process, where each pixel is enlarged and superposed by another pixel, and the superposition between adjacent pixels causes an undesirable interference effect in the output image.

5.4 Pixel-to-pixel mapping method

To solve this quality degradation problem, another CIIR method using a pixel to pixel mapping technique was proposed in [SY09], the principle of this approach is described in figure 5.8. It removes the magnification process of previous CIIR techniques, generates a dot-pattern ROI by projecting each pixel of the elemental images into the output plane, and then fills the empty pixels in this dot-pattern image by a linear interpolation method. Because of the exact generation of pixel position, this method can remove the interference effect between adjacent pixels and improve the quality of the reconstructed image. However, due to the repetitive projecting of each pixel in elemental images, the calculation complexity is still very important.

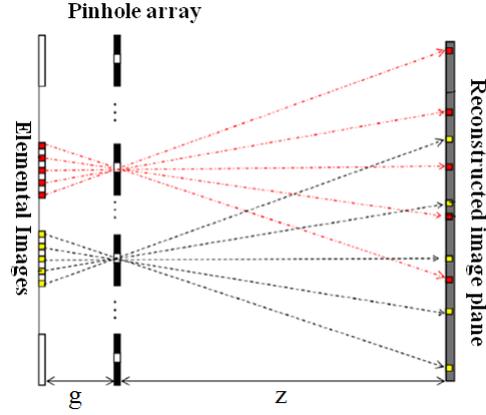


Figure 5.8: Reconstruction principle of pixel-to-pixel mapping CIIR method

5.5 Proposed plane-based CIIR methods

In this section, two novel CIIR methods are proposed to resolve the mentioned problems in previous CIIR techniques.

5.5.1 Parallel-group Projection (PGP)-based CIIR method

In both of the two existing plane-based CIIR approaches, the reconstructed image is generated from an integral image by projecting each pixel of the EIs one by one onto a predefined output plane. Because of the repetitive projection for each pixel and the supplementary transform operation, applying conventional CIIR in the proposed CII system is relatively time-consuming.

To reduce the calculation cost, we propose a new plane-based CIIR approach using the parallel-group projection method. Based on the same principle as the parallel group rendering method in the SI generation process, this new reconstruction approach projects sub-images instead of original elemental images to build the reconstructed output image. Since SIs record the information of parallel light rays from the given 3D scene, the reverse mapping rays of each pixel from the SI are also mutually parallel. Therefore, we considered reconstructing the ROI by parallel-group projection instead of pixel-to-pixel projection to exploit this mapping regularity. This new CIIR technique can prominently increase the efficiency of previous CIIR methods.

As illustrated in figure 5.9, the projected image of the original EIs through a computer synthesized pinhole array can be represented as the accumulation of the mapped result of each SI through the same virtual pinhole array. Due to the known parallel character of the mapping light rays, all the mapped pixels from one SI are arranged as a grid with a given interval in the reconstructed plane. Therefore, from the location of one projected pixel we are able to estimate the projected coordinate of every pixel by simply displacing the first mapped pixel with a specific offset, and eventually the ROI is formed by accumulating the projected result of every SI on the reconstructed plane.

Figure 5.10 illustrates the detailed calculation process of this CIIR method, which

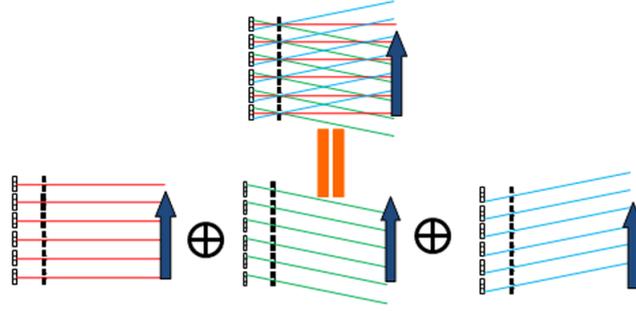


Figure 5.9: Concept of proposed CIIR method

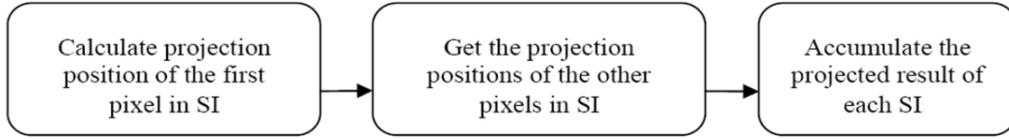


Figure 5.10: Block diagram of proposed CIIR method

is carried out in three steps. In the beginning, we project the first pixel of an SI into the reconstructed image plane following the geometric optical rule shown in figure 5.11. According to this diagram, the projected coordinates of one pixel from an SI can be calculated by the following equation:

$$(X_0, Y_0) = \{[(n-1) - x] \times M, [(n-1) - y] \times M\} \quad (5.1)$$

$$M = \frac{Z \times \tan(\theta) \times 2}{(n-1) \times phi} \quad (5.2)$$

where (X_0, Y_0) expresses the mapped pixel coordinates in the ROI, (x, y) represents the index of the projecting SI, which is ranged as $(0, 1, 2, \dots, (n-1))$, n is the number of SIs on the horizontal or vertical axis. M is the magnification factor representing the quotient between the distance from the virtual lens array to the reconstructed image plane (denoted by z in figure 5.11) and the distance from the elemental image plane to the pinhole array (denoted by g in figure 5.11). ϕ is the size of a pixel in the SI. θ is the viewing angle of the marginal SI.

Once we get the correct projected position of the first pixel in each SI from formulas 5.1 and 5.2, the second step is to deduce the mapped coordinates of the other pixels in this SI through the regular layout of projected pixels, which can be represented simply by the following equation:

$$(X, Y) = (p \times \phi \times n + X_0, q \times \phi \times n + Y_0) \quad (5.3)$$

where p and q are the pixel coordinate in the SI, arranged as $(0, 1, 2, \dots, (m-1))$, m is the number of pixels in one SI. n still represents the number of SIs on the horizontal or vertical axis.

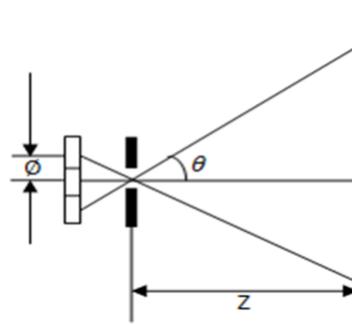


Figure 5.11: Computational reconstruction

Finally, the projected outcomes of each sub-image are superimposed on each other to construct the displayed ROI at the determined position of the output plane. Besides its high computational efficiency derived from parallel group projection, this approach further reduces calculation complexity by removing the transformation process from SI to II required in previous CIIR techniques.

5.5.2 DIBR-based-CIIR

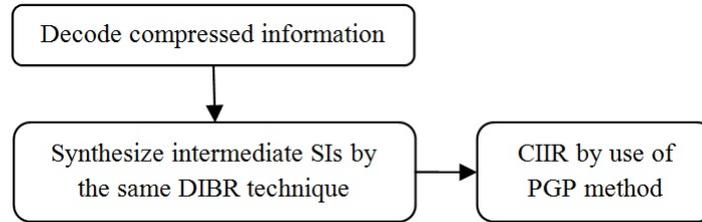


Figure 5.12: Proposed computational integral imaging system

As mentioned in Section 3.1, the resolution of the reproduced 3D image is determined by the number of EIs, while the viewing angle density is dependent on the pixel number in each EI. Meanwhile, the number of pixels in one EI is equal to the number of SIs and the number of EIs is equal to the pixel count in one SI. Therefore, increasing the number of SIs is a possible approach to improving the viewing angle density for the viewer [aSPHSK06] [HPSK08].

The scheme of this method is given in figure 5.12. Similar to the prediction process in compression operation, in the reconstruction part, the DIBR technique introduced in section 4.4.2 is again used to synthesize numbers of virtual SIs from the decoded reference SI and its corresponding depth map. In this way, the displayed 3D image can be composed of a larger number of SIs than the original recorded ones, and the parallax fluency of the reconstructed 3D image is consequently increased.

Cumulating the performance of the proposed PGP and virtual SI synthesizing approaches, our CII system can reconstruct a better 3D effect with lower calculation cost. Simulation results are reported in next section.

