

國立交通大學

資訊科學與工程研究所

碩士論文

應用於三維立體視訊之低複雜度深度圖

壓縮演算法

A Low-Complexity Depth Map Coding Algorithm
for 3-D Videos

研究生：楊復堯

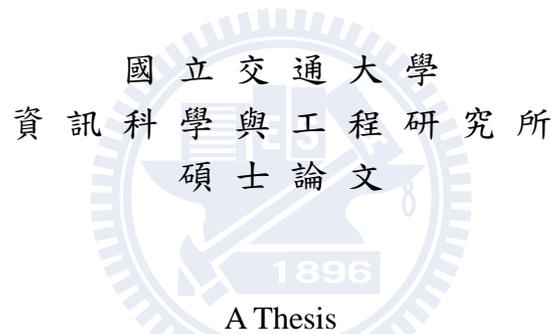
指導教授：彭文孝 助理教授

中華民國九十九年九月

應用於三維立體視訊之低複雜度深度圖壓縮演算法
A Low-Complexity Depth Map Coding Algorithm for 3-D Videos

研究生：楊復堯 Student: Fu-Yao Yang

指導教授：彭文孝 Advisor: Dr. Wen-Hsiao Peng



A Thesis
Submitted to Institute of Computer Science and Engineering
College of Computer Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
in
Computer Science

August 2010

Hsinchu, Taiwan, Republic of China

中華民國九十九年九月

應用於三維立體視訊之低複雜度深度圖壓縮演算法

學生：楊復堯

指導教授：彭文孝 博士

國立交通大學資訊科學與工程研究所碩士班

中文摘要

MPEG 自由視角電視標準 (FTV) 之架構，包含深度資訊；若深度資訊經由現有的壓縮標準 (MVC) 壓縮，則會產生兩個問題：一、在物體邊界處容易造成深度資訊之失真，並影響合成虛擬影像之主、客觀品質；二、MVC 有兩種預測方式，分別為 Intra 與 Inter 之預測方式，由於壓縮後深度圖與自然影像之資料量相差懸殊，即使利用運算複雜度較 Intra 高的 Inter 預測方式也顯得沒有壓縮效率。本論文闡釋一個利用深度圖之特性，設計簡單且有效率之深度圖壓縮演算法。

本篇論文將分析單張深度圖之特性並使用眾數演算法挑選需要壓縮之深度資訊。其次，將需要壓縮之深度資訊以預測模式之方式，估測該深度資訊，企圖減少需壓縮之資訊量。最後，經由壓縮重建後的深度圖使用線性權重之過濾器以修補深度圖。實驗結果顯示，本文提出的方法在不影響合成結果之 PSNR 情況下，其需要的網路頻寬較 MVC 少 4~13%，且運算複雜度也較 MVC 之 Intra 預測方式少 5% ~ 15%。

A Low-Complexity Depth Map Coding Algorithm for 3-D Videos

Student : Fu-Yao Yang

Advisor : Dr. Wen-Hsiao Peng

Institute of Computer Science and Engineering

College of Computer Science

National Chiao Tung University

Abstract

In the framework of MPEG Free-viewpoint Television (FTV), the depth information is included. Once the depth information is coded by Multi-view Video Coding (MVC), two kinds of problems may occur. The first one is the depth errors around the boundaries of objects, which might affect the quality of rendered virtual view. The other is that inter prediction cannot provide high coding efficiency for depth map. The size of compressed image is much higher than depth map. Therefore, a simple and efficient coding method is proposed based on the characteristic of the depth map.

The proposed method analyzes the characteristic of a single depth map and uses mode determination to calculate the depth information. Then, the depth information is predicted by mode decision to increase the compression efficiency and linear weighted filters are used to refine the reconstructed depth map. The experiments show our methods save 4-13% bandwidth with the same PSNR and the computation complexity is less than the one needed for Intra prediction.

誌謝

回顧這兩年的碩士生涯，首先我要先感謝我的指導教授—彭文孝 博士，彭老師對於學術研究實事求是、追根究底的態度，已成為日後我在學習與研究的典範。其次，我要感謝我的學長—陳俊吉 博士，不怕任何的麻煩與辛勞和我討論，並適時地給予建議修正已偏差的研究方向，讓我在這兩年的研究可以更加順利。謹此對於兩人獻上由衷的感謝。

有幸成為多媒體架構與處理實驗室的成員之一，在這個充滿學習氣氛的實驗室，可以和其他成員一起討論相關研究與任何對我有幫助的知識，這是我在大學畢業後過的最充實的一段時間。感謝我的學長們—陳漪紋 博士、陳俊吉 博士、詹家欣 博士、林哲永與陳建穎，在研究上給我適時的意見與建議；感謝我的同學們，蔡閔旭、王澤瑋與吳思賢，不管在課業或是研究上，都可以給予幫助；感謝我的學弟妹們，吳崇豪、曾于真、陳孟傑與黃嘉彥，提供任何的協助。

最後我要感謝我的家人—楊奇岳 先生、蔡治 女士的栽培，可以在沒有任何的煩惱之下，全力爭取碩士資格。感謝我的兄長—楊竣能與楊竣閔，適時地提供生活方面的協助。感謝我的女友—許嘉琪，這幾年來辛苦的陪伴、體貼和關心。最後再一次的感謝我的老師、家人、朋友們，有你們的鼓勵，我才可以獲得碩士學位，謝謝你們。

目錄

| | |
|---|-----|
| 中文摘要 | i |
| Abstract | ii |
| 誌謝 | iii |
| 目錄 | iv |
| 圖目錄 | v |
| 表目錄 | vii |
| 第一章、緒論 | 1 |
| 1.1 背景介紹 | 1 |
| 1.2 研究動機 | 3 |
| 1.3 研究貢獻 | 4 |
| 1.4 論文編排 | 5 |
| 第二章、相關研究 | 6 |
| 2.1 Depth-Image-based Rendering | 6 |
| 2.2 多重視角深度圖壓縮演算法 | 7 |
| 2.3 改善 MVC 之移動向量架構 | 8 |
| 2.4 Platelet | 9 |
| 2.5 Geometry-based Block Partitioning for Intra Prediction Method | 10 |
| 2.6 簡易修補深度圖演算法 | 11 |
| 第三章、深度圖之分析 | 12 |
| 第四章、研究方法 | 19 |
| 4.1 Mode Determination、Quantization、Residual Extraction | 19 |
| 4.2 Residual Coding | 22 |
| 4.2.1 Line-based Residual Segmentation | 24 |
| 4.2.2 Removal of Isolated Segments | 24 |
| 4.2.3 Symbol Compaction | 24 |
| 4.2.4 Entropy Coding | 27 |
| 4.3 Depth Map Refinement | 27 |
| 第五章、實驗結果與討論 | 30 |
| 第六章、結論與未來展望 | 33 |
| 附錄 | 41 |
| 參考文獻 | 42 |

圖目錄

| | | |
|------|--|----|
| 圖 1 | (a)自然影像；(b) 與(a)對應之深度圖 | 1 |
| 圖 2 | (a) 原始深度圖及其合成影像；(b) 重建深度圖及其合成影像 | 2 |
| 圖 3 | 演算法之流程圖..... | 5 |
| 圖 4 | 根據參考影像及其對應之深度圖，合成虛擬影像..... | 7 |
| 圖 5 | MVC 之壓縮架構，紅色箭頭為 Inter-view prediction；黑色箭頭為 Temporal prediction [9] | 8 |
| 圖 6 | 經由 Platelet 之壓縮方法產生的切割區塊 [9]..... | 9 |
| 圖 7 | 預測欲壓縮區塊之分割方式，左圖為參考左邊之區塊，右圖為參考右邊 之區塊 [10]..... | 10 |
| 圖 8 | 修補深度值之範例..... | 11 |
| 圖 9 | 不同影片之合成虛擬影像品質比較。..... | 13 |
| 圖 10 | 比較 Intra 與階層 B 壓縮架構之合成虛擬影像品質 | 14 |
| 圖 11 | 比較 Intra 與 階層 B 影像壓縮架構之主觀視覺。(a) Intra；(b) 階層 B 影像壓縮架構 | 15 |
| 圖 12 | 比較 Intra 與階層 B 影像壓縮架構之深度圖與自然影像資料量比例關 係。(a) 原圖；(b) 紅色框框之放大圖 | 16 |
| 圖 13 | 統計深度圖與自然影像的區塊出現不同值的個數..... | 18 |
| 圖 14 | Residual 之分布圖..... | 20 |
| 圖 15 | 公式 3 之 TH1 設定對主觀視覺之影響。(a) TH1 條件較嚴苛；(b) TH1 條件較寬鬆..... | 21 |
| 圖 16 | 眾數填入對應區塊產生的深度圖..... | 21 |
| 圖 17 | Residual Coding 流程圖..... | 22 |
| 圖 18 | 深度圖產生 Residual 之示意圖。(a) 原始深度圖；(b) 眾數深度圖；(c) Residual 深度圖；(d) DC (非黑色) 與 Non-DC (黑色) 片段深度圖；(e) 傳輸至接收端之資訊；(f) Residual 片段之預測方向 | 23 |
| 圖 19 | Non-DC 片段分割圖..... | 25 |
| 圖 20 | Iso-residual segments 被移除而產生錯誤..... | 25 |
| 圖 21 | 重建後深度圖失真之區域。(a) 非灰色區域為本篇論文產生失真之區域； (b) 非灰色區域為 MVC 產生失真之區域..... | 28 |
| 圖 22 | 深度圖之區塊經由壓縮重建後產生失真。(a) 藍色虛線為原始深度圖之 訊號，紅線重建後之訊號為；(b) 比較不同大小之修補過濾器，(1) 為較 小之過濾器，(2) 為較大之過濾器..... | 28 |
| 圖 23 | 修補深度值之範例..... | 29 |
| 圖 24 | 修補深度圖之前後，其合成虛擬影像之主觀視覺之比較。(a) 修補前； (b) 修補後..... | 29 |

| | | |
|------|---|----|
| 圖 25 | MVC 與本篇論文演算法合成虛擬影像之 R-D 曲線..... | 34 |
| 圖 26 | Champagne Tower, 深度圖與其合成虛擬影像之主觀視覺比較。(a) 原始深度圖 ; (b) MVC Intra Prediction ; (c) Proposed | 36 |
| 圖 27 | Book Arrival, 合成虛擬影像之主觀視覺比較。(a) Original ; (b) MVC All Intra Prediction ; (c) MVC Inter Prediction ; (d) Proposed | 37 |
| 圖 28 | Door Flowers, 合成虛擬影像之主觀視覺比較。(a) Original ; (b) MVC All Intra Prediction ; (c) MVC Inter Prediction ; (d) Proposed | 38 |
| 圖 29 | Leaving Laptop。(a) Original ; (b) MVC All Intra Prediction ; (c) MVC Inter Prediction ; (d) Proposed | 39 |
| 圖 30 | Lovebird1, 合成虛擬影像之主觀視覺比較。(a) Original ; (b) MVC All Intra Prediction ; (c) MVC Inter Prediction ; (d) Proposed | 39 |
| 圖 31 | Newspaper, 合成虛擬影像之主觀視覺比較。(a) Original ; (b) MVC All Intra Prediction ; (c) MVC Inter Prediction ; (d) Proposed | 40 |



表目錄

| | | |
|-----|---|----|
| 表 1 | 比較 Intra 與 階層 B 影像壓縮架構之平均運算複雜度 (單位：秒) | 15 |
| 表 2 | DC 片段之預測模式 | 26 |
| 表 3 | 合成虛擬影像之參數設定 | 30 |
| 表 4 | JMVC 之參數設定 | 30 |
| 表 5 | 本篇論文演算法與 MVC 之平均運算複雜度比較 (單位：秒) | 32 |
| 表 6 | 本篇論文演算法與 MVC 之平均運算個數比較 | 32 |
| 表 6 | 深度圖限制在高資料量的情況下，R.I. 預測模式的比例關係 | 41 |
| 表 7 | 深度圖限制在低資料量的情況下，R.I. 預測模式的比例關係 | 41 |

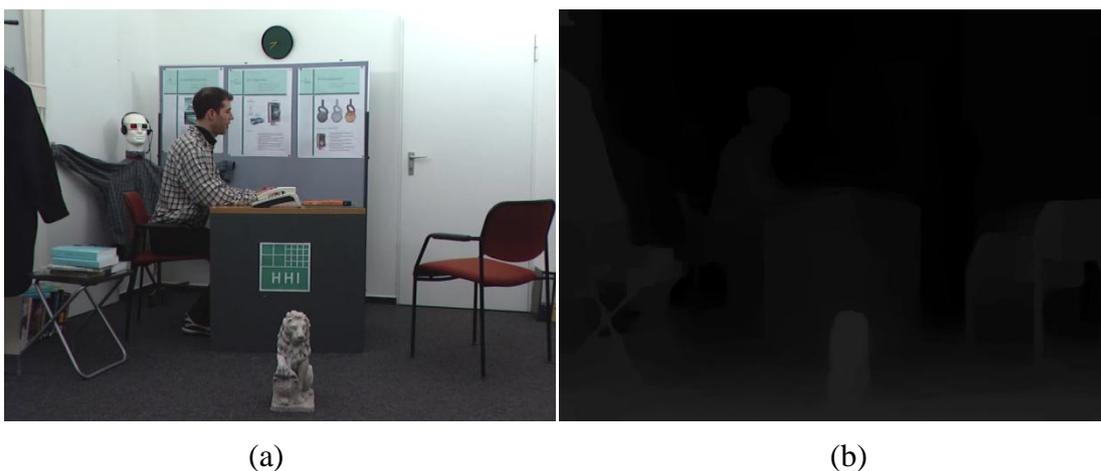


第一章、緒論

1.1 背景介紹

多重視角影片是目前熱門的應用之一，和過去不同是，傳統影片不能隨使用者的喜好任意切換視角；而現在卻可以讓使用者依照自己的喜好，選擇自己想要觀賞場景的視角。為了達到此目的，廠商需要在任一個視角加裝攝影機以獲得該角度之影片；後果是除了會大量增加硬體成本之外，在傳輸影片的網路頻寬上也面臨極大的挑戰。因此，如何減少硬體成本和傳輸影片所需要的網路頻寬便是目前需要解決的課題，而多重虛擬視角合成技術可以解決上述之問題。

近幾年，MPEG 組織為了能夠有效減少硬體成本且達到可以觀賞任意視角影片的目標，而訂立新的資料格式 — Multi-view video plus depth [1]，此格式用來表示多重視角影片及任何視角之像素都各自擁有深度值；深度值為相機與實際物體之深度關係，此深度值可描述 3D 場景之幾何空間資訊。一旦具備 3D 幾何空間的資訊，即可利用現有的影片及其對應之深度值以合成任意視角的影片，這個技術稱為 Depth-image-based rendering (DIBR) [2][3][4][5]。每個像素對應的深度值，若以影像的方式儲存之，則稱之為深度圖。因此，深度圖是以二維影像表示該自然影像在 3D 場景之深度值，提供合成虛擬影像之用，如圖 1。



(a)

(b)

