

以 H.264 壓縮技術為基礎之環物影片壓縮與即時播放

黃彥碩^{1,2} 黃春融¹ 黃文吉²

1. 中央研究院資訊所 2. 國立師範大學資訊系

E-mail: {greg, nckuos}@iis.sinica.edu.tw, whwang@csie.ntnu.edu.tw

摘要

目前許多博物館致力建構 3D 數位化虛擬博物館，而以影像式描繪技術建立虛擬博物館可提供與實物接近的影像品質。其中增添式環場的技術提供了高擬真與高互動性的展示環境，然而在網際網路之展示中，增添式環場內龐大的影像資料要如何即時的傳遞到使用者的電腦，同時又能即時的依據使用者與文物的互動播放便是一重要問題。在本論文中，我們利用 H.264/AVC 壓縮技術，降低增添式環場中環物影片的資料量，並針對環物影片之特性對 H.264/AVC 壓縮和解壓縮兩方面做效能上的改進，以達到減少壓縮時間與即時播放之目的。除此之外，進一步結合了多層次的快取架構(Multi-Layer Caching Structure)，以達到快速隨機存取任意環物影像之目的。

關鍵詞：虛擬博物館，H.264/AVC 壓縮技術、增添式環場、環物影片。

1. 前言

世界上的博物館往往保存著大量的古蹟或收藏品，為了確保物品被妥善地保存與維護，加上運送、維護與展覽空間等因素的限制，使得人們必須在限定的場所與限定的時間，透過櫥窗來觀賞這些物品，減少了這些收藏品呈現在廣大的群眾面前的機會。近幾年隨著虛擬實境(Virtual Reality, VR)和增添式實境(Augmented Reality, AR)技術的發展，許多博物館致力於透過 VR/AR 技術建構 3D 數位化虛擬博物館，透過虛擬博物館展示(Virtual Museum Exhibitions)可以讓不同地方的觀賞者觀賞這些世界各地的珍貴古物或收藏品，同時提供觀賞者與文物有別於傳統展示所無法達到的互動效果，此外進一步透過網際網路，3D 虛擬博物館展示也成為一個超越時間與空間的理想展示平台。為了建立數位化 3D 虛擬博物館，首先面臨的問題便是在文物數位化的工作，如何取得高品質的 3D 數位化文物資料，處理並儲存這些資料，更進一步展示這些數位化文物是目前許多研究致力的方向[1][2][3][4][5]。

3D 數位化文物資料最常遭遇到的問題，就是如何提供與實物接近的影像品質。不論由電腦圖學(Computer Graphics)角度出發，所研討的虛擬實境(Virtual Reality)或由電腦視覺角度出發所研究的增添式虛擬實境(Augmented Reality)，都是朝這個方向努力。在目前的研究中，相對於傳統的三維圖學技術，影像式描繪技術(Image-Based Rendering,

IBR)[6][7][8]，例如環物影片(Object Movie)[9]可提供觀賞者在觀看 3D 數位化文物時有如照片般真實的視覺效果。然而一個良好的虛擬展示環境應如真實的展覽館環一樣，可提供多個展覽室來展示不同類型的陳列品，並可讓使用者對整個展示環境進行互動與瀏覽。因此 Hung et. al.[5]的研發了一個可將環物影片(Object Movie)以符合幾何一致性的方式，逼真地合成於環場影像(Panorama)中的方法與技術稱為增添式環場(Augmented Panorama)。透過此技術可將博物館空間延伸至虛擬的數位展覽空間中，使用者透過此技術，不但可有實景般的視覺效果，且使用者也可與展示環境進行互動[1]。

雖然增添式環場提供了高擬真與高互動性的虛擬博物館展示環境，但是由於環物影片是由許多二維陣列影像所組成，而且在一個增添式環場中通常會有多個環物影片，因此應用增添式環場在網際網路之虛擬展示時便會面臨到一重大問題：龐大的影像資料要如何即時的傳遞到使用者的電腦，同時又能即時的依據使用者與文物的互動播放。為解決此問題，本論文採用目前最新的 H.264/AVC (MPEG4-part 10)[10][11][12]視訊壓縮技術來壓縮環物影片。並根據環物影片的特徵，本論文進一步提出改進 H.264 壓縮與解壓縮效能的方式與 cache 的架構與機制，使增添式環場技術能夠在網路上達到即時互動播放之虛擬展示。希望藉由有效的壓縮能夠降低環物影片的資料量，使其在短時間內就能夠透過網路傳輸到觀賞者的電腦上；同時，藉由快速地解壓縮與克服隨機存取(random access)的問題，讓觀賞者能夠與虛擬物品做順暢的互動式瀏覽。本篇論文所討論的主題正是如何針對影像式描繪技術中的數位化資料進行有效地壓縮與解壓縮之處理。