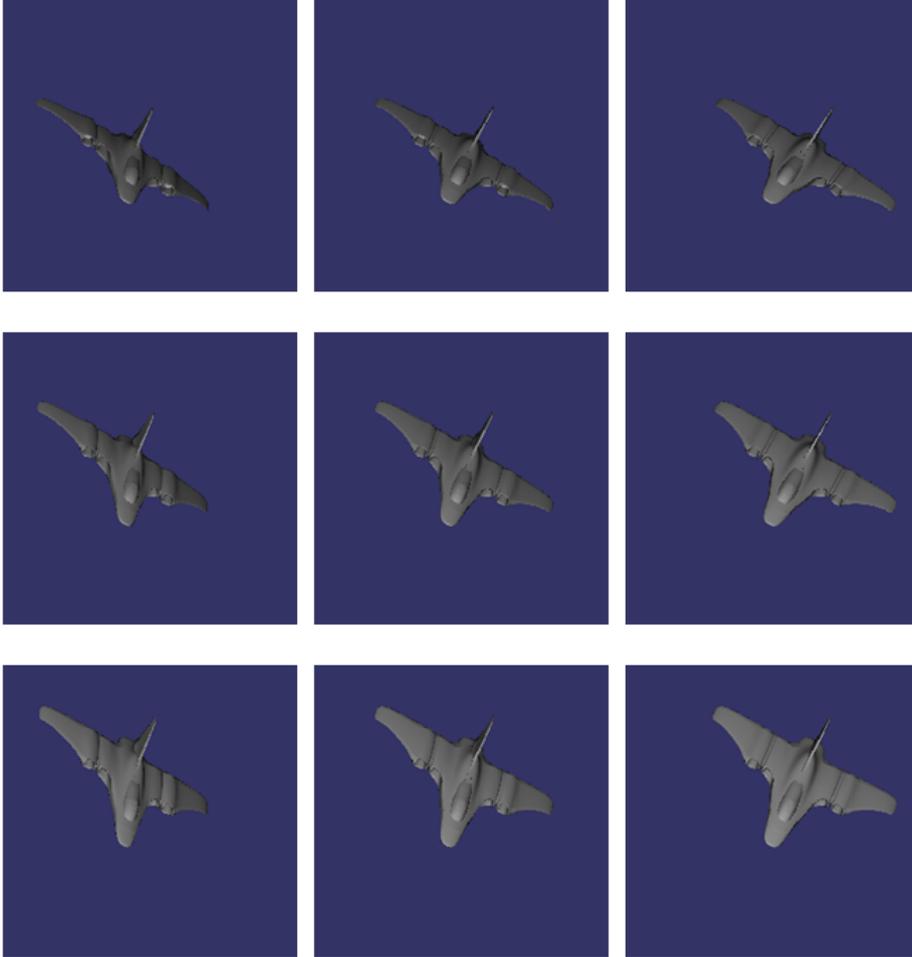


Figure 5.13: Generated SIs for 3D scene ‘spaceship’

5.6 Experimental results

Two simulations are designed to evaluate the PGP method and the DIBR-based view synthesis method respectively. As in the previous chapters, two 3D scenes of ‘cube’ and ‘spaceship’ are applied to demonstrate the effect of the proposed approaches.

5.6.1 Evaluation of PGP-based CIIR

In the first experiment, three different plane-based CIIR methods are tested for both ‘spaceship’ and ‘cube’ scenes, including the traditional CIIR method, the point to point-based reconstruction method and the proposed parallel-group projection method. Similar to the experiment in chapter 4, the original sub images are acquired by the parallel-group rendering method introduced in section 3.2, which is realized by the openSceneGraph (OSG) toolkit with orthographic projective cameras.

Figure 5.13 shows the generated nine sub images for the 3D scene ‘spaceship’ from distinguishable viewing angles. The resolution of each SI is 256×256 , the number

of SI in horizontal and vertical axis is 3, and the range of viewing angle is 30° (from -15° to 15°) in both horizontal and vertical directions.

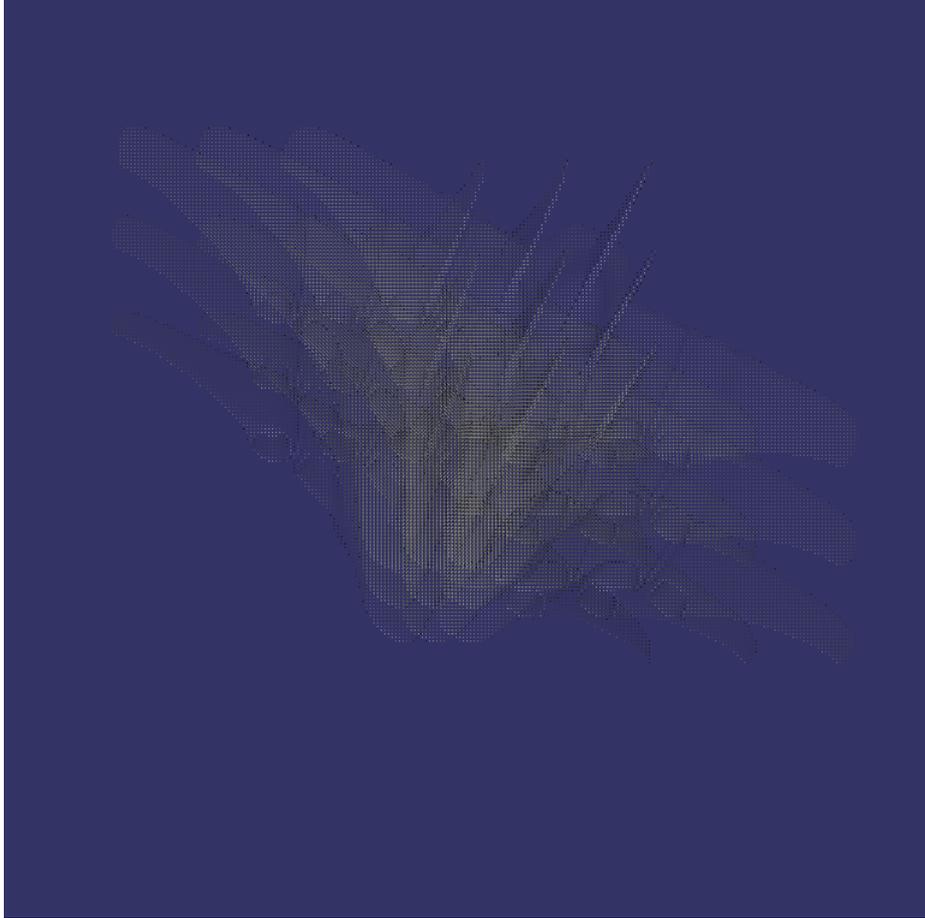
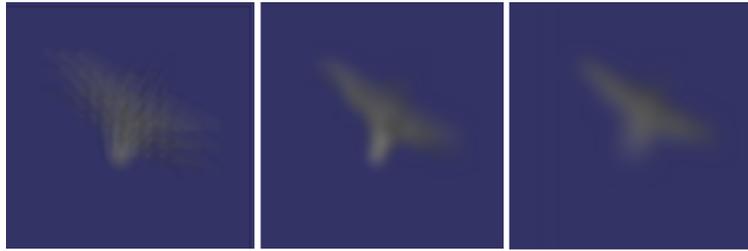


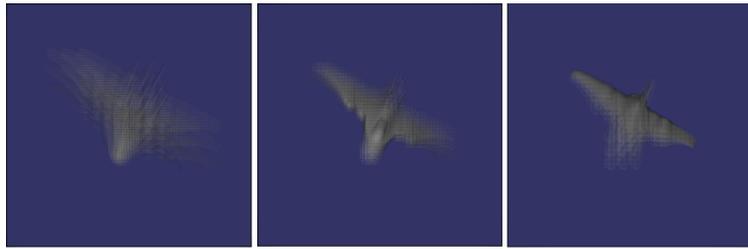
Figure 5.14: II composed by the generated SIs for 3D scene ‘spaceship’

With the generated SI shown above, an integral image with 768×768 pixels is created, illustrated in figure 5.14. Then the traditional and pixel to pixel CIIR algorithms are tested by use of this integral image.

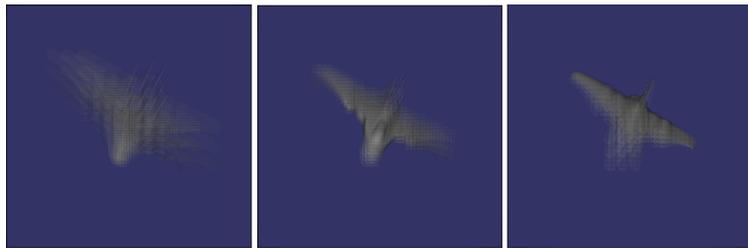
In our experiment, the performance of the tested plane-based CIIR techniques is evaluated in terms of image quality and calculation efficiency. To illustrate the change of reconstructed image along with the movement of the output plane, the distance of the reconstructed plane is moved from 20 to 100. According to the known parameters and formula 5.2, we can get the magnification factor M form 6 to 28.



(a) ROI in conventional method with interference effect



(b) ROI in pixel-to-pixel mapping method with higher visual quality



(c) ROI in proposed method with higher visual quality

Figure 5.15: Comparison of reproduced ROI of three different methods for 3D scene ‘spaceship’

Figure 5.15 shows the reproduced ROI locats 20, 60 and 100 from lens array. Figure 5.15(a) illuminates the ROI reconstructed by the conventional method with a severe interference effect. The ROI reconstructed by the pixel-to-pixel mapping method and the proposed PGP-based CIIR method are illustrated in figure 5.15(b) and Figure 5.15(c) respectively. Compared to previous methods, the proposed plane-based CIIR approach can eliminate the interference effect in the traditional method (cf. figure 5.15(a) and 5.15(c)), and provides a high quality ROI as the pixel-to-pixel mapping method (cf. figure 5.15(b) and 5.15(c)).