圖 1 (a)自然影像；(b) 與(a)對應之深度圖

DIBR 在內容格式之要求上，需要參考影像及其對應的深度圖。在接收端，深度圖的取得方式有兩種：一、接收端獨立產生各個參考視角之深度圖，其缺點為運算複雜度非常可觀，無法即時產生合成虛擬視角之影像，而延遲播放影片之速度。二、深度圖經由傳送端傳輸至接收端，但是每個像素深度值是由 8 個位元所組成，所以深度圖佔全部資料量的 25%，如果深度圖資訊不經過壓縮的話，勢必在網路頻寬的限制下，亦無法即時合成虛擬影像以提供使用者觀賞。

為了減少網路頻寬的使用，MPEG 組織致力於多重視角影片格式的壓縮，以 H.264/AVC 為基礎，結合 Temporal 和 Inter-view 的移動補償及視差補償估測技術，發展出多重視角影片壓縮（Multi-view Video Coding，簡稱 MVC）之標準，此壓縮標準亦適用於深度圖之壓縮。

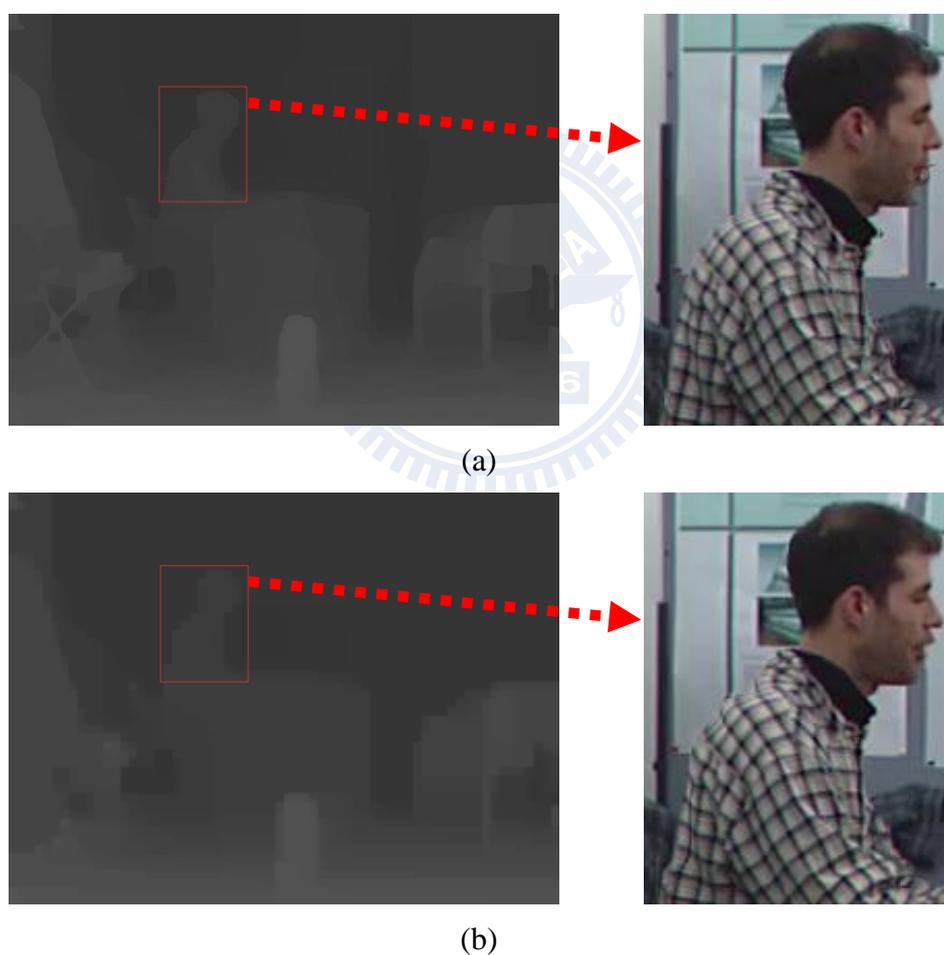


圖 2 (a) 原始深度圖及其合成影像；(b) 重建深度圖及其合成影像

1.2 研究動機

由於現今的影像壓縮技術是針對人眼視覺對自然影像之邊界及細緻的地方較不敏銳之特性所設計，而 MVC 亦是如此，並且在壓縮深度圖上達到不錯的壓縮效率。MVC 之特點是原本由空間定義上的 Residual 值經過離散餘弦轉換後成為頻率定義的資料型態，接著再量化訊號而產生失真；但是，深度圖是表示 3D 場景的幾何空間之資訊，對於物體邊界及細緻的地方不能產生失真，否則會產生嚴重的合成影像瑕疵。因此，深度圖和自然影像之特性是完全不同，需要有一套針對深度圖壓縮之演算法；如果選用 MVC 來壓縮深度圖，則深度圖容易在高、低頻訊號產生幾何資訊上的失真，尤其高頻訊號之失真會在合成虛擬影像時產生不可預期的瑕疵。如圖 2，深度圖的人之背部經由 MVC 壓縮並重建後，產生失真，後果是在合成虛擬影像時，人之背部產生明顯的合成錯誤。

為了合成較高品質的虛擬影像，經實驗結果顯示，參考影像之品質不能太差，所以本篇論文採用高品質的自然影像為參考影像，提供合成虛擬影像之用途。根據實驗統計，分別使用 MVC 的 Intra 和階層 B 影像壓縮架構（Hierarchical B Picture Coding Structure）之預測方式來壓縮深度圖，Intra 之深度圖與自然影像資料量的比例為 4%~9%，階層 B 影像壓縮架構之深度圖與自然影像資料量的比例為 1%~3%，由此可以得知兩者之深度圖資料量遠小於自然影像資料量；不過階層 B 影像壓縮架構的運算複雜度卻為 Intra 之數十倍，雖然使用階層 B 影像壓縮架構預測模式可以提高資料壓縮效率，不過其運算複雜度卻遠超過資料壓縮之好處。因此，考慮深度圖資料量遠不及自然影像資料量之情況下，選用 Intra 之預測方式來壓縮深度圖，企圖降低深度圖之運算複雜度。

除此之外，對於主觀視覺或是客觀分析的比較上，顯示深度圖資訊較原始影像單純。因此，本篇論文希望在不影響主、客觀的品質之下，針對深度圖之特性提出簡單且更有效率的方式，壓縮深度圖。

1.3 研究貢獻

本篇演算法與 MVC 之 Intra 預測方式不同的地方最主要可以分成三種：

1. 預測深度值部份，MVC 採用相鄰區塊的像素深度值估測目前區塊的像素深度值；本篇論文在每個區塊採用眾數演算法，挑選代表該區塊之深度值。
2. 資料量化器設計部份，MVC 採用純量量化器精簡資料輸出值的種類數至一個較小的集合裡；本篇論文則以門檻量化器將不同於眾數之像素深度值，取代為眾數深度值。
3. 資料型態部份，MVC 將空間定義的 Residual 值經過離散餘弦轉換後成為頻率定義之資料型態；本篇論文直接採用空間定義的資料型態處理像素深度值。

因此，本篇論文提出壓縮深度圖之架構，如圖 3，其最主要之貢獻如下：

1. 傳送端壓縮深度圖之前，每個區塊經由眾數演算法，挑選出眾數及需要傳輸至接收端之剩餘深度資訊，本篇論文稱此為 Residual，部份與眾數不同的像素深度值會經由門檻量化器而產生失真。
2. 傳送端使用 1D 掃瞄的方式將剩餘深度資訊排列，並比較相鄰深度資訊間是否可以經由估測之方式重建該深度值。最後資訊以無失真壓縮之方式傳輸至接收端。
3. 當接收端重建深度圖之後，根據前處理端之眾數演算法產生的幾何資訊失真，設計線性權重的過濾器以重建該像素之深度值。

實驗結果顯示，在不影響客觀品質評量 PSNR 的情況下，相較於 MVC Intra 之預測方式，本論文之演算法不僅可以再下降 4%~13% 的資料量，且運算複雜度也會下降 5%~15%。在同樣的網路頻寬的情況下，合成虛擬影像之主觀視覺也較 MVC Intra 之預測方式好。

1.4 論文編排

本篇論文的編排如下:第二章為 DIBR 以及現今深度圖壓縮演算法之簡介;第三章為針對深度圖特性作分析;第四章為介紹本篇論文所提出的深度圖壓縮演算法;第五章為 MVC 與本篇論文演算法的主、客觀品質之比較;第六章為本篇論文之結論與未來展望。

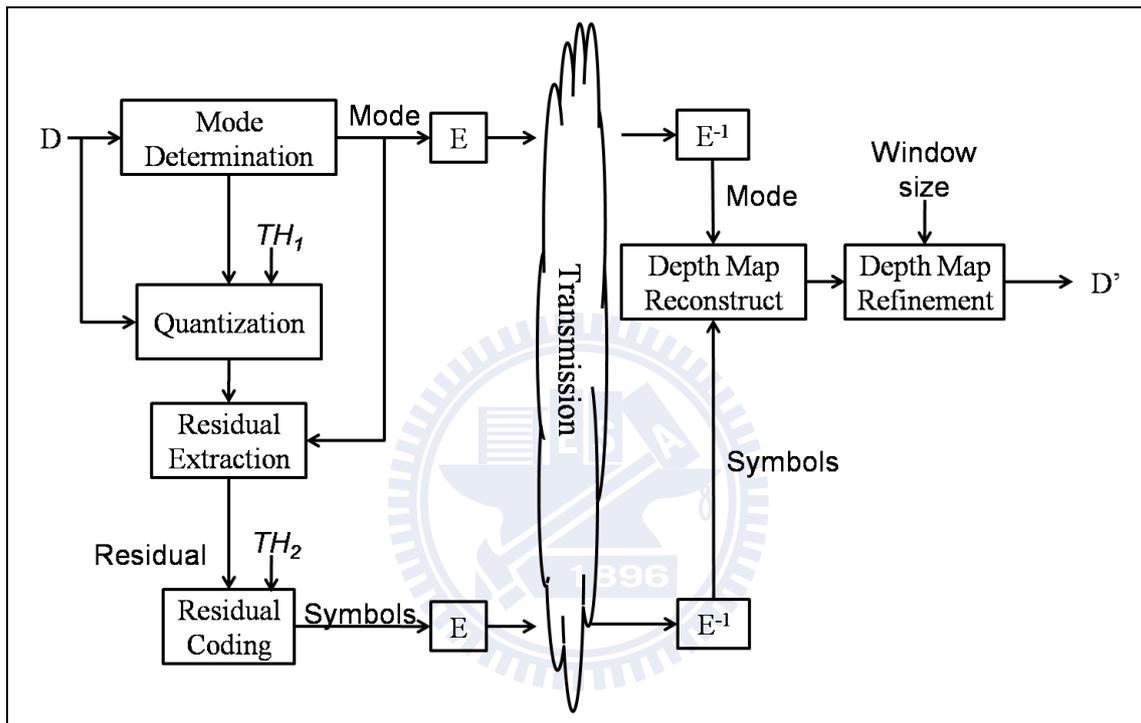


圖 3 演算法之流程圖

第二章、 相關研究

2.1 Depth-Image-based Rendering

Depth-Image-based Rendering (DIBR) 是一門合成影像的技術，透過參考視角之影像及其對應的深度值，合成虛擬影像。依序步驟為：參考視角之影像利用其對應的深度值，投影至 3D 幾何空間，最後再投影至欲合成之位置。本篇論文假設每個視角之相機位置為互相平行，因此投影及合成動作可以用向量的式子 ψ 來表示參考視角之影像的任一點 $\mathbf{p} = [x, y]^T$ ，對應在合成影像上的特定點 $\mathbf{p}' = [x', y']^T$ ，其式子如下：

$$\psi: \begin{bmatrix} \mathbf{p}' \\ 1 \end{bmatrix} = A' R A^{-1} \begin{bmatrix} \mathbf{p} \\ 1 \end{bmatrix} + \frac{1}{Z_p} A' T \quad (1)$$

其中 R 和 T 分別為旋轉和位移矩陣，是參考影像與虛擬影像之間的相對位置關係； A' 和 A 分別為虛擬和參考相機之內部參數； Z_p 是參考影像之像素 \mathbf{p} 對應的深度值。

透過公式 1，使用者只要有相機參數、虛擬和參考視角的相對位置關係及深度值，便可以合成任何視角的虛擬影像；但實際上，在合成虛擬影像時，會面臨 Occlusion 之問題：由於物體與物件之間會有遮蔽效應，所以如果在合成影像可以看到的物體，但在參考影像卻找不到對應的物體，則會在合成影像產生「洞」，此現象會隨著合成影像與參考影像之距離增加而更加明顯。

為了解決 Occlusion 的問題，MPEG 組織使用二個以上的參考影像實現合成虛擬影像之目的。由圖 4 中可以發現，在合成影像中的大部份的像素，將會對應到任何參考影像之特定點，而參考影像之對應點對虛擬影像之像素的影響力，會與參考視角及虛擬視角之間的距離呈反比關係，所以透過虛擬與參考影像之間距離的權重關係，決定合成虛擬影像的像素值為何；至於少部份的像素因為 Occlusion 的關係而形成「洞」，該像素值則使用鄰近已合成的像素值內插之。

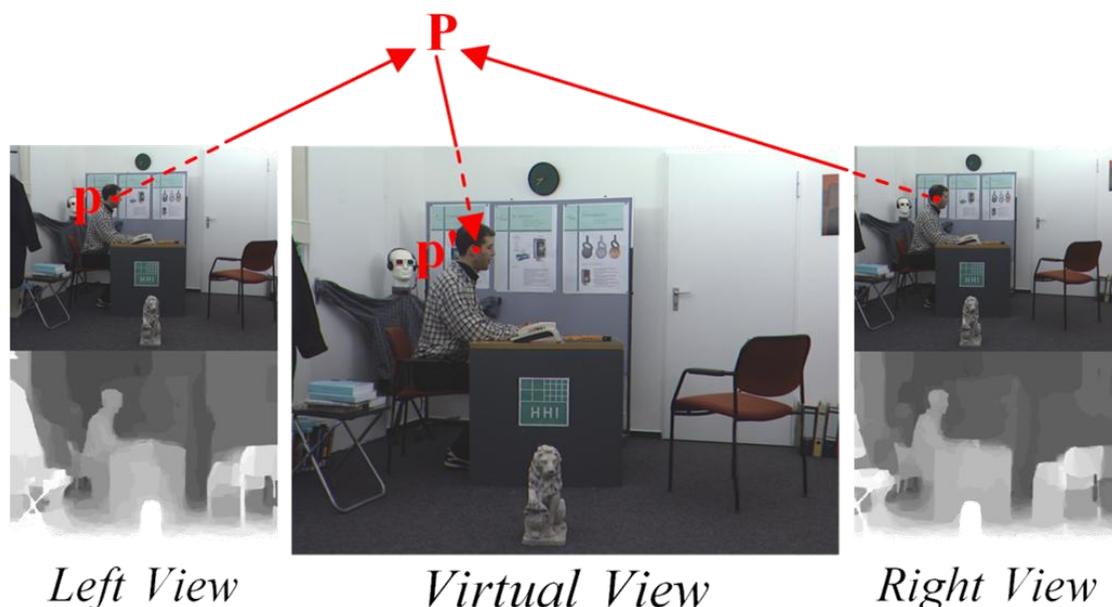


圖 4 根據參考影像及其對應之深度圖，合成虛擬影像

2.2 多重視角深度圖壓縮演算法

AVC 常用於單一視角影片的壓縮標準，此壓縮演算法在時間排列影像的統計上具有高度相關性，因此發展 AVC Temporal 的壓縮架構，分別為 Intra (I)、Predictive (P) 以及 Bi-predictive (B) 影像組成。對於 I 影像來說，該張畫面內容只能由該張畫面作預測，P 和 B 影像則可以參考其他畫面的資訊。為了能夠有效增加壓縮效率，AVC 發展出階層 B 影像壓縮架構，其 B 影像內容可以參考其它 I、P 和 B 影像的資訊，經由估測後再重建該 B 影像。