本論文的架構組織如下：第二章我們提出如何將根據環物影片之特性，進行 H.264 快速壓縮與解壓縮，同時又兼顧 random access 的需求的方法與系統架構；第三章為實驗結果；第四章為結論。

2. 主要內容

透過增添式環場(Augmented Panorama)建立 3D 虛擬博物館，可以自然地融合場景和物體的影像來呈現逼真的虛擬展示效果。然而為了能提供觀賞者與文物良好的互動性，增添式環場往往需要大量的影像資料來表示展覽的文物，譬如一個環物影片(object movie)常常需要上百張的影像，倘若一個展場有多個環物影片時，整體的資料量就相當可觀

了。這麼龐大的影像資料要如何快速地傳遞到使用者的電腦，同時又能夠即時地根據使用者與文物的互動播放，都是增添式環場應用在虛擬網路博物館展示所會遭遇到的問題。

一般來說，有兩種方式可以壓縮環物影片達到減少資料量大小的目的，第一種方式是將每張環物影像以靜態影像的壓縮方式獨立處理，如 JPEG、PNG... 等等，然而這樣的方式無法利用環物影像之間的高相似性導致壓縮的效率有限；第二種方式將環物影片視作一般的視訊影片，利用視訊的壓縮方法來處理，如 MPEG-2、H.263... 等等，這種方式的優點是有效利用影像之間的相似性，能夠大幅降低檔案的大小。但是利用視訊壓縮的方式來處理環物影片會遇到下列三個主要的問題：1. 如何將環物影片從二維的排列關係轉換成一維的影像序列，以方便用來做視訊壓縮；2. 如何能夠快速地從壓縮的 bitstream 中解出指定的影像；3. 克服觀賞環物影片時隨機存取(random-access)的問題。

為了克服以上的問題，本論文根據一般使用者習慣，提出一套環物影片排列方法，以利於目前最新的 H.264/AVC 視訊壓縮技術來壓縮與解壓縮環物影片。同時為達到 Just-In Time Rendering，本論文也提出了應用暫存(cache)的技術來快速解出任意一張壓縮好的環物影像，此外並根據環物影片的特徵，進一步提出改善的壓縮與解壓縮效能的方法，使得增添式環場系統能應用於網路瀏覽與提供使用者更加順暢之互動感。

2.1 環物影片排列方式

環物影片乃是具有空間上二維關係的影像集合(如Figure 1)，但是 H.264 視訊壓縮則是將一依照時間順序排列的一維影像序列進行壓縮，如何將二維關係的影像集合排列成一維的影像序列，是進行視訊壓縮之前先要解決的問題。視訊壓縮技術會使用影像間的相似性來提升壓縮的效能，經由影像之間的 motion estimation/compensation，以減少壓縮後影像所需的 bits。早期的壓縮標準通常只允許參考前/後一個 frame，但 H.264 具有多重參考 frame 運動補償 (multiple reference frame motion compensation) 的特性，在壓縮一張影像的時候可以同時參考多張的 frame，藉此提升影像的壓縮品質，然而要達到這樣的目的編碼器與解碼器必需儲存多張的參考影像並維持相同的順序，會造成解壓縮時的負擔。因此如果能把越相似的影像排列在一起，藉由 multiple reference frame motion compensation 的特性使壓縮之後的影像品質會更好，所以要找出一個比較好的排列順序來壓縮環物影片。排列方式的不同會影響壓縮後影像品質 (PSNR) 的不同。

有別於一般影片依照時間順序播放的方式，環物影片的播放是根據使用者觀看的角度，然後從 bitstream 解出相對應的影像。這樣的過程中除非每張影像均採用靜態影像的壓縮法，或是該張影像恰為 I-frame，不然大多數的情況下無法直接解出所需要的影像。因此除了 I-frame 之外，不管是 P-frame 或 B-frame 都需要參考的影像已經被解壓縮之後才