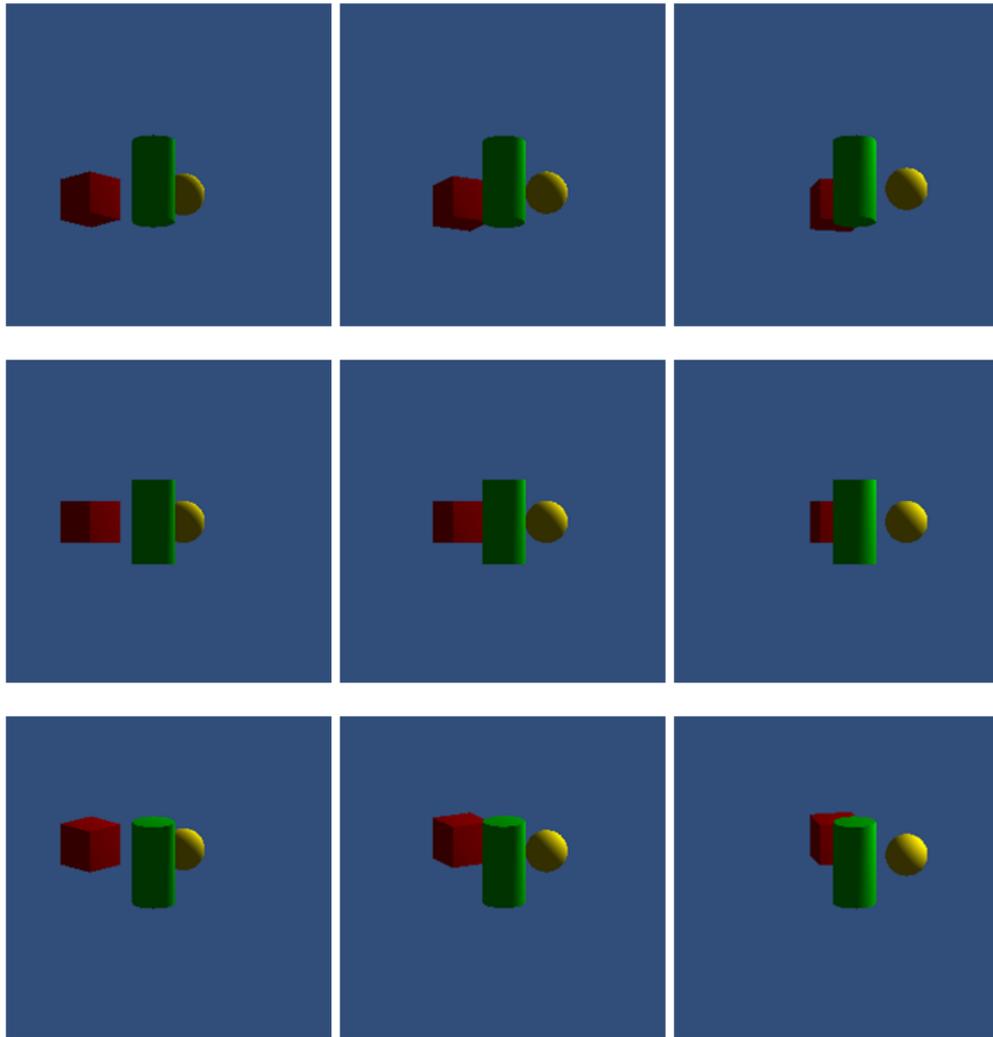


Figure 5.16: Generated SIs for 3D scene ‘cube’

Another 3D object named ‘cube’ is also applied to prove the performance of the proposed PGP-based CIIR method. The same acquisition process is employed with identical generation parameters. 9 sub-images are created by use of the orthographic camera in the OSG toolkit, see figure 5.16. As in the first experiment, the sub-images are recorded from three different viewing directions of $(-15^\circ, 0^\circ, 15^\circ)$ in the horizontal and vertical axes. The resolution of each SI is 256×256 .

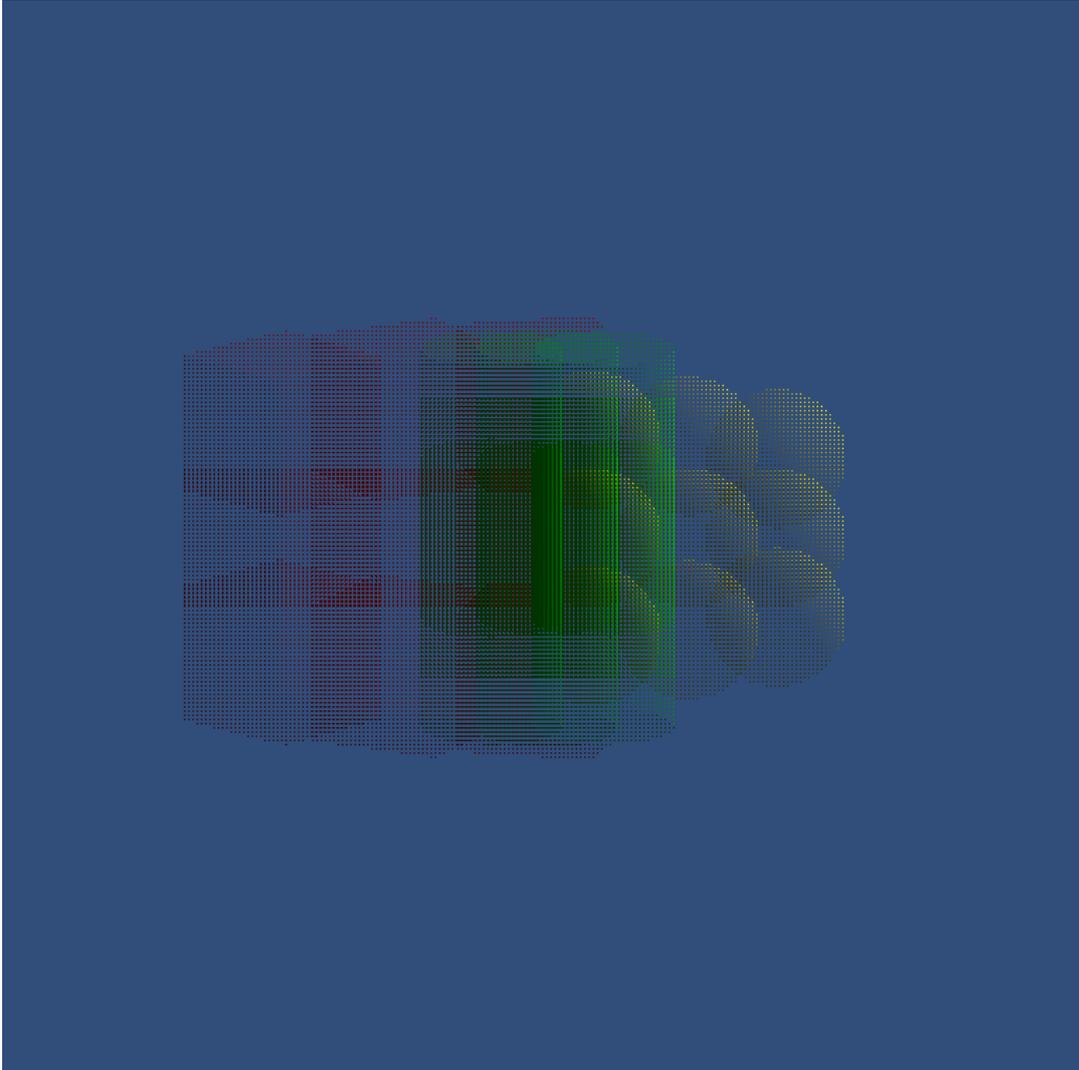
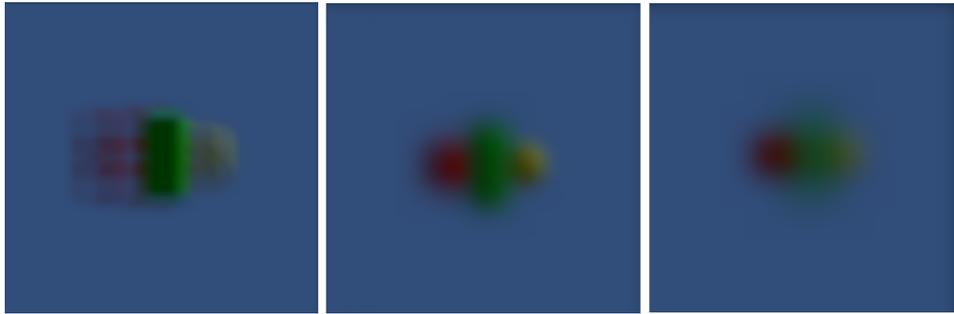
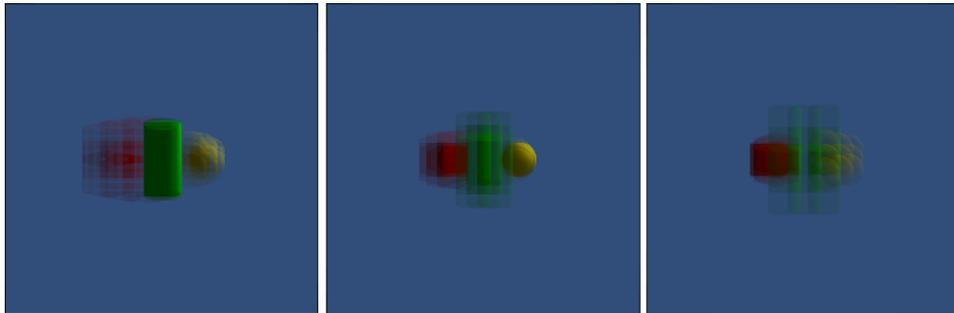


