

國立交通大學

電機與控制工程學系

碩士論文

區域式主動表現模型演算法應用於人臉特徵匹配

Local Active Appearance Models Using For Face

Feature Matching

研究生：林育弘

指導教授：林進燈 博士

中華民國九十六年 七 月

區域式主動表現模型演算法應用於人臉特徵匹配

Local Active Appearance Models Using For Face Feature
Matching

研 究 生：林育弘

Student : Yu-Hung Lin

指導教授：林進燈 博士

Advisor : Dr. Chin-Teng Lin

國立交通大學

電機與控制工程學系



Submitted to Department of Electrical and Control Engineering
College of Engineering and Computer Science
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of Master
in
Electrical and Control Engineering
July 2007
Hsinchu, Taiwan, Republic of China

中華民國 九十六 年 七 月

區域式主動表現模型演算法應用於人臉特徵匹配

學生：林育弘

指導教授：林進燈 博士

國立交通大學電機與控制工程研究所

中文摘要

以主動表現模型為基礎，提出一個新的模型建立方法：區域式主動表現模型。主動表現模型目前應用廣泛，它結合形狀與紋理的資訊，使用主成分分析法建立一個統計模型。在主動表現模型特徵匹配中，使用模型紋理與影像紋理資訊的差異預測模型參數的走向而得到最佳的匹配。

由觀察得知，人臉主要變化的區域為眼睛與嘴巴，而且彼此間變化大致上獨立。所以在擷取形狀與紋理資料後，我們將形狀與紋理資料分成三個區域：眼睛、嘴巴、其他區域。每個區域都有自己的形狀與紋理資訊。將這三個區域分別利用主成分分析法各自建立三個獨立模型：眼睛、嘴巴與其他區域模型。將三個模型結合即為我們提出的區域式主動表現模型。

區域式主動表現模型使用區域式的紋理相減差異去預測模型參數，各模型間的收斂獨立運作。而位移參數則用全域紋理相減差異來預測。由實驗可得 1) 區域式主動表現模型模擬效果表現較主動表現模型佳。2) 因為區域式主動表現模型的區域各自獨立，不需考慮到資料間變化的排列組合，所以需要的訓練資料較少即可達到一樣的模擬效果，且在訓練資料較少的情形下，需要挑選的特徵向量較少，所需提供給模型參數儲存的空間也可以減少，即可用比較小的空間達到一樣的模擬程度。3) 區域式主動表現模型的模型參數各自獨立，所以在人臉特徵匹配後，若要再做嘴巴眼睛的表情辨識可直接將相關的模型參數直接用來做分類器的輸入。

Local Active Appearance Models Using For Face Feature Matching

Student: Yu-Hung Lin

Advisor: Dr. Chin-Teng Lin

Department of Electrical and Control Engineering
National Chiao Tung University

Abstract

We propose a new model based on active appearance model (AAM), called local active appearance model (LAAM). A popular approach, active appearance model, uses a combined statistical model (PCA) of shape and texture. The AAM searches use the texture residual between the model and the target image to predict improved model parameters to obtain the best possible match.

After observation, we find the eyes and mouth regions of face vary a lot and change nearly independently. After extract facial shape data and texture data from manual labeled images, we divide shape and texture data into three major regions : eyes, mouth, and other region. Each region has their shape and texture information. We form their statistical model of shape and texture. Model of eyes, model of mouth, and model of other are establish. Combining three models, we can get the local active appearance model (LAAM).

LAAM model predict improved model parameters using local texture difference between the model and the target image. The prediction of translation parameter uses global texture difference between the model and the target image. After experiment, we can get conclusion as follow: 1) The LAAM modeling performs better than AAM modeling. 2) Because of the independence character of eyes, mouths and other models. We don't need to consider about the relation of each region's expression. For the same performance the training data number can reduce. For the fewer of data number, we can choose fewer number of eigenvector to get the same performance. 3) The LAAM model parameter is independent of each other. If we want to do the classification of the expression of eyes or mouth, the model parameter can be sent to a classifier directly.

致 謝

本論文的完成，首先要感謝指導教授林進燈博士這兩年來的悉心指導，讓我學習到許多寶貴的知識，在學業及研究方法上也受益良多。另外也要感謝口試委員們的建議與指教，使得本論文更為完整。

其次，感謝實驗室鶴章學長、得正學長與所有學長姊、學弟妹們在研究過程中所給我的鼓勵與協助，在理論及程式技巧上給予我相當多的幫助與建議，讓我獲益良多。

感謝我的父母親對我的教育與栽培，並給予我精神及物質上的一切支援，使我能安心地致力於學業。此外也感謝姊姊對我不斷的關心與鼓勵。

謹以本論文獻給我的家人及所有關心我的師長與朋友們。



章節目錄

| | |
|-----------------------------------|-----------|
| 中文摘要 | ii |
| English Abstract..... | iii |
| 致謝..... | iv |
| 章節目錄 | v |
| 表格目錄 | vii |
| 圖片目錄 | viii |
| 1 第一章 緒論 | 1 |
| 1.1 研究動機..... | 1 |
| 1.2 研究目的..... | 1 |
| 1.3 系統流程..... | 2 |
| 1.4 論文大綱..... | 4 |
| 2 第二章 相關研究 | 5 |
| 3 第三章 區域性主動表現模型建模 | 8 |
| 3.1 形狀資訊擷取..... | 8 |
| 3.1.1 形狀校準..... | 9 |
| 3.2 紋理校正..... | 11 |
| 3.2.1 紋理資料點校準..... | 11 |
| 3.2.2 紋理資料點亮度校準..... | 13 |
| 3.3 區域性主動表現模型建模..... | 14 |
| 3.3.1 主成分分析..... | 14 |
| 3.3.2 區域性主動表現模型建模..... | 16 |
| 4 第四章 區域式主動表現模型定位演算法 | 21 |
| 4.1 人臉影像差異..... | 21 |
| 4.2 模型更新演算法..... | 27 |
| 5 第五章 實驗結果 | 31 |
| 5.1 實驗設計..... | 31 |
| 5.1.1 人臉表情定義..... | 31 |
| 5.1.2 人臉資料特色..... | 32 |
| 5.2 區域表現模型之模型建立測試..... | 33 |
| 5.3 區域表現模型之人臉特徵定位測試..... | 42 |

| | |
|---------------------|----|
| 6 第六章 結論與未來展望 | 48 |
| 6.1 結論..... | 48 |
| 6.2 未來展望..... | 50 |
| 參考文獻 | 51 |



表格目錄

| | |
|---------------------------------|----|
| 表 1：主動表現模型與區域式主動表現模型比較 | 35 |
| 表 2：主動表現模型與區域式主動表現模型比較(二) | 36 |