Figure 1. 環物影片的例子

有辦法跟著解出。也就是說當解壓縮某一張影像時，它所參考到的 frame 必需存在才能進行 motion estimation (ME) 與 motion compensation (MC)，如果參考影像不存在則必需先解出來。譬如 H.264 reference model software 預設前面 5 張影像都可以當作所參考影像，所以為了維持跟編碼端相同的參考順序，每次解壓縮時前面的 5 張影像都必需先被解壓縮，即使該張影像並沒有參考到這麼多張。這些額外被解壓縮的影像就可能造成播放上的延遲，而這個延遲會受到兩個因素影響，第一是 I-frame 的個數，當 I-frame 的數量越多的時候，這個延遲就越短但 bitstream 的資料量也會增加；第二是 bitstream 中影像排列的方式是否符合使用者操作環物影片的模式，如果使用者觀看的影像排列越相近則延遲也就會縮短，但我們很難事先預測使用者觀賞的方式，而且每次觀看的次序也不見相同，因此要有效衡量排列的好壞是一件困難的工作。

以 QTVR 所採取的方式為例，環物影片的排列是以行優先 (row major)，也就是相同垂直角度 (tilt) 的影像優先串接在一起，這樣的優點是水平轉動環物時，可以獲得平順的播放效果。但是觀賞者突然從某一列跳到另一列的時候，可能因為 P 或 B frame 的出現造成播放延遲的現象。因此排列方式的不同會影響隨機存取之結果。

本論文為了解決上述的問題，同時配合 H.264 rate-distortion 的機制，根據一般使用者在瀏覽環物影片時會先觀看對應於物體某一面之影像與其附近的影像之習慣，提出一套環物影片排列方法，根據 I-frame，利用分群的技術 (clustering) 將環物影片以相似性分成若干群，然後再根據每張影像對群內其他影像執行 motion compensation 的影響將影像排成一維序列，如此一來，藉由把越相似的影像排列在一起，使得壓縮之後的影像品質會更好。

2.1.1 Image Clustering

為了透過 H.264 壓縮環物影片，必需先將環物影片的二維關係透過適當的方式轉換成一維序列。假設將環物影片中每張影像都壓縮成 I-frame 的話，雖然不管影像序列如何排列都不會影響到壓縮的品質或播放時解壓縮的速度，但如此一來 bitstream 的大小勢必增加許多，不利於數位博物館的網路虛擬展示。相對來說，若序列的 I-frame 的個數太少，則 P-frame/B-frame 解壓縮時因為參考影像所造成的播放延遲勢必會增加。在此兩難的情況下，如何在 I-frame 的個數與環物影片播放時的順暢度之前能夠取得平衡是我們所需要解決的問題。

在 H.264 如果壓縮時影像中每個 slice 或 macroblock 都採用 intra prediction 的話，則該張影像為 I-frame，H.264 進一步支援一種特別的 I-frame 形式稱之為 instantaneous decoding refresh (IDR) picture，當 IDR 影像出現時會把目前在 reference list（參考影像序列，編碼端或解碼端都需要維持相同的序列）裡的影像會標示成“unused for reference”，則在 IDR 影像之後的 P-frame 不能使用 IDR 影像之前的影像當作參考影像。我們利用此一特性，在選定適當的 I-frame 個數之後，把每個 I-frame 設定成 IDR picture，則每個 I-frame 之後所帶領的數個 P-frame 可以視為獨立的群組，於是就把整個環物影片分成數個群組。接下來將討論如何透過 clustering 使得每個群組內的影片彼此之間具有較高的相似性，讓後續的壓縮時 ME / MC 能夠獲得較好的結果而增加壓縮的效能。

在 clustering 的過程中，首先我們先由環物影片中選定 k 張影像當作 IDR picture，然後再根據相似性評估(similarity measure)，在剩下的環物影片中找尋與這些 IDR pictures 相似的影像。在此 similarity measure 是採用 H.264 motion estimation/compensation 的結果來評估。將非 IDR picture 的環物影像對 IDR picture 進行 motion compensation 之後的 residual 當作跟該 IDR picture 之間的相似值，所有 micro block (MB) 的 residual 加起來的和越小表示該張影像跟 IDR picture 的相似度越高，把相似度高的影像分成同一群，將來壓縮時這些相似性高的影像就有可能互相參考，而得到較佳的壓縮結果。詳細的演算法如下：