Figure 5.17: II composed by the generated SIs for 3D scene ‘cube’

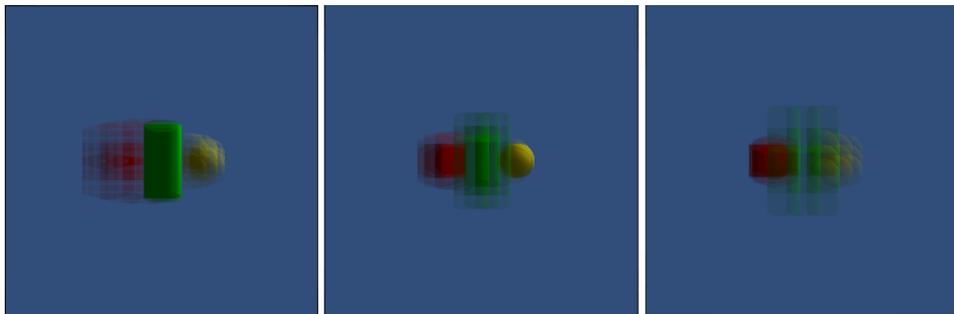
Then, figure 5.17 shows the composed integral image with 256×256 elemental-images, where each EI contains 3×3 pixels.



(a) ROI in conventional method with interference effect



(b) ROI in pixel-to-pixel mapping method with higher visual quality



(c) ROI in proposed method with higher visual quality

Figure 5.18: Comparison of reproduced ROI of three different methods for 3D scene ‘cube’

Using the same reconstruction parameters as in the previous experiment, we can obtain the reconstructed image moving from 20 to 100, and the magnification factor M changes from 6 to 28. To demonstrate the effect of the proposed plane-based CIIR using PGP method, the contrasts of three different methods are shown in figure 5.18. According to the illustrated images, we can find that the pixel-to-pixel and proposed PGP-based mapping method removed the interference effect present in the conventional reconstruction approach and thus improves the quality of reconstructed images.

As we explained before, besides the improvement of reconstructed image quality, the proposed CIIR technique based on the parallel group projection method can also greatly increase the efficiency of the reconstruction process. Tables 5.1 and tables 5.2

give the computation time of the three different plane-based CIIR techniques as a function of magnification factor M .

Viewing distance	Traditional CIIR	CIIR using pixel-to-pixel mapping	CIIR using parallel group projection
20	30094	1157	328
40	59141	2282	328
60	103453	4563	328
80	230421	6844	328
100	410438	9188	328

Table 5.1: Comparison of the computation time of three CIIR methods for ‘spaceship’

Viewing distance	Traditional CIIR	CIIR using pixel-to-pixel mapping	CIIR using parallel group projection
20	30203	1157	359
40	59047	2328	359
60	103063	4657	359
80	229828	7000	359
100	407891	9313	359

Table 5.2: Comparison of the computation time of three CIIR methods for ‘cube’

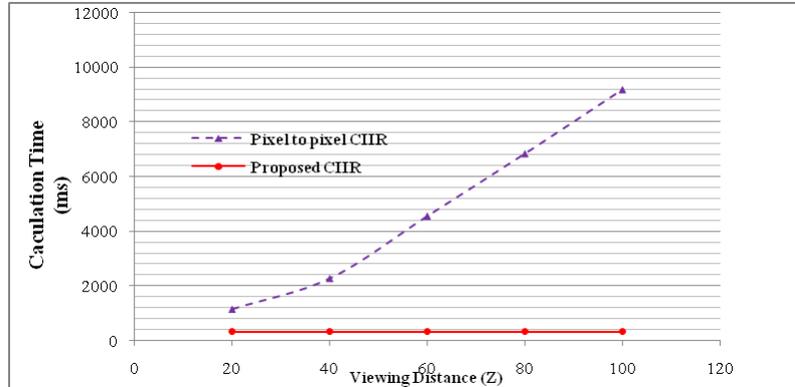


Figure 5.19: Computation cost of pixel-to-pixel and parallel-group projection CIIR for the 3D scene ‘spaceship’

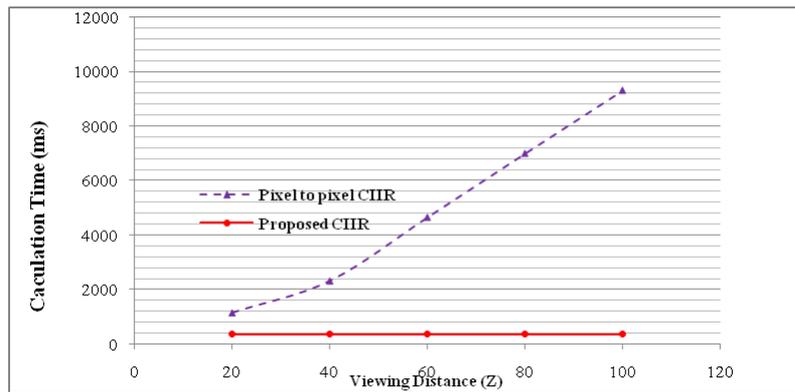


Figure 5.20: Computation cost of pixel-to-pixel and parallel-group projection CIIR for the 3D scene ‘cube’

The results in these two tables are illustrated in figure 5.19 and 5.20. According to the diagram, we can detect that for a reconstruction process with a given magnification factor, the proposed PGP-based CIIR method shows a significant improvement in efficiency compared to the pixel-to-pixel mapping method, and its calculation time is not affected by increasing the magnification factor M .

5.6.2 Evaluation of DIBR-based CIIR

Applying the DIBR technique to synthesize the virtual SI is another method to improve reconstructed image quality. As shown in figure 5.21, a reference sub-image from viewing angle equal to 0° with its corresponding depth map is generated by the OSG toolkit from a given 3D scene named 'spaceship'. The size of the generated SI and its associated depth map is 512×512 .

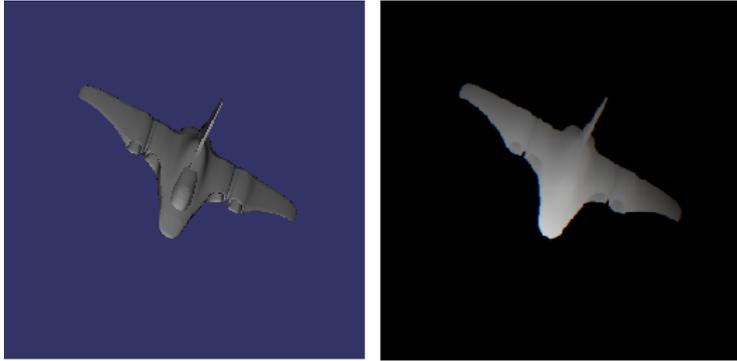
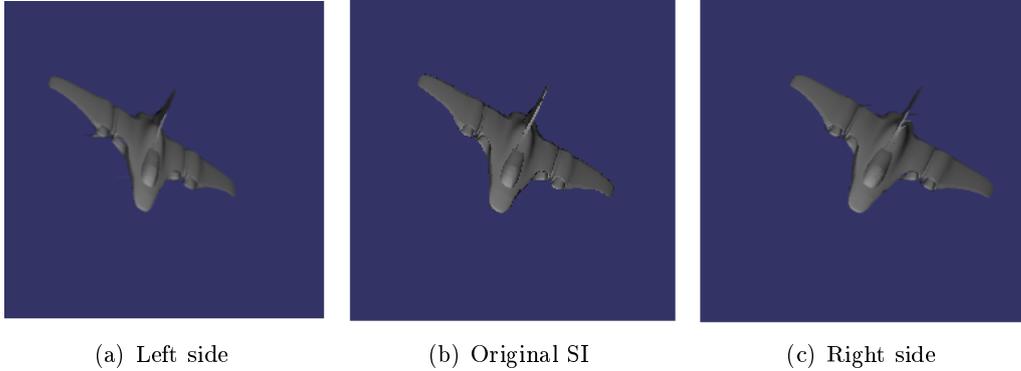