圖片目錄

| | |
|---|----|
| 圖 1-1: 區域式主動表現模型定位結果 | 2 |
| 圖 1-2: 系統架構 | 3 |
| 圖 3-1 : 四個相同形狀的笑臉, 但不同的位置、大小、角度 | 8 |
| 圖. 3-2: (左)定義的形狀資料點分布 (右)實際手工標定的結果 | 9 |
| 圖. 3-3: (左)未處理的形狀資訊 (右)校正過後的形狀資訊 | 11 |
| 圖. 3-4: 臉部標定點(左) 人臉三角形化網狀圖(右) | 12 |
| 圖 3-5 : 三角形之間的轉換 | 12 |
| 圖 3-6 : 三角形化紋理資訊擷取 | 13 |
| 圖 3-7 : 亮度對模型的影響: (第一列)輸入圖形 (第二列)模擬結果 | 14 |
| 圖 3-8: 區域式主動表現模型流程圖 | 16 |
| 圖 3-9: 手動標定影像取出形狀與紋理資訊 | 16 |
| 圖 3-10: 眼部、口、其它部分三個區域的形狀與紋理資訊 | 17 |
| 圖 3-11: 區域式主動表現模型 眼睛區域 前四個參數 $c_{eye} \pm 3\sigma_{eye}$ 的影響 | 19 |
| 圖 3-12: 區域式主動表現模型 嘴巴區域 前四個參數 $c_{mouth} \pm 3\sigma_{mouth}$ 的影響 | 19 |
| 圖 3-13: 區域式主動表現模型 其他區域 前四個參數 $c_{other} \pm 3\sigma_{other}$ 的影響 | 20 |
| 圖 3-14: (比較)主動表現模型前四個參數 $c \pm 3\sigma$ 的影響 | 20 |
| 圖 4-1: 區域式主動表現模型 - 模型預測因子 R 的訓練流程圖 | 25 |
| 圖 4-1: 區域式主動表現模型 - 模型預測因子 R 的訓練流程圖(二) | 26 |
| 圖 4-3: 區域式主動表現模型 - 人臉特徵定位演算法流程圖 | 29 |
| 圖 4-4: 區域式主動表現模型 - 人臉特徵定位演算法流程圖(二) | 29 |
| 圖 5-1: 臉部表情定義 | 31 |
| 圖 5-2: 人臉標定資料 | 33 |
| 圖 5-3: 區域主動表現模型維度 | 34 |
| 圖 5-4: 模型測試流程圖 | 35 |
| 圖 5-5: LAAM AMM 模擬差異 | 38 |
| 圖 5-6: 嘴巴區域形狀模擬誤差 | 39 |
| 圖 5-7: 嘴巴區域紋理模擬誤差 | 39 |
| 圖 5-8: 眼睛區域形狀模擬誤差 | 40 |
| 圖 5-9: 眼睛區域紋理模擬誤差 | 40 |
| 圖 5-10: 其他區域形狀模擬誤差 | 41 |
| 圖 5-11: 其他區域紋理模擬誤差 | 41 |

| | |
|---|----|
| 圖 5-12：人臉偵測結果 | 42 |
| 圖 5-13：區域式主動表現模型 - 人臉特徵定位演算法結果 | 44 |
| 圖 5-14：區域式主動表現模型 - 人臉特徵定位演算法結果(人臉一) | 45 |
| 圖 5-15：區域式主動表現模型 - 人臉特徵定位演算法結果(人臉二) | 45 |
| 圖 5-16：區域式主動表現模型 - 人臉特徵定位演算法結果(人臉三) | 46 |
| 圖 5-17：區域式主動表現模型 - 人臉特徵定位演算法結果(人臉四) | 46 |
| 圖 5-18：區域式主動表現模型 - 人臉特徵定位演算法結果(人臉五) | 46 |



第一章

緒論

1.1 研究動機

與人臉相關的研究在電腦視覺領域中佔有重要的地位，如人臉辨識技術應用在非侵入式生物認證、保全門禁系統、犯罪資料庫搜尋…等。表情辨識技術應用於人類與電腦間的互動。在人臉辨識或表情辨識之前，需要先定位人臉五官精確位置才能提供正確的資料給辨識系統。。一般會直接以人臉偵測系統的輸出結果當作人臉辨識或表情辨識的輸入資料，但由於臉型與五官比例的差異性，假若我們直接使用人臉偵測系統的輸出結果當作訓練資料，在同樣的資料區段描述的不一定是相同的五官位置，會隨著人臉的不同而有位置上的偏移差異。一般的人臉偵測系統只能給定人臉的大略位置與大小，所以在人臉偵測過後仍需要人臉特徵點定位系統的輔助找到五官的精確位置。得到人臉五官精確位置能夠幫助人臉辨識、人臉建模擷取特徵的正確性。所以研究人臉特徵點定位是研究人臉辨識、人臉建模、表情辨識的一個基礎，所以人臉特徵點定位是一個關鍵性的技術，應用廣泛，人臉特徵點定位的精確與否，會影響之後應用的效果。

1.2 研究目的

在人臉偵測系統偵測出大致人臉位置與大小後能夠精確的鎖定人臉特徵(眼睛、嘴巴、鼻子…等)與特徵的輪廓。利用可變形的形狀與紋理模型(在此，紋理是指人臉位置上的亮度值)，對一個新影像做匹配追蹤。實驗中，使用人臉為偵測的物體，此方法能同時在靜態影像與動態影片中自動地找到人臉形狀特徵，如：

眼睛、嘴巴四周...等，正確地匹配五官位置還有人臉紋理特徵。如圖 1-1 所示，有一等待定位的人臉影像，由初始圖(左一)，一開始只知道人臉大略的大小與位置(由人臉偵測系統得知)，將我們之前建立的平均模型投影至初始偵測位置，經由人臉特徵匹配程式收斂，由圖中(左二)程式第三次迴圈、(左三)程式第八次迴圈到(右三)程式收斂，可以觀察到模型整體位置與模型的人臉形狀有逐漸與新影像匹配正確。所以此程式可正確定位出物體的形狀(右三)與內部紋理的資訊(右二)。如此可得到精確的五官位置。

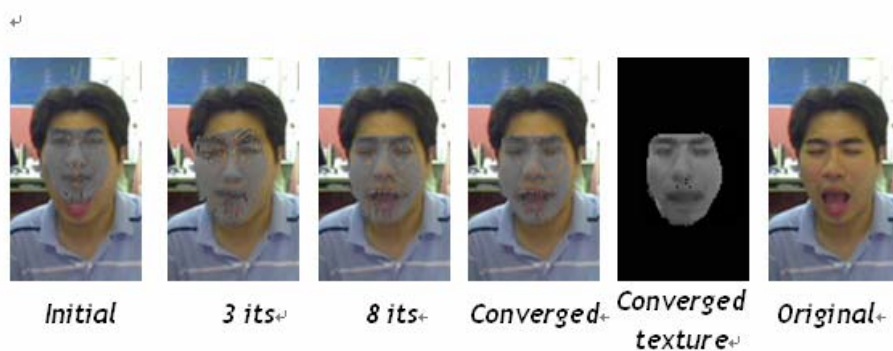


圖 1-1：區域式主動表現模型定位結果

1.3 系統流程

主動表現模型演算法(Active Appearance Models : AAM) [1]合併形狀與紋理的資訊建立一個統計模型，主動表現模型演算法的匹配搜尋以統計模型為基礎，使用模型與目標影像的差異去預測模型與位移參數如何改變去得到最佳匹配。我們以主動表現模型演算法為基礎，在統計模型方面提出區域性主動表現模型(Local Active Appearance Models : AAM)，對於物體中趨近於獨立變化部分，如：人臉中的眼睛與嘴巴。將這些區域各自獨立(眼睛區域、嘴巴區域與其他區域)以主成分分析法(PCA)製作區域式統計模型，再做人臉上的結合形成區域式主動表現模型，區域式主動表現模型比主動表現模型更能夠精確的模擬出人臉的形狀