MVC 包含多個單一視角影片的壓縮標準，所以 MVC 壓縮標準是以 AVC 為基礎而設計之壓縮演算法。MVC 除了有單一視角影片之特性之外，不同視角的影像在統計上也具有高度相關性。因此，MVC 分別設計移動補償與視差補償之估測方法，企圖提高壓縮效率。如圖 5，為 MVC 標準之壓縮架構，對於 cam1 本身來說，是三個階層的 Temporal 估測，其 Group of Pictures (GOP) 大小為 8 來實現階層 B 畫面之壓縮架構，黑色箭頭表示在同一視角影片當中，該張影像內容由其他參考影像之內容做估測；另外，cam1 與其他視角使用 Inter-view 之估測方式企圖減少需要壓縮的資訊，紅色箭頭表示該視角影像之內容由其它參考視角影像做估測。

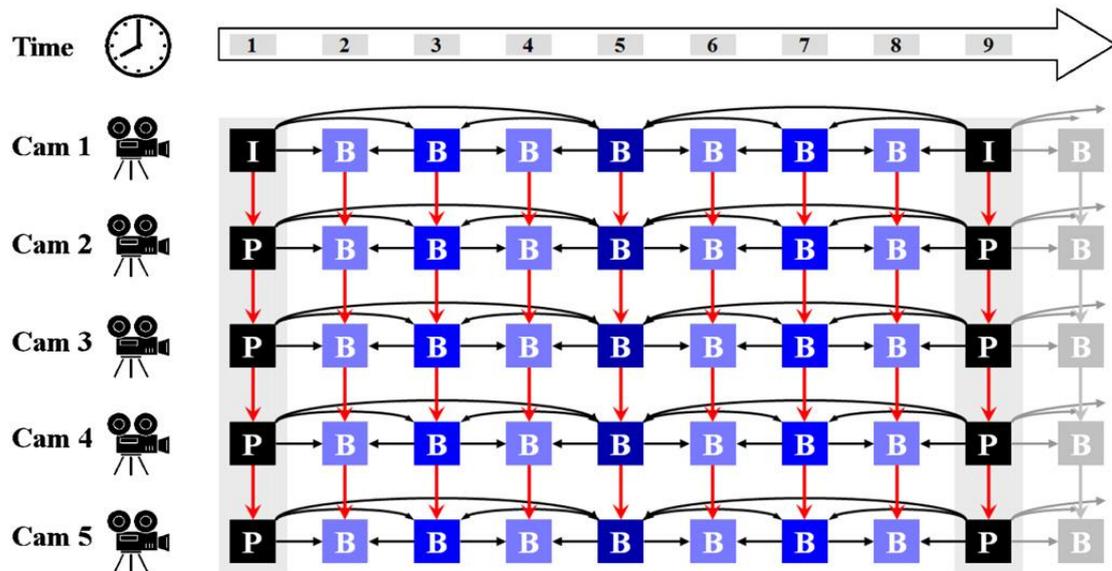


圖 5 MVC 之壓縮架構，紅色箭頭為 Inter-view prediction；黑色箭頭為 Temporal prediction [9]

2.3 改善 MVC 之移動向量架構

以 MVC 架構為基礎，也有人提出一些不同的方法來壓縮深度圖。[6] 和 [7] 嘗試將原本 MVC 的 2D 移動預測增加為 3D 移動預測，其向量表示分別為 X、Y 和 Z (深度值)。X 和 Y 和原來的 2D 移動預測一樣，為影像最小 Rate-Distortion 的移動向量；Z 為移動向量 X 和 Y 決定後，選擇最佳的深度值偏移量，使得該選擇區塊之深度值與需要壓縮區塊之深度值的失真量為最小。因此，3D 移動預測可以降低需要壓縮 Residual 的資訊。[8] 利用自然影像之移動向量以預測深度圖之移動向量，其原因為自然影像和深度圖之移動向量具有某種程度上的相關性，好處為深度圖可以直接使用自然影像之移動向量，不需要再額外傳輸深度圖之移動向量至接收端。

2.4 Platelet

[9] 提出 Platelet 的壓縮方法。首先，深度圖以固定大小的方式分割成若干區塊，區塊內所有的像素深度值以公式 2 進行估測，其中公式 2 之 x 、 y 為像素在深度圖之位置；參數 A 、 B 、 C 為區塊以 Least-Square 之最佳化方式得之。接著，計算此區塊經由公式 2 所估測的像素深度值與實際深度值之間的誤差，若誤差超過使用者所設定的門檻條件，則此區塊需要再細分為四個小區塊，四個小區塊分別再做一次 Least-Square 之最佳化，得到公式 2 之 A 、 B 、 C 參數，直到滿足該門檻條件，如圖 6，為深度圖經由 Platelet 方式所切割後之結果。最後，需將深度圖切割的資訊與公式 2 之 A 、 B 、 C 參數傳輸至接收端，以供重建深度圖之用

$$F(x, y) = Ax + By + C \quad (2)$$

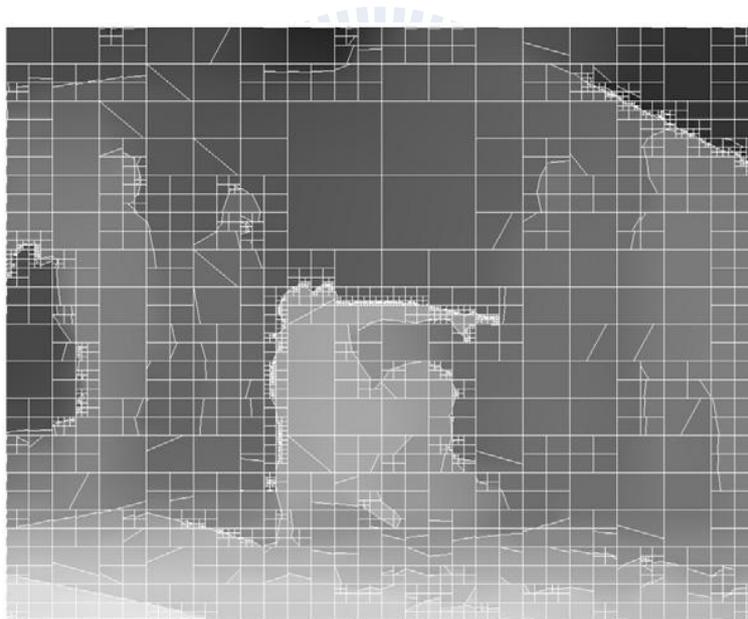


圖 6 經由 Platelet 之壓縮方法產生的切割區塊 [9]

2.5 Geometry-based Block Partitioning for Intra Prediction

Method

[10] 以 AVC 壓縮架構為基礎，利用相鄰的區塊和目前需要壓縮的區塊在空間定義上有高度相關之特性，提出 Geometry-based Block Partitioning for Intra Prediction Method。其作法為：首先，相鄰區塊（左方及上方）之分割線經由 Edge Detection Filter 產生，並以極座標的方式表示該分割線之位置；之後，藉由相鄰區塊產生之分割線段往目前需要壓縮的區塊作延伸，得到目前區塊的預測分割方式，並與傳統 AVC 所有 Intra 預測模式比較 Rate-Distortion Cost，以決定最適合該區塊之預測模式。此方法之特色為傳送端不需要傳送相鄰區塊由極座標產生分割線段之資訊至接收端，因為相鄰區塊為已重建的相鄰區塊。透過目前區塊決定的分割方式來預測該區塊之像素深度值，如圖 7，Partition 0 的區域由相鄰區塊的深灰色深度值做預測；Partition 1 的區域由淺灰色深度值做預測。

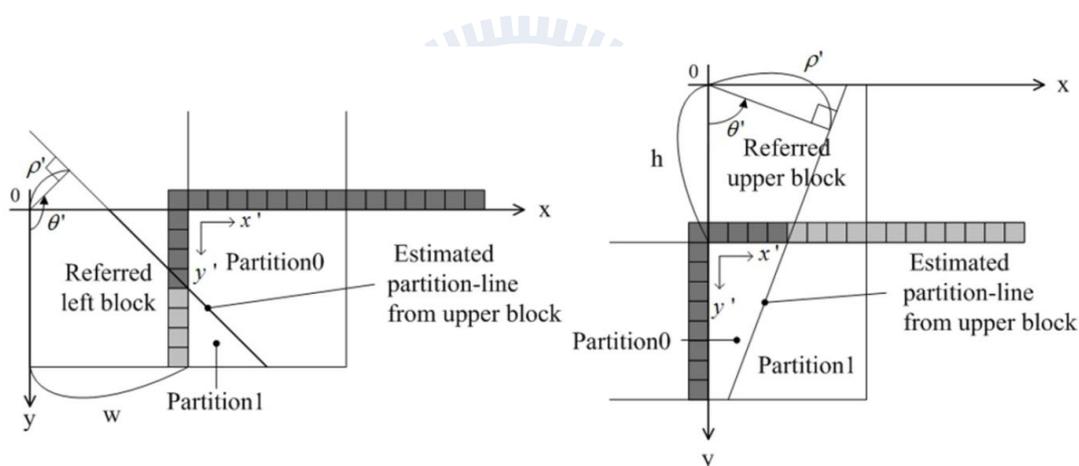


圖 7 預測欲壓縮區塊之分割方式，左圖為參考左邊之區塊，右圖為參考右邊之區塊 [10]

2.6 簡易修補深度圖演算法

修補深度圖方面，[11] 嘗試不改變現有 AVC 壓縮架構之情況下，在接收端設計簡單的過濾器以修補重建後的深度圖，企圖降低合成影像產生之瑕疵。其方法為：作者假定深度圖裡所有像素深度值需要被修補，因此分別以像素深度值為中心，以固定區塊大小之方式，找出該區塊之中位數。接著，將區塊內所有深度值分成大於中位數和小於中位數兩個集合，之後兩集合以眾數方式決定該集合所代表的深度值並與原本需要修補的深度值做比較。最後，選擇相差絕對值較小的眾數深度值而取代原本像素深度值。

以圖 8 為例，15 為需要修補之深度值，12 為此區塊深度值之中位數，深度值比 12 小的集合裡有 9、11、11、11 之元素；深度值比 12 大的集合裡有 15、21、21、21 之元素，兩集合之眾數分別為 11 和 21。由於 15 和 11 之絕對值差小於 15 和 21，所以最後修補之深度值為 11。

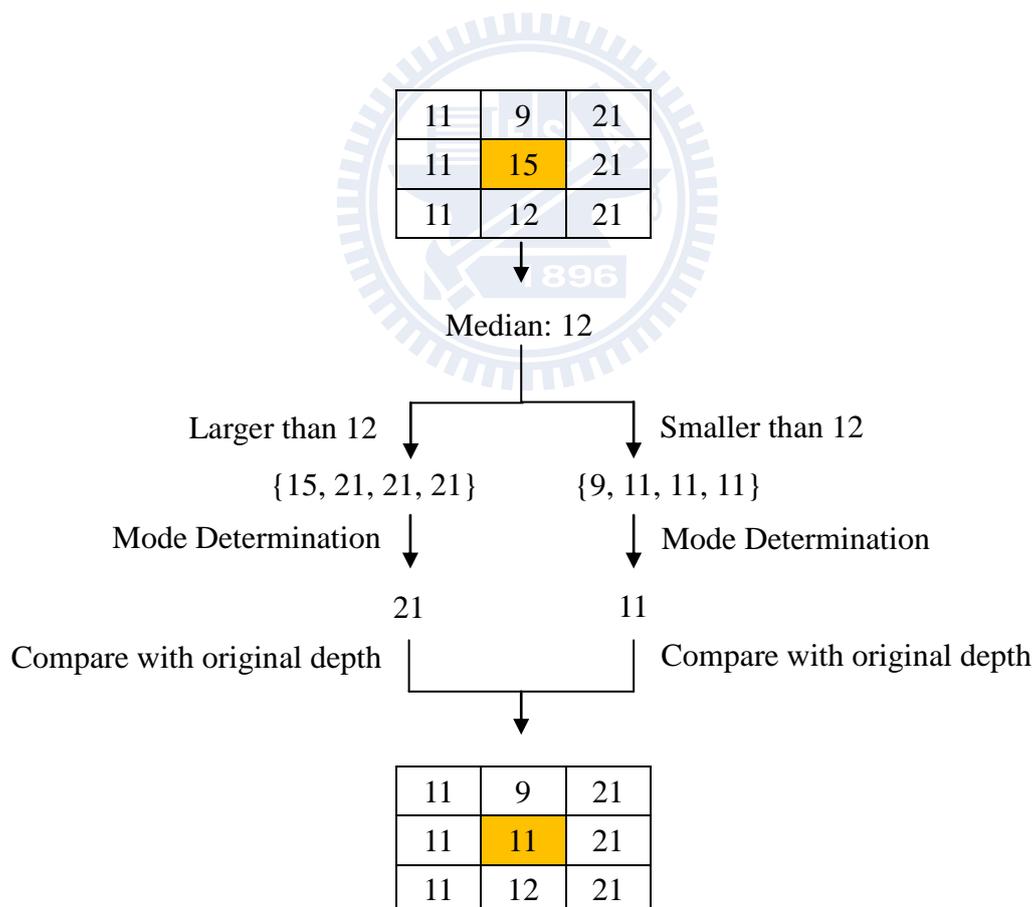


圖 8 修補深度值之範例

第三章、深度圖之分析

首先，本篇論文探討深度圖、自然影像與合成虛擬影像之關係，如圖 9。圖 9 不同曲線代表相同品質的自然影像，利用不同品質的深度圖，其合成虛擬影像之客觀品質為何。由圖 9 可以獲得三個結論：

1. 觀察不同的曲線，可以觀察出不同影片在相同資料量時，為了讓合成虛擬影像之客觀品質較好，分配給自然影像之資料量應該比深度圖之資料量多，如圖 9 不同影片之紅色虛線。本篇論文的研究在於深度圖壓縮，因此為了只讓深度圖影響合成虛擬影像之主、客觀品質，將採用高品質的自然影像以提供合成之用。
2. 觀察單一的曲線，除了某些特定影片(如 Door Flowers)之外，其餘影片的深度圖品質即使受到嚴重的破壞，影響合成虛擬影像之客觀品質還是有限。對於 Door Flowers 來說，在相同資料量時，為了讓合成虛擬影像之客觀品質較好，分配給深度圖之資料量應該比自然影像之資料量多，如圖 9 中 Door Flowers 影片之藍色虛線；如果一味的增加自然影像之品質，而不增加深度圖之品質，則會發現合成虛擬影像之品質幾乎不變，如圖 9 中 Door Flowers 影片之綠色虛線。本篇論文之研究會著重在深度圖失真後會嚴重影響合成虛擬影像之客觀品質，才得以比較與 MVC 之間的優劣。
3. 為了以最少的資料量達到最好的合成虛擬影像之品質，將所有影片之自然影像與深度圖最佳比例描繪出來，如圖 9 之紫色曲線，會發現在使用者可以接受的合成品質的情況下，大多數深度圖和自然影像的資料量比例相當低，因此可以得出深度圖需要在低資料量的限制下壓縮。