Figure 5.21: Reference SI from viewing angle equals to 0° and its corresponding depth map



(a) Left side

(b) Original SI

(c) Right side

Figure 5.22: Synthesized virtual SIs

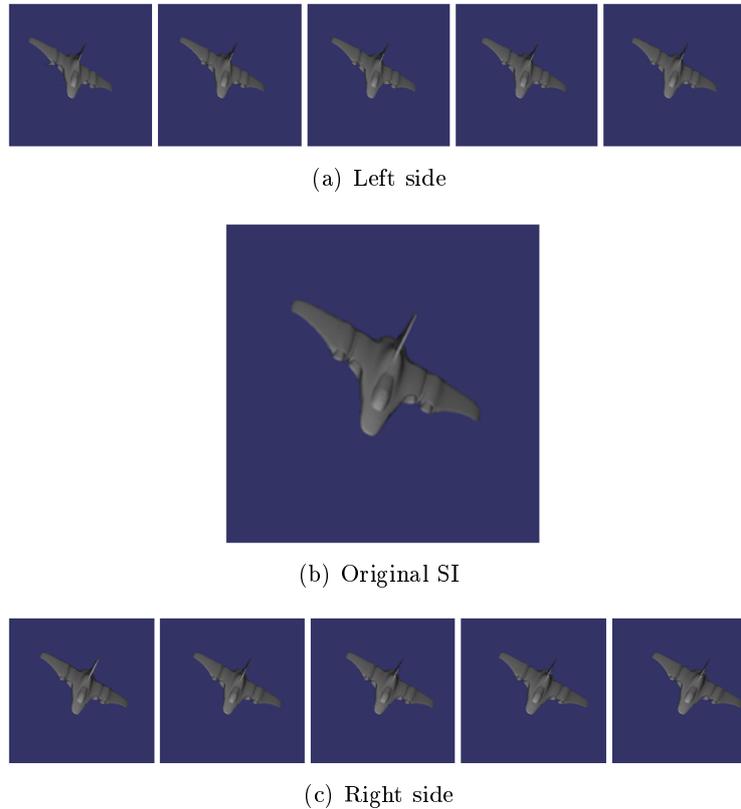


Figure 5.23: Synthesized virtual SIs

In order to clarify the influence of sub-image number to reconstructed image quality, two tests with different numbers of interpolated sub-images are executed separately in this experiment.

The first test synthesizes only two virtual SIs for reproducing the 3D image, the proposed DIBR method explained in chapter 4 is reapplied to synthesize the virtual SIs with the given viewing angle equal to -15° and 15° respectively. The synthesized virtual SIs are shown in pictures 5.22(a) and 5.22(c).

In contrast, another test synthesizing 10 virtual SIs from the same reference SI and its corresponding depth image is also implemented. The synthesized SIs are shown in figure 5.23, where the viewing angle of these SIs is also ranged from -15° to 15° .

To highlight the improvement of viewing angle resolution, we decided to evaluate the reconstructed result by the view-based CIIR approach introduced in section 5.2.2.

Figure 5.24 gives the integral image composed from three sub-images illustrated in figure 5.22. The generated integral-image contains 512×1 elemental images and each EI is made up of 3×1 pixels.

Then, the view-based CIIR method introduced in section 5.3.1 is applied to reproduce the perceived images from different latitude directions. Reconstructed images from -15° , 0° and 15° are generated and illustrated in figure 5.25. From the displayed images, it is easy to find that the quality of reconstructed image is relatively low due to the limited numbers of employed sub-images (three SIs are used in this test).

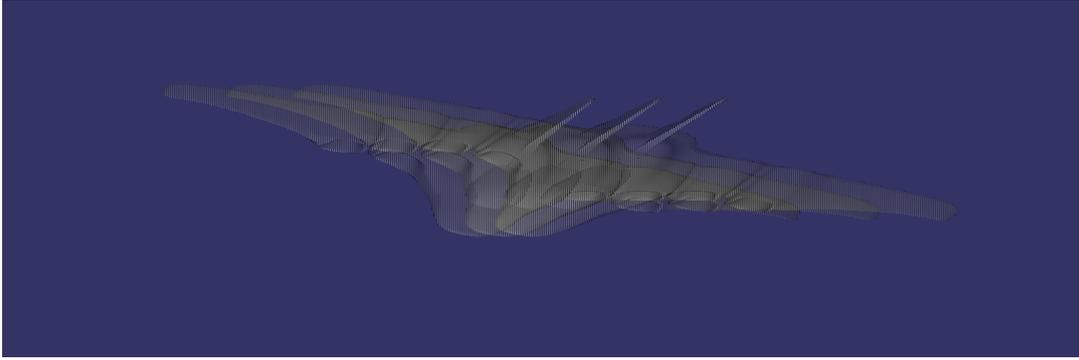


Figure 5.24: Integral image composed by three sub-images shown in figure 5.22



Figure 5.25: Reconstructed images from -15° , 0° and 15°

Another test using a larger number of sub-images is implemented with the same procedure. At first, the 11 sub-images illustrated in figure 5.26 are consolidated together to form an integral image. This II contains the same number of elemental images as the previous one shown in figure 5.24, but for each EI it possesses 11×1 pixels which is much more than the EI in previous II.

Using the new II shown in the picture above, the same view-based reconstruction technique is applied again to generate the perceived image for different viewing angles. As for test 1, three reconstructed images from -15° , 0° , and 15° are generated in this test.

From reconstructed images shown in figure ??, we can find that the quality of the reconstructed image is significantly improved compared to the viewed image in the first test (see figure 5.25). Furthermore, according to the recorded video along the latitude direction, the parallax effect of this II is more natural and fluent than the previous one. Table 5.3 demonstrates the improvement of reconstructed image quality by PSNR value comparison.

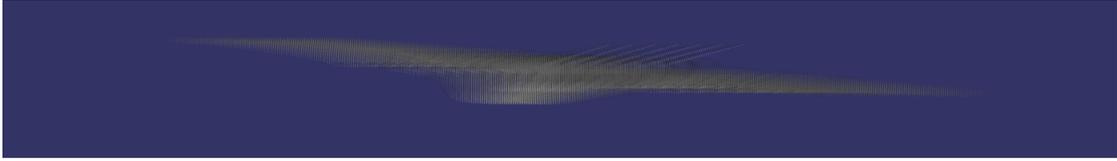


Figure 5.26: Integral image composed by 11 sub-images shown in figure 5.23

	viewed image reconstructed by 3 SIs	viewed image reconstructed by 11 SIs
PSNR(dB)	30.23	32.84

Table 5.3: PSNR comparison of 3D image reconstructed by three sub images and eleven sub images

5.7 Conclusions

In this chapter we proposed two novel CIIR techniques to improve the efficiency and performance of traditional CIIR methods.

The first method is based on the parallel-group projection technique. In this method, we transform the computation process from the elemental image domain to the sub-image domain, and project an entire sub-image into the output plane at one time instead of calculating the projected position of each pixel in the elemental images array one by one. Compared to the traditional CIIR method, the proposed approach is able to reduce the reconstruction complexity significantly.

The second CIIR method is provided to improve the reconstructed 3D image quality. The principle of this method is to render additional numbers of virtual SIs via the same DIBR technique applied in the compression step, then apply the newly composed integral image with enhanced resolution to reconstruct the 3D scene. Thanks to the enlarged number of sub-images, the quality and parallax continuity of the observed image is improved substantially by this approach.

Chapter 6

Conclusion

After the emergence of high definition television, 3DTV is presented as the next revolution of the television industry. Compared to current 2D color TV, three dimensional television can provide an extraordinary viewing experience to spectators. Along with the rapid development of electrical and computer graphic technology, 3DTV has experienced significant progress. Nowadays, 3DTV is no longer a scientific conception, but a concrete and attractive product for TV users. A large variety of alternative 3D video recording and display techniques such as (auto)-stereoscopic, multi-view, integral imaging, volumetric and holographic have been demonstrated and refined over the last decades. Among all these different 3D technologies, integral imaging is considered the most suitable method for future 3DTV due to its capacity of providing full and continuous motion parallax to the viewer by the use of a set of 2D images, as well as its compatibility with classical 2DTV systems. Despite these attractive advantages, its practical application is still obstructed by some inherent limitations such as poor image quality, huge data volume and high processing complexity. In this manuscript we have presented and analyzed these challenging issues individually, and proposed corresponding solutions in order to ameliorate the viewers' 3D experience when using the integral imaging system.

6.1 Summary of Contributions

This thesis tried to improve the performance of different aspects of the integral imaging system including content generation, coding and reconstruction. The main contributions in this thesis are summarized as follows:

6.1.1 A new method for computational integral imaging generation

The first part concentrated on integral image generation. The problem of limited viewing range and restricted depth range are addressed and resolved in this part.