1. 從環物影片中選定 m 張影像當作 IDR pictures。
2. 根據公式(1)計算某影像 n 與 IDR pictures 的相似度 J_n ，找出相似度最大的 IDR picture k ，則該影像 n 屬於第 k 群。 MC_k 為 H.264 計算 motion estimation 的函式。若該群的個數超過預設的值 I ，則將該群中相似度最低的影像移出該群，並從執行第二步找出最佳的 IDR picture。

$$J_n = \min_k \sum_{mb_i \in n} MC_k(mb_i), n \in K \quad (1)$$

依序執行完所有影像，使得每一張影像都會與某 IDR picture 同一群，且每一群包含的影像個數相同。此外由於每一群包含的影像個數相同，所以我們也可預期最長解壓縮時間並可調整 IDR picture 個數，以使得檔案大小與環物影片播放時的順暢度之間能夠取得平衡。

2.1.2 Linear-Order Sequence

在上一節中，我們已經將環物影片根據與 IDR picture 的相似性分為 k 個群組，接下來我們將探討如何將每個群組的影像轉換成一維的序列。為了能夠得到最好的 MC/ME 的結果，在此我們考慮將對於其他的影像能夠有較佳 MC/ME 結果的影像優先排列，也就是說，排列越前面的影像對於後面的影像能夠提供良好的 MC/ME 結果。在一維序列中第一張影像就是 IDR picture，假設某張影像對於除了 IDR picture 之外，對剩下的每張影像能夠提供較佳的 MC/ME 結果，也就是說以此張影像當作參考影

像去預測其他影像所得到的 residual 的和為最小，因此排列在序列的第二個位置，以此類推直到完成整個序列為止。詳細的演算法如下

1. 在某一個群組 k 中假設所有影像屬於集合 E ，首先將 IDR picture 提出來當成壓縮序列 L 的第一張。接下來從 E 中選定一張影像 e_n ，計算其他影像以該張影像為參考影像 MC/ME 之後的 residual，並將所有的 residual 加總起來得到 R_n 。

$$R_n = \sum_{i=1, i \neq n}^N MC_n(i), i \in E \quad (2)$$

2. 對集合 E 中所有影像依序執行第一步，將 R_n 最小的影像移出集合 E 並放入序列 L 中。
3. 反覆執行 1 跟 2 直到所有影像都放入序列 L 中。如此一來，由於相似的影像都排列在一起，所以相互參考的效果也會變好。同時由於相近的影像都歸類於同一群中，當使用者瀏覽這群影像中的某一張影像時，這張影像與需要解壓縮這張影像所需要之 IDR picture 與 reference frame 都已經解好，當使用者在瀏覽附近的環物影像時，便有機會可以再使用這些已經解好的影像，而不需要再次由 IDR picture 重新解壓縮。

2.2 Multi-Layer caching structure

前面一節提出了排列環物影片的方式，使得二維的環物影片轉為一維之影像序列，同時又能提升壓縮後的影像品質與考量環物影片播放時的順暢度。此外，為了達到環物影片即時描繪(just-in-time rendering, JIT)，本節將提出一個 cache 的架構，透過此一架構能夠進一步降低解碼器的負擔，讓使用者可以更順暢地觀賞環物影片。

環物影片與一般視訊影片觀賞上最大的不同在於環物影片是依據與觀賞者的互動來決定播放的影像，並沒有一定的順序是屬於隨機播放；而視訊影片則為循序性的播放，雖然可以倒轉或快轉但通常只能移動到特定的 I-frame。此外，視訊影片播放過後就可以丟棄，不需要存放在記憶體中，然而環物影片播放過的影像很有可能馬上又會需要播放。如果每次都需要從 bitstream 中重新解壓縮或者把所有解過的影像都存放在記憶體裡的話，對於使用者的電腦而言都會是個沈重的負擔。為了達到環物影片 JIT rendering 的播放，環物影片對於快速隨機存取(random access)能力的要求要比一般視訊影片高上許多，一般視訊影片的隨機存取同常只提供特定幾張影像的存取，但是使用者觀賞環物影片時有可能會觀賞到任意一張的影像，因此解碼器必需能夠從 bitstream 裡隨機存取某一張影像的能力。

在介紹 cache 架構之前，先簡單的回顧 H.264 解壓縮的流程，H.264 的 bitstream 是由許多 NAL unit 所構成，當解某一張影像時必需先在 bitstream 裡找到該影像所在的 NAL unit 位址，再經由 CABAC 或 CAVLC entropy decoding 解出壓縮參數和 MB 係數，然後透過 inverse quantization 和 inverse transform 得到 residual coefficient，最後跟參考的影像做 motion compensation 之後就可以還原出完整影像。此外，H.264 為了消除 block distortion 情形，在影像輸出之前還做了 in-loop deblocking。根據[14]