與紋理表情。

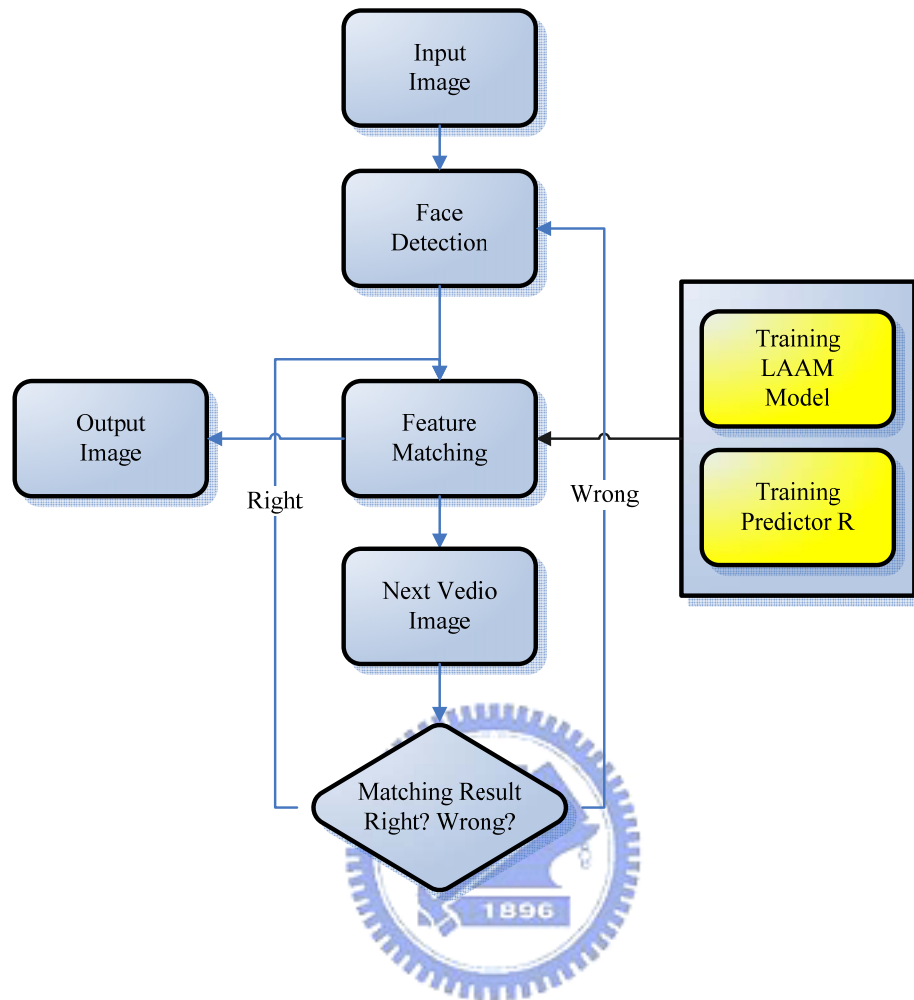


圖 1-2：系統架構

在追蹤特徵形狀與紋理部份，區域式主動表現模型使用區域式的紋理相減差異去預測模型參數，而位置、大小、旋轉參數則用全域紋理相減差異來預測其收斂數值。在預測的步驟中由於速度上的需求，在預測各參數的改變大小的最佳化步驟使用的預測因子 R 由訓練資料計算得知。

系統在實際執行的流程如圖 1-2 所示，輸入的初始人臉影像經由人臉偵測系統偵測出大致的人臉大小與位置，再經由特徵匹配系統得到精確的人臉形狀與紋理資訊。經由錯誤判斷決定此次的結果是否有真正收斂。若有真正收斂則以此次的結果當作下一張影像的初始值繼續由特徵匹配系統得到精確的人臉形狀與紋理資訊。若錯誤太大則下一張影像回到初始的人臉偵測系統偵測大略的人臉大小

與位置再經由特徵匹配系統得到精確的人臉形狀與紋理資訊。而在特徵匹配系統中，需要先訓練區域式主動表現模型與訓練預測因子 R 才能夠建立特徵匹配系統。

1.4 論文大綱

本文章節依序為第二章介紹相關研究，回顧人臉特徵點定位的相關論文。第三章介紹區域性主動表現模型建模，第四章敘述區域性主動表現模型匹配追蹤演算法，第五章為研究結果，第六章則是結論。



第二章

相關研究

許多學者發表過建立一個表現統計模型並對一影像中的物體做匹配搜尋的方法。Nastar 等人 [2] 利用三維可變形亮度模式建立一個形狀與亮度變化的模型，模型維度為XYI，模型建立使用主成分分析法。他們利用最接近點表面匹配演算法(closest point surface matching algorithm)混合光流追蹤法與可變形形狀匹配法執行追蹤。使得在變形物體的影像匹配上更穩定而且更有效率，對於影像中因為遮蔽物遮蔽而被隱藏的區域可以模擬出原本的圖像。因為他們利用最接近點表面匹配演算法，對於初始值的比重佔很大的因素。Lantis 等人 [3] 與 Edwards 等人 [4] 用主動式形狀模型(Active Shape Model)找到最好的形狀，然後利用此形狀去與物體的紋理模型做比對。[3][4]使用標定點分佈模型(Points Distribution Models)提供了一個在任何情況下能完整參數化地敘述形狀，可以應用在不同解析度下，主動式形狀模型(Active Shape Model)搜尋中的形狀定位。Poggio 等人 [5],[6]除了形狀與紋理上的預測，額外預測人臉姿勢的變化情形。使用合成網路模組模擬人臉的姿勢，並根據隨機梯度遞減演算法(stochastic gradient descent algorithm)開發適合合成網路模組應用的合成分析演算法收斂人臉姿勢與人臉形狀紋理資料。Vetter [7] 使用一般的最佳化方法讓高解析度的人臉模型與影像做匹配。其中使用3D人臉模型的線性物體分類法(linear object class approach)產生多方向的人臉，形狀與紋理資訊分開處理，形狀方面使用線性形狀近似調整人臉視角方向，紋理方面使用線性紋理近似加入紋理的資訊。最後兩個方法因為問題維度過高與模型品質過高使得執行速度較慢。在追蹤領域中，快速的模型匹配演算法是一個很熱門的議題，Gleicher [8] 利用一個可變形(利用很多變換如放射，投影等)單一模板，選擇參數使誤差最小，此方法會先計算參數變化時對可變形單一模板的差分方程式。在追蹤上提出誤差分解法(difference decomposition method):最佳化為基礎的方法，對於非線性上的追蹤有較佳的穩定性，對於非線性上的追蹤搜尋的影像區域較小而且可

