

# Project 1 (HDR) Report

R10922124 林家毅 (Group 46)

## 大綱

- 照片拍攝、前處理
- 實作 Ward's MTB Algorithm (**Bonus**)
- 重建 response curves
- 建立 HDR radiance map
- 實作 Reinhard's global & local tone mapping (**Bonus**)
- Reference

## 照片拍攝、前處理

這次 HDR recovering 是利用 FUJIFILM X-T10 在快門時間 1 ~ 1/64 秒之間拍攝的 7 張照片作為原始圖片，檔案格式為 .JPG format，如下圖。



為了方便後續 HDR 重建流程，在讀取照片的時候先 resize 成高度 1024 pixels 的同比例照片，並分別以 RGB、YCbCr 格式讀入。

## 實作 Ward's MTB Algorithm (**Bonus**)

由於拍攝照片過程中難免因為按下快門的輕微晃動使得每一張照片之間有細微差異，因此可以利用 MTB Algorithm[1] 將所有照片以某一張 reference image 為基準做 alignment，而在這次的實作是以第 4 張照片為 reference image。

為了將每一張照片跟 reference image 做 MTB alignment 求出最佳的 translation，要先取出代表 intensity 的 ‘Y’ channel，並利用 median threshold 將它們變成 intensity bitmap (高於 median 為 1，反之為 0)，同時，也會將 intensity 在 median  $\pm 4$  範圍內的 pixels 加入 exclusion bitmap (在範圍中為 0，反之為 1)，如此一來，將兩張照片的 intensity bitmap 做 element-wise XOR，接著再跟兩個 exclusion bitmap 做 element-wise AND，最後加總每個 element 就代表兩張圖片的 difference。

假設兩張照片的在每個方向的位移最多是 32 pixels，要找到最佳的 translation 使得照片跟 reference 之間的 difference 最小，通常會使用 pyramid 利用 coarse-to-fine 的概念求得，實作上是先把 reference image 和要被 align 的 image (relative image) 縮小 32 倍，找尋最佳 translation 為  $(-1, -1), (-1, 0), (-1, 1), \dots, (1, 1)$  九種中的哪一種，然後回到縮小 16 倍的圖片中，並將上一層得到的最佳 translation 乘以兩倍，然後在這層找九個可能的最佳 translation，依此類推，最後可以得到在原尺寸上最佳 translation。

最終，這個演算法的 output 會是每張 relative image align 到 reference image 的最佳 translation，故在這個 case 中會有 6 組 (x 方向 translation, y 方向 translation)，因此可藉由這些 translation 完成 MTB Alignment，結果如下。

Translation:  $[(-2, 3), (-1, 2), (-1, 1), (0, 0), (1, -1), (1, -1), (-2, 1)]$

## 重建 response curves

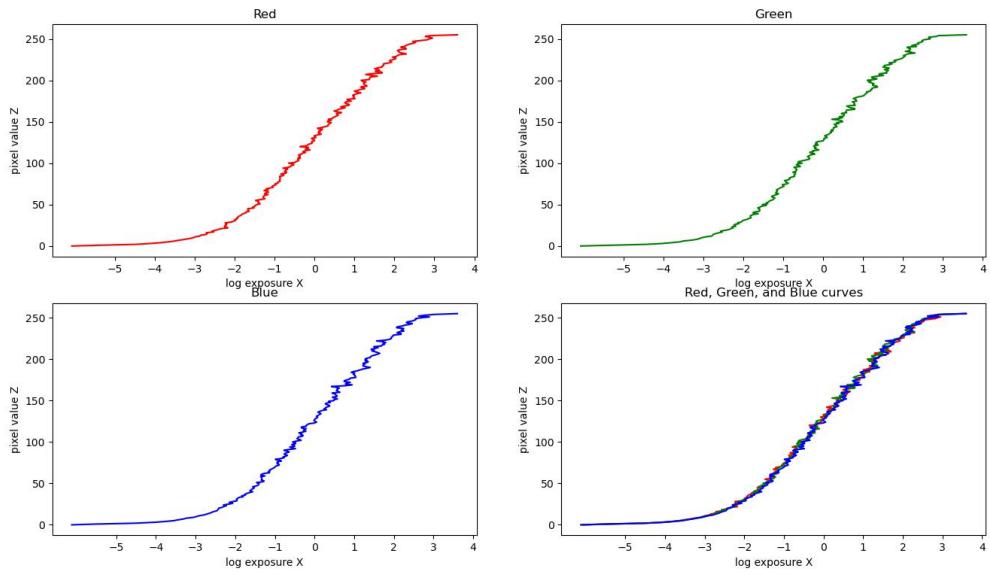
利用不同曝光時間拍攝的 7 張照片，可以選一些點來還原出相機的 response curves，經過實測，隨機在照片抽 1000 個點可以還原出蠻平滑的 response curves。

根據老師在上課所解釋的Debevec's algorithm，我們需要求以下 objective function 的最小值，

$$\mathcal{O} = \sum_{i=1}^N \sum_{j=1}^P [g(Z_{ij}) - \ln E_i - \ln \Delta t_j]^2 + \lambda \sum_{z=Z_{min}+1}^{Z_{max}-1} g''(z)^2$$

$$g''(z) = g(z-1) - 2g(z) + g(z+1)$$

實作上，會將上式轉換成 over-determined linear system，再用 Python `numpy` package 中的 `linalg.lstsq()` 求出 least-square solution，最終 recover 的 response curves 如下圖。