的分析，上述的解碼流程最花時間的部分為 motion compensation，其次為 integer transform 和 dequantization，因此會以節省這些步驟做為設計 cache 架構的依歸。

基於以上分析，我們建立了一個多層的暫存架構(multi-layer caching structure)來克服環物影片隨機存取的問題，並減輕解碼器的負擔，當隨機存取環物影片時也能達到 JIT rendering 播放的目的。因為環物影片的觀賞跟一般的影片不同，使用者很可能會隨時觀看之前看過的影像，因此把已經重建後的影像暫存起來，當使用者需要觀看時會先到 cache 中搜尋是否有重建後的影像存在，若沒有則再要求解碼器重新解壓縮該張影像。如此將會大幅減少解壓縮的時間提供更順暢的觀賞效果。但是存放所有重建後的影像需要大量的記憶體，需要視不同的使用者環境做調整。因此在我們的架構中最上層的 cache 直接暫存 deblocking 之後的 reconstructed frame，暫存一張影像需要 786432 bytes (512x512x3)。

此外，為了縮短 motion compensation 和 inverse transform、inverse quantization 的時間，最簡單的作法是將這些結果都先暫存起來，但是如此一來所需要的記憶體空間相當於儲存一張完整的影像。因此，中間層的 cache 僅會儲存曾經解出影像的 motion vector 和 reference index 節省一些解壓縮的時間，每個 MB 的係數則不予儲存。如此一來，為了要重新解壓縮某張影像時，必需先到 bitstream 中找到影像所在 NAL unit，而為了能夠快速在 bitstream 中找到所需的 NAL unit，最底層的 cache 會儲存每張影像所存放的 NAL unit，並透過 index table 記錄每個 NAL unit 的位址。以上 3 層的 cache structure 的完整示意圖如所示。

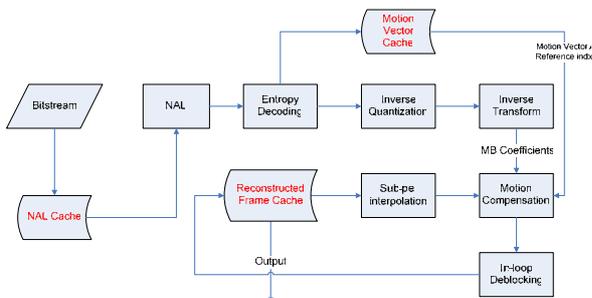


Figure 2. Multi-layer Caching Structure

2.3 前景與背景的壓縮

在前面提到增添式環場(Augmented Panorama)乃是將環場影像(Panorama)與環物影片(Object Movie)透過符合幾何一致性的方式自然的融合在一起[5]，給使用者有如真實一般的 3D 虛擬展示效果。將環物影片融合在環場影像之前，必需先對影像資料做去背處理(segmentation)，使得物體的前景與背景分離，藉由 alpha 值可以將物體與背景作 alpha blending。但是 H.264 只支援的 YCbCr video color space，Y 代表影像的亮度(luma)，CbCr 代表影像的藍、紅的色度(chroma)。雖然在 H.264 的延伸規格 Fidelity Range Extensions 提出了 auxiliary

pictures[16]，用以輔助影片產生 alpha blending 的效果，但以目前編碼器跟解碼器架構無法同時壓縮或解碼環物影片和其 auxiliary pictures。考量整體播放效率與處理便利性上，本文暫不考慮 auxiliary pictures。在此我們採用了另一種能夠同時壓縮環物影片和 alpha mask 影像的方式如下。在 H.264 新的規格中支援了更高精度的 color space，例如 YCbCr 422 和 YCbCr 444，利用這個特性首先將環物影片和 alpha mask 影像從 RGB 轉換成 YCbCr 422。而 alpha mask 影像只是一張灰階影像，轉換成 YCbCr 之後只需要保留 Y channel 的 luma 的資訊就足以用來做 alpha blending。因此，將環物影像的 CbCr 兩個 channel 合併單一個 channel，加上 mask 影像的 Y channel 合併成一個新的 YCbCr444 的影像(如Figure 3)。如此一來，將環物影片與 alpha mask 影像合併成一個 YCbCr444 格式的檔案後就可直接經由 H.264 的編碼器壓縮，同時也減少了影像資料量的大小。以這種方式編排過的環物影片，以一般的 H.264 解碼器解壓縮後影像的顏色會混亂且有鬼影，對於博物館在虛擬展示時可以提供保密功能。然而，這樣合併的缺點在於 U、V 兩個 channel 的影像不同，壓縮之後的 PSNR 會略微降低，實際測試降低的幅度約 0.05dB，對於實際觀看的品質並不會有太大的影響。