A new computation integral image generation (CIIG) method has been proposed to control the depth range and expand the viewing range of the displayed 3D image. On one hand, the method selects an interesting plane in the 3D scene, and makes all the virtual cameras overlap in this plane so as to improve the resolution of distant objects around this depth value. On the other hand, in order to expand the viewing

range for the observer, the proposed method simulates the curve recording panel by changing the viewing direction of each elemental image (EI) as a function of its own coordinate. The results of our experiments demonstrate that the proposed CIIG method is able to control the depth range and enhance the viewing range of traditional CIIG methods without increasing the size of source image.

6.2 Three new techniques for Integral Image compression

The second part concerned the compression issues of the integral image. Three different compression approaches were proposed and evaluated. The first compression method is based on the multi-view video coding (MVC) technique. It is applied on sub-images (SI), which is an alternative form of 2D image transformed from original elemental images. Each SI represents the 3D scene from parallel viewing directions and contains superior compression capabilities than the original captured EI. In this method, we arrange the group of SIs following the format of multi-view video (MVV) and then encode the generated MVV using the MVC standard. This approach improves the compression efficiency when compared to the traditional MPEG-4/AVC method.

The second compression method is inspired by the Multi-View Video plus Depth (MVD) representation. It is also applied on sub-images. The principle of this method is to use the 3D information existing in depth maps to predict the SI for other directions and then remove the inter-view correlation between original SIs. Thanks to the 3D geometrical information in the depth map, this method outperforms the MVC-based scheme mentioned before.

Besides the MVC and MVD based compression approaches, we also considered applying the discrete wavelet transform (DWT) for integral image compression. Differing from the DCT, the DWT is able to transform the original 2D information from the pixel domain into the pixel-frequency domain, but still conserves the spatial domain character of the original data. In the proposed approach, the input IIs are filtered multiple times by a determined high pass and low pass filter until the correlation within one EI and between adjacent EIs are completely eliminated. According to the simulation results, the DWT works better than the MVC-based method in low-debit situations.

6.2.1 Two novel methods for computational integral imaging reconstruction

Finally, we investigated the reconstruction process and developed two novel plane-based computational integral imaging reconstruction techniques in order to improve the performance and reduce the complexity of the basic CIIR. The first method leverages the Parallel-Group Projection (PGP) technique. Differing from previous works on CIIR techniques which project each point of integral image to the reconstructed plane pixel by pixel, the proposed method reconstructs the 3D image by mapping a series of sub images onto the reconstructed plane successively. Because of the parallelization between the light rays in each SI, this approach can simplify the pixel mapping complexity in the reconstruction process.

The second design aims to improve the quality of the reconstructed 3D image by increasing the number of pixels in each EI. In this method, numbers of virtual sub-images are synthesized via a proposed depth image rendering algorithm. By use of these synthesized SIs, the definition of the EI is consequently increased which brings a significant improvement for the observed image quality and parallax fluency. Moreover, this method also brings benefits for data transmission. Instead of transmitting multiple color images, it requires only one color image and one depth map to reproduce the 3D scene, which can improve the transmission efficiency of the integral imaging system.

6.3 Perspectives

This research has significantly improved the end-users' 3D experiences in the integral imaging system. Noticeable ameliorations have been achieved in various aspects such as viewing range, image depth range, and viewing resolution. However, there are remaining issues to be resolved in many directions. The proposal for applying integral imaging technology in 3DTV still needs to progress.

6.3.1 Computational integral image generation

In this thesis, we have proposed a method to control the recorded depth range of the 3D scene, which is able to select an interesting region in the depth direction and record this region with higher resolution. However, this approach cannot intrinsically enlarge the total length of the depth range. It improves the quality of a certain range by sacrificing the quality of other depth ranges. Therefore, in our future work, we intend to explore a true depth range enhancement approach, which is able to expand the recording range of the integral image without reducing the viewing quality of other depth ranges.

Another interesting research topic is applying the CIIG technique to generate a hologram, which can be used to provide an alternative to the conventional complex hologram acquisition process.

6.3.2 Integral image compression

All the proposed compression methods are demonstrated with very simple 3D scenes. Evaluations conducted from more complex 3D scenes are necessary to further highlight the benefits of our contributions.

For the MVC-based compression method, a more pertinent disparity estimation algorithm should be found in order to reduce the residual error of the predicted image.

For the MVD-based compression method, we could envision a new DIBR algorithm with lower complexity so as to provide a more efficient encoding approach.

For the DWT-based compression method, different wavelets will be tested in following work in order to compare their efficiency with the Haar wavelet.

6.3.3 Computational Integral image reconstruction

For the CIIR process, besides the presented contributions in plane-based reconstruction, we are planning to propose new methods to improve the efficiency of view-based reconstruction.

Depending on these perspective works and continuing efforts, we suppose that the application of integral imaging will be extended from the academic domain into industrial applications in the near future.

Bibliography

- [Agg11a] A Aggoun. Compression of 3d integral images using 3d wavelet transform. *Journal of Display Technology*, 7(11):586–592, 2011. 53
- [Agg11b] A. Aggoun. compression of 3d integral images using 3d wavelet transform. *Journal of Display Technology*, 7(11):586–592, 2011. 55
- [aSPHSK06] ae Sung Park, Dong-Choon Hwang, Dong-Hak Shin, and Eun-Soo Kim. Enhanced-resolution computational integral imaging reconstruction using an intermediate-view reconstruction technique. *Optical Engineering*, 45(11), 2006. 91
- [ATGA08] M. Oguz Biciand Anil Aksay, Antti Tikanm, Atanas Gotchev, and Gozde Bozdagi Akar. Stereo video broadcasting simulation for dvbh. In *NEM*, 2008. 21
- [BAB95] Adrian S. Bruce, David A. Atchison, and Harsha Bhoola. Accommodation-convergence relationships and age investigative ophthalmology visual science. *Invest. Ophthalmol. Vis. Sci.*, 36(2):406–413, 1995. 23
- [BE12] Pressigout M-Morin L Bosc E, Riou P. Bit-rate allocation between texture and depth: influence of data sequence characteristics. In *3DTV Conference: the True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, 2012. 19
- [BFD95] Michale Brewin, Matthew C. Forman, and Neil A. Davies. Electronic capture and display of full parallax 3d images. In *SPIE Stereoscopic Displays and Virtual Reality Systems II*, 1995. 24
- [CCJW13] Huan Deng Da-Hai Li Chao-Chao Ji, Cheng-Gao Luo and Qiong-Hua Wang. Tilted elemental image array generation method for moire-reduced computer generated integral imaging display. *Optics Express*, 21(17):19816–19824, 2013. 31
- [CG11] Ortega A Cheung G, Velisavljevi V. On dependent bit allocation for multiview image coding with depth-image-based rendering. *IEEE Trans Image Process*, 20(11):3179 – 3194, 2011. 19
-

-
- [CKH11] Daniel Herrera C., Juho Kannala, and Janne HeikkilAccurate and practical calibration of a depth and color camera pair. *Computer Analysis of Images and Patterns*, 6855(1):433–475, 2011. 14
- [CP12] Young-Sic Choi and Rae-Hong Park. Mpeg-2 stereoscopic video coding technique using adaptive bandwidth control. *IEEE Transactions on Multimedia*, 14(1):157–167, 2012. 18
- [CW93] Shenchang Eric Chen and Lance Williams. View interpolation for image synthesis. In *SIGGRAPH*, 1993. 61
- [DMY88] Neil Davies, Malcolm McCormick, and Li Yang. Three dimensional imaging systems: A new development. *Applied Optics*, 27(21):4520–4528, 1988. 12
- [DN10] Sachin Deshpande and Jeonghun Noh. Liquid crystal active glasses for 3d cinema. *Journal of Display Technology*, 6(10):522–530, 2010. 3
- [EB13] Vincent Jantet Paul Riou Muriel Pressigout Luce Morin-Vincent Jantet Emilie Bosc, Fabien RacapA study of depth/texture bit-rate allocation in multi-view video plus depth compression. *Annals of telecommunications - annales des tcommunications*, pages 1 – 10, 2013. 19
- [ESHJ07] E.Elharar, Adrian Stern, Ofer Hardar, and Bahram Javidi. A hybrid compression method for integral images using discrete wavelet transform and discrete cosine transform. *Journal of Display Technology*, 3(3):321–325, 2007. 53, 55
- [ETKS13] N. P. Sgouros E. T. Koufogiannis and M. S. Sangriotis. Perspective rectification of integral images produced using arrays of circular lenses. *Applied Optics*, 52(20):4959–4968, 2013. 36
- [FDH⁺01] Gregg E. Favalora, Rick K. Dorval, Deirdre M. Hall, Michael Giovinco, and Joshua Napoli. Volumetric three-dimensional display system with rasterisation hardware. In *SPIE 4297, Stereoscopic Displays and Virtual Reality Systems VIII*, 2001. 25
- [FE05] Favalora and Gregg E. Volumetric 3d displays and application infrastructure. *Computer*, 38(8):37–44, 2005. 24
- [Feh04] Christoph Fehn. Depth-image-based rendering (dibr), compression and transmission for a new approach on 3d-tv. In *SPIE 5291*, 2004. 17
- [FP08] Yasutaka Furukawa and Jean Ponce. Accurate camera calibration from multi-view stereo and bundle adjustment. In *5Computer Vision and Pattern Recognition*, 2008. 13
- [FS11] Mei Yu Ken Chen Feng Shao, Gangyi Jiang. Asymmetric coding of multi-view video plus depth based 3-d video for view rendering. *Proceedings of the IEEE*, 99(4):643 – 656, 2011. 19
-

