



國立台灣科技大學  
資訊工程研究所

---

## 碩 士 論 文

3D 點雲與 SIFT 特徵點室內影像定位之研究

3D Point Cloud and SIFT Descriptor Indoor Localization  
Research

研 究 生：陳致良

學 號：M9915057

指 導 教 授：項天瑞博士

中 華 民 國 一 百 零 一 年 七 月 十 日





# 3D 點雲與 SIFT 特徵點室內影像定位之研究

學生：陳致良

指導教授：項天瑞博士

國立台灣科技大學資訊工程研究所

## 摘 要

本篇論文探討如何利用單張彩色影像來重建出三維人臉模型。我們的方法是利用彩色影像來取代灰階值去建立張量模型 (tensor model)，在人臉資料庫中，是用典型相關分析 (Canonical correlation analysis) 是來建立彩色影像與深度資訊的對應關係，一旦建立好屬於各自的對應關係後，在重建人臉的過程中，就只需要單張彩色影像即可利用典型相關分析的對應關係來推算出正確的深度資訊。實驗中，我們的方法可以在不同的光線環境跟人臉角度得到不錯的效果。

# 3D Point Cloud and SIFT Descriptor Indoor Localization Research

Student: ZZhi-Liang Chen

Advisor: Dr. Tien-Ruey Hsiang

Submitted to Department of Computer Science and Information  
Engineering

College of Electrical Engineering and Computer Science  
National Taiwan University of Science and Technology

## ABSTRACT

This paper develops a tensor-based 3D face reconstruction approach from a single color image. Instead of the grayscale image, we also consider additional color factors in constructing the tensor model. Canonical correlation analysis is applied to establish the relationship between the color image and the depth information in the face database. During the face reconstruction, given a single color face image, the depth estimation is computed from the CCA-based mapping between the tensor models. Experimental results show our approach is better suited under different lighting conditions and poses.

# 誌

# 謝

本論文能夠完成，首先要感謝的是指導教授項天瑞老師。老師嚴謹的治學態度，讓我不但在學術研究上學習到更謹慎的思考，也在日常生活上獲益良多。

感謝實驗室的同學們，實驗室的生活有苦有樂，有你們才讓我能撐得下這三年漫長的時間。感謝建群、松翰學長給我的指導。感謝益偉、崇峰、承誌、誠儀、慶豪學長以及恩緯、盈樽、嘉駿、世寬學長給我的指導。感謝實驗室的學弟妹們，訓哲、立昂、致良、青緯、庭耀、宗博、雅筑，常常幫了我不少忙。更要感謝我的同梯們，貴彥、薇穎、冠佑，我們同甘共苦，一起奮鬥，特別是在弄計畫的過程中，冠佑幫助了我很多，有你們在我才能走到這一步。也感謝好朋友志強跟我一起修正文法與用詞。

最後我要把我最深的感謝留給我的家人。謝謝我的爸媽跟兩位姊姊，你們讓我沒有經濟壓力地讀完這這碩士學位，也常常給我很多鼓勵，今天我終於拿到這個學位，終於可以讓你們放下心上的一塊大石頭了。

# Table of Contents

論文指導教授推薦書 . . . . .	i
考試委員審定書 . . . . .	ii
摘要 . . . . .	iii
Abstract . . . . .	iv
誌謝 . . . . .	v
Table of Contents . . . . .	vi
List of Tables . . . . .	vii
List of Figures . . . . .	viii
1 相關研究 . . . . .	1
1.1 Simulated Localization And Mapping . . . . .	1
1.2 尺度不變特徵向量 (Scale Invariant Feature Transform) . . . . .	1
Bibliography . . . . .	5

# List of Tables



# List of Figures

1.1	高斯差分與影像金字塔 . . . . .	2
1.2	尺度極值偵測 . . . . .	2
1.3	特徵點向量方為與強度示意直方圖 . . . . .	4

# Chapter 1 相關研究

## 1.1 Simulated Localization And Mapping

## 1.2 尺度不變特徵向量 (Scale Invariant Feature Transform)

尺度不變特徵向量 (SIFT) 一開始由 Lowe [1] 所提出，目的是尋找兩張圖片中的相似特徵向量來比對兩張圖片的相對關係，主要分成四個階段：

- (1) 區域空間極值分布
- (2) 特徵點定位與篩選
- (3) 特徵點方向分配
- (4) 特徵點描述向量建立

第一階段 (1) 區域空間極值篩選，先利用不同尺度間的高斯金字塔選擇區域中的最大極值，其高斯分布式子如下：

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp(-(x^2 + y^2)/2\sigma^2) \quad (1.1)$$

