



## Pixel 手機相機技術大解密 會後重點整理

### 近年來手機相機的發展趨勢，不僅在數量上快速成長，在軟硬體設計上也不斷優化

- 目前全世界有數十億支手機相機，其數量仍在快速成長中
- 擁有更佳的像素、鏡頭和更大的光圈
- 多相機鏡頭成為發展趨勢
- 具有深度感應器 (Depth Sensor)

### 在手機相機系統的發展趨勢

- 以軟體定義 (Software-defined) 為主的相機：在圖像計算上不再單靠具固定功能的硬體，運用計算攝影 (Computational Photography)，整合連拍的幀數
- 導入機器學習：運用機器學習技術取代傳統的演算法，且在訓練機器學習的過程中獲得許多訓練用的數據，有助提升執行的準確度
- 更公開的科技與技術：以Google為例，Google具備軟體和機器學習的優勢，而在發展手機相機技術的過程中，不斷公開對外分享技術成果，不僅有助於驅動更快速的創新，也藉此吸引更多博士新秀加入Google AI研究團隊

### 手機相機應用程式執行時需遵守的基本原則

- 執行上必須要很迅速
  - 即時觀景器 (Live Viewfinder) 大於15 幀率 (fps)
  - 快門延遲時間小於150毫秒
  - 照片成像時間小於5秒：即為4秒，並且不讓手機因執行過程中受阻而升高表面溫度
- 預設模式絕不能失敗：須具有可靠的曝光、對焦和白平衡，並且避免鬼影或視覺殘影
- 重視消費者在攝影時遇到的特殊情況：過去Google的機器學習偵測器可以捕捉九成五以上出現在場景中的人臉，但是卻會忽視有戴墨鏡的人 (現在已經解決此問題)
- 在特殊模式下，可以接受偶發的失敗，尤其是幽默的結果

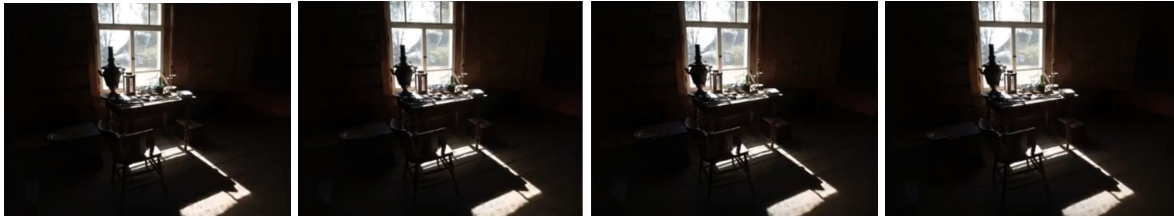


### 案例一：Nexus/ Pixel手機上的HDR+功能

- HDR (High Dynamic Range Imaging, 高動態範圍成像) 的典型作法：
  - 曝光定標試驗 (Exposure Bracketing)：
    - 捕捉不同曝光程度的影像，將從曝光時間短的影像所捕捉的亮處細節，與從曝光時間長的影像所捕捉到的暗處細節結合，呈現出最終的影像



- 在相機晃動和移動中物件的拍攝情況下，難以對準影像：當這些影像透過手持攝影而非使用固定三腳架拍攝時，由於每一個影像的曝光程度不同，雜訊也不同，猶如不同的影像，因此Google不使用曝光定標試驗達到HDR
- Google Pixel手機相機應用中的HDR+
  - 捕捉曝光不足下的連拍影像：透過捕捉在相同曝光程度下的連拍影像，並且避免亮處過度曝光



- 進行校準與合併：
  - 相同曝光程度下的連拍影像，由於影像相似因此更容易校準
  - SNR (Signal-to-noise ratio, 信號雜訊比) 和連拍的幀數的平方根成正比
  - 減少陰影處的雜訊
- 採用色調映射 (Tonemap)增強陰影，降低亮處：犧牲整體色調和對比度，以保留局部對比度