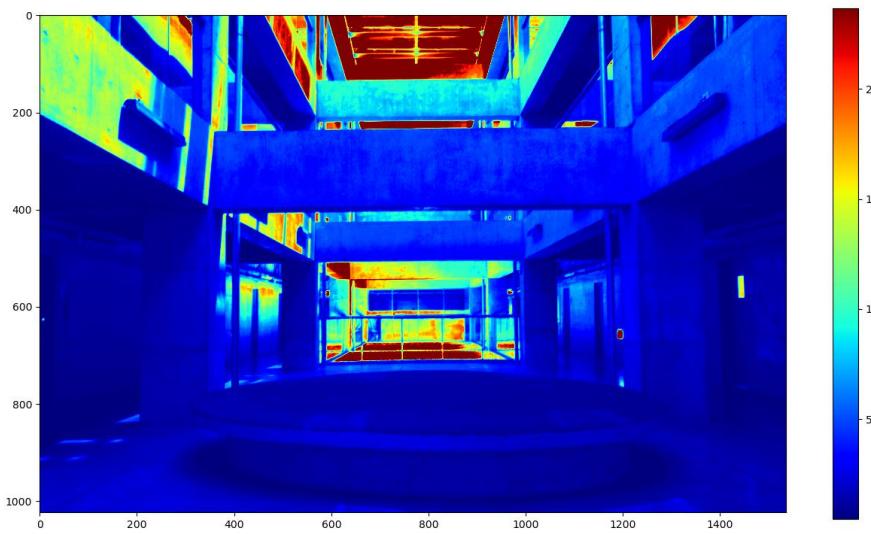


## 建立 HDR radiance map

有了 RGB 三個 channels 分別的 response function 之後，可以將 7 張照片的每個 pixel 利用以下式子求出 HDR radiance map

$$\ln E_i = \frac{\sum_{j=1}^P w(Z_{ij})(g(Z_{ij}) - \ln \Delta t_j)}{\sum_{j=1}^P w(Z_{ij})}$$

將 HDR image 以視覺化方式呈現 (並非真的 HDR image，因為它實際上是實數影像，視覺化只能呈現出能量高低分佈)，如下圖，可以看到紅色區域代表能量較高的地方 (光源強烈)、藍色區域代表能量較低的地方 (陰影為主)，跟實際場景看到的情況吻合。



## 實作 Reinhard's global & local tone mapping (Bonus)

為了將 HDR radiance map 轉換成跟人類視覺接近的圖片，我們需要做 tone mapping，在這邊共實作兩種 tone mapping，分別是 Reinhard's Photographic Tone Reproduction[2] 的 global mapping 和 local mapping 版本。

無論哪一種，都會利用以下式子將原本的 HDR radiance map  $L_w$  轉換成一個中繼的 map  $\bar{L}_w$ ，其中在論文中提到  $a$  是一個 key 值，較高的 key 值會使得最終結果較亮，而論文中以  $a = 0.18$  為 low key、 $a = 0.5$  為 high key，實作上，我認為  $a = 0.5$  在這個 case 上所呈現的結果較接近人類視覺。

$$\bar{L}_w = \exp\left(\frac{1}{N} \sum_{x,y} \log(\delta + L_w(x,y))\right)$$

$$L_m(x,y) = \frac{a}{\bar{L}_w} L_w(x,y)$$

### Global Tone Mapping

在做 global mapping 時，將整張圖片最亮的值設為  $L_{white}$ ，並利用以下式子得到  $L_d$

$$L_d(x, y) = \frac{L_m(x, y) \left( 1 + \frac{L_m(x, y)}{L_{white}^2(x, y)} \right)}{1 + L_m(x, y)}$$

最後將  $L_d$  乘以 255 倍取整數後即可得到 global tone mapping 的結果，如下圖。



## Local Tone Mapping

在做 local mapping 的時候，主要是希望可以找出每個 pixel 周圍最大平坦區域的平均亮度，來作為調整 pixel intensity 的依據，所以需要利用以下式子算出當 Gaussian kernel size  $s = 1, 3, 5, \dots, 43$  的情況下分別的  $L_s^{blur}$  以及  $V_s$ ，進而為每個 pixel 找出  $L_{s_{max}}^{blur}(x, y)$

$$L_s^{blur}(x, y) = L_m(x, y) \otimes G_s(x, y)$$

$$V_s(x, y) = \frac{L_s^{blur}(x, y) - L_{s+1}^{blur}(x, y)}{2^{\phi} a/s^2 + L_s^{blur}}$$

$$s_{\max} : |V_{s_{\max}}(\mathbf{x}, \mathbf{y})| < \varepsilon$$

接著就可以利用以下式子得到 local mapping 的  $L_d$

$$L_d(x, y) = \frac{L_m(x, y)}{1 + L_{s_{\max}}^{blur}(x, y)}$$

最後將  $L_d$  乘以 255 倍取整數後即可得到 local tone mapping 的結果，如下圖。



## Reference

1. Ward, G. (2003). Fast, robust image registration for compositing high dynamic range photographs from hand-held exposures. *Journal of graphics tools*, 8(2),

17-30.

2. Reinhard, E., Stark, M., Shirley, P., & Ferwerda, J. (2002, July). Photographic tone reproduction for digital images. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques* (pp. 267-276).