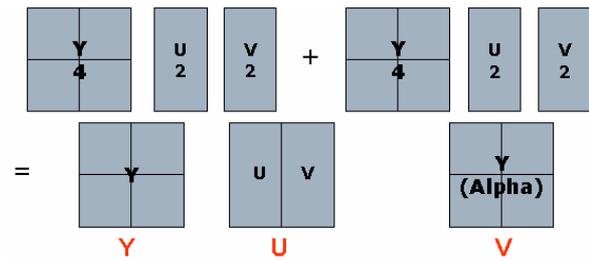


Figure 3. 環物影片與 mask 影像合併示意圖

2.4 增進環物影片編碼與解碼的效能

在上一節中，我們解決了環物影片跟 alpha mask 影像如何同時壓縮的問題。在本節中，我們將進一步利用環物影片的 alpha 資訊來加速編碼與解碼的過程，尤其 H.264 壓縮演算法的複雜度比之前的視訊壓縮標準都要高，如果可以降低壓縮與解壓縮的時間，將有助於環物影片的製作與播放。

新的視訊壓縮標準(如 MPEG-4 或 H.264)在壓縮影片時若能獲知其前景與背景的資訊，則通常可以利用該資訊針對前景與背景作不同的處理以獲得更高的壓縮效能，像 MPEG-4 的 shape coding 或是 H.264 的 Arbitrary slice ordering (ASO) 在每個 slice 配合不同的 QP。例如在[15]，Lin et al.提出一個藉由設定 RD optimization 不同的 λ 值達到控制前景與背景壓縮品質的目的。

觀察環物影片可以很明顯的發現，物體前景的部分只部分影像，以 H.264 這種 block-based 的壓縮標準來處理環物影片時，表示有將近一半的時間是用在編碼或解碼背景的部分，而這些背景 MB 處理的結果對於虛擬展示系統來說是一點都不重要，因為背景會在 alpha blending 之後被略去。所以背景的 block 對於我們的系統來說最好的處理方是就是完全忽略不壓縮，不但可以減少壓縮的時間，也可以

增加解壓縮的速度。在 H.264 規格中有定義了 skipped macroblock 型態，將沒有任何資料需要壓縮的 MB 標示成 skipped 不予處理，但在 reference software 僅有 B-frame 能夠包含 skipped macroblock。我們改良 reference software [13] 不管是哪種 frame type 都能夠將屬於背景的 MB 直接標示成 skipped。壓縮的時候會根據每個 MB 的 alpha 值是否超過一個 threshold，若大於則表示為前景反之為背景。同時將這個資訊寫入 MB 檔頭，如此解壓縮時就能夠知道每個 MB 是屬於前景還是背景。屬於背景的 MB 不管在壓縮或解壓縮都不需要再做處理。

2.5 Optimization Decoder

H.264 為了達到在低 bit-rates 還能夠維持良好的視訊品質，採用了許多複雜的演算法，以目前 reference software 的執行效率來看，無法滿足虛擬博物館對於 just-in-time rendering 的要求，在本論文中，我們藉由重新編寫了 H.264 JVT reference model software 9.0，包括最佳化程式模組和 memory alignment，同時結合 Intel® Integrated Performance Primitives library (Intel® IPP)，透過 Intel® IPP library 中的函式，我們可利用 Intel MMX 或 SSE 等硬體技術來改善 inverse transform、inverse quantization 和 deblocking filter 執行效率，獲得 decoding 效能大幅的改善。

3. 實驗結果

在實驗中，我們準備了 3 組環物影片資料如 Figure 4 所示，均為數位典藏計畫[17]的成果資料，其中，兩個為歷史博物館的唐三彩文物。每組環物影片均為 360 張，水平每 10 度一張，垂直每 10 度一層，每張大小為 512x512 pixels。所有實驗數據是在 Pentium 4 CPU 3.2GHz 和 1 GB ram 的機器上執行所得。我們最佳化的 H.264 decoder 是根據 reference software 9.0 修改而來，因此一些測試數據也是與 reference software 9.0 的版本作比較。