不同尺度的高斯分布利用摺積 (Convolution) 將影像模糊化。 $I(x, y)$  代表原始影像， $G(x, y, \sigma)$  代表高斯函數：

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1.2)$$

再利用每組影像相鄰的高絲模糊影像進行高斯差分 (Difference-Of-Gaussian)，目的用於在集合內 4 組高斯差分影像中找出極值，式子如 (1.3) 所示：

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (1.3)$$

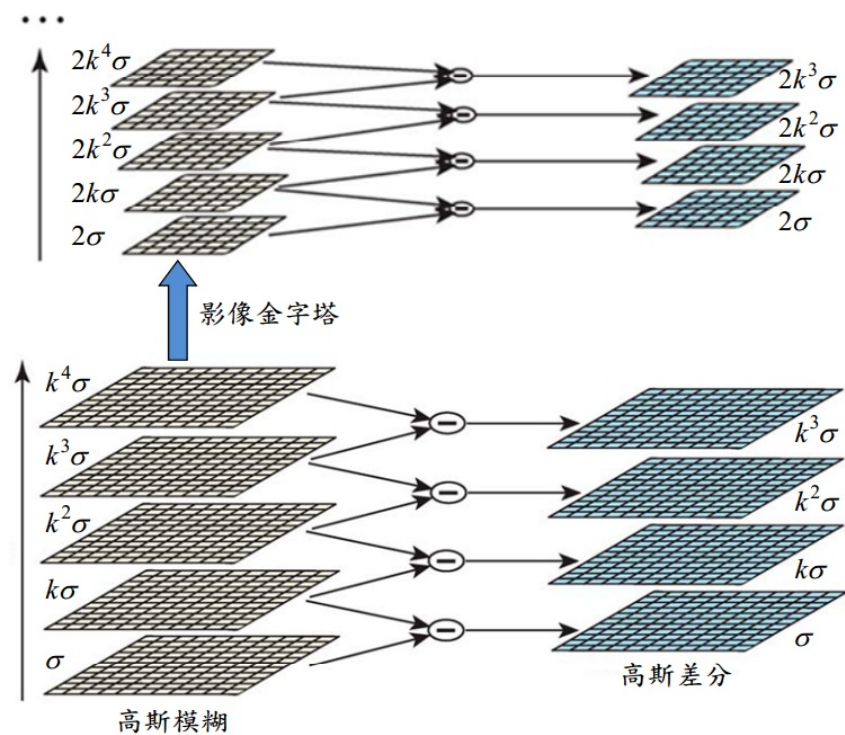


Figure 1.1: 高斯差分與影像金字塔

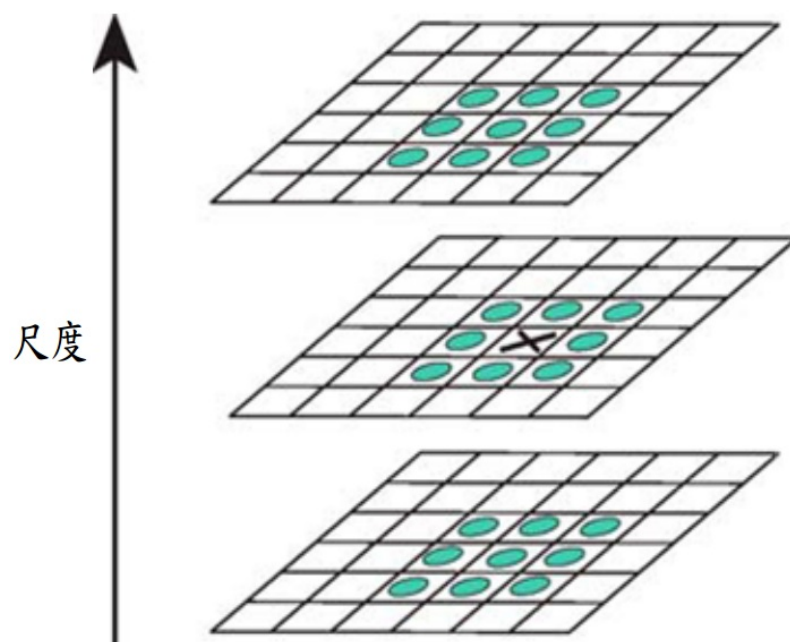


Figure 1.2: 尺度極值偵測

在此  $k$  為高斯模糊的尺度比值，設為  $\sqrt{2}$ ，若某個像素的極值為 26 個相鄰的像素中最大或最小的話，則此像素的位址即為區域極值的所在。