接著，本篇論文以 MVC 之壓縮標準分析深度圖使用 Intra 和階層 B 影像預測模式之間的差異，其方法為分別比較合成虛擬影像之客觀品質、壓縮效率與壓縮運算複雜度。

從合成虛擬影像之客觀品質來看，分別使用 Intra 和階層 B 影像預測模式壓縮深度圖，其合成虛擬影像之誤差值平均為 0~0.3 之間，如圖 10，再以主觀視覺觀察兩者，其合成虛擬影像也幾近相同，如圖 11。

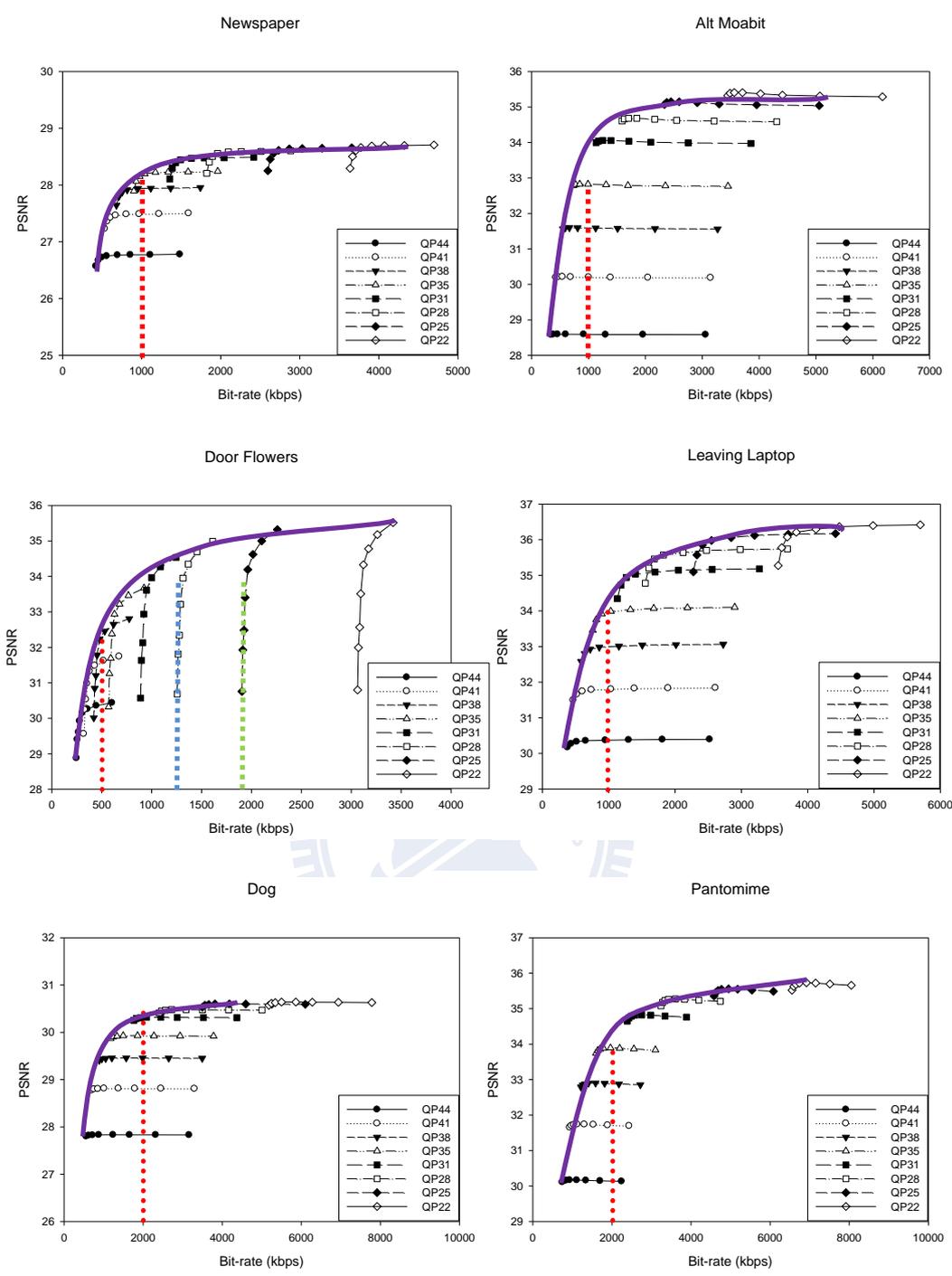


圖 9 不同影片之合成虛擬影像品質比較。

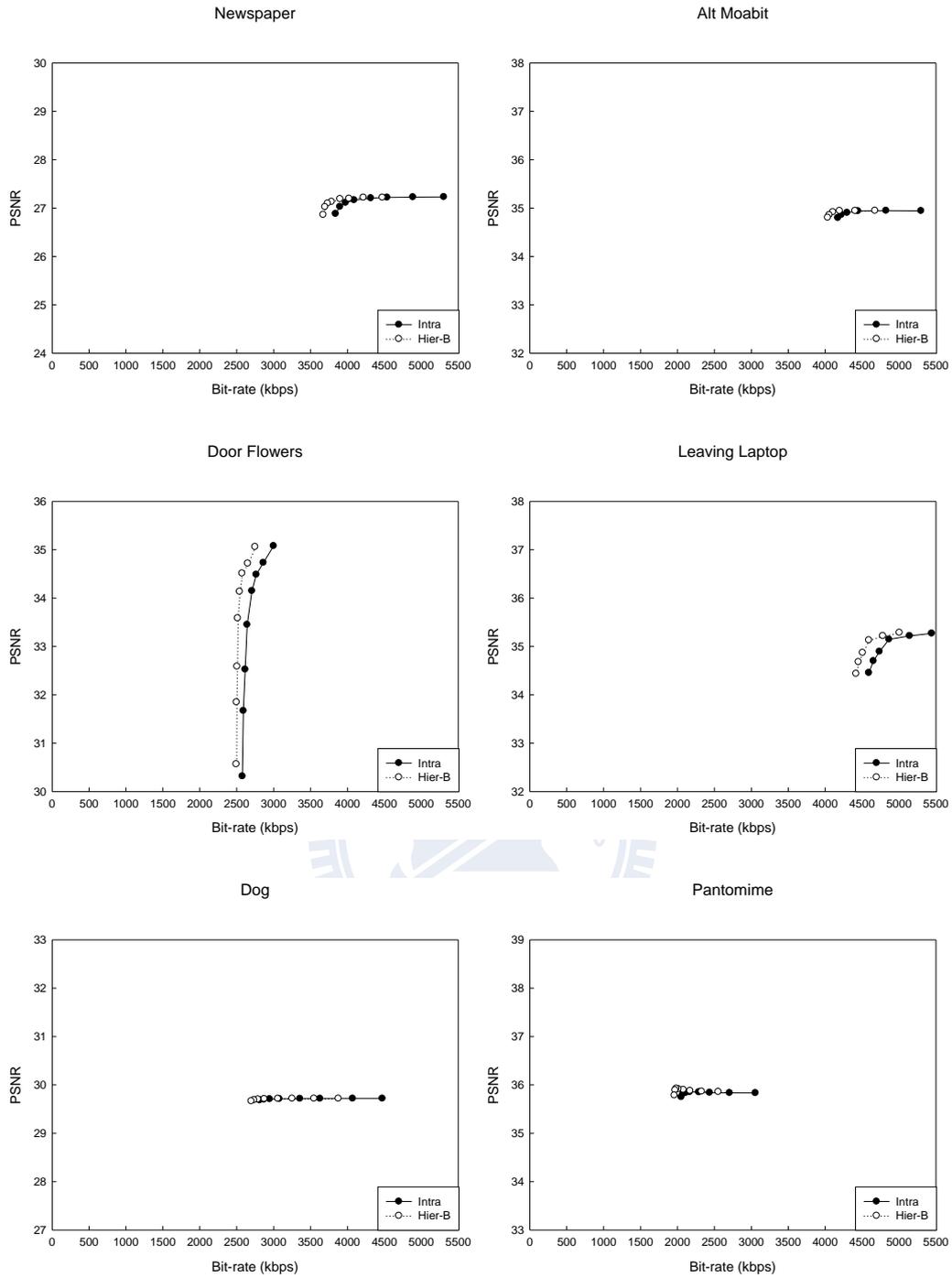


圖 10 比較 Intra 與階層 B 壓縮架構之合成虛擬影像品質



(a)

(b)

圖 11 比較 Intra 與 階層 B 影像壓縮架構之主觀視覺。(a) Intra；(b) 階層 B 影像壓縮架構

比較 Intra 與階層 B 影像壓縮架構之壓縮效率與運算複雜度方面，在圖 9 可以得知，絕大多數的影片不需要高品質的深度圖以提供合成虛擬影像，所以本篇論文從圖 9 找出各個影片之自然影像與深度圖資料量的最佳比例，如圖 12 (a) 紅色星星表示之。由圖 12 (a) 中可以發現，使用階層 B 影像架構的壓縮效率是 Intra 的數倍；不過將圖 12 (a) 紅色框框放大來看，如圖 12 (b)，會發現兩者與自然影像資料量的比例差不超過 6%，Intra 和階層 B 影像與自然影像資料量平均比例分別為 4% ~ 9% 與 1% ~ 3%，但階層 B 影像壓縮架構之運算複雜度遠大於 Intra，如表 1。因此，本篇論文之深度圖壓縮演算法以壓縮單張深度圖出發，目的是除了希望讓運算複雜度較 MVC 的 Intra 預測模式小之外，還可以提高其壓縮效率。

表 1 比較 Intra 與 階層 B 影像壓縮架構之平均運算複雜度 (單位:秒)

| | Lovebird1 | Newspaper | Door Flowers | Dog |
|------------------------|-----------|-----------|--------------|-----|
| Intra | 18 | 20 | 14 | 32 |
| Hierarchical B picture | 206 | 235 | 184 | 316 |