以分隔較多張後的影像也能穩定追蹤到。並且提出分段式的放射可變形單一模板較單一放射可變形單一模板更能較佳的趨近變形的圖形。Hager 與 Belhumeur [9] 承襲 Gleicher，加上光影變化的模型使穩健性加強，其最佳化以 (sum-of-squared different SSD) 為基礎，優點為可以克服光影變化所造成的追蹤錯誤。Sclaroff 與 Isidoro [10] 使用有限元素低能量法，模型建立方面使用彩色的紋理投影三角形網狀模型(texture-mapped triangular mesh model)，並使用主動式塗抹法(Active Blob Method)，追蹤以彩色為基礎紋裡豐富之物體為目標，此方法已經用來追蹤頭部[11] 使用不變形的圓錐頭部模型。還有很多電腦視覺技術例子結合形狀與紋理的模型然後與新的影像做匹配 [1][12][13][14] T.F. Cootes 等人 [1] 提出的主動式表現模型(Active Appearance Model)演算法，使用合併形狀與紋理的統計模型，接著在追蹤上，使用模型與目標影像的差異去預測模型參數如何改變去的到最佳匹配，實驗上以人臉為追蹤目標。被廣泛應用於各方面。Hou XW 等人 [16] 以主動式表現模型為基礎提出直接式表現模型，直接使用紋理資訊直接預測形狀與紋理與位置，而且在新的人臉位置與紋理上的預測基礎使用模型與目標影像差異的主成分分析。Song Gang 等人 [15] 提出將

Lucas-Kanade 光流跟蹤演算法與人臉特徵點定位的統計模型 DAM(direct appearance model) 在 Bayesian 框架下結合起來，提出了視頻中人臉特徵定位與跟蹤的一種混合模型方法。其中使用兩層直接表現模型 DAM(direct appearance model)，分別為高解析度的直接表現模型與低解析度的直接表現模型，高解析度的直接表現模型對於人臉的細節如五官變化... 等能夠精確的模擬。而低解析度的值皆表現模型能夠較穩定的追蹤人臉的位置。所以若高解析度的直接表現模型若追蹤失敗，會以低解析度的直接模型追蹤找回失去的人臉再交由高解析度的直接表現模型追蹤，增加人臉追蹤的穩定性。D. Cristinacce 等人 [17] 提出樣板選取追蹤演算法(template Selection Tracker TST)，建立一組形狀與紋理資料庫首先以紋理為基礎對資料庫做 Nearest Neighbor Selection，選取最接近的幾個資料，使用 Nelder-Mead 提出的 A simplex method for function minimization 作形狀參數的最佳化，主要為儲存特徵樣板之後以形狀限制搜尋技術收斂形狀參

數定位人臉五官的位置。

本論文以主動式表現模型為基礎，在建立模型部分，將物體中變化較多樣且可分別運作的部份獨立出來(如人臉中的眼睛與嘴巴部分)，分別去做主動式外表模型。在追蹤演算法中，因為執行速度上的要求，最佳化步驟中需要的預測因子 R 是由訓練資料得到。修改其中追蹤的步驟預測物體位置、大小、旋轉角度與形狀與外表。其技術方法在第三章與第四章中有詳細的介紹。



第三章

區域性主動表現模型建模

在本章裡，說明區域性主動表現模型建立模型的方法，主要步驟為取出校正過後(位置、尺寸、角度)的形狀資料與校正(數量、亮度)過後的紋理資料，將形狀與紋理參數分離成眼睛、嘴巴與其他部分三個不同的資料群，接著利用主成份分析法將各部份選取重要資訊，建立模型。

3.1 形狀資訊擷取

D.G. Kendall [18] 在Statistical Shape Analysis.中對於形狀的定義如定義一所示：

定義一： 形狀 是一個物體將位置、大小、旋轉的效應都濾除剩下來所有的幾何資訊。



圖 3-1：四個相同形狀的笑臉，但不同的位置、大小、角度

圖3-1有四個笑臉，由定義一我們可以知道這四個笑臉有相同的形狀，差異只在她們有不同的位置、大小與角度。所以想要單純擷取物體形狀資訊那我們就要濾除位置、大小、角度上的干擾。對一個形狀的描述，我們可以標定一定數目的點在圖形的輪廓周圍，並以那些點的座標值來代表此圖形的形狀。而對一個n個點，k維度的形狀，可將每個維度結合成一個kxn的向量。對一個平面形狀的描述將會是：

$$X = [x_1, y_1, x_2, y_2, \dots, x_n, y_n]^T \quad (3.1)$$

由定義一可知，我們在擷取形狀資訊，必須將位置、大小、旋轉因素消除，才能

真正算是形狀的資訊，所以我們由一張影像中擷取出標定點後還需對位置、大小、旋轉因素做調整的動作。

我們首先定義形狀資料點的分布，人臉的特徵主要來自於眉毛、眼睛、鼻子、嘴巴與人臉輪廓。所以我們將這些區域以58個資料點標定出來。如圖3-2（左）為我們定義的形狀資料點分布，圖3-2(右)為實際手工標定的結果。

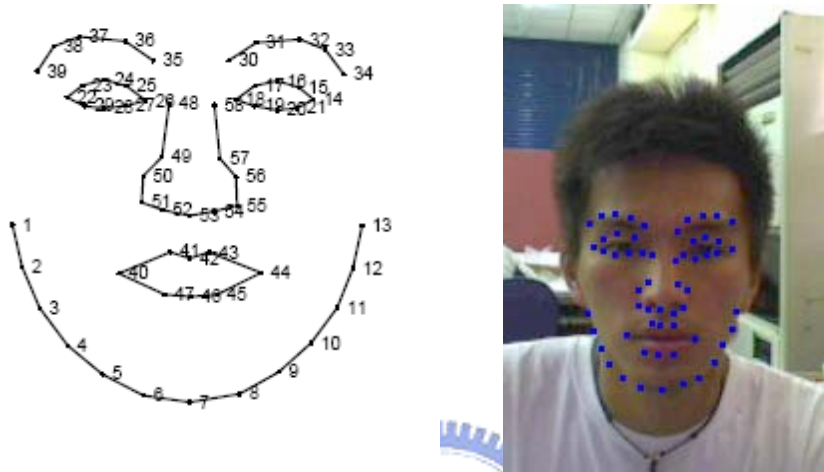


圖. 3-2: (左)定義的形狀資料點分布 (右)實際手工標定的結果

3.1.1 形狀校準

執行形狀校準時，依序為這四個步驟：

1. 計算每個形狀的中心點。
2. 調整每個形狀至一定大小。
3. 調整原點座標至形狀的中心點。
4. 校正角度誤差。

有關中心點的計算如下：

$$(\bar{x} - \bar{y}) = \left(\frac{1}{n} \sum_{j=1}^n x_j, \frac{1}{n} \sum_{j=1}^n y_j \right) \quad (3.2)$$

為了要進行第二個步驟，我們需要建立一個形狀尺寸的量測標準。

定義二：形狀尺寸的量測標準 $S(X)$ 一個形狀向量帶入是一個恆正實數的函數，而且符合下面的特性：

$$S(aX) = aS(X) \quad (3.3)$$

Frobenius norm 用來當作形狀尺寸的量測標準：

$$S(X) = \sqrt{\sum_{j=1}^n [(x_j - \bar{x})^2 + (y_j - \bar{y})^2]} \quad (3.4)$$

另一種狀尺寸的量測標準：

$$S(X) = \sum_{j=1}^n \sqrt{[(x_j - \bar{x})^2 + (y_j - \bar{y})^2]} \quad (3.5)$$