透過Google Pixel 手機相機應用中的HDR+所呈現出的影像效果



左圖為沒有開啟HDR+模式的影像，右圖為開啟HDR+的影像



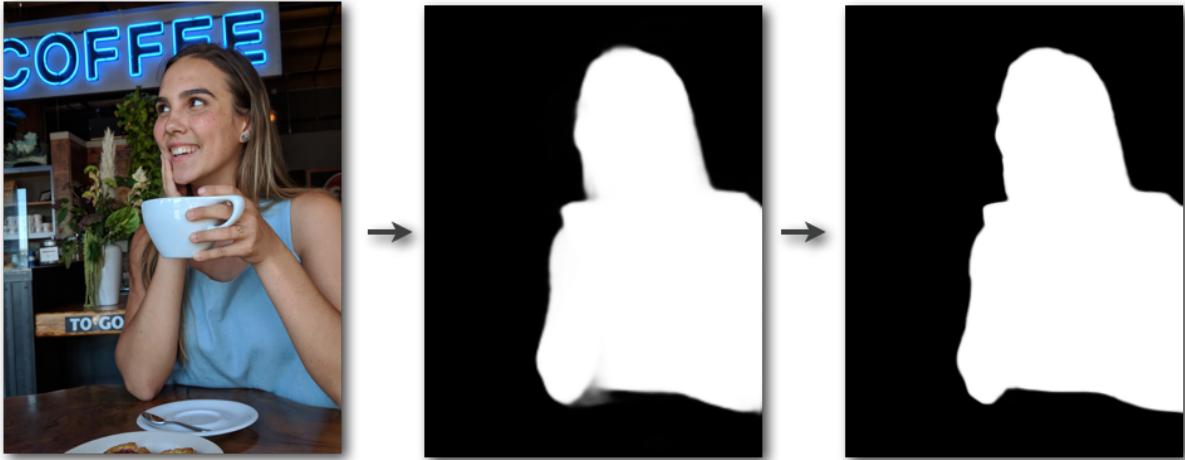
在低光源的情境下，透過HDR+也可以呈現清晰的影像 (如右圖)

- 在2016年，HDR+讓Google Pixel手機相機獲得當年DXO評分第一名 (DXO 89)

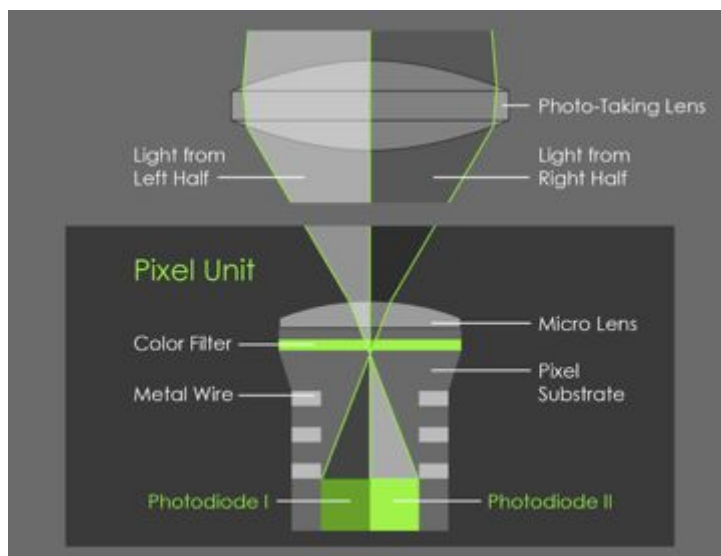
### 案例二：Google Pixel 2 肖像模式

- 肖像模式：將焦點集中在主要的人物或物品上，使其從背景中凸顯而出
- Google Pixel 2 肖像模式優秀的相機表現，獲得2017年DXO評分最高的智慧型手機，也獲得DPRReview Awards Innovation of the Year 年度創新獎的肯定
- **合成淺景深 (Synthetic Shallow Depth of Field)**：這是目前大多數的智慧型手機製造商所採用的方式，以呈現出照片背景模糊的效果。該作法為運用作假的景深達到真實景深的效果，而大部分的使用者其實難以辨別差異
  - 製造合成淺景深的典型作法包含：
    - 步驟一：利用雙鏡頭相機，在相似的焦點下捕捉兩個影像
    - 步驟二：運用立體匹配演算法 (Stereo Matching) 去計算深度圖 (Depth Map)
    - 步驟三：在場景中選擇一個平面以保持銳利影像
    - 步驟四：模糊距離該平面更近或更遠的影像特徵，讓背景呈現失焦的影像
- Google Pixel 2 和Pixel 3 手機在僅有單一個後置鏡頭的條件下，透過以下兩項技術達到景深的效果：
  - **利用機器學習技術分割影像中的人物**：當使用Pixel手機自拍時，前置鏡頭雖無搭載「雙像素 (Dual-pixels)」科技，但仍可透過機器學習執行肖像模式