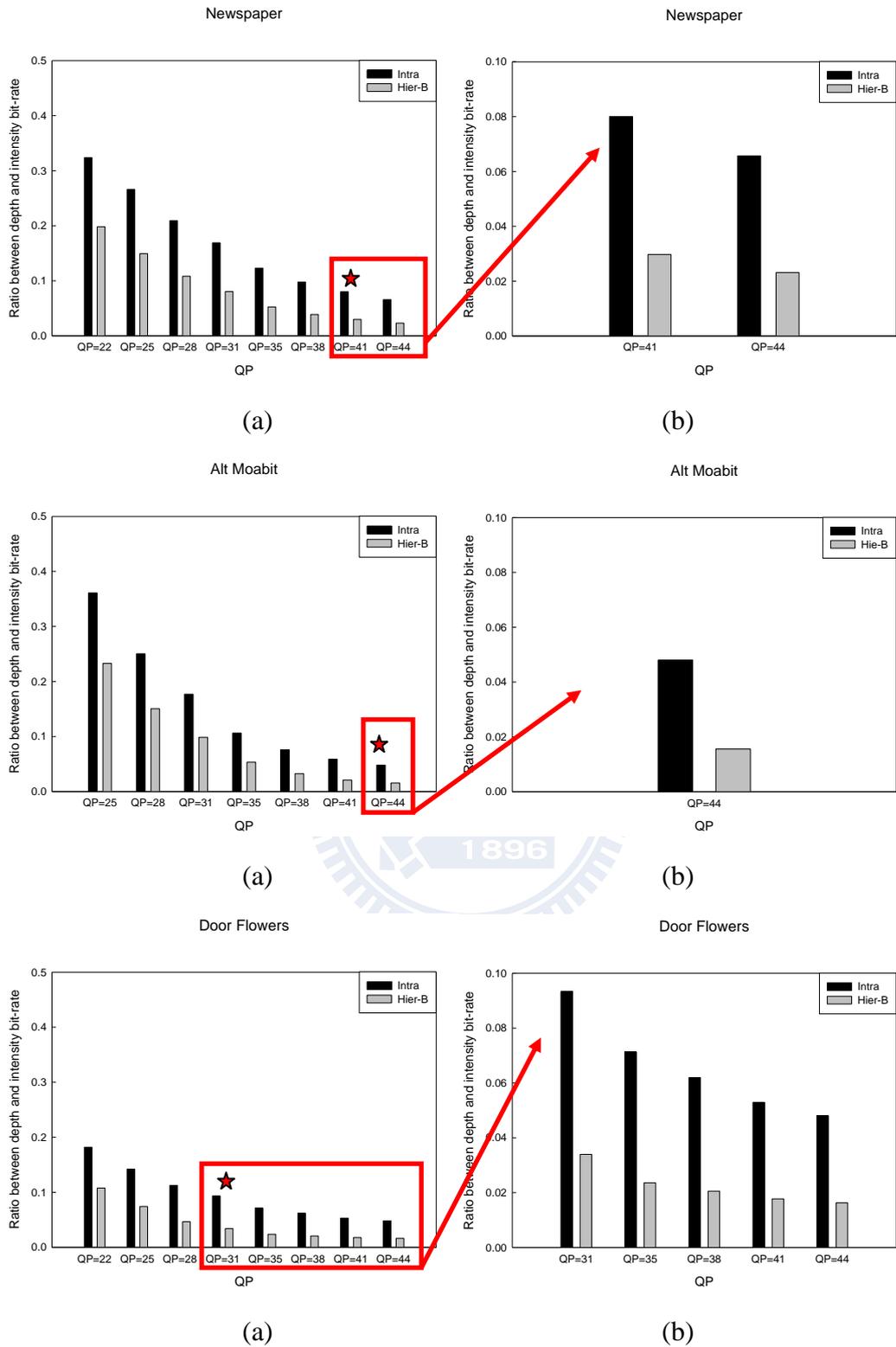


圖 12 比較 Intra 與階層 B 影像壓縮架構之深度圖與自然影像資料量比例關係。(a) 原圖；(b) 紅色框框之放大圖

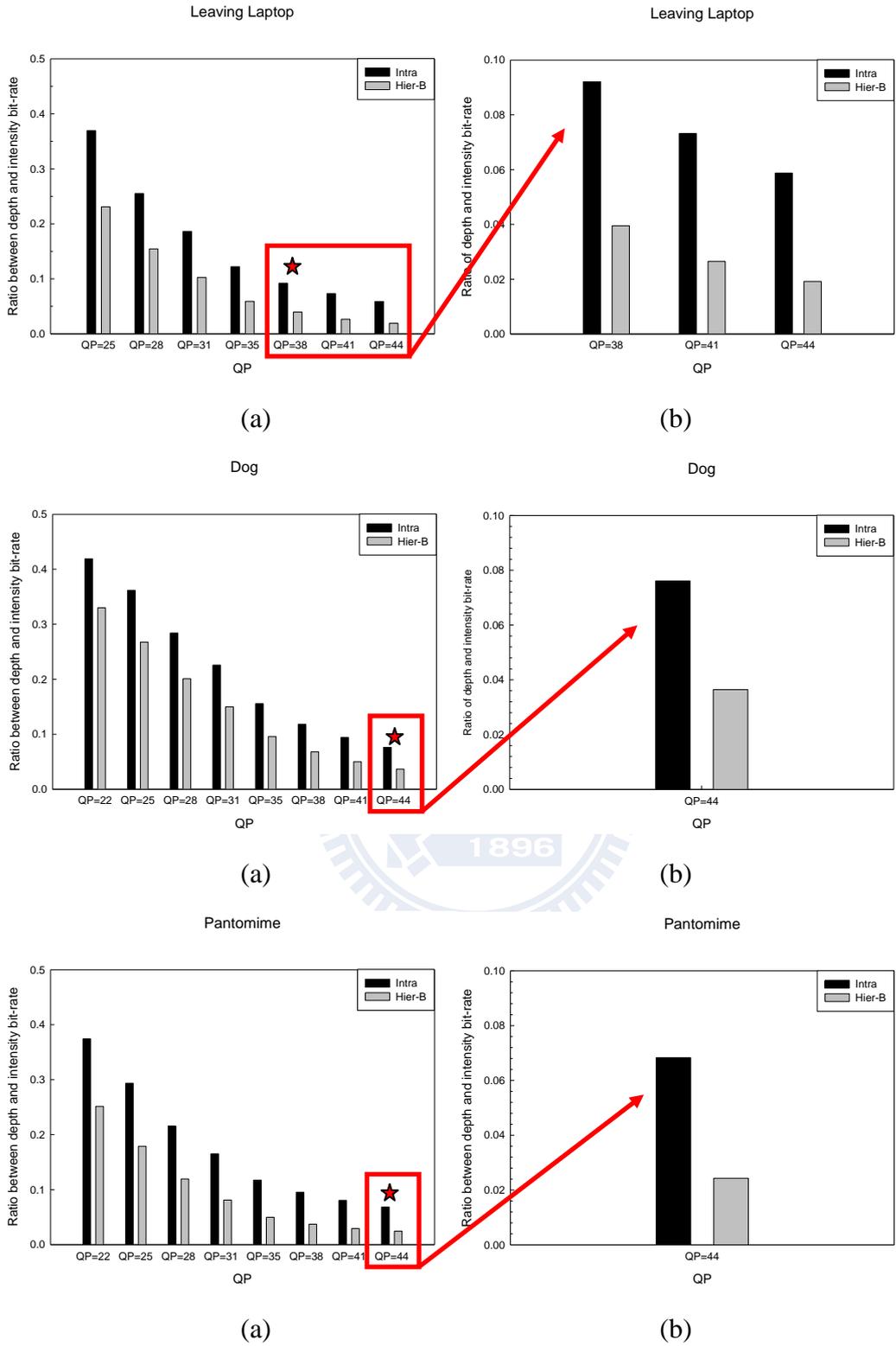


圖 12 (Continued)

為了盡可能提高深度圖之壓縮效率，需要了解並分析深度圖與自然影像之間不同的特性，因此本篇論文將深度圖及自然影像分割成若干區塊，每個區塊大小為 16x16，並統計各個區塊分別為由幾個不同的深度值加以組成。對於深度圖來說，絕大多數的區塊都由少數幾個深度值組成，少部份區塊的深度值分布非常廣，如圖 13 (a)；對於自然影像來說，大部份區塊的亮度值分布非常廣，如圖 13 (b)。另外，對於深度圖來說，接近一半以上的區塊由一個深度值所組成，此為深度圖所獨有之特性。

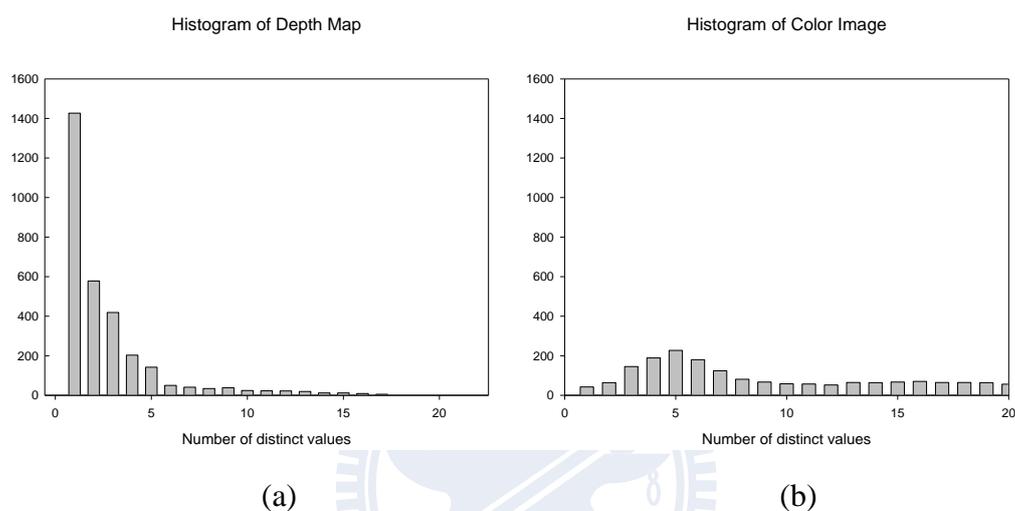


圖 13 統計深度圖與自然影像的區塊出現不同值的個數

第四章、研究方法

根據第三章之深度圖分析結果，本篇論文提出深度圖壓縮演算法，其流程圖為圖 3。

4.1 Mode Determination、Quantization、Residual Extraction

以區塊大小為 16x16 分割深度圖，每個區塊採取眾數演算法決定該區塊代表的眾數深度值，其目的為盡可能降低需要傳送至接收端的資料量；所有區塊之眾數決定後，即可將之壓縮至接收端。對於其它與眾數不同的像素深度值必須經由壓縮並傳送至接收端，判斷條件如下：

$$D_i = \begin{cases} M, & \text{if } |D_i - M| < TH_1; \\ D_i, & \text{else} \end{cases} \quad (3)$$

其中 D_i 為該區塊某一點的深度值、 M 為該區塊之眾數值、 TH_1 為使用者設定門檻之常數，若該像素深度值與眾數之絕對值差小於使用者所訂立的門檻，則該像素深度值可被眾數取代。 TH_1 為控制深度圖品質之用，其效果類似於 AVC 中的 QP 設定， TH_1 越大，雖然可以增加深度圖之壓縮效率，但是也降低深度圖之品質。

由此可知道本篇論文採用的深度值預測方式為眾數，和其他影像壓縮演算法的平均數的不同之處在於深度圖與自然影像特性不同，一般壓縮演算法採用的是最小均方根預測該像素值。若深度圖以平均數來預測深度值的話，雖然可以滿足最小均方根之特性，不過深度圖每個像素深度值都會有些許的誤差，造成合成虛擬影像會有不可預期之瑕疵。若像素深度值採用眾數預測的話，與眾數相差極大的部分會額外編碼，企圖維持深度圖之幾何特性。

剩餘未被眾數取代的像素深度值，本篇論文將之稱為 Residual，如圖 14。由圖 14 可以觀察出大多數 Residual 分佈在物體邊界處上；少部份的 Residual 分佈在平坦區域內，而分佈在平坦區的 Residual 可以被鄰近的深度值取代，因為並不會造成合成影像之誤差上有很大的影響。

以圖 18 為例，圖 18 (a) 為原始深度圖，其畫面大小為 8x8，分割區塊大小為 4x4，經由眾數演算法各自區塊的眾數深度值分別為 20、60、90 和 20，如圖 18 (b)；經過公式 3 計算後，圖 18 (c) 彩色部份為 Residual 之分布。

不同影片之深度圖，其公式 3 之門檻 TH_1 設定也會有所不同，本篇論文並沒有深入探討如何將 TH_1 加以模組化，而直接依據該深度圖之像素深度值分布決定該門檻 TH_1 之設定，以達到減少壓縮深度資訊之目的。

如果 TH_1 之門檻設定非常寬鬆時，原本該區塊有明顯的前後景之深度相對關係，會變成只有單一前景或後景之深度值，導致無法透過之後的深度圖修補演算法 (§4.3) 來解決深度資訊失真之問題，並嚴重影響合成虛擬影像之主、客觀品質，如圖 15 中的海報板與桌子之間，深度圖之深度值由原本前景的深度值變成後景之深度值，因此在合成虛擬影像時，該位置產生極大的合成誤差；若該區塊有 Residual 之資訊時，使用 (§4.3) 的深度圖修補演算法則可以有效的修補挑選 Residual 所造成深度資訊的失真，降低合成虛擬影像之誤差。

若將每個區塊之眾數分別填入對應的區塊，如圖 16，以主觀視覺的方式評斷，可以發現由眾數產生的深度圖在空間定義上具有高度相關性；不過眾數和 Residual 兩者之間並沒有相關性，因此分別壓縮眾數以及 Residual 時，其壓縮效率遠比將眾數及 Residual 合起來壓縮還來的高。

最後將 Residual 與眾數壓縮並傳送至接收端，Residual 壓縮方法的部份將由下一個章節闡述之。



圖 14 Residual 之分布圖



(a)

(b)

圖 15 公式 3 之 TH_1 設定對主觀視覺之影響。(a) TH_1 條件較嚴苛；(b) TH_1 條件較寬鬆



圖 16 眾數填入對應區塊產生的深度圖

4.2 Residual Coding

本節將介紹 Residual 資訊之壓縮，由圖 14 可以觀察出，Residual 在深度圖的分布上非常的散亂，如何有效壓縮 Residual 資訊，便是目前需要解決之課題。為了方便本節介紹 Residual 壓縮演算法，以圖 18 假設深度值分布之情況，Residual 壓縮之流程圖如圖 17。

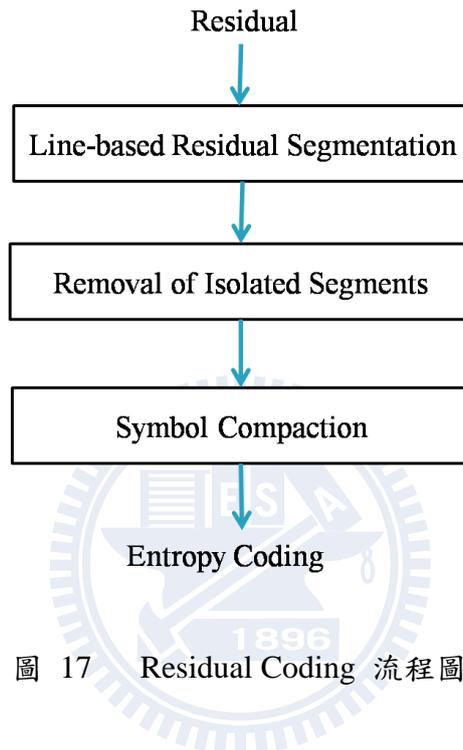


圖 17 Residual Coding 流程圖

| | | | | | | | |
|----|----|----|----|----|----|----|----|
| 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| 60 | 60 | 20 | 20 | 20 | 60 | 60 | 60 |
| 90 | 90 | 30 | 30 | 45 | 60 | 60 | 60 |
| 90 | 90 | 60 | 50 | 55 | 60 | 60 | 60 |
| 90 | 90 | 60 | 60 | 60 | 60 | 60 | 60 |
| 90 | 90 | 20 | 20 | 20 | 20 | 20 | 20 |
| 90 | 90 | 90 | 20 | 20 | 20 | 20 | 20 |
| 90 | 90 | 90 | 20 | 20 | 20 | 20 | 20 |

(a)

| | | | | | | | |
|----|----|----|----|----|----|----|----|
| 20 | 20 | 20 | 20 | 60 | 60 | 60 | 60 |
| 20 | 20 | 20 | 20 | 60 | 60 | 60 | 60 |
| 20 | 20 | 20 | 20 | 60 | 60 | 60 | 60 |
| 20 | 20 | 20 | 20 | 60 | 60 | 60 | 60 |
| 90 | 90 | 90 | 90 | 20 | 20 | 20 | 20 |
| 90 | 90 | 90 | 90 | 20 | 20 | 20 | 20 |
| 90 | 90 | 90 | 90 | 20 | 20 | 20 | 20 |
| 90 | 90 | 90 | 90 | 20 | 20 | 20 | 20 |

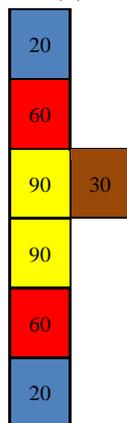
(b)

| | | | | | | | |
|----|----|----|----|----|----|----|----|
| 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| 60 | 60 | 20 | 20 | 20 | 60 | 60 | 60 |
| 90 | 90 | 30 | 30 | 45 | 60 | 60 | 60 |
| 90 | 90 | 60 | 50 | 55 | 60 | 60 | 60 |
| 90 | 90 | 60 | 60 | 60 | 60 | 60 | 60 |
| 90 | 90 | 20 | 20 | 20 | 20 | 20 | 20 |
| 90 | 90 | 90 | 20 | 20 | 20 | 20 | 20 |
| 90 | 90 | 90 | 20 | 20 | 20 | 20 | 20 |

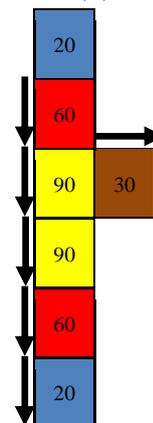
(c)

| | | | | | | | |
|----|----|----|----|----|----|----|----|
| 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| 60 | 60 | 20 | 20 | 20 | 60 | 60 | 60 |
| 90 | 90 | 30 | 30 | 45 | 60 | 60 | 60 |
| 90 | 90 | 60 | 50 | 55 | 60 | 60 | 60 |
| 90 | 90 | 60 | 60 | 60 | 60 | 60 | 60 |
| 90 | 90 | 20 | 20 | 20 | 20 | 20 | 20 |
| 90 | 90 | 90 | 20 | 20 | 20 | 20 | 20 |
| 90 | 90 | 90 | 20 | 20 | 20 | 20 | 20 |

(d)



(e)



(f)

圖 18 深度圖產生 Residual 之示意圖。(a) 原始深度圖；(b) 眾數深度圖；(c) Residual 深度圖；(d) DC (非黑色) 與 Non-DC (黑色) 片段深度圖；(e) 傳輸至接收端之資訊；(f) Residual 片段之預測方向

4.2.1. Line-based Residual Segmentation

本篇論文採用 1D 的方式掃描 Residual，目的是為了將每個不同深度值的 Residual 片段分開，以不影響合成影像的主、客觀品質為第一優先。若 Residual 裡像素位置為相鄰，且兩者之深度值相同，則判定兩者像素為相同片段，並繼續判斷下一個像素是否合乎上述之條件；若違反上述兩者條件其中之一的话，則記錄原本蒐集的片段之起始位址、長度和深度值，之後重新判斷下一個像素是否合乎上述之條件，直到判斷至最後之 Residual 深度值。記錄區段資訊時，需再判斷其長度是否超過 TH_2 ，此條件是為避免 Residual 之長度過短，增加需要傳送給接收端之資訊，一般設定為 4。