第二階段 (2) 特徵點定位篩選，其主要的目的在於找出真正有用的特徵點，在此特徵點的精準度必須要達到次像素的精度。有的特徵點其極值為低對比度的點，這時候這些低對比度的特徵點就會不予採用，剩下的特徵點即可為下一階段所使用。作法首先將 (1.3) 利用泰勒展開得到 (1.4):

$$D(x) = D + \frac{\delta D^T}{\delta X} X + \frac{1}{2} X^T \frac{\delta^2 D}{\delta X^2} X \quad (1.4)$$

式中  $X$  為極值  $(x, y, \delta)^T$ ， $D$  為高斯差分後的結果，再將 (1.4) 對  $X$  作偏微分可得  $\vec{X}$  算出  $X$  為極值點的的偏移量。

$$\vec{X} = -\frac{\delta^2 D^{-1}}{\delta X^2} \frac{\delta D}{\delta X} \quad (1.5)$$

若是  $\vec{X} \geq 0.5$ ，或是  $\sigma > k/2$ ，表示此區域極值點較靠近相鄰的點位，則需要再將此點移至相鄰的極值再經 (1.5) 計算後得到最佳的位置。若將  $\vec{X}$  帶入 (1.4) 中，可得 (1.6) 我們所用來篩選的式子:

$$D(\vec{X}) = D + \frac{1}{2} \frac{\sigma D^T}{\sigma X} \vec{X} \quad (1.6)$$

利用 (1.6) 將求出的絕對值與其他絕對值相比，可將對比度小的特徵點刪除以達到過濾的效果。

第三階段 (3) 特徵點方向分配，目的在於當對比的圖片有旋轉或者是尺度上的變化，相同的特徵點為了保有相同方向的特性，必須賦予每個特徵點一組特定的方向。其做法則是利用統計的方式，將所有的梯度值以角度每 10 個單位做方位直方圖記錄，並且記錄每個梯度的強度，以 (1.7)(1.8) 表示：

$$\theta(x, y) = \tan^{-1}\left(\frac{\delta L}{\delta y} / \frac{\delta L}{\delta x}\right) \quad (1.7)$$

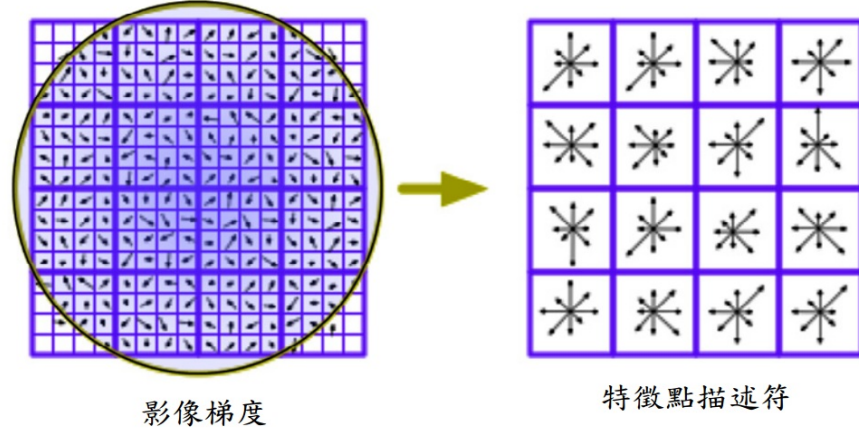


Figure 1.3: 特徵點向量方為與強度示意直方圖

$$m(x, y) = \sqrt{\left(\frac{\delta L}{\delta x}\right)^2 + \left(\frac{\delta L}{\delta y}\right)^2} \quad (1.8)$$

第四階段 (4) 徵點描述向量建立，最後一個階段為提供特徵點作為依據。做法上先將影像旋轉，使其特徵點向量與畫面主方巷一致之後，再以特徵點為中心，作出一個  $16 \times 16$  的視窗，並加一個尺度為  $0.5\sigma$  的高斯函數為權重。把每個區塊區分成  $4 \times 4$  的大小，切割成 16 個區塊，在每個區塊統計出梯度方位  $\theta(x, y)$  以及強度值  $m(x, y)$ ，而後分別每個區塊有 8 個區間，代表 8 個不同的方向，如 1.3 所表示。

每個特徵點有 16 個方位的直方圖，每個方位直方圖內有 8 個梯度強度值，因此共有  $16 \times 8 = 128$  個特徵值，這些特徵點則為我們所需要的描述向量。

# Bibliography

- [1] D. G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
- [2] H. Du, P. Henry, X. Ren, M. Cheng, D. B. Goldman, S. M. Seitz, and D. Fox, “Interactive 3D modeling of indoor environments with a consumer depth camera,” *Proceedings of the 13th international conference on Ubiquitous computing - UbiComp '11*, p. 75, 2011.