

一種結合影像拼接與對比強化的環景照建構方法
**A Combined Method of Image Stitching and Contrast
Enhancement for Constructing Panoramas**

學生：XXX

指導教授：楊肅煜

國立中央大學數學系

February 14, 2025

1 前言

我們最常接觸的照片應該是從手機上的相機功能所拍攝的，一般拍照鏡頭是朝單一方向，因為視角有限，無法把四面八方周圍的環境同時拍攝進來，所拍攝到的只是當前方向的景物。環景照 (panorama) 或全景攝影顧名思義是用極廣的視角所拍攝的影像，在過去要完成大角度全景攝影是需要許多特殊的拍攝技巧與輔助設備，甚至使用軟體功能才能完整呈現。環景影像是較為完整的影像，可以捕獲完整的周圍環境而盡可能不會丟失任何信息，它們廣泛用於娛樂目的、顯示系統和虛擬環境，以提供身臨其境的體驗。然而，環景照是一項極具挑戰性的任務，在計算機視覺領域中被視為不適定的問題。儘管現代的智慧型手機搭載多顆鏡頭，但仍然存在有限的視野，無法捕捉周圍環境的完整影像，會丟失部份信息。

2 影像拼接

在本節中，我們將使用 SIFT 做為提取特徵的手法，SIFT 是計算機視覺中的一種特徵檢測的斑點檢測算法，用於檢測和描述影像中的局部特徵，它是由 David Lowe 於1999年發表於第七屆 IEEE 計算機視覺國際會議論文集 [?]。SIFT 對於影像有尺度不變性和旋轉不變性，而且對於以下的變化有強健性：仿射失真、雜訊、以及影像明亮的變化。在 SIFT 中對於輸入的影像給出多個 128 維的特徵向量，而這些特徵向量就是這個影像的特徵描述子。

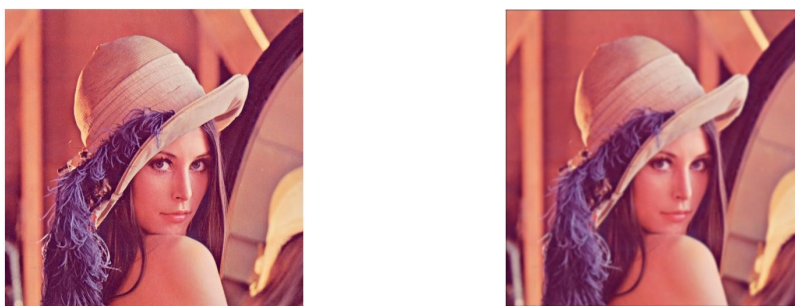


圖 2.1. (左) Lena 原始影像；(右) 經過高斯模糊 $\sigma = 1.5$ 處理後的結果

所謂一張灰階影像 $I(x, y)$ 的尺度空間 $L(x, y, \sigma)$ 定義為 I 與可變的尺度高斯核 $G(x, y, \sigma)$ 的卷積生成，也就是

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y), \quad (2.1)$$

其中 $*$ 代表捲積，尺度則由參數 σ 控制，而函數 G 可表示成

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}. \quad (2.2)$$

一般在計算 $G(x, y, \sigma)$ 的離散近似時，大概在 3σ 距離之外的像素都可以看作不起作用，這些像素的計算也就可以忽略。通常影像處理程序時只需要計算 $N \times N$ 的矩陣且 $N = \lceil 6\sigma \rceil$ 向上取最鄰近的奇數，再對每個 i, j 歸一化 $\hat{G}(i, j, \sigma) = G(i, j) / \sum G_{N \times N}(i, j)$ ，這時高斯卷積核即可達到所預期的效果；如圖 2.2 示意。

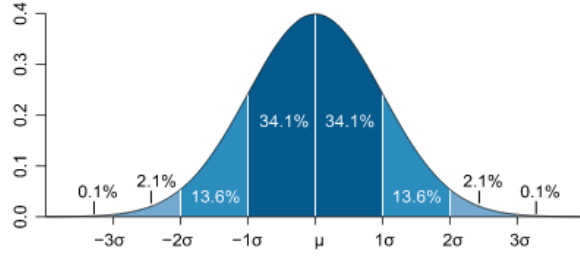


圖 2.2. 一維高斯模糊左右各取 3σ

觀察圖2.3 中特徵匹配可以發現還是有些特徵點對匹配錯誤，接下來會針對配對到的特徵點，找尋一個座標轉換矩陣以達到影像拼接的效果。我們將運用隨機抽樣一致法 (RANSAC) 找尋合適的特徵點對，去建立座標轉換矩陣也就是單應性矩陣。