4.2.2. Removal of Isolated Segments

經由 (§4.2.1) 計算後，Residual 最主要分成兩種：一種是以單一深度值表示的片段 (以 DC 片段稱之)，如圖 18 (d) 的非黑色像素點；另一種是無法以單一深度值表示或長度未超過 4 的片段 (以 Non-DC 片段稱之)，如圖 18 (d) 的黑色像素點。為了提高壓縮效率，本篇論文將該 Non-DC 片段分成兩段，之後與兩邊的片段結合，最主要分成三種狀況，如圖 19。若 Non-DC 片段之兩邊都為眾數之片段，則 Non-DC 以眾數取代之，不加以編碼壓縮，如圖 19 (a)；若 Non-DC 片段之一邊為眾數片段，一邊為 Residual 片段，則 Non-DC 片段一分為二，分別被眾數和 Residual 取代之，而 DC 片段資訊需要重新計算，如圖 19 (b)；若 Non-DC 片段之兩邊都為 Residual 之片段，則 Non-DC 片段一分為二，分別被兩邊之 Residual 值取代，而原本兩邊的 DC 片段資訊需要重新計算，如圖 19 (c)。將原本 Non-DC 片段以眾數或鄰近的 Residual 取代會造成合成虛擬影像之瑕疵，如圖 20，紅線為原本經由眾數量化而產生失真，紫色區域為 Non-DC 片段還未失真，經由本篇論文將 Non-DC 片段一分為二，形成綠線之訊號而產生失真，所以原本該地方之深度值會遞增或遞減，現在卻被本演算法處理為 DC 片段；不過本篇論文 (§4.3) 會提供修補演算法將此誤差修補之。

4.2.3. Symbol Compaction

Residual 片段決定後，有三個資訊需要傳輸至接收端，分別是該片段的起始位置、長度和深度值，最後將此資料結構向左集中，如圖 18 (e)，每一格代表 Residual 片段的三個資訊。由圖 14 觀察 Residual 片段，雖然都集中於物體之邊界上，不過鄰近 Residual 之間還是有一定程度的相關性，因此，希望透過預測的方式來增加壓縮之效率，預測之方向如圖 18 (f)，本篇論文將預測模式分為三種，如表 2。

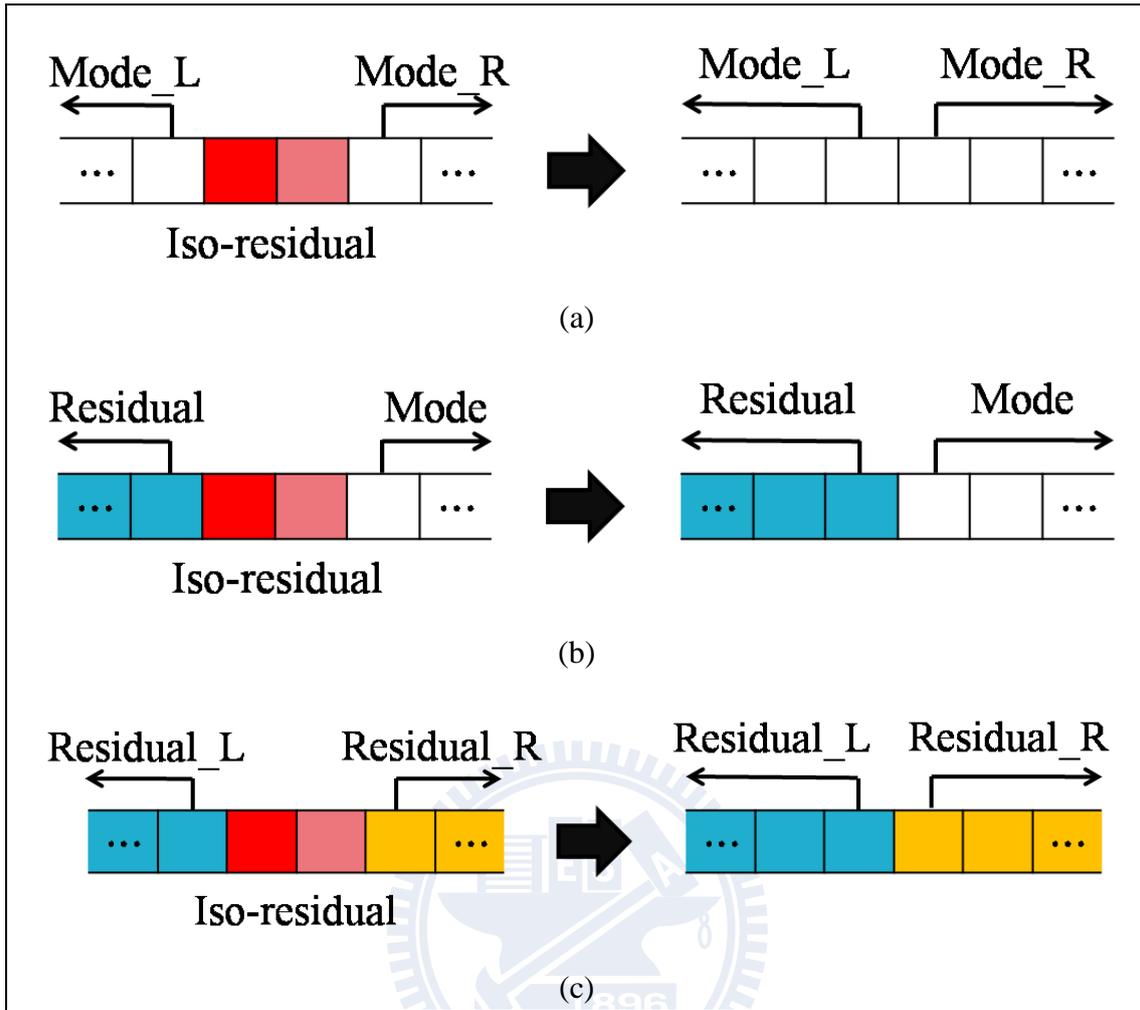


圖 19 Non-DC 片段分割圖

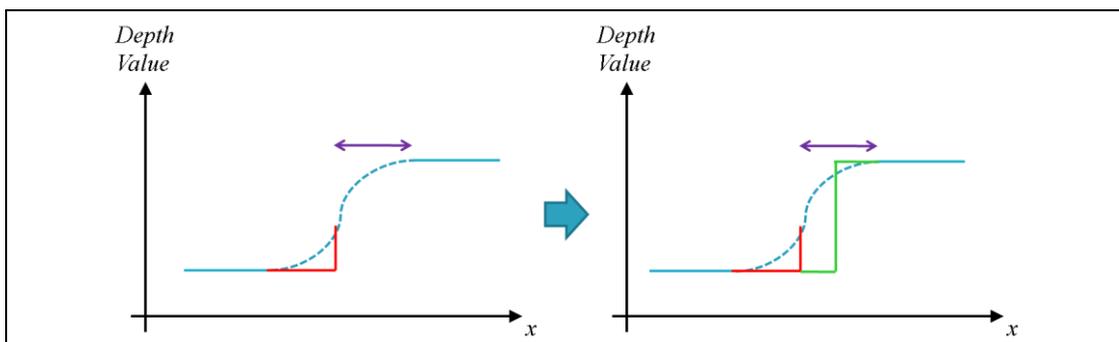


圖 20 Iso-residual segments 被移除而產生錯誤

表 2 DC 片段之預測模式

| 模式 | 描述 |
|----|--------------------------|
| 1 | 相鄰 R.I. 片段之起始位置、長度和深度值相同 |
| 2 | 相鄰 R.I. 片段之深度值相同，其餘不相同 |
| 3 | Otherwise |

若增加預測模式，則各自機率需達到多少才可以比完全將三個參數傳送至接收端時更好。假設 Residual 片段總共為 Z 個；R.I. 片段的模式花費 2 個位元；深度值、起始位址和長度各花用 8、16、16 位元，且模式 1、2 和 3 各占的機率分別是 P_1 、 P_2 、 P_3 ，則式子可以如下：

$$2 \times (P_1 + P_2 + P_3) \times Z + 32 \times P_2 \times Z + 40 \times P_3 \times Z < 40 \times Z \quad (4)$$

公式 4 經由簡單的整理後，則可變為公式 5：

$$16 \times P_2 + 20 \times P_3 < 19 \quad (5)$$

不同影片之在高、低資料量之限制下所占的比例如附錄，將所有之影片的機率代進去，都符合上述之判斷式。因此，採用 Residual 片段預測模式來估測對於壓縮之效率是有好處。

本篇論文不採用其它 Residual 片段預測模式之原因有兩個：(1) 附錄可以觀察出，在高資料量的情況下，Residual 片段的預測模式 3 佔的比例較重，原因是為了提高 Residual 片段之準確度，絕大多數相鄰的 Residual 片段的起始位置、長度和深度值不盡相同；(2) 在第三章深度圖之分析有提到，深度圖一般會要求在低資料量的限制下進行壓縮，盡可能不和自然影像競爭網路頻寬，且在附錄可以觀察出，大部份的 Residual 片段預測模式都集中在模式 1 和模式 2。根據上述的原因，若其他 Residual 片段之預測模式加入演算法，由於其餘的預測模式佔整體比例過小，反而額外增加選擇預測模式之負載，造成整體的壓縮效率不增反減，因此不增加額外 Residual 片段之預測模式。

由上面的分析可以得知，若提供不同的 Residual 片段預測方式來估測該 Residual 片段之資訊，可以減少傳送至接收端所需要的資訊（包括模式、長度、起始位置和深度值）。由附錄可以觀察出，Residual 片段的長度、起始位置和深度值與相鄰 Residual 片段完全吻合占整體的 65~93%，這些 Residual 片段只要傳送一個預測模式之資訊即可，其餘的資訊可以不用傳送至接收端，因為可藉由相鄰的 Residual 片段作估測；若只有深度值和鄰近 Residual 片段一樣，則必須將片段之模式、起始位置和長度傳至接收端，此模式佔全部的 0.5~9.3%；其餘

需要傳送所有的資訊，佔 6.5~25.7%。

4.2.4. Entropy Coding

深度圖經過眾數演算法挑選眾數 (§4.1) 後，剩餘的 Residual 需採無失真壓縮，Residual 對於深度圖相當重要，如果沒有經過無失真壓縮的話，不僅容易造成合成虛擬影像的誤差，且對於 (§4.3) 之修補深度圖產生重大的影響，原因為本身 Residual 資訊已不具可靠性。

由於 Residual 片段之模式、起始位置、長度及深度值在實驗之統計上並沒有任何的相關性，因此本演算法認定這四種資訊之相關性為相互獨立而採取分開壓縮之方式傳至接收端，其壓縮效率會比將這四種資訊組合起來一起壓縮還來的高。

4.3 Depth Map Refinement

深度圖經由眾數挑選 Residual，會產生失真，其原因在於公式 3 所採用 TH_1 門檻之設定會使得與該區塊眾數不同的深度值也被判定為眾數。此失真會在物體與物體之間的邊界處出現，如圖 21 (a)，非灰色區域為深度圖產生失真之區域，若不修補經由壓縮重建後之深度圖，會影響合成虛擬影像之主、客觀品質。因此，在接收端處提供簡易的方法來修補失真的深度值。

(§4.2) 區分 Residual 片段是以不影響合成影像之主、客觀品質所設計，所以本篇論文假設重建 Residual 資訊的可靠度很高，透過 Residual 資訊以修補深度圖產生失真的區域。以圖 22 為例，假設圖 22 (a) 之藍線為原始深度圖之訊號，經過 (§4.1) 之演算法挑選該區塊之眾數並量化原始訊號以後，紅線為量化後之深度圖訊號。

由於接收端無法得知哪些像素深度值已產生失真，所以修補的對象為眾數之區域。為了盡可能修補邊界失真的像素深度值，避免對原來是正確的像素深度值做錯誤的修正，本篇論文採用的方式為：與 Residual 相鄰的像素深度值為首要修補的對象，並以大小為 7x7 的過濾器進行修補深度值，已修補的深度值納入 Residual 之集合；接著，再重覆搜尋與 Residual 相鄰的眾數進行修補，直到該區塊的深度值不會改變或整個像素深度值已被修補。而過濾器之設計方法為將所包含的深度值 (包括 Residual 資訊及眾數之深度值)，經由平均權重之關係，產生應該修補之深度值。以圖 23 為例子，區塊的起始狀態 (Initial) 之黃色區域為 Residual，白底為眾數。在 Iteration 1 時會找出與 Residual 相鄰之眾數，如圖 23 之灰色區域，這些像素深度值會以過濾器修補；之後已修補的像素深度值

設定為可靠之像素深度值，再找出與之相鄰的眾數深度值並修補，直到結束。如圖 24 為深度圖經由修補後，其合成虛擬影像之主觀視覺。

圖 22 (b) 之不同虛線代表使用的過濾器的大小不同，若過濾器的大小為 0 時，則修補後的訊號和原本量化後訊號一樣；若過濾器設定太小的話，則應修補訊號過少導致無法修補至與原來量化前差不多，如圖 22 (b) 之 (1)；若過濾器設定太大的話，大部份的訊號都會被修補，導致與原來量化前差很多，圖 22 (b) 之 (2)。因此過濾器的大小會決定修補演算法之優劣。

此修補方式的好處是由於權重設計的關係，當修補的像素離原來 Residual 越遠時，經由過濾器所計算之深度值會與原來像素之深度值越接近而達到收斂。

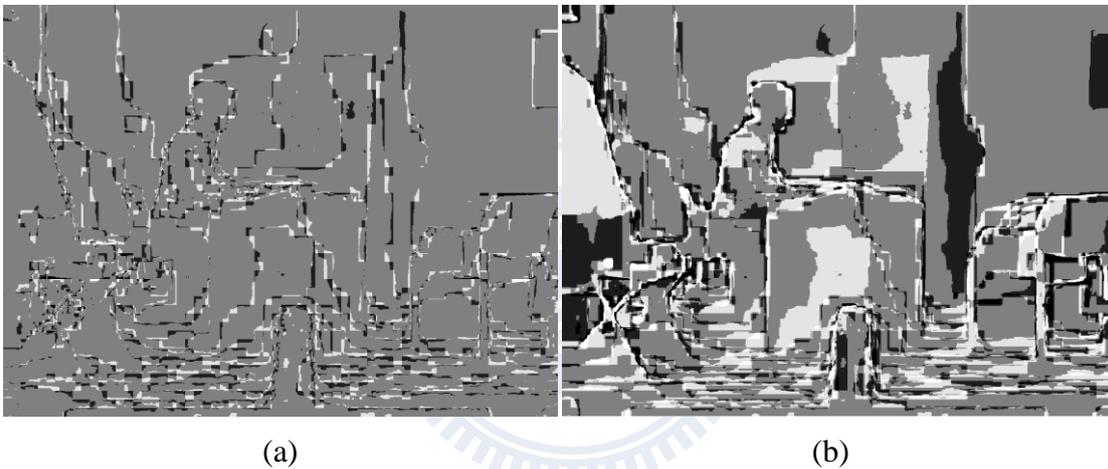


圖 21 重建後深度圖失真之區域。(a) 非灰色區域為本篇論文產生失真之區域；(b) 非灰色區域為 MVC 產生失真之區域

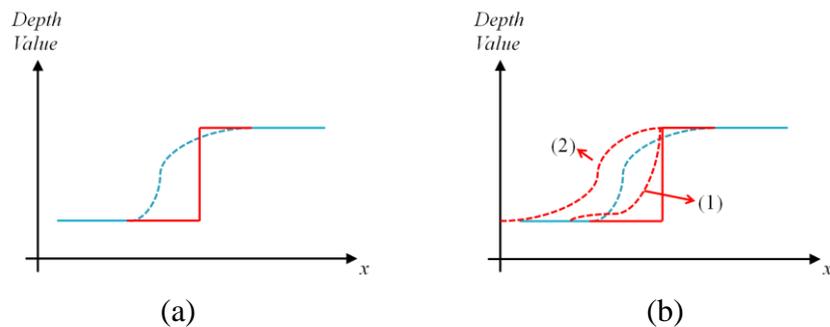


圖 22 深度圖之區塊經由壓縮重建後產生失真。(a) 藍色虛線為原始深度圖之訊號，紅線重建後之訊號為；(b) 比較不同大小之修補過濾器，(1) 為較小之過濾器，(2) 為較大之過濾器