第四步驟使用奇異值分解(Singular Value Decomposition : SVD)：

1. 將已經完成尺寸與位置校正的兩個形狀 X_1 與 X_2 轉換成 $n \times k$ 的矩陣(在平面形狀中 $k=2$)
2. 計算 $X_1^T X_2$ 的 SVD, UDV^T
3. 用來根據 X_2 調整 X_1 的旋轉矩陣就是 VU^T

$$VU^T = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \quad (3.6)$$

圖 3-3 中綠色資料點代表眉毛形狀、紅色資料點代表眼睛形狀、淺藍色資料點代表鼻子形狀、黃色資料點代表嘴巴形狀、藍色資料點代表人臉輪廓形狀。圖 3-3(左)顯示訓練資料的形狀資訊在未處理前的資料分佈情形，其分布點非常雜亂，許多區域的分布相互重疊，如：眼睛與眉毛區域、眼睛與鼻子區域、鼻子與嘴巴區域、嘴巴下半與人臉輪廓區域。經由上面所述執行過位置、角度、大小的校正後，結果如圖 3-3(右)所示各點的分布情形很清楚。



圖. 3-3: (左)未處理的形狀資訊 (右)校正過後的形狀資訊

3.2 紋理校正

以下介紹擷取紋理資訊的方法，將形狀標定點作三角形化，由每個三角形取出固定量的紋理資訊點數，對亮度做調整。

3.2.1 紋理資料點校準

由於要用統計模型製作人臉模型，人臉形狀與紋理資料數目大小必須固定。一般普遍使用的方法即為將人臉周圍用一個四方形定義出範圍，利用一個遮罩將周圍背景資訊錄除，再將四方形依照某個一定比例正規化，即可得到一定的紋理資料點數。(如圖 3-6 所示)。此方法的缺點是：1. 背景資料不易完全去除 2. 人臉的形狀不同，有長形臉與寬型臉的分別，直接對人臉的框做正規化人臉會有所變形 3. 隨著人臉表情的不同，一些區域(如嘴巴的張開與閉合)的總紋理資料數目也會不同，這會造成之後建立模型時產生資訊的重疊而降低模型的精準度。所以我們提出另一種紋理資料點數目校正的方法。形狀上的資料量在定義形狀參數時即可固定(數目為 $n \times k$)，我們將人臉標定點定義為 58 個如圖 3-2(左)所示。而紋理資料會因為人臉形狀上的不同再同一區域中其灰階值數目不盡相同(如：嘴巴張開與閉合，三角臉與國字臉...等的差異，同一區域中灰階值數目不相同)，所以要做紋理資料擷取數目上的校正。我們需要利用形狀上的資訊，由於形狀標定點內部即為我們需要的紋理資訊，將形狀標定點作三角形化如圖 3-2(右)所示，得到一個人臉的三角形化網狀圖。如此以來，對於不同的人臉我們都可以得到相同數量的三角形，再利用一個一般表情的人臉來定義出每個三角形內部應該取出紋理資訊的點數與相對位置。即可將人臉紋理特徵點大小值確定。

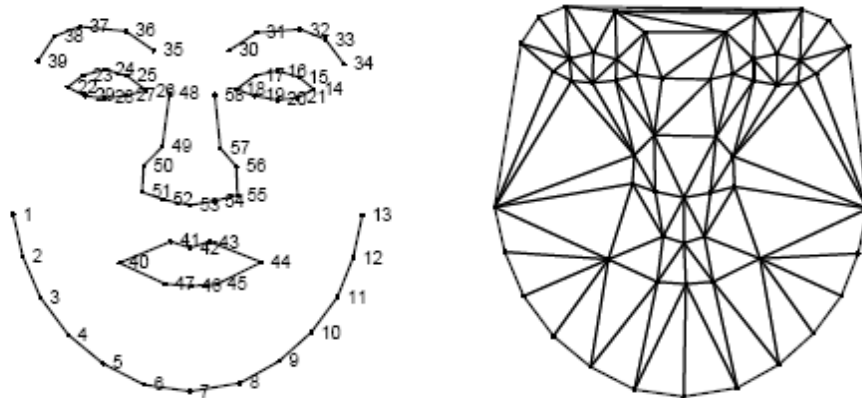


圖. 3-4：臉部標定點(左) 人臉三角形化網狀圖(右)

兩個人臉經由三角形演算法三角形化後，相互對應的三角形取出紋理資料點的方法如下，如圖 3-3 所示，若已知 A1、A2、A3、A4、B1、B2、B3 的座標值要求得 B4 座標值的方法為先計算式子(3.7)中的 a、b 值再帶入式子(3.8)中即可得到 B4 座標值。圖 3-4 為紋裡資料的擷取成果。

$$\begin{bmatrix} x_{A2} - x_{A1} & x_{A3} - x_{A1} \\ y_{A2} - y_{A1} & y_{A3} - y_{A1} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x_{A4} - x_{A1} \\ y_{A4} - y_{A1} \end{bmatrix} \quad (3.7)$$

$$\begin{bmatrix} x_{B4} \\ y_{B4} \end{bmatrix} = \begin{bmatrix} x_{B2} - x_{B1} & x_{B3} - x_{B1} \\ y_{B2} - y_{B1} & y_{B3} - y_{B1} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} x_{B1} \\ y_{B1} \end{bmatrix} \quad (3.8)$$

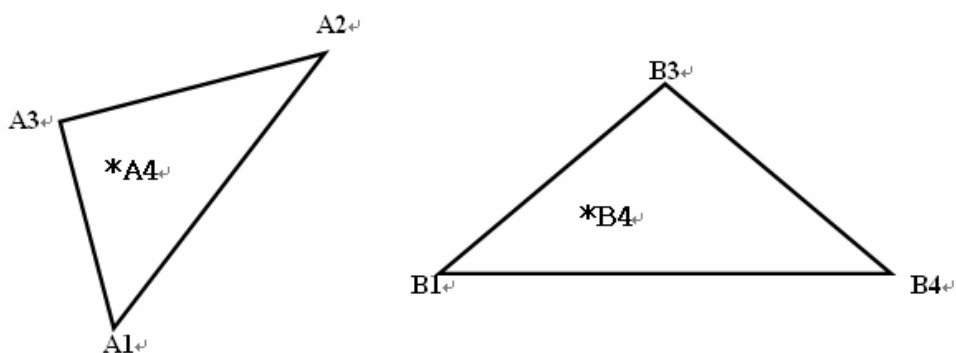


圖 3-5：三角形之間的轉換



圖 3-6：三角形化紋理資訊擷取



3.2.2 紋理資料點亮度校準

得到紋理資訊後，為了使模型更穩定，不易受亮度的影響，會對其值做調整，假設調整前紋理資訊為 g_b ，調整後紋理資訊為 g_f ，調整方法為：

$$g_f = g_b - \bar{g}_b + 100 \quad \text{if } g_{fi} > 255 \quad g_{fi} = 255 \quad (3.9)$$

$$\text{if } g_{fi} < 0 \quad g_{fi} = 0$$

g_{fi} 代表 g_f 向量內各點的資訊。

圖 3-7 正中央為建立模型的資料之一，模型建立後，若將資料亮度做整體上的調整，送入模型重新我們會發現亮度差異越大，模型模擬出來的結果會越差。所以雖然亮度調整為一個簡單的動作但是，對模型好壞的影響很大。