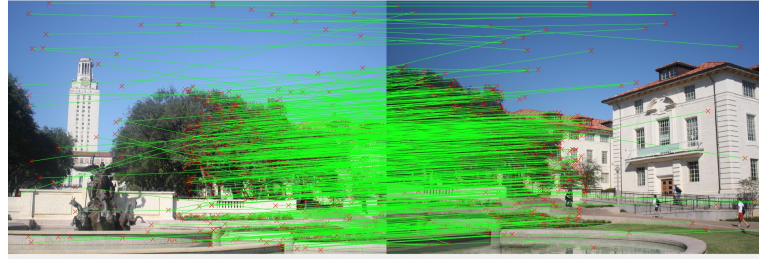


圖 2.3. 特徵匹配

單應性矩陣 (homography matrix) 是將一個四邊形經由矩陣轉換變成任意四邊形。令

$$\lambda \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}. \quad (2.3)$$

固定其中 $h_{33} = 1$ ，即

$$\lambda \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix},$$

我們可得

$$x'_i = \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + 1}, \quad y'_i = \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + 1}. \quad (2.4)$$

也就是

$$h_{11}x_i + h_{12}y_i + h_{13} - h_{31}x_ix'_i - h_{32}y_ix'_i = x'_i, \quad (2.5)$$

$$h_{21}x_i + h_{22}y_i + h_{23} - h_{31}x_iy'_i - h_{32}y_iy'_i = y'_i. \quad (2.6)$$

任意取 4 個特徵點對 $(x_i, y_i), (x'_i, y'_i), i = 1, 2, 3, 4$, 我們可以得到

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x'_1x_1 & -x'_1y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y'_1x_1 & -y'_1y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x'_2x_2 & -x'_2y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -y'_2x_2 & -y'_2y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x'_3x_3 & -x'_3y_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -y'_3x_3 & -y'_3y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x'_4x_4 & -x'_4y_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -y'_4x_4 & -y'_4y_4 \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{bmatrix} = \begin{bmatrix} x'_1 \\ y'_1 \\ x'_2 \\ y'_2 \\ x'_3 \\ y'_3 \\ x'_4 \\ y'_4 \end{bmatrix}.$$

所以單應性矩陣可表示成

$$\mathbf{A} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix}. \quad (2.7)$$

因為單應性矩陣有 8 個自由度，且一個特徵點對可以帶來兩個線性獨立的方程，由此可知給定 4 個特徵點對能決定一個單應性矩陣。如何找尋合適的 4 個特徵點對變成接下來要討論的重點。在這裡隨機抽樣一致法 (RANSAC) 就可以很好的解決此問題。隨機抽樣一致法概念就是重複隨機選取 4 個特徵點對盡可能生成出合適的單應性矩陣，使得單應性矩陣對所有特徵點對有較高匹配次數的一種迭代方法。

3 對比強化

多張原始影像在經由第二章的拼接流程後，我們會得到拼接完整的環景照影像。但是整張環景照可能會因為是多張影像拼接而造成某些區域亮度偏亮而某些區域偏暗，以及某些區域對比度不足而導致紋理結構較為不清晰，這些問題並不是因為拼接的不完全而造成的問題，而是因為各張影像在拍照時或取得時就已經亮度不一，針對這個問題，我們將在本章中引入一種對影像的對比強化 (contrast enhancement) 變分模型 [6]，做進一步處理。

4 數值實驗

本章將提供一系列的影像拼接和對比強化數值實驗，以驗證本文所探討的方法對建構環景照的有效性。為了完成實際影像的初步拼接及後續的對比強化，主要的步驟大致上包含從原始影像出發、圓柱投影、特徵選取、特徵匹配、RANSAC、多頻段融合、對比強化、及最後的最簡易色彩平衡。



圖 4.1. 德克薩斯大學校園影像 - 兩張待拼接原圖 (1024 × 683)

5 結語

本文研究結合影像拼接與對比強化技術的環景照建構方法。關於影像拼接，一開始每張影像都需要先進行圓柱投影，這會讓拼接後的影像視覺上更為自然，對每張影像使用尺度不變特徵轉換演算法 (SIFT) 提取影像特徵，接著對各張影像的特徵進行特徵匹配，並且使用隨機抽樣一致法，消除錯誤匹配的特徵點對，並篩選出合適的單應性矩陣，做為拼接的座標轉換矩陣，最後再進行多頻段影像融合。

參考文獻

- [1] A. B. Petro, C. Sbert, and J.-M. Morel, Automatic correction of image intensity non-uniformity by the simplest total variation model, *Methods and Applications of Analysis*, 21 (2014), pp. 91-104.
- [2] D. G. Lowe, Object recognition from local scale-invariant features, *Proceedings of the International Conference on Computer Vision*, 2 (1999), pp. 1150-1157.
- [3] J.-M. Morel, A. B. Petro, and C. Sbert, Screened Poisson equation for image contrast enhancement, *Image Processing On Line*, 4 (2014), pp. 16-29.
- [4] T. Lindeberg, Scale-space theory: a basic tool for analyzing structures at different scales, *Journal of Applied Statistics*, 21 (1994), pp. 225-270.
- [5] M. Brown and D. G. Lowe, Recognising panorama, *Proceedings of the 9th International Conference on Computer Vision*, (2003), pp. 1218-1225.
- [6] P.-W. Hsieh, P.-C. Shao, and S.-Y. Yang, Adaptive variational model for contrast enhancement of low-light images, *SIAM Journal on Imaging Sciences*, 13 (2020), pp. 1-28.