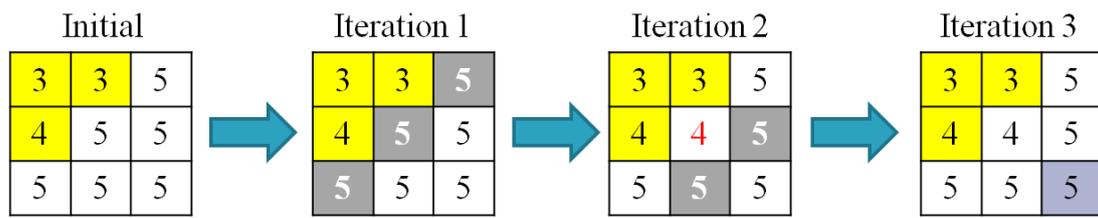


圖 23 修補深度值之範例



(a)

(b)

圖 24 修補深度圖之前後，其合成虛擬影像之主觀視覺之比較。(a) 修補前；(b) 修補後

第五章、實驗結果與討論

本篇論文採用 MPEG FTV 提供的影片，由 MPEG 組織提供的 DERS 2.0 產生深度圖；並由 VSRS 3.5 產生合成虛擬影像。本篇論文比較的對象為 JMVC 6.1，此為現今常用於深度圖壓縮之標準。本篇論文的演算法採用單張深度圖之壓縮，並沒有考慮 Temporal 和 Inter-view 在統計上有相關性，為了能和本篇論文可以公平比較，在 JMVC 的設定上，不使用 Temporal 和 Inter-view 的預測方法，只能由該張畫面現有的資訊來做估測；而本篇論文演算法產生的參數，採用現有的無失真壓縮標準 7 zip 來壓縮。最後，比較兩者之優劣標準採用影片 30 張之合成虛擬影像並計算其平均 PSNR；其中，表 3 和表 4 分別為 VSRS 及 JMVC 的設定。

表 3 合成虛擬影像之參數設定

| | S | NL-NR | Precision |
|-----------------|----|-------|-----------|
| Lovebird1 | 6 | 5-8 | 2 |
| Lovebird2 | 6 | 5-8 | 2 |
| Newspaper | 4 | 3-6 | 2 |
| Alt Moabit | 9 | 10-7 | 2 |
| Book Arrival | 9 | 10-7 | 2 |
| Door Flowers | 9 | 10-7 | 2 |
| Leaving Laptop | 9 | 10-7 | 2 |
| Champagne Tower | 39 | 38-42 | 2 |
| Dog | 39 | 38-42 | 2 |
| Pantomime | 39 | 38-42 | 2 |

表 4 JMVC 之參數設定

| | |
|-----------------------|----------------------------|
| Intra Period | 1 |
| CABAC | On |
| 8x8 Transform | On |
| Basis QP | 22, 25, 31, 35, 38, 41, 44 |
| Inter-view Prediction | Off |
| Search Mode | 4 (Fast Search) |
| Motion Search Range | ± 32 |

一般在壓縮深度圖和自然影像時，由於網路頻寬的限制下，自然影像通常會被分配到較多的網路頻寬，其原因是如果自然影像的品質很差時，即使分配給深度圖再多網路頻寬也不會讓合成虛擬影像的主、客觀品質變高。因此，深度圖通常會在網路頻寬較低的情況下，比較不同壓縮深度圖演算法之優劣。

比較重建深度圖之失真，本篇論文失真為圖 21 (a) 之非灰色區域，MVC 失真為圖 21 (b) 之非灰色區域。由於本篇論文會保留區塊之眾數及 Residual 區域，因此失真會在物體之邊界上；然而 MVC 的作法為將訊號之 Residual 值經由離散餘弦轉換並量化，其失真不僅會在物體邊界上，在平坦區域也會產生失真，雖然平坦區域之失真影響合成虛擬影像之品質並不大，不過其幾何特性較本篇論文還差。另外，本篇論文演算法也不會有 Blocking Artifact，其原因為 MVC 採用區塊的 Residual 離散餘弦轉換及量化，所以在低頻訊號上容易產生失真；但本篇論文的 Residual 除了在 (§4.2.2) 部份被量化之外，其他並不會被量化而產生失真，所以本篇論文不會有 Blocking Artifact。

圖 25 為表示本篇論文與 MVC 之 PSNR 比較，由圖 25 中可以觀察到絕大多數的影片在網路頻寬較低的限制下，本篇論文之演算法的效能較 MVC 來的好；少部份的影片較 MVC 差，以 Champagne Tower 之影片來說，如圖 26，雖然合成虛擬影像之主、客觀視覺較 MVC 差；不過比較原始之深度圖、MVC 產生的深度圖及本篇論文產生的深度圖，本篇論文之深度圖與原始深度圖比較接近，但原始深度圖和本篇論文之深度圖的合成影像品質較 MVC 差，初步分析之結果可能是原始深度圖之深度資訊不正確；經由 MVC 壓縮演算法重建深度圖後，反而修正為正確的深度資訊，才導致有此現象發生。在主觀視覺上，如圖 27、28、29、30、31，也會發現本篇論文合成影像之瑕疵較 MVC 少。因此，在同樣的客觀品質底下，本篇論文所提出的演算法需要的網路頻寬較 MVC 少 4~13%。

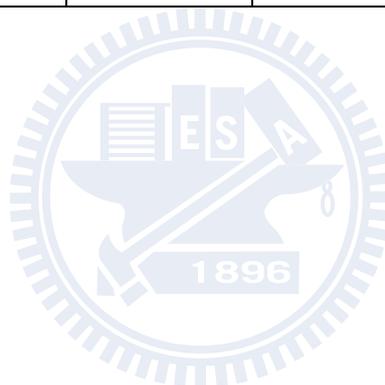
最後，本篇論文與 MVC 之運算複雜度比較，如表 5、表 6，本篇論文運算複雜度較 MVC Intra 之預測方式少 5%~15%；運算個數也較 MVC 少。

表 5 本篇論文演算法與 MVC 之平均運算複雜度比較 (單位：秒)

| | Proposed | MVC Intra | MVC Inter |
|--------------|----------|--------------|--------------|
| Door Flowers | 13 | 14 (+7.7%) | 184 (+1315%) |
| Book Arrival | 13 | 15 (+15.38%) | 241 (+1853%) |
| Dog | 29 | 32 (+10.3%) | 316 (+990%) |
| Pantomime | 19 | 20 (+5.2%) | 260 (+1368%) |

表 6 本篇論文演算法與 MVC 之平均運算個數比較

| MVC | | | Proposed | | |
|--------------|------------|-----------|--------------|------------|-----|
| | ADD | MUL | | ADD | MUL |
| Prediction | $4n^2$ | 0 | Mode | n^2+256 | 0 |
| DCT | n^3 | n^3 | Quantization | n^2 | 0 |
| Quantization | n^2 | n^2 | Residual | n^2 | 0 |
| Total | n^3+5n^2 | n^3+n^2 | Total | $3n^2+256$ | 0 |



第六章、 結論與未來展望

本篇論文針對深度圖與自然影像做分析，並根據深度圖之特性加以設計低複雜度壓縮演算法。藉由深度圖在空間定義上有高度相關之特性，在前處理端使用眾數演算法，將需要的深度資訊壓縮並傳輸至接收端；接收端重建深度圖後，再經由簡單的線性權重之過濾器以修補壓縮重建後深度圖，提高合成影像之主、客觀品質。因此，本篇論文在不影響合成虛擬影像的主、客觀品質下，不僅對網路頻寬的需求較 MVC 少 4% ~ 13%，運算複雜度也較 MVC 少 5% ~ 15%。

本篇論文在 Entropy Coding 的部分採用現今無失真資料壓縮演算法，此壓縮方法是針對一般資料之特性所設計，並不是針對本篇論文產生之參數所設計，所以未來應針對本篇論文之參數特性設計一套 Entropy Coding。

由於本篇論文之演算法的一些門檻之參數(如公式 3 之 T_{h1})並沒有依據不同的影片的資料分布與特性，而加以模組化，目前是以訓練參數之方式，找出適合不同影片的門檻之參數。在重建深度圖的後處理上，僅利用前處理產生深度值失真的原因，做簡單的修補，若可以偵測哪些像素深度值需要被修補，便可以增加合成虛擬影像之主、客觀品質。

本篇論文所提供的演算法，僅利用該張深度圖資訊設計演算法，未來應可以再擴充 Temporal 及 Inter-view 在統計上有相關之特性，設計深度圖壓縮演算法。