-
- [FTS⁺09] R. Fukushima, K. Taira, T. Saishu, Y. Momonoi, M. Kashiwagi, and Y. Hirayama. Effect of light ray overlap between neighboring parallax images in autostereoscopic 3d displays. In *SPIE 7237, Stereoscopic Displays and Applications*, 2009. 24
- [FW10] Zeng Fanfeng and Feng Wei. Hole filling algorithm based on contours information. In *2nd International Conference on Information Science and Engineering (ICISE)*, 2010. 16
- [GL12] Eun-Soo Kim Gen Li, Seung-Cheol Kim. Viewing quality-enhanced reconstruction of 3-d object images by using a modified computational integral- imaging reconstruction technique. *3D Research*, 3(3):1–10, 2012. 82
- [HFFD02] Philip V. Harman, Julien Flack, Simon Fox, and Mark Dowley. Rapid 2d-to-3d conversion. In *SPIE 4660 Stereoscopic Displays and Virtual Reality Systems IX*, 2002. 16
- [HHKK12] Ju-Han Lee Ho-Hyun Kang and Eun-Soo Kim. Enhanced compression rate of integral images by using motion-compensated residual images in three-dimensional integral-imaging. *Optics Express*, 20(5):5440–5459, 2012. 53
- [Hir09] Yuzo Hirayamai. One-dimensional integral imaging 3d display systems. In *3rd International Universal Communication Symposium*, 2009. 35
- [HJ01] H. Arimoto and B. Javidi. Integral three-dimensional imaging with digital reconstruction. *Optics Letters*, 26(3):157–159, 2001. 82
- [HJJ04] Seung-Hyun Hong, Ju-Seog Jang, and Bahram Javidi. Three-dimensional volumetric object reconstruction using computational integral imaging. *Optics Express*, 12(3):483–491, 2004. 24, 82
- [HM93] Isono H. and Yasuda M. Autostereoscopic 3-d display using lcd-generated parallax barrier. *Electronics and Communications in Japan*, 76(7):77–84, 1993. 22
- [HPSK08] Dong-Choon Hwang, Jae-Sung Park, Kong-Hak Shin, and Eun-Soo Kim. Depth-controlled reconstruction of 3d integral image using synthesized intermediate sub-image. *Optics Communications*, 281(24):5991–5997, 2008. 91
- [JTLC92] Moor J.R., A.R.L. Travis, S.R. Lang, and O.M. Castle. The implementation of a multi-view autostereoscopic display. In *IEEE Colloquium on Stereoscopic Television*, 1992. 22
- [Kak08] Hideki Kakeya. Formulation of coarse integral imaging and its applications. In *SPIE 6803, Stereoscopic Displays and Applications XIX*, 2008. 32
-

- [Kaw02] Takashi Kawai. 3d displays and applications. *Displays*, 23(1):49–56, 2002. 2
- [KCKY12a] Munkh-Uchral Erdenebat Ji-Seong Jeong Jeong-Hun Choi Nam Kim Jae-Hyeung Park Young-Tae Lim Ki-Chul Kwon, Chan Park and Kwan-Hee Yoo. High speed image space parallel processing for computer-generated integral imaging system. *Optics Express*, 20(2):732–740, 2012. 32
- [KCKY12b] Munkh-Uchral Erdenebat Ji-Seong Jeong Jeong-Hun Choi Nam Kim Jae-Hyeung Park Young-Tae Lim Ki-Chul Kwon, Chan Park and Kwan-Hee Yoo. High speed image space parallel processing for computer-generated integral imaging system. *Optics Express*, 20(2):732–740, 2012. 81
- [KHL10] Youngmin Kim, Keehoon Hong, and Boungho Lee. Recent researches based on integral imaging display method. *3D Research*, 1(1):17–27, 2010. 49
- [KKH⁺12] Youngmin Kim, Hongshin Kim, Keehoon Hong, Hee Kyung Yang, Jae-Hyun Jung, Heejin Choi, Sung-Wook Min, Jong-Mo Seo, Jeong-Min Hwang, and Byoungho Lee. Accommodative response of integral imaging in near distance. *Journal of Display Technology*, 8(2):70–78, 2012. 32
- [KMW11] Philipp Merkle Karsten Muller and Thomas Wiegand. 3-d video representation using depth maps. *Proceedings of the IEEE*, 99(4):643 – 656, 2011. 13
- [KPC⁺04] Yunhee Kim, Jae-Hyeung Park, Heejin Choi, Sungyong Jung, Sung-Wook Min, and Byoungho Lee. View-angle enhanced integral imaging system using a curved lens array. *Opt. Express*, 12(3):421–429, 2004. 35
- [KSK10] Ho-Hyun Kang, Dong-Hak Shin, and Eun-Soo Kimi. Efficient compression of motion-compensated sub-images with karhunen-loeve transform in three-dimensional integral imaging. *Optics Communications*, 283(6):920–928, 2010. 53, 56
- [KUO10] Takafumi Koike, Kei Utsugi, and Michio Oikawa. Analysis for reproduced light field of 3d displays. In *3DTV-Conference*, 2010. 49
- [LCP⁺06] Bin-Na-Ra Lee, Yongjoo Cho, Kyoung Shin Park, Sung-Wook Min, Joa-Sang Lim, Min Cheol Whang, and Kang Ryong Park. Design and implementation of a fast integral image rendering method. In *5th international conference on Entertainment Computing*, 2006. 33
- [LK00] Yung-Kai Laia and C.-C.Jay Kuob. A haar wavelet approach to compressed image quality measurement. *Journal of Visual Communication and Image Representation*, 11(1):17–40, 2000. 64
-

-
- [Lue11] Ernst Lueder. *3D Displays*. John Wiley Sons, 2011. 1, 21, 22
- [MB95] Leonard McMillan and Gary Bishop. Plenoptic modeling: An image-based rendering system. In *SIGGRAPH*, 1995. 61
- [McMCS⁺08] Manuel Martinez-corrall, Raul Martinez-Cuenca, Genaro Savedra, Hector Navarro, Amparo Pons, and Bahram Javidi. Progresses in 3d integral imaging with optical process. *Journal of Physics*, 139(1), 2008. 31
- [MCND⁺11] M.Martinez-Corrall, H. Navarro, A. Dorado, G.Saavedra, A.Pons R.Martinez-Cuenca, and B.Javidi. Lightfield recording and reconstruction by integral imaging. In *10th Euro-American Workshop on Information Optics (WIO)*, 2011. 34
- [Md06] Yannick Morvan and Peter H. N. de. Platelet-based coding of depth maps for the transmission of multiview images. *Stereoscopic Displays and Applications*, 6055:93–100, 2006. 19
- [MFN⁺96] Keiichi Maeno, Naoki Fukaya, Osamu Nishikawa, Koki Sato, and Toshio Honda. Electro-holographic display using 15mega pixels lcd. In *Proc. SPIE 2652, Practical Holography*, 1996. 26
- [MHKL05] Sung-Wook Min, Minsoo Hahn, Joochwan Kim, and Byoungcho Lee. Three-dimensional electro-floating display system based on integral imaging technique. *Optics Express*, 13(12):4358–4369, 2005. 29, 32
- [MIS04] Lydia M. J. Meesters, Wijnand A. IJsselsteijn, and Pieter J. H. Seuntjens. A survey of perceptual evaluations and requirements of three-dimensional tv. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(3):381–391, 2004. 9
- [MMIW09] Lambooij Marc1, Fortuin Marten, Heynderickx Ingrid, and IJsselsteijn Wijnand. Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Technology*, 53(3):1–14, 2009. 10
- [MP04] W. Matusik and H. Pfister. 3d tv: A scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scene. In *ACM SIGGRAPH*, 2004. 21
- [MWH97] Adam B. Kropp Michael W. Halle. Fast computer graphics rendering for full parallax spatial displays. In *Proc. SPIE Practical Holography XI and Holographic Materials III*, 1997. 31
- [N.94] Davies N. Design and analysis of an image transfer system using microlens arrays. *Optical Eng.*, 33(11):3624–3633, 1994. 23
- [NYH01] Takeshi Naemura, T. Yoshida, and H. Harashima. 3-d computer graphics based on integral photography. *Optics Express*, 8(4):255–262, 2001. 31
-

