

High Dynamic Range Imaging

Yu-Ting Wu (吳昱霆)

9527



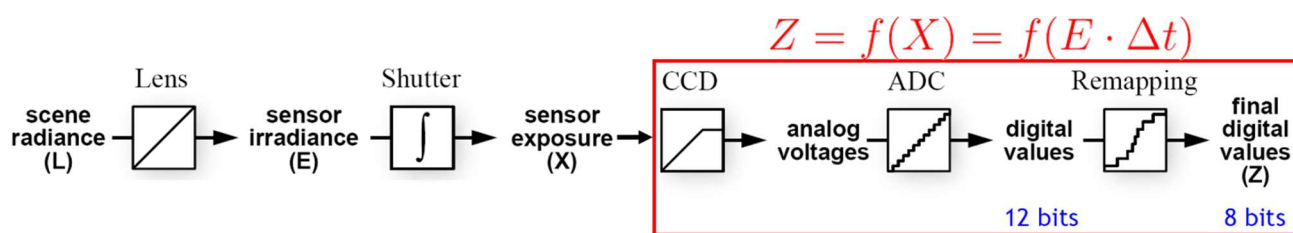
圖一：低動態範圍成像（LDR）與高動態範圍成像（HDR）的比較。低動態範圍成像因為相機的硬體限制，僅能捕捉真實世界中部份區段的能量，無法完整記錄完整的能量分布。例如在左圖中我們無法同時呈現較暗的室內與較亮的室外細節，當選擇捕捉戶外的細節時，室內門上的繪畫便會因能量過低，導致過暗而無法辨識。透過利用高動態範圍成像技術，我們可以重建真實世界的完整能量分布，如右圖中能同時呈現各種亮度的影像細節。

概述

動態範圍（Dynamic Range）指的是場景中最亮的點和最暗的能量比例。真實世界場景的動態範圍能夠到達數百萬倍，但一般相機因為硬體的限制，僅能記錄數百倍的動態範圍。高動態範圍成像（High Dynamic Range Imaging，HDRI）即是為了不受到相機感光元件的限制，能捕捉真實世界完整能量的分布並記錄於一張影像中而發展出來的技術。由於近年來高動態範圍成像已逐漸成為現代相機的標準配置技術，並即將成為未來超高畫質電視需要支援的標準技術之一，因此在多媒體領域是個很重要的題目。此篇報告將會從一般相機無法捕捉真實世界的完整能量分布的原因開始談起，接著描述高動態範圍成像的技術原理，最後說明其最新的技術發展。

1. 問題描述

現代數位相機的運作原理在於透過其內的感光元件捕捉真實世界的能量，並紀錄成影像。圖二展示了現代相機產生影像的過程：真實世界中各個方向的光線在經由透鏡折射後，會到達感光元件（CCD）並累積能量，在經過一段時間的曝光後，感光元件會將光能轉換成類比電壓訊號。為了儲存與操作上的方便性，此類比電壓會再被轉換成 12bit 至 16bits 的數位訊號(ADC)。最後為了讓拍攝出來的照片更貼近人眼視覺的結果，數位訊號會再被相機廠商設定的轉換曲線（non-linear remapping）以非線性的方式轉換成 8 bits 的數值並儲存成最後的影像。



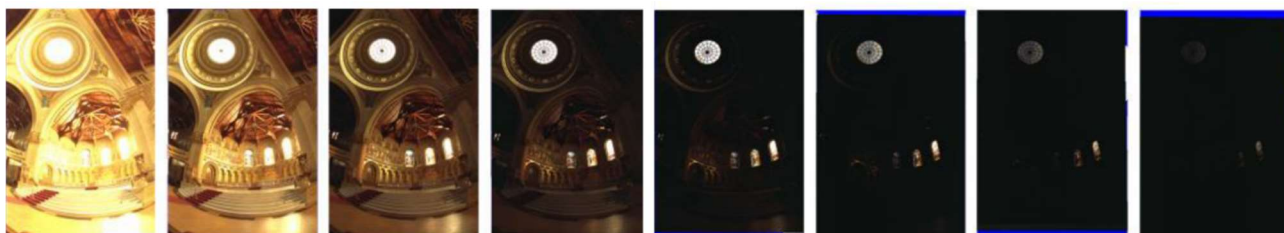
圖二：現代相機產生照片的流程。由於感光元件的硬體限制，真實世界的場景能量會受到截斷、量化、與壓縮等步驟，最終能表示的動態範圍僅為 0 至 255 的範圍。