Figure 4. (a) 故宮博物院轉心瓶 (b) 歷史博物館唐三彩人俑 (c) 歷史博物館唐三彩馬

3.1 環物影片壓縮率的比較

最常見的環物影片壓縮方式是藉由 QTVR 套裝軟體，QTVR 提供 Photo JPEG、Cinepak 和 Sorenson 三種方式，表 1 為 H.264 跟這三種方式的壓縮率和壓縮品質的比較，明顯看出 H.264 的壓縮率大約為 QTVR 的 3-8 倍。而且經由 H.264 壓縮後的影像品質比 Cinepak 和 Sorenson 都要來的好。

3.2 環物影片壓縮與解壓縮效能之比較

表 2 與表 3 分別比較了我們的 decoder、encoder 和 reference software 在效能上的增進。重新改寫後的 reference software 每張解壓縮的時間比原先要快上 5-6 倍，如果再加上環物影片的 alpha 資訊省略掉不需要處理的 MB 之後，速度可以快上 6-8 倍之多，而 PSNR 並沒有因為速度的增加而降低，大致上維持與原本 reference software 相同的結果。就 decoder 來說，我們已經能夠達到 12 fps 以上的水準，再結合 cache 機制已經能夠使虛擬博物館提供 JIT rendering 的播放效果。藉由環物影片的 alpha 資訊，每組環物影片幾乎平均可以省略掉 50% 以上的 MB，使得環物影片的整體壓縮時間大幅的減少，除了唐三彩馬減少了 39.4% 的時間之外，其他兩組均節省了 50% 以上的時間。此外，因為省略掉的 MB 可以視為完全不失真，所以壓縮後的 PSNR 有微幅的提升。

4. 結論

在本論文中，我們依照環物影片之特性，提出將環物影片由二維影像序列根據所選擇之 IDR pictures 重新排列成適合 H.264 壓縮之一維影像序列之方法，在重新排序的同時，我們也考量了如何排序能使得隨機選取環物影像之效果較好。同時為達到 Just-In Time Rendering，本論文也提出了應用暫存(cache)的技術，在考量記憶體與解壓縮速度的情況下，來快速解出任意一張壓縮好的環物影像，此外，由於環物影片含有 alpha mask 影像，我們將環物影片與 alpha mask 影像合併成一個 YCbCr444 格式以減少了影像資料量的大小，同時利用 alpha mask 定義出可省略不處理之 micro block，來進一步提升壓縮與解壓縮環物影片之效能。藉由我們所提出的架構以 H.264 壓縮環物影片，不僅可達到比目前所常見之方法更加優秀的壓縮比，還可以符合環物 Just-In Time Rendering 與 random access 之性質，使得增添式環場系統更有機會能應用於網路瀏覽與提供使用者更加順暢之互動感。