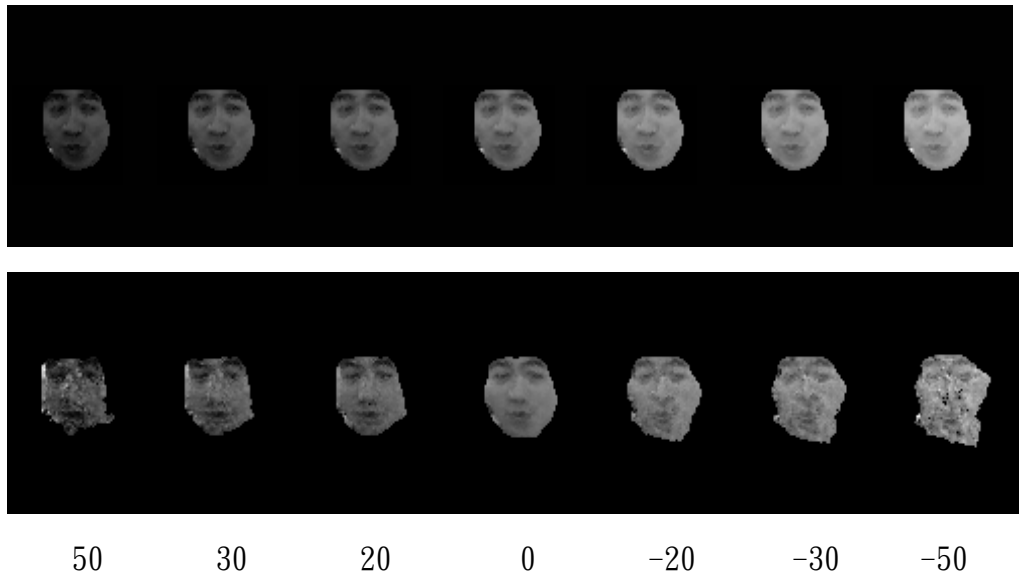


圖 3-7：亮度對模型的影響：(第一列)輸入圖形 (第二列)模擬結果

3.3 區域性主動表現模型建模



3.3.1 主成分分析

主成份分析法用於資料的分析已經有一段時間了。主要是用來將多個變量經過線性變換以選出較少個數重要變量的方法。其中每個變量在數學上的意義 j 為與其他變量正交，也就是不具有交互相關性。於是，我們可以用此方法將原本包含大量變數的資料用較少的主成份所表示，不但可以達到降維的作用，同時也能對於一些雜訊（比較不重要的成份）有所過濾。

使用主成份分析的方法，事實上就是在利用奇異值分解法(Singular Value Decomposition) 解出資料的相關變異矩陣(Covariance Matrix) 的特徵值與特徵向量。參考資料[20]介紹基本的數學概念。假設有一個一維的矩陣 X ，計算它的相關變異就等於計算它的自身變異量。

$$\text{cov}(X) = \text{var}(X) = \frac{\sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})}{(n-1)} \quad (3.10)$$

在此， X_i 表示矩陣 X 的第 i 列，且

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad (3.11)$$

於是，此矩陣的自身的相關變異矩陣為

$$C^{n \times n} = (c_{i,j}, c_{i,j} = \text{cov}(\text{Dim}_i, \text{Dim}_j)) \quad (3.12)$$

在此， $C^{n \times n}$ 表示這是一個具有 n 行跟 n 列的矩陣。 Dim_x 表示第 x 維。

一旦我們得到了一個方形的相關變異矩陣後，即可計算其特徵值與特徵向量。這些特徵向量彼此互相垂直，但更重要的，它們同時說明了資料量的一些資訊。我們將這些這些特徵向量與其所對應的特徵值按照大小順序排列，較大的特徵值所對應的特徵向量包含了資料集中較為重要的資訊。於是，我們可以捨棄掉一些包含比較少資訊，也就是較不重要的成份，並且不會失去原資料集的大部分資訊。假設一個資料集最初為 n 維，於是計算出 n 個特徵值與特徵向量。若是我們只選取 p 個特徵向量做為主成份，則原本 n 維的資料，經過處理後，可用 p 維來表示。

3.3.2 區域性主動表現模型建模

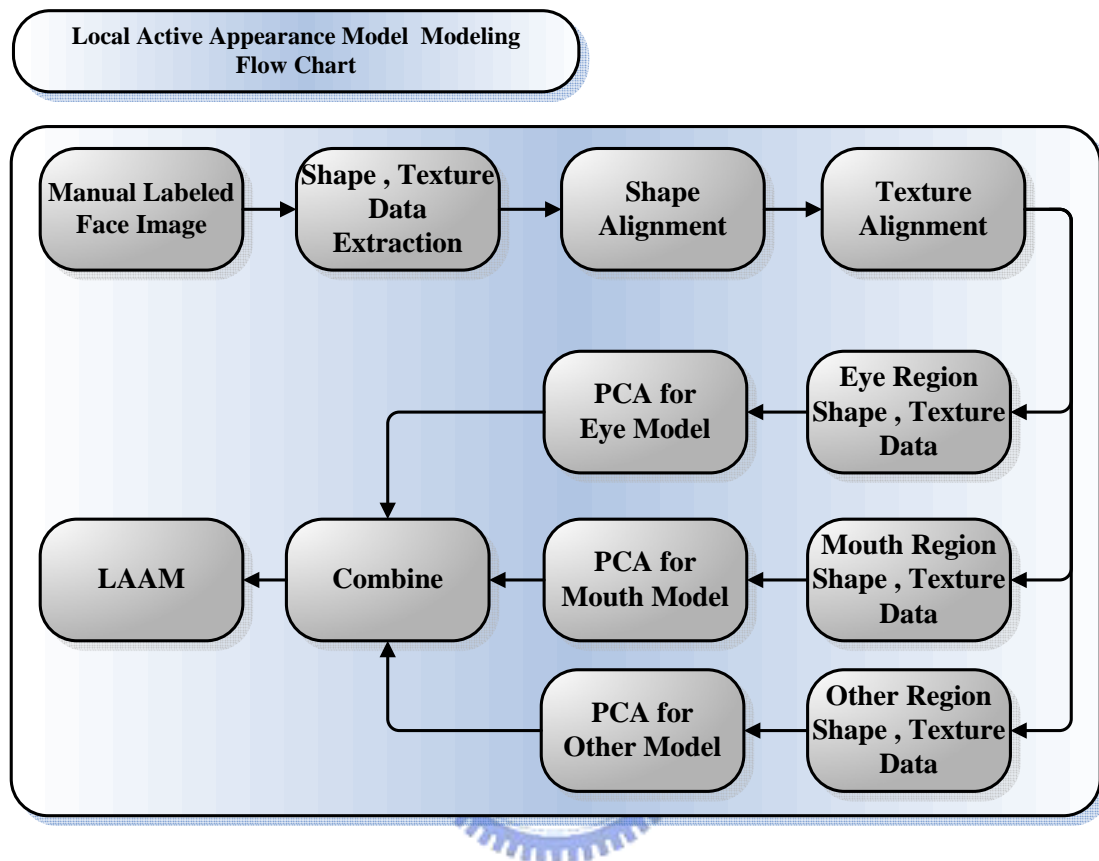


圖 3-8：區域式主動表現模型流程圖



圖 3-9：手動標定影像取出形狀與紋理資訊

區域主動表現模型的模型建立方法流程圖如圖 3-9 所示，由手工標定的人臉圖形中擷取形狀與紋理資訊，對形狀與紋理資訊做校準，人臉變化性較大的部分為嘴巴與眼睛，而且其動作為獨立的，所以我在再將人臉整體形狀與紋理的資訊分成三個部份，再利用主成分分析模擬出眼睛、嘴巴、其他部份的模型