在圖二展示的相機拍照流程中，有數個硬體限制導致相機拍攝的影像沒辦法捕捉真實世界的完整能量。首先，感光元件能夠承受的能量有其上限，若是超過其能夠紀錄的範圍能量就會被截斷 (clamping)。再者，將類比電壓轉換成數位訊號時必須執行量化 (quantization) 的動作，也會造成能量的誤差。最後但最重要的，儲存的圖片僅使用 8bits 來儲存每個像素的能量，代表其將能量表示在 0 至 255 的範圍，僅能表示約三百倍的能量差，但真實世界的能量分布很大，能量差距可以達到數百萬倍。因此當使用相機拍攝同時亮度變化很大的場景時，相機僅能透過調整曝光時間與光圈大小，選擇性的紀錄某個能量區段，無法在一次的拍照設定中同時捕捉完整的能量分布。如圖一左圖的例子，為了能捕捉室外高能量的場景，我們必須縮短曝光時間，以避免得到室外過曝的影像，曝光時間過短同時造成亮度較低的室內場景能量累積不夠，導致室內場景過暗而完全喪失細節；反之，如果我們為了凸顯室內細節而選擇用長時間的曝光，則室外的部分便會因為能量超過感光元件能儲存的範圍而過曝。由此可見，因為相機的限制，我們無法在一次的相機設定中同時捕捉到明亮與陰暗範圍的細節。

2. 高動態範圍技術

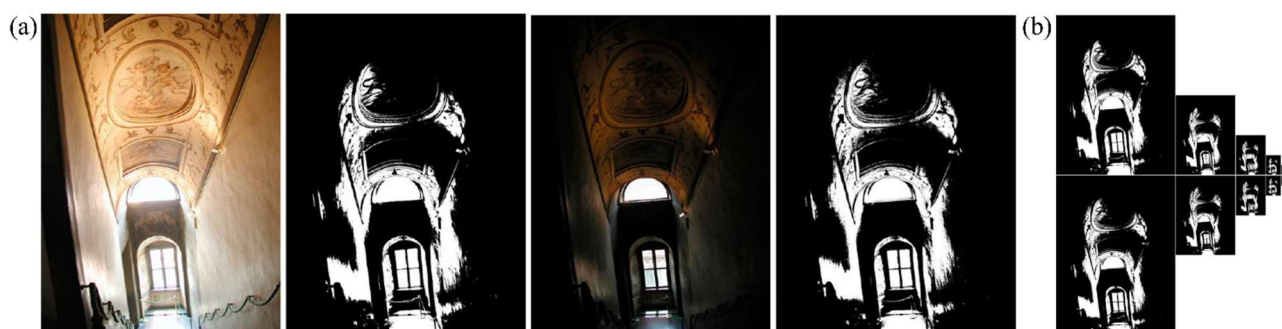
高動態範圍成像技術是為了不受到相機感光元件的限制，能捕捉真實世界完整能量的分布並記錄於一張影像中而發展出來的技術，最早是由 Paul Debevec 於 SIGGRAPH 1997 年提出 [1]。其主要的想法為在同一個場景以不同的曝光時間拍攝多張影像（如圖三），每張影像能紀錄某一個能量區段，再透過整合這些能量區段重建場景的真實能量。以圖二來說明，如果用 f 代表相機將真實場景的能量值 E 轉換成拍攝影像的像素值 Z 的函式，即 $Z = f(E \cdot \Delta t)$ ，其中 Δt 為曝光時間，高動態範圍成像技術的目的是透過求解 f 以得到每個像素 Z 對應的 E 。其流程為：

- 一、對同一個場景（相機需維持不動），以不同的曝光時間拍攝 N 張影像。
- 二、使用影像對齊 (image alignment) 演算法將 N 張影像對齊。
- 三、在對齊的每張影像中取樣 P 個點，用 $N \times P$ 個點估計相機拍照的函數。
- 四、使用額外的演算法移除鬼影 (ghost) 或是耀斑 (flare)。