- 運用卷積神經網路 (Convolutional Neural Network, CNN) 預估圖像中每一個像素的人物；卷積神經網路可以在TensorFlow中執行
  - 利用超過100萬張標籤過的人物和配件照片進行機器學習模型訓練
  - 使用數學計算中的邊緣感知雙邊求解器(Edge-aware Bilateral Solver)進行細化，讓人物分割的邊緣更加銳利



- 運用相機中的硬體技術「雙像素 (Dual pixels)」預估深度圖 (Depth Map)：當使用Pixel 手機對一般物件進行拍攝時，因物件非人物，故無法使用機器學習的方式，但仍可透過擁有「雙像素」的後置鏡頭技術完成
  - 關於雙像素 (Dual-pixels)
    - 即「相位對焦(Phase Detection Auto-focus, PDAF)」
    - 為較新單眼相機在錄影時對焦所使用
    - 每一個單一像素上的影像可分為兩部分：每個像素的右半部通過左半部鏡頭看世界，每個像素的左半部則通過右半部鏡頭看世界 → 事實上會得到兩種具些微差異的成像
    - 透過非常小的基線 (Baseline) (約1mm) 讓影像立體化



圖片來自Markus Kohlpaintner, 經許可轉載

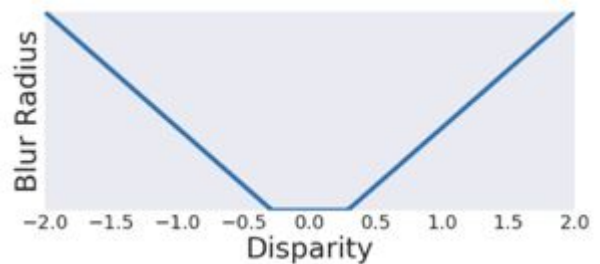
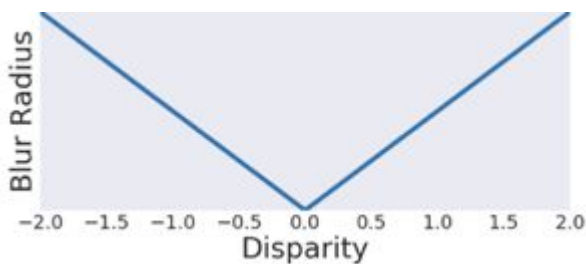


雙像素 (Dual-pixels) 所產生的深度圖

- 為了保持人物影像銳利，從機器學習人物分割技術中，根據人物與背景的距離按比例模糊化，並且將人物所在位置前後一段距離的影像也保持銳利，以讓人物影像能清楚呈現。（雖然非物理上正確，但有助於攝影初學者透過Pixel手機肖像模式拍攝人物）



- 透過專業相機手動對焦，焦點之外的景物會隨與焦點距離的長短而呈現不同程度的模糊化 (如左下圖)；Pixel手機相機則是將焦點擴展，對於攝影初學者而言較容易拍攝出具有景深的人物肖像 (如右下圖)