以下簡述本實驗所使用的方法：

先從手動標定的影像取出人臉整體形狀與紋理的資訊(如圖 3-9 所示)，我們在直覺上的判斷，人臉變化性較大的部分為嘴巴與眼睛，而且其動作趨近於獨立的，所以我在再將人臉整體形狀與紋理的資訊分成三個部份，分別是眼睛部份的形狀與紋理的資訊，嘴巴部份的形狀與紋理的資訊，其它部份的形狀與紋理的資

訊，這三個區塊(如圖 3-10 所示)，再利用主成分分析法分別對這三個區域計算特徵值與特徵向量挑選。

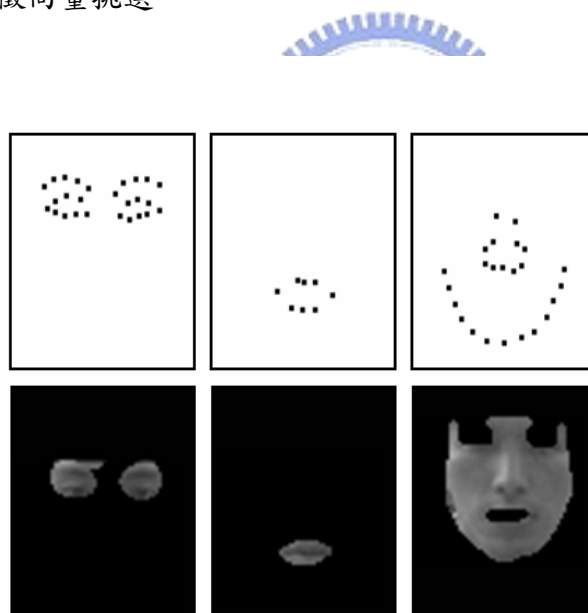


圖 3-10：眼部、口、其它部分三個區域的形狀與紋理資訊

1. 取得 M 張訓練影像，先由手動標的地影像將人臉整體形狀與紋理的資訊取出(如圖 3-11 所示)
2. 假設形狀資訊 x : 維度 $(2B, 1)$ ，紋理資訊 g : 維度 $(A, 1)$ ，結合排列為 $N \times 1$ 的向量 Γ_i ($=[\Gamma_{i1}, \Gamma_{i2}, \dots, \Gamma_{iN}]^T$)，因為有 M 張影像，則會產生一個 $N \times M$ 的矩

陣。

3. 計算平均向量 Ψ 與標準差 σ

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad , \quad \sigma_k = \sqrt{\frac{1}{M} \sum_{i=1}^M (\Gamma_{ik} - \Psi_k)^2} \quad (3.13)$$

4. 將每個 Γ_i 與平均向量 Ψ 與標準差 σ 做正規化，得到的結果為 Φ_i

5. 計算其相關變異矩陣 C

$$C = \frac{1}{M} \sum_{n=1}^M \Phi_n \Phi_n^T = AA^T \quad (N^2 \times N^2 \text{ matrix}) \quad (3.14)$$

where $A = [\Phi_1 \Phi_2 \dots \Phi_M]$ ($N^2 \times M$ matrix)

6. 計算 AA^T 的特徵向量 u_i ，基本上會產生 N^2 個特徵向量，將其從對應的特徵值由大到小排列。

7. 保留對應前 K 個最大的特徵值的特徵向量。

8. 將訓練影像以這些特徵向量為基底加以表示：

每張影像經過正規化之後的 Φ_i ，可以此 K 個特徵向量作一線性組合來表示。

$$\Phi_i = \sum_{j=1}^K w_j u_j, \quad (w_j = \Phi_i^T u_j) \quad (3.15)$$

將三者(眼睛部份、嘴巴部分、其他部分的形狀與紋理)，以 1.~8. 的步驟分別製作出各自的模型，再將三者合成，即為我們所要介紹的區域式主動表現模型

區域式主動表現模型也可以轉換成以下表示法，

x_{eye} 、 x_{mouth} 、 x_{other} ：眼睛、嘴巴、其它部份的形狀資訊。

\bar{x}_{eye} 、 \bar{x}_{mouth} 、 \bar{x}_{other} ：眼睛、嘴巴、其它部份的形狀資訊平均。

g_{eye} 、 g_{mouth} 、 g_{other} ：眼睛、嘴巴、其它部份的紋理資訊。

\bar{g}_{eye} 、 \bar{g}_{mouth} 、 \bar{g}_{other} ：眼睛、嘴巴、其它部份的紋理資訊平均。

Q_{s_eye} 、 Q_{s_mouth} 、 Q_{s_other} ：眼睛、嘴巴、其它部份有關形狀的特徵向量。

Q_{g_eye} 、 Q_{g_mouth} 、 Q_{g_other} ：眼睛、嘴巴、其它部份有關紋理的特徵向量。

c_{eye} 、 c_{mouth} 、 c_{other} ：眼睛、嘴巴、其它部份模型的參數，主成分分析模擬的係數。

$$\begin{aligned} x_{eye} &= \bar{x}_{eye} + Q_{s_eye} c_{eye} \\ g_{eye} &= \bar{g}_{eye} + Q_{g_eye} c_{eye} \end{aligned} \quad (3.16)$$

$$\begin{aligned} x_{mouth} &= \bar{x}_{mouth} + Q_{s_mouth} c_{mouth} \\ g_{mouth} &= \bar{g}_{mouth} + Q_{g_mouth} c_{mouth} \end{aligned} \quad (3.17)$$

$$\begin{aligned} x_{other} &= \bar{x}_{other} + Q_{s_other} c_{other} \\ g_{other} &= \bar{g}_{other} + Q_{g_other} c_{other} \end{aligned} \quad (3.18)$$

如圖 3-13、圖 3-14、圖 3-15、所示，分別改變區域式主動表現模型的 c_{eye} 、 c_{mouth} 、 c_{other} ，取各自的前四個模型參數各自 $\pm 3\sigma$ 。圖 3-15 為主動表現模型，我們做個比較對照。