致謝

本論文感謝國科會數位典藏計畫 NSC95-2422-H-001-007 與 NSC95-2422-H-001-008 之支持。

參考文獻

- [1] C.-R. Huang, C.-S. Chen, and P.-C. Chung, "Tangible Photo-Realistic Virtual Museum," *IEEE Computer Graphics and Applications*, vol. 25, no.1, pp. 15-17, 2005.
- [2] R. Wojciechowski, K. Walczak, M. White and W. Cellary, "Building Virtual and Augmented Reality Museum Exhibitions", *3D technologies for the World Wide Web*, pp. 135-144, 2004.
- [3] R. Walczak, W. Cellary, and M. White, "Virtual Museum Exhibitions", *IEEE Computer*, vol. 39, pp. 93-95, 2006.
- [4] B Hemminger, G Bolas, D Schiff, "Visiting Virtual Reality Museum Exhibits", *Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries*, pp. 423, 2004.
- [5] Y.-P. Hung, C.-S. Chen, Y.-P. Tsai, and S.-W. Lin, "Augmenting Panoramas with Object Movies by Generating Novel Views with Disparity-Based View Morphing," *J. Visualization and Computer Animation*, vol. 13, no. 4, 2002, pp. 237-247.
- [6] H. Y. Shum, S. B. Kang, and S. C. Chan, "Survey of Image-Based Representations and Compression Techniques", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 11, pp. 1020- 1037, 2003.
- [7] X. Tong, R.M. Gray, "Interactive Rendering From Compressed Light Fields", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 11, pp. 1080-1091, 2003.
- [8] H. Y. Shum, K.T. Ng, and S.C. Chan, "A Virtual Reality System Using the Concentric Mosaic: Construction Rendering, and Data Compression", *IEEE Transactions on Multimedia*, vol. 7, no. 1, pp. 85-95, 2005.
- [9] S. C. E. Chen, "QuickTime VR – An Image-Based Approach to Virtual Environment Navigation", *Proc. Siggraph, ACM*, Press, pp. 29-38, 1995.
- [10] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, "Draft ITU-T recommendation and final draft international standard of joint video specification", JVTG050, 2003.
- [11] T. Wiegand, G.J. Sullivan, G. Bjntegaard, A. Luthra, "Overview of the H.264/AVC video coding standard", *IEEE Trans. Circuits Syst. Video Tech.*, vol. 13, no.7, pp 560-576, 2003.
- [12] A. Luthra, G.J. Sullivan, T. Wiegand, "Introduction to the special issue on the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 13, no.7, pp. 557-559, 2003.
- [13] H.264/AVC reference software, <http://iphome.hhi.de/suehring/tml/>
- [14] X. Zhou, EQ Li, and Y.-K. Chen, "Implementation of H.264 Decoder on General Purpose Processors with Media Instructions", *SPIE Conference on Image and Video Communication and Processing*, pp. 224-235, 2003
- [15] Z.-K. Lin, and H.-H. Lin, Y.-H. Chen and J.-H. Chuang, "On The Performance Improvement of H.264 Through Foreground and Background Analyses", *In Proc. Of International Conference on Multimedia and Expo, ICME2005*, pp. 848-851, 2005.
- [16] G.J. Sullivan, P. Topiwala, and A. Luthra, "The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions", *SPIE Conference on Applications of Digital Image Processing*, pp. 454-474, 2004.
- [17] National Digital Archives Program, Taiwan, http://www.ndap.org.tw/index_en.php

表 1 H.264 與 QTVR 壓縮率比較表

(單位：KB)		唐三彩馬	轉心瓶	唐三彩人俑
Raw Data		276,480	276,480	276,480
QTVR Photo JPEG	壓縮率	32,918 (8:1)	27,071 (10:1)	22,429 (12:1)
	PSNR	44.723	46.921	46.494
QTVR Cinepak	壓縮率	24,086 (11:1)	24,091 (11:1)	24,087 (11:1)
	PSNR	37.343	39.094	38.4
QTVR Sorenson	壓縮率	10,153 (27:1)	7,983 (35:1)	7,493 (37:1)
	PSNR	37.396	37.201	40.623
H.264 (include alpha mask)	壓縮率	3,966 (70:1)	3,899 (71:1)	2,528 (109:1)
	PSNR	40.49	42.63	42.96

表 2 Decoder 效能比較表

		Reference software JM9.0	Our decoder	Our decoder (with MB skipped)
轉心瓶	Avg. time(ms)	486	78	58
	PSNR(Y)	42.48	42.48	42.48
	PSNR(U)	50.05	49.50	49.67
	PSNR(V)	42.55	42.33	42.36
唐三彩人俑	Avg. time(ms)	484	90	72
	PSNR(Y)	42.47	42.47	42.58
	PSNR(U)	48.76	48.69	48.75
	PSNR(V)	46.14	45.80	45.86
唐三彩馬	Avg. time(ms)	504	94	80
	PSNR(Y)	40.06	40.04	40.15
	PSNR(U)	47.05	46.95	47.01
	PSNR(V)	44.73	44.42	44.48

表 3 Encoder 效能比較表

		Reference software JM9.0	Our encoder
唐三彩馬	Total time(sec)	12180.6	7378.5(39.4%)
	Total bytes	4060653	4050276
	Avg. Skip MB	-	493.0 (48.1%)
	PSNR(Y)	40.06	40.16
	PSNR(U)	47.05	47.07
	PSNR(V)	44.73	44.75
轉心瓶	Total time(sec)	11229.5	5442.6(51.5%)
	Total bytes	3998430	3992374
	Avg. Skip MB	-	620.6 (60.6%)
	PSNR(Y)	42.48	42.48
	PSNR(U)	50.05	50.11
	PSNR(V)	42.55	42.57
唐三彩人俑	Total time(sec)	11322.6	5507.8(51.3%)
	Total bytes	2587975	2591339
	Avg. Skip MB	-	595.2 (58.1%)
	PSNR(Y)	42.47	42.57
	PSNR(U)	48.76	48.79
	PSNR(V)	46.14	46.18