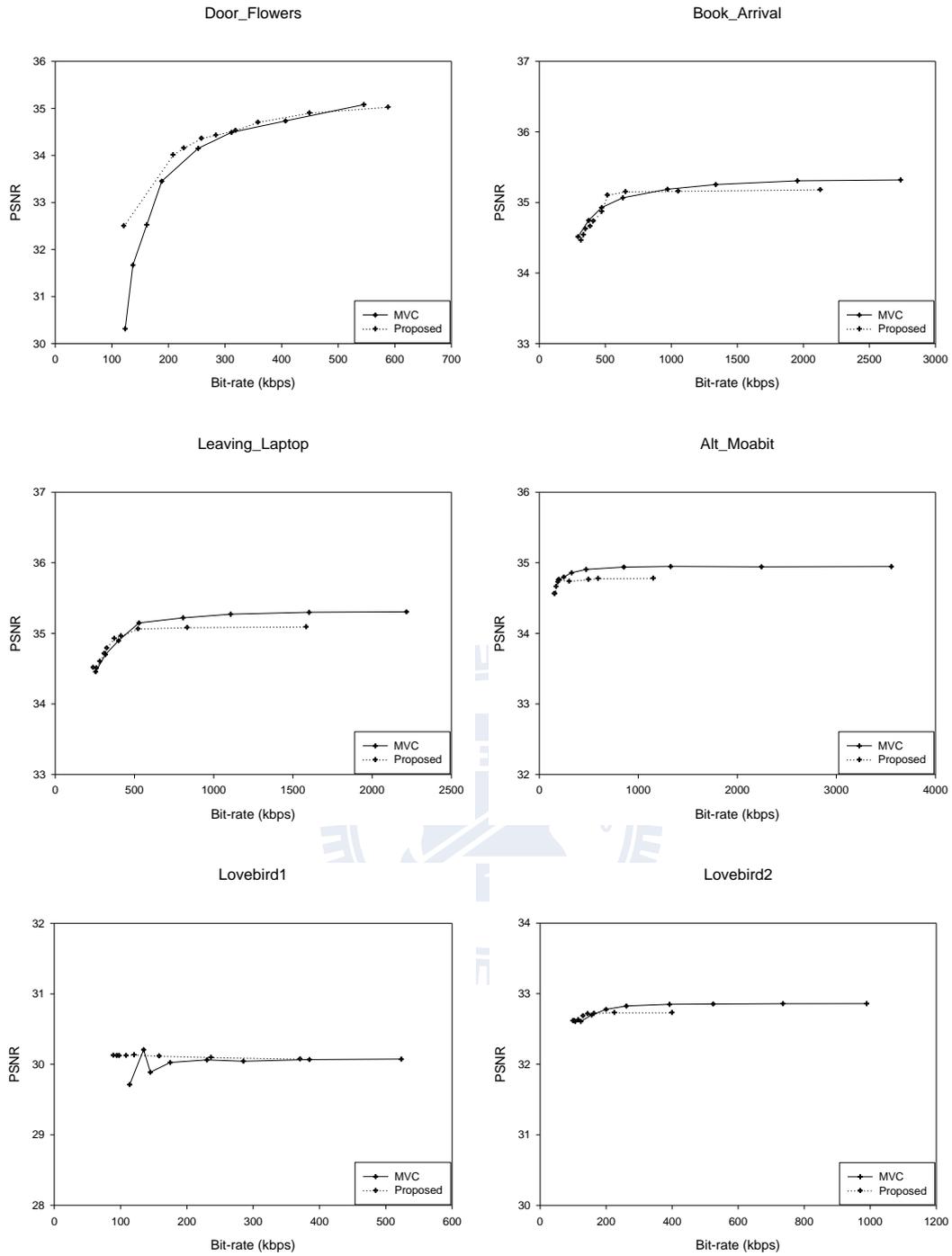


圖 25 MVC 與本篇論文演算法合成虛擬影像之 R-D 曲線

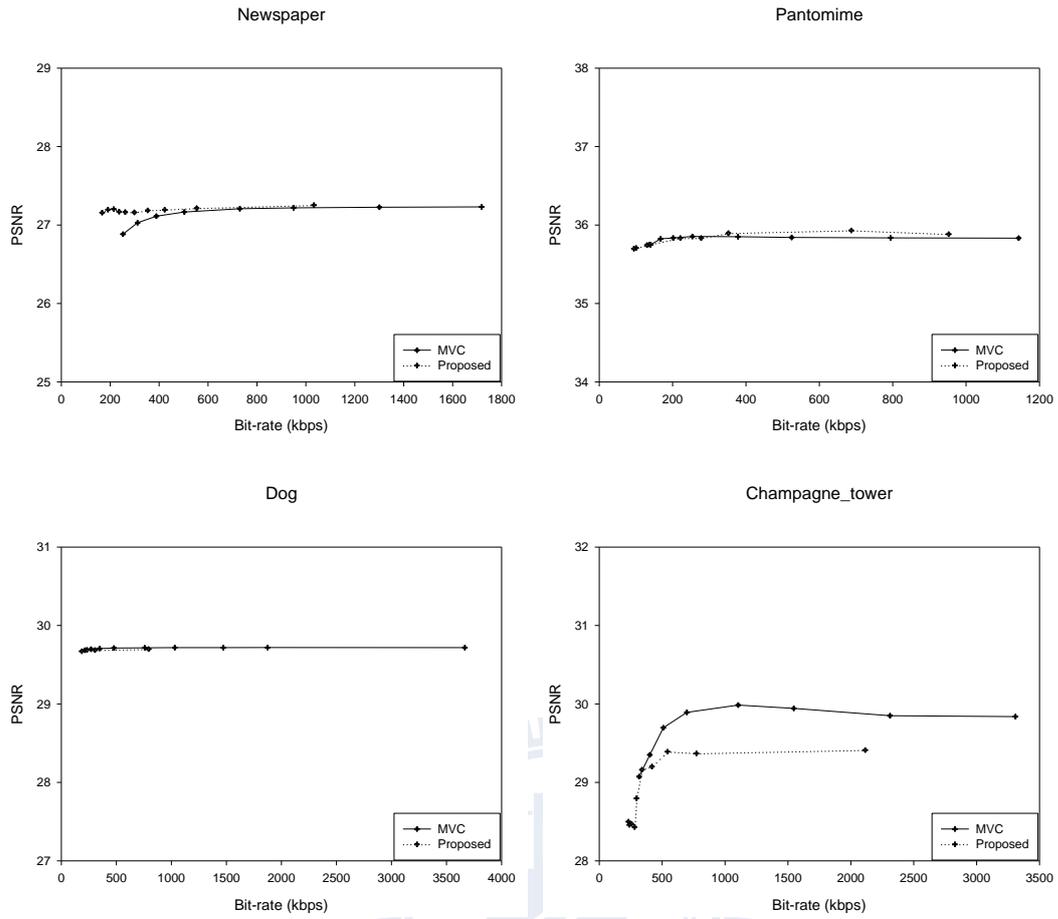
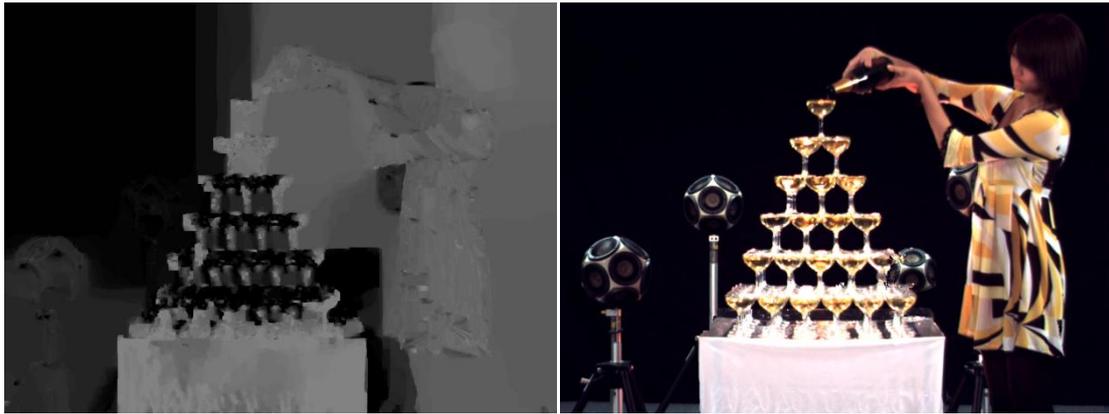
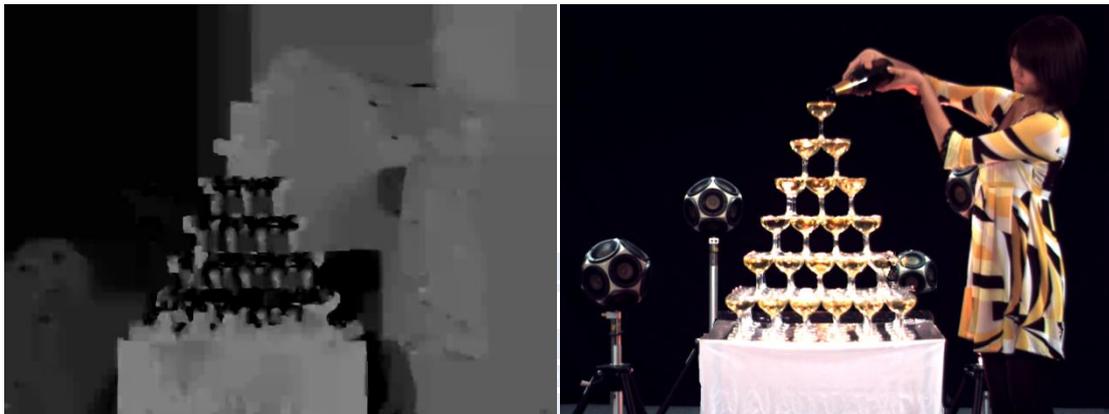


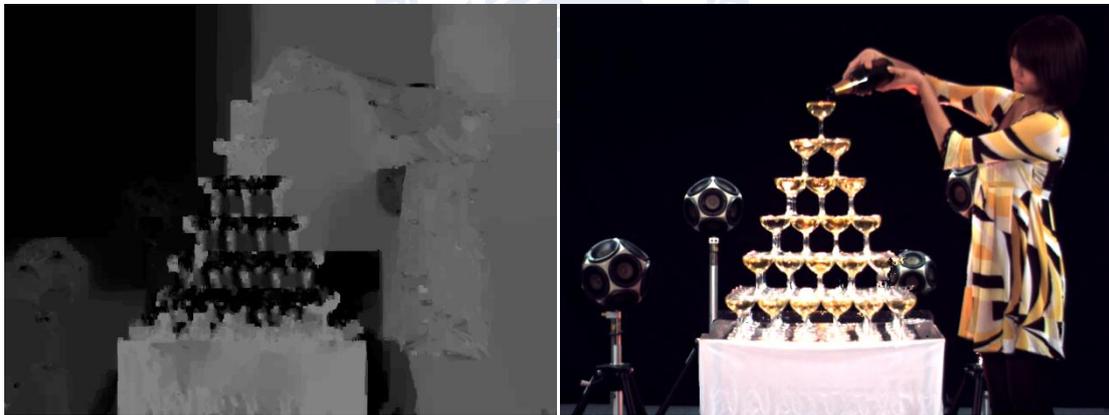
圖 25 (Continued)



(a)



(b)



(c)

圖 26 Champagne Tower，深度圖與其合成虛擬影像之主觀視覺比較。(a) 原始深度圖；(b) MVC Intra Prediction；(c) Proposed



(a)

(b)

(c)

(d)

圖 27 Book Arrival，合成虛擬影像之主觀視覺比較。(a) Original；(b) MVC All Intra Prediction；(c) MVC Inter Prediction；(d) Proposed



(a)

(b)

(c)

(d)

圖 28 Door Flowers, 合成虛擬影像之主觀視覺比較。(a) Original ; (b) MVC All Intra Prediction ; (c) MVC Inter Prediction ; (d) Proposed



圖 29 Leaving Laptop。 (a) Original； (b) MVC All Intra Prediction； (c) MVC Inter Prediction； (d) Proposed



圖 30 Lovebird1，合成虛擬影像之主觀視覺比較。(a) Original； (b) MVC All Intra Prediction； (c) MVC Inter Prediction； (d) Proposed

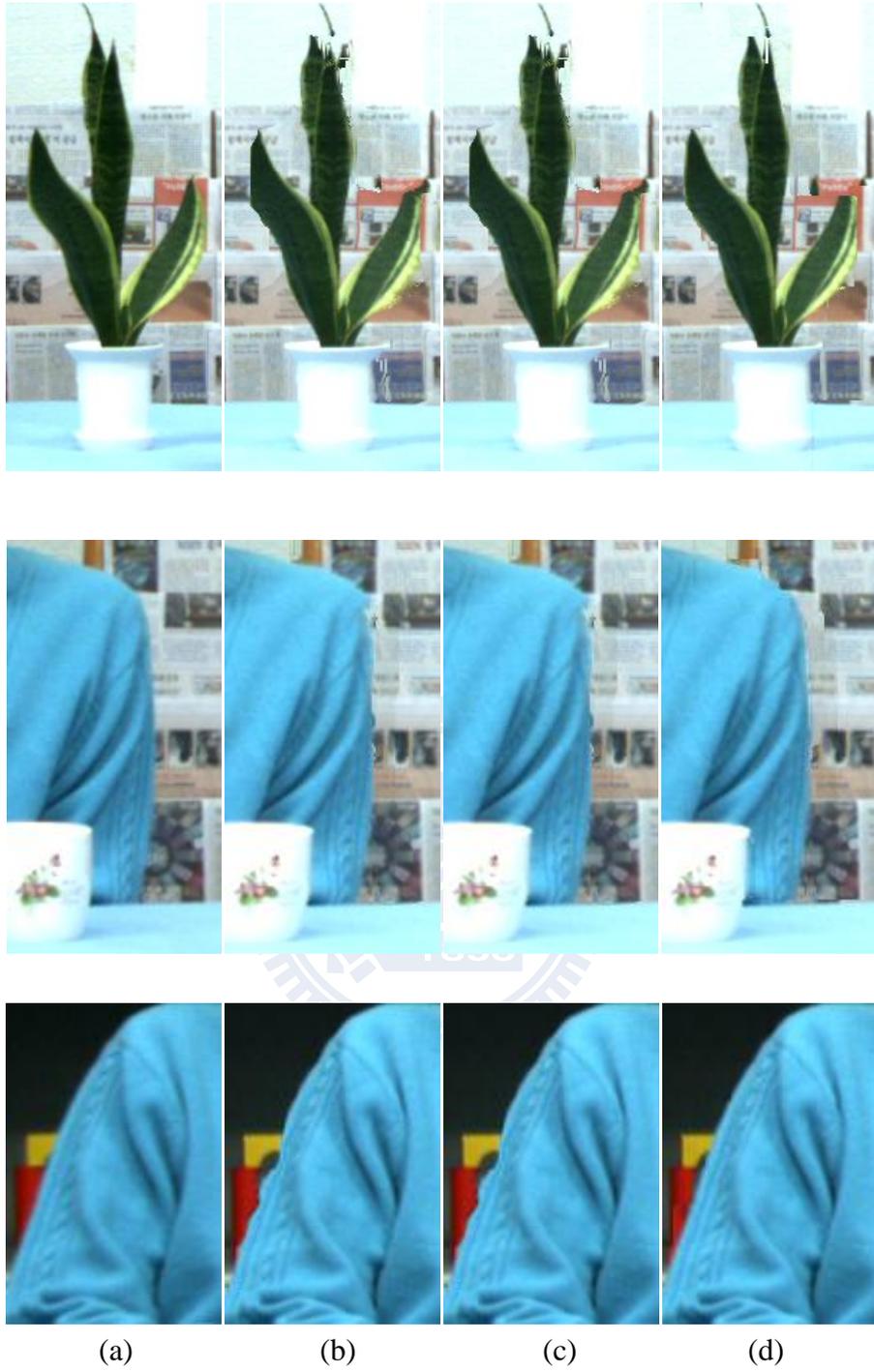


圖 31 Newspaper，合成虛擬影像之主觀視覺比較。(a) Original；(b) MVC All Intra Prediction；(c) MVC Inter Prediction；(d) Proposed

附錄

表 6 深度圖限制在高資料量的情況下，R.I. 預測模式的比例關係

| | 模式 1 | 模式 2 | 模式 3 |
|-----------------|-------|-------|-------|
| Lovebird1 | 36.7% | 13.6% | 49.7% |
| Lovebird2 | 36.7% | 18.3% | 45.0% |
| Newspaper | 41.2% | 9.10% | 49.7% |
| Alt Moabit | 13.6% | 17.8% | 68.6% |
| Book Arrival | 39.6% | 8.50% | 51.9% |
| Door Flowers | 31.1% | 12.1% | 56.8% |
| Leaving Laptop | 45.7% | 7.90% | 46.4% |
| Champagne Tower | 24.3% | 12.1% | 63.6% |
| Dog | 45.5% | 16.2% | 38.3% |
| Pantomime | 28.3% | 20.8% | 50.9% |

表 7 深度圖限制在低資料量的情況下，R.I. 預測模式的比例關係

| | 模式 1 | 模式 2 | 模式 3 |
|-----------------|-------|------|-------|
| Lovebird1 | 80.8% | 2.4% | 16.8% |
| Lovebird2 | 79.0% | 2.4% | 18.6% |
| Newspaper | 72.1% | 2.9% | 25.0% |
| Alt Moabit | 93.1% | 0.5% | 6.40% |
| Book Arrival | 70.2% | 4.5% | 25.3% |
| Door Flowers | 65.0% | 9.3% | 25.7% |
| Leaving Laptop | 70.3% | 4.8% | 24.9% |
| Champagne Tower | 81.0% | 2.0% | 17.0% |
| Dog | 74.3% | 7.0% | 18.7% |
| Pantomime | 85.4% | 4.6% | 10.0% |

参考文献

- [1] C. Fehn, R. Barre, and R. S. Pastoor, "Interactive 3-DTV: Concepts and Key Technologies," *Proceedings of the IEEE*, vol. 94, pp. 524-538, March 2006.
- [2] C. Fehn, "A 3D-TV Approach Using Depth-Image-Based Rendering (DIBR)," *Proceedings of Visualization, Imaging, and Image Processing*, September 2003
- [3] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High-Quality Video View Interpolation Using a Layered Representation," *ACM Transactions on Graphics*, vol. 23, pp. 600-608, August 2004.
- [4] A. Smolic, K. Muller, K. Dix, P. Merkle, P. Kauff, and T. Wiegand, "Intermediate View Interpolation based on Multiview Video plus Depth for Advanced 3D Video Systems," *IEEE Int'l Conf. on Image Processing*, October 2008.
- [5] E. Cooke, P. Kauff, and T. Sikora, "Multi-view Synthesis: A Novel View Creation Approach for Free Viewpoint Video," *Signal Processing: Image Communication*, vol. 21, pp. 476-492, July 2006
- [6] B. Kamolrat, W. Fernando, M. Mrak, and A. Kondo, "3D motion estimation for depth image coding in 3D video coding," *IEEE Transactions on Consumer Electronics*, vol. 55, 2009, pp. 824-830.
- [7] B. Kamolrat, W. Fernando, and M. Mrak, "3D motion estimation for depth information compression in 3D-TV applications," *Electronic Letters*, vol. 44, Oct. 2008, pp. 1244-1245.
- [8] S. Grewatsch and E. Müller, "Sharing of Motion Vectors in 3D Video Coding," *Int. Cong. Image Processing*, pp. 3271-3274, 2004

- [9] P. Merkle, Y. Morvan, A. Smolic, D. Farin, K. Müller, P. de With, and T. Wiegand, "The effects of multiview depth video compression on multiview rendering," *Signal Processing: Image Communication*, vol. 24, Jan. 2009, pp. 73-88.
- [10] Min-Koo Kang, Jaejoon Lee, Jin Young Lee, Yo-Sung Ho, "Geometry-based Block Partitioning for Efficient Intra Prediction in Depth Video Coding," *Proc. SPIE*, vol. 7543 75430A, 2010, pp.1-11.
- [11] Kwan-Jung Oh, Sehoon Yea, A. Vetro, Yo-Sung Ho, "Depth Reconstruction Filter and Down/Up Sampling for Depth Coding in 3-D Video," *IEEE Signal Processing Letters*, vol. 16, 2009, pp. 747-750.