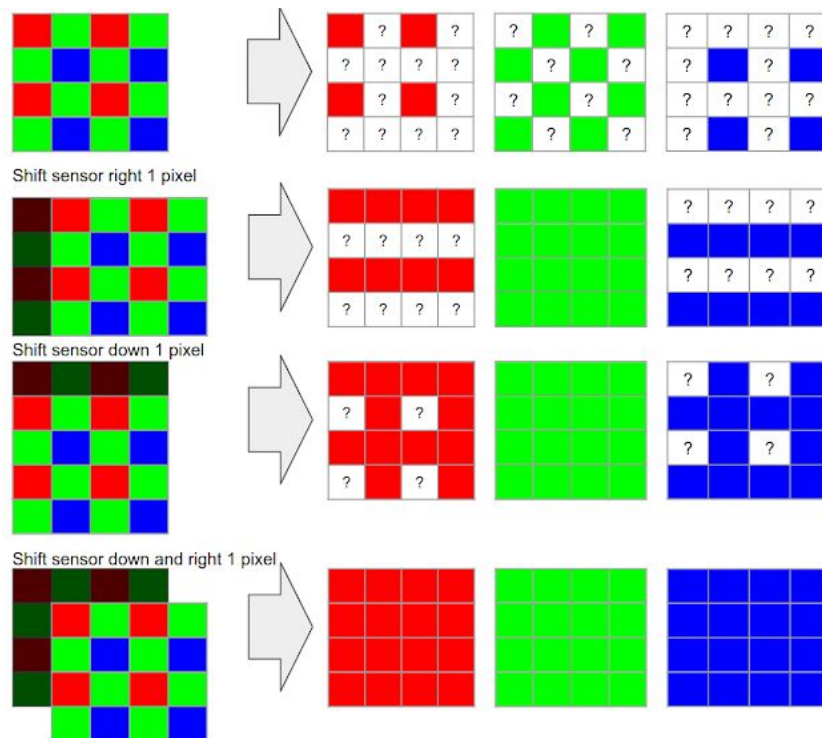


- Pixel 3 手機相機肖像模式的優化：Google進一步訓練機器學習藉由輸入RGB顏色、左和右，就能得到深度圖，取代在立體空間中搜尋、計算每一個二分之一像素中影像呈現是來自現實世界中的左半部或右半部
  - Google運用特殊的相機裝置，拍攝成千上萬的的照片，透過機器學習中的Ground Truth來訓練Dual-pixels更好產生深度圖 (Depth Map)。

### 案例三：高解析變焦 (Super Res Zoom)

Google在Pixel 3 手機相機中推出高解析變焦功能，雖然此功能背後並無搭載人工智慧及機器學習技術，但它可將多層影像合成為更高像素的照片，不僅提升照片細節的呈現之外，解析度與許多搭載2x光學變焦鏡頭的智慧型手機不相上下。

- **像素移位 (Pixel Shifting) 為基本原理**：影像是由一個個像素所組成，而每一個影像擁有紅色、綠色和藍色三種顏色資訊。高解析變焦透過連續拍攝多幀照片，以彌補缺失的顏色資訊以合並成更高解析度的數位照片，數位照片中的三分之二的顏色資訊都是重建出來的



當相機同時拍攝了4幀畫面，其中三幀分別於水平、垂直，還有水平與垂直同時位移一個像素，組合這4幀畫面後會發現原先缺失的像素都被填滿，不必再執行去馬賽克

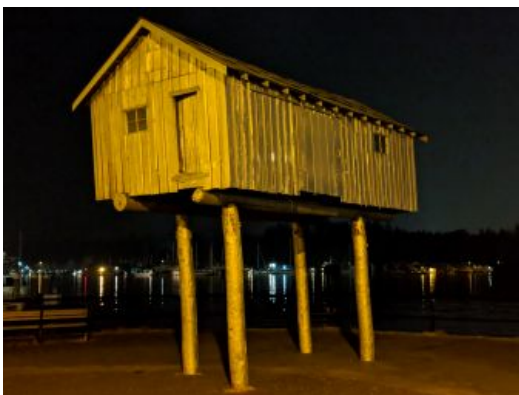
- 當手機或相機完全固定 (例如放置於三腳架上) 且不受手震影響時，仍可藉由強制開啟光學影像穩定功能 (Optical Image Stabilization)，製造鏡頭「搖晃」來模擬手震。此畫面的「搖晃」幅度非常小，並不會干擾正常照片。



左圖：無開啟高解析變焦功能；右圖：開啟高解析變焦功能

#### 案例四：夜視模式 (Night Sight)

- 夜視模式的技術：
  - 在按下快門後，觀景器將會靜止不動幾秒鐘，高速捕捉最多15幀畫面
  - 移動測量：根據手震程度與畫面中是否有移動物件來衡量所需拍攝時間
    - 當手震越強或是拍攝移動物件時，會縮短曝光時間，導致幀數越少、畫質越不清晰。反之，拍攝靜止物件時，曝光時間就越長，因此幀數增加，照片成果也更清晰
  - 結合其他Pixel手機相機功能：
    - Pixel 3 會結合高解析變焦功能
    - Pixel 1 / Pixel 2 / Pixel 3a 則是結合HDR+功能
  - 透過學習式白平衡 (Learning-based White Balancing) 調整影像色調



啟發式白平衡器  
(Heuristics-based White Balancer)



學習式白平衡器  
(Learning-based White Balancer)

- 藉由計算攝影的自動調整，仍可保持景物最真實的色調：
  - 在夜晚，景物的顏色依然存在，只不過人類的肉眼無法辨識，因此如何讓夜間拍攝的影像仍保有真實的色彩是一大挑戰
  - Google從古典的畫作中發現三大特徵，作為開發夜視功能的考量要素：
    - 強化對比度

- 將陰影投射成黑色
- 用黑暗環繞場景



Joseph Wright of Derby, A Philosopher Lecturing on the Orrery (1766)

- 夜視功能能藉由計算攝影進行色調映射 (Tone Mapping), 讓影像保有原本的色彩