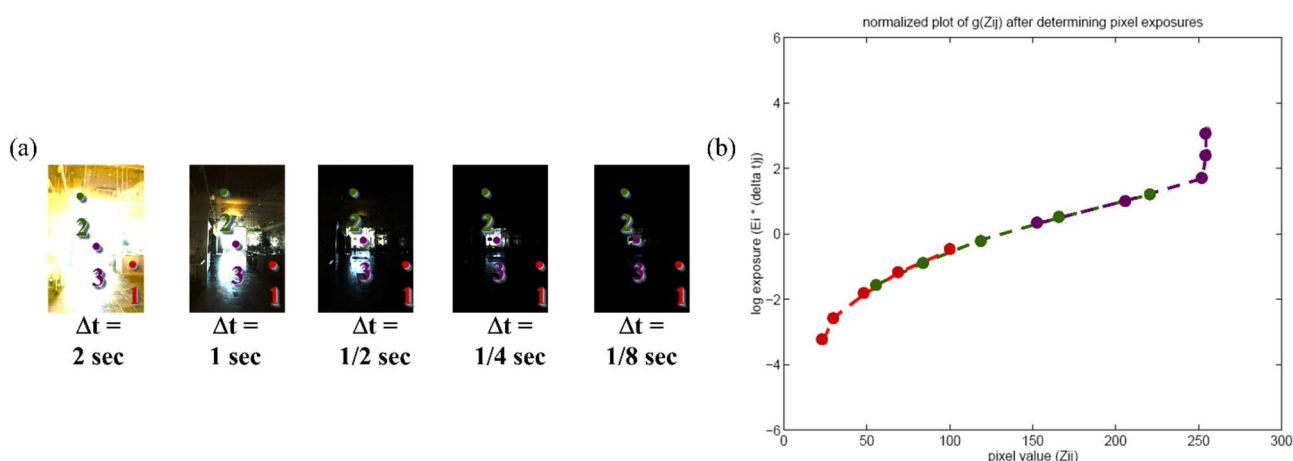
圖三：使用不同的曝光時間拍攝同一場景，以捕捉不同區段的能量。

在第一個步驟中，為了盡可能地維持相機不動，通常會使用腳架拍攝不同曝光時間的影像，同時為了之後的計算方便，每次曝光時間應以兩倍的方式遞增。不過即便是使用腳架固定，常常仍無法完全避免些微的晃動，因此步驟二的目的即是使用額外的演算法確保影像有對齊。由於一般的影像對齊演算法的假設是兩張要對齊的影像具有差不多的顏色，但在此應用中，不同曝光時間拍攝的影像顏色差距往往很大，因此會採用比較特殊的影像對齊演算法如 Median Threshold Bitmap (MTB) [3]處理。此方法的特色為首先計算影像明暗的中間值，並以此為閾值將像素二分，由圖四 (a) 所示，不同曝光時間的影像即使原本顏色差異很大，在二分法後會得到很接近的結果。為了減少計算量，MTB 演算法會對要對齊的影像計算圖像金字塔(image pyramid)，從最小的解析度逐步往上計算，如圖四 (b) 所示。



圖四：Median Threshold Bitmap 演算法。(a)透過計算每張影像亮度的中間值並以此為閾值對影像二分，即使是不同曝光時間的影像也會得到很類似的結果。(b)使用圖像金字塔減少對齊所需的計算量。

第三個步驟為高動態範圍成像技術最重要的一個步驟，目的為利用同一個場景下不同曝光值的觀察，還原相機拍照的函數，也就是圖二中的 f 。實際上，由於我們真正需要的是從照片的觀察值 Z 去回推其原能量 E ，因此我們真正需要找的是 f 的反函式 $g = f^{-1}$ 。如圖五 (a) 所示，首先在具有不同曝光時間的 N 張影像中選取相同的 P 個點，這 P 個點因為是屬於場景中同一個點，所以可以假設其能量 E 是相同的，不同的是曝光時間 Δt ，我們可以利用此關係去當成限制，根據下面兩個限制求取相機函數：



圖五：藉由在不同曝光的影像上取樣以求解出相機的拍照函數。(a) 相鄰照片的曝光時間應為兩倍間隔，圖中舉例為在每張圖上選取三個取樣點。(b) 所有的取樣點應當落在一條平滑的曲線上，此曲線即為相機的拍照函式曲線。

一、根據我們拍照時的設定，相鄰的曝光時間間隔為兩倍遞增，因此我們可以確定圖五 (b) 中相同顏色且相鄰的點能量差為兩倍。

二、所有的照片都是由同一台相機拍攝出來的，因此取樣的 $N \times P$ 個點都應落在相機函數的曲線上。

當我們取樣的點夠多時，我們可以以最佳化的方法找出一條曲線滿足上面的兩個限制條件，此曲線即為相機函數 f 的反函數 g 。藉由將 g 套用在所有的像素上，我們就可以還原整張影像完整的能量分布。