-
- [OAMO06] Fumio Okano, Jun Arai, Kohji Mitanni, and Makoto Okui. Real-time integral imaging based on extremely high resolution video system. *Proceeding of the IEEE*, 94(3):490–501, 2006. 32
- [OSO⁺06] L. Levent ONURA, A. Thomas SIKOR, N. Jm OSTERMAN, C. Aljoscha SMOLI, M. Reha Civanlar, and N. John WATSO. An assessment of 3dtv technologies. *NAB BEC*, pages 456–467, 2006. 10
- [Pas91] Richard V. Abadiand Eve Pascal. Visual resolution limits in human albinism. *Vision Research*, 31(7):1445–1447, 1991. 9, 10
- [PBK⁺08] Jae-Hyeung Park, Ganbat Baasantseren, Nam Kim, Gilbae Park, Jin-Mo Kang, and ByoungHo Lee. View image generation in perspective and orthographic projection geometry based on integral imaging. *Optics Express*, 16(12):8800–8813, 2008. 34
- [PHL09] Jae-Hyeung Park, Keehoon Hong, and ByoungHo Lee. Recent progress in three-dimensional information processing based on integral imaging. *Applied Optics*, 48(34):77–94, 2009. 4, 23
- [PJKL08] Gilbae Park, Jae-Hyun Jung, Joohwan Kim, and ByoungHo Lee. Computer-generation method for elemental image of integral floating display using virtual integral imaging system. In *Digital Holography and Three-Dimensional Imaging*, 2008. 33
- [PKL05] Jae-Hyeung Park, Joohwan Kim, and ByoungHo Lee. Three-dimensional optical correlation using a sub-image array. *Optical Express*, 13(13):5116–5126, 2005. 57
- [PLL11] Gilbae Park, Mu-Chieh Lo, and ByoungHo Lee. Pseudo-curved integral imaging pickup system with realtime view image generation. In *3DTV Conference*, 2011. 35
- [PMFN10] Melania Paturzo, Pasquale Memmolo, Andrea Finizio, and Thomas J. Naughton. Synthesis and display of dynamic holographic 3d scenes with real-world objects. *optics Express*, 18(9):8806–8815, 2010. 25, 26
- [PRA11] Raymond Phan, Richard Rzeszutek, and Dimitrios Androustos. Semi-automatic 2d to 3d image conversion using scale-space random walks and a graph cuts based depth prior. In *2011 18th IEEE International Conference on Image Processing (ICIP)*, 2011. 16
- [RBBHS01] Krishnamurthy Ravi, Chai Bing-Bing, Tao Hai, and Sethuraman Sri-ram. Compression and transmission of depth maps for image-based rendering. In *International Conference on Image Processing*, 2001. 19
- [Red02] Andrdert. Attest: Advanced three-dimensional television system technologies. In *3D data Processing Visualization and Transmission*, 2002. 9, 13
-

-
- [RHF10] Stephan Reichelt, Ralf Hsler, and Gerald Ffterer. Depth cues in human visual perception and their realization in 3d displays. In *SPIE 7690*, 2010. 10
- [rKG⁺12] V.Srinivasa rao, Dr P.Rajesh Kumar, G.V.H.Prasad, M.Prema Kumar, and S.Ravichand. Discrete cosine transform vs discrete wavelet transform: An objective comparison of image compression technique for jpeg encode. *International Journal of Advanced Engineering Applications*, 1(5), 2012. 64
- [SCA⁺08] N. P. Sgouros, D. Chaikalis, S. S. Athineos, D. Maroulis, and N. Theofanous. Performance considerations for a real-time integral image camera in ray tracing environment. In *5th international conference on Image Analysis and Recognition*, 2008. 24
- [SCCKES11] Kim Seung-Cheol, Kim Chang-Keun, and Kim Eun-Soo. Depth-of-focus and resolution-enhanced three-dimensional integral imaging with non-uniform lenslets and intermediate-view reconstruction technique. *3D Research*, 2(2):1–9, 2011. 30
- [SGwHS98] Jonathan Shade, Steven Gortler, Li wei He, and Richard Szeliski. Layered depth images. In *SIGGRAPH*, 1998. 61
- [SLsK06] Dong-Hak Shin, ByoungHo Lee, and Eun soo Kim. Improved viewing quality of 3-d images in computational integral imaging reconstruction based on lenslet array model. *ETRI journal*, 28(4):521–524, 2006. 37, 40, 81
- [SMsK05] D.-H. Shin, K.-c. Park M.Cho, and E. s. Kim. Computational technique of volumetric object reconstruction in integral imaging by use of real and virtual image fields. *ETRI Journal*, 27(6):708–712, 2005. 86
- [SS99] I Sexton and P. Surman. Stereoscopic and autostereoscopic display systems. *Signal Processing Magazine*, 16(3):85–99, 1999. 12, 21
- [SSG11] Gioia Patrick Shi Shasha and Madec Gerard. Efficient compression method for integral images using multi-view video coding. In *18th IEEE International Conference on Image Processing (ICIP)*, 2011. 58
- [SY09] Dong-Hak Shin and Hoon Yoo. Computational integral imaging reconstruction method of 3d images using pixel-to-pixel mapping and image interpolation. *Optics Communications*, 282(14):2760–2767, 2009. 88
- [TIY95] Motoki T., H. Isono, and I. Yuyama. Present status of three-dimensional television researc. *Proceedings of the IEEE*, 83(7):1009–1021, 1995. 22
- [Umb06] Eden Ashley Umble. Making it real: the future of stereoscopic 3d film technology. *ACM SIGGRAPH Computer Graphics Homepage archive*, 40(1), 2006. 3
-

- [VA08] Vetro and Anthony. Toward the light field display: Autostereoscopic rendering via a cluster of projectors. *IEEE Transaction on Visualization and Computer Graphics*, 14(1):1–12, 2008. 23
- [WPM02] W.A.IJsselstein, P.J.H.Seuntings, and L.M.J. Meesters. State-of-the-art in human factors and quality issues of stereoscopic broadcast television. In *ATTEST Project Deliverable 1*, 2002. 1, 10, 14
- [YHLJ11] Ruigang Yang, Xinyu Huang, Sifang Li, and Christopher Jaynes. 3d-tv content storage and transmission. *IEEE Transactions on Broadcasting - Special Issue on 3D-TV Horizon: Contents, Systems and Visual Perception*, 57(2):384–394, 2011. 21
- [YM12] Young Min Kim ; Joon Kyu Yim and Sung-Wook Min. Pre-distorted elemental image generation for off-axis integral floating system. In *Optics and Photonics for Information Processing VI*, 2012. 31
- [YSJ04] Sekwon Yeom, Adrian Stern, and Bahram Javidi. Compression of 3d color integral images. *Optical Express*, 12(8):1632–1641, 2004. 53
- [ZAM02] R. Zaharia, A. Aggoun, and M. McCormick. Adaptive 3d-det compression algorithm for continuous parallax 3d integral imaging. *Signal Processing: Image Communication*, 17(3):231–242, 2002. 53, 54
- [ZW12] Piao Yan Zheng Wanying. Research on integral (3d) image compression technology based on neural network. In *2012 5th International Congress on Image and Signal Processing (CISP)*, 2012. 49
-

Publications

International Conferences

- [1] Shasha SHI, Patrick GIOIA, Gérard MADEC, 'High quality computational integral imaging reconstruction using view synthesis rendering", in *IMPVT'2011: first joint 3DIM/3DPVT conference*, Hangzhou, China, 16-19 may 2011.
- [2] Shasha SHI, Patrick GIOIA, Gérard MADEC, 'High efficient computational integral imaging reconstruction based on parallel-group projection (PGP) method", in *2nd Annual SMPTE : International Conference on Stereoscopic 3D for Entertainment & Media*, New York, United States, 21-22 June 2011.
- [3] Shasha SHI, Patrick GIOIA, Gérard MADEC, 'Efficient compression method for integral images using multi-view video coding", in *ICIP 2011: International Conference on Image Processing*, Brussels, Belgium, 11-14 September 2011.

Journal

- [1] Shasha SHI, Patrick GIOIA, Gérard MADEC, High performance computational integral imaging system using multi-view video plus depth representation", in *«3D Research», Topic issue: Real 3D Imaging & Display*.
-

-O-

OII	Optical Integral Imaging
OIRR	Optical Integral Image Reconstruction
OSG	OpenSecneGraph

-P-

PGP	Parallel-Group Projection
PRR	Point Retracing Rendering
PGP	Parallel Group Projection
ROI	Reconstructed Output Image

-S-

SI	Sub-Images
----	------------

-V-

VCEG	Video Coding Experts Group
------	----------------------------

-Z-

ZPD	Zero Parallax Distance
-----	------------------------