上述演算法的一個問題在於拍攝多張不同曝光的影像沒辦法在瞬間完成，如圖六 (a)，若場景中有物體在這段時間中移動，則重建出來的高動態範圍影像就會產生鬼影。如圖六 (b) 所示，我們可以在對每 N 張原始曝光影像反推場景能量後，計算 N 張影像中每個像素的變異量，當變異量太大時我們就不採信此像素。利用此方法我們就能大致移除重建影像中的鬼影，得到如圖六 (c) 的影像。



圖六：鬼影的移除。(a) 拍攝多張不同曝光時間的影像時若場景有會移動的物體，重建出來的高動態範圍影像就會有鬼影發生。(b) 透過將每張不同曝光時間的影像反推回場景能量並計算每個像素的變異量，我們可以移除有鬼影的像素，得到如 (c) 的結果。

3. 進階的高動態範圍影像技術

由於高動態範圍為非常實用的多媒體技術，在其發表之後，後續又有許多新的研究發表。例如 SIGGRAPH 2012 年發表的 patch-based HDR [3] 移除了靜態場景的限制，能夠產生動態場景中每張影格的高動態範圍影像 (圖七)。而 CVPR 2020 的方法 [4] 更進一步的移除傳統方法需要多張不同曝光影像的條件，藉由深度學習的方法預測任意一張影像的完整能量分布 (圖八)。

引用

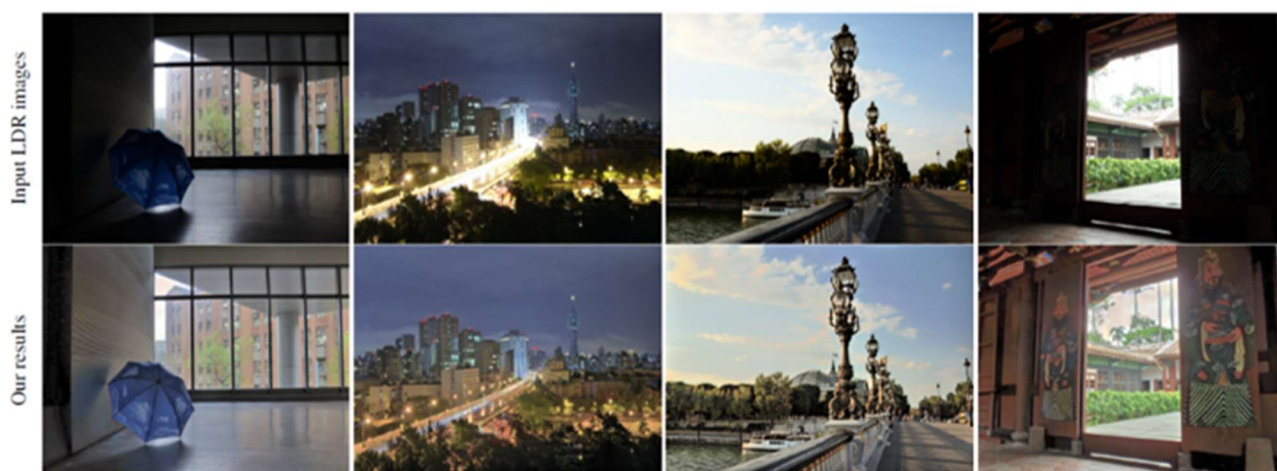
- [1] P. E. Debevec, J. Malik, “Recovering High Dynamic Range Radiance Maps from Photographs”, SIGGRAPH 1997.
- [2] G. Ward, “Fast Robust Image Registration for Compositing High Dynamic Range Photographs from Hand-held Exposures”, Journal of Graphics Tools, vol. 8, no. 2, 2003.
- [3] P. Sen, N. Kalantari, M. Yaesoubi, S. Darabi, D. Goldman, E. Shechtman, “Robust Patch-Based

HDR Reconstruction of Dynamic Scenes”, SIGGRAPH Asia 2012.

- [4] Y. Liu, W. Lai, Y. Chen, Y. Kao, M. Yang, Y. Chuang, J. Huang, “Single-Image HDR Reconstruction by Learning to Reverse the Camera Pipeline”, CVPR 2020.



圖七：SIGGRAPH 2012 發表的 Patch-based HDR，能夠產生動態場景下的高動態範圍影像。



圖八：CVPR 2020 發表的 Single-image HDR，以深度學習的方法，僅由單一一張影像輸入預測其完整的能量分布。