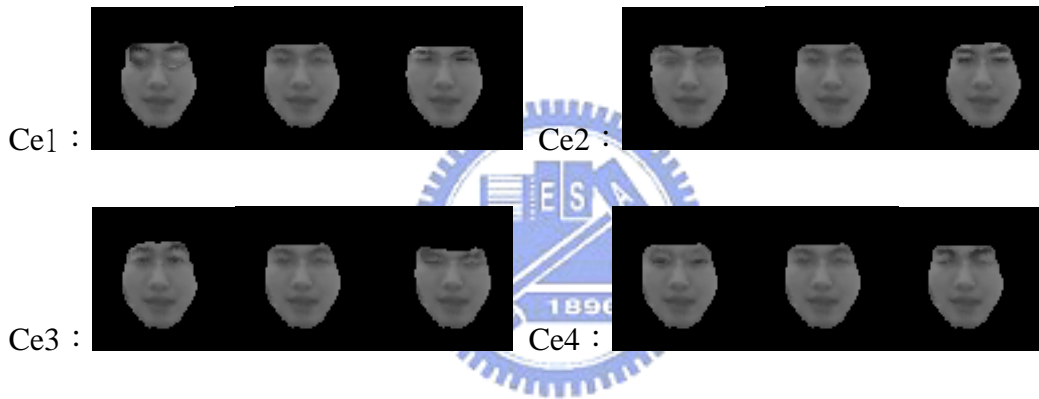


圖 3-11：區域式主動表現模型眼睛區域前四個參數 $c_{eye} \pm 3\sigma_{eye}$ 的影響

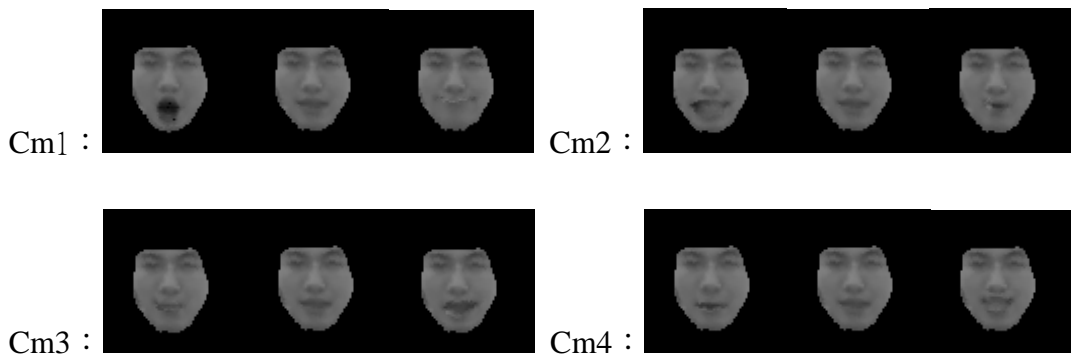


圖 3-12：區域式主動表現模型嘴巴區域前四個參數 $c_{mouth} \pm 3\sigma_{mouth}$ 的影響

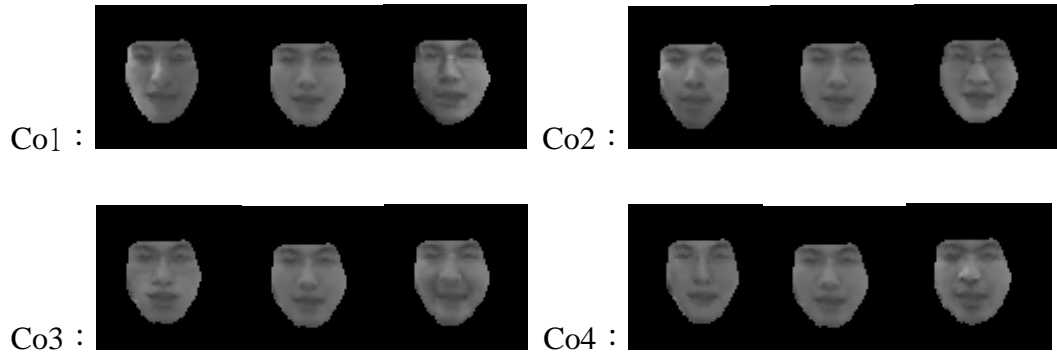


圖 3-13：區域式主動表現模型其他區域前四個參數 $c_{other} \pm 3\sigma_{other}$ 的影響

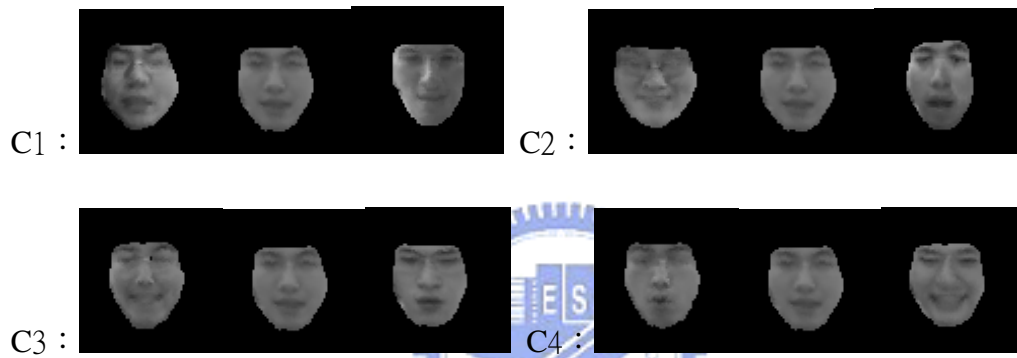


圖 3-14：(比較)主動表現模型前四個參數 $c \pm 3\sigma$ 的影響

第四章

區域式主動表現模型定位演算法

現在有 1. 一張新的人臉影像作人臉形狀定位與紋理模擬，2. 一個區域式主動表現模型如上一章所敘述，3. 大概的人臉位置、角度、大小資訊。其中大概的人臉位置、角度、大小資訊是模型放置的初始點。我們需要一個有效率的演算法去調整模型參數 c_{eye} 、 c_{mouth} 、 c_{other} 與位移參數 t ， $t = (t_x, t_y, \theta, s)^T$ 。 t_x ：人臉 X 方向位置、 t_y ：人臉 y 方向位置、 θ ：人臉旋轉角度、 s ：人臉大小。對需要特徵定位的人臉影像做最佳的匹配。

4.1 人臉影像差異

$$x_{eye} = \bar{x}_{eye} + Q_{s_eye} c_{eye} \quad (3.16)$$

$$g_{eye} = \bar{g}_{eye} + Q_{g_eye} c_{eye}$$

$$x_{mouth} = \bar{x}_{mouth} + Q_{s_mouth} c_{mouth} \quad (3.17)$$

$$g_{mouth} = \bar{g}_{mouth} + Q_{g_mouth} c_{mouth}$$

$$x_{other} = \bar{x}_{other} + Q_{s_other} c_{other} \quad (3.18)$$

$$g_{other} = \bar{g}_{other} + Q_{g_other} c_{other}$$

我們現在重新定義問題，有一張新的人臉影像需要作人臉形狀定位與紋理模擬，已經根據手工標定的訓練資料製作出區域式主動表現模型(如第三章所敘述)，先利用人臉偵測程式得到大概的人臉位置、角度、大小資訊。其中大概的人臉位置、角度、大小資訊是模型放置的初始點。已知大概人臉位置、角度、大小資訊，將模型初始值根據人臉大概尺寸調整其大小，放入初始位置中，如圖 1-1 中的 initial 所示，目前所要求得的參數為區域式表現模型的參數 c_{eye} 、 c_{mouth} 、 c_{other} ，與位移參數 t ， $t = (t_x, t_y, \theta, s)^T$ 。 t_x ：人臉 X 方向位置、 t_y ：人臉 y 方向位置、 θ ：人臉旋轉角度、 s ：人臉大小。在位移參數 t 上的預測，我們以 $x_{total} (x_{eye} + x_{mouth} + x_{other})$ 在人臉影像中取得所有人臉紋理 g_{s_total} (眼睛, 嘴巴與其

他區域紋理資訊的總和)與模型模擬之所有人臉紋理 g_{m_total} ($g_{eye} + g_{mouth} + g_{other}$) 相減的差異向量 $r_{total}(t)$ (4.1) 做預測分析。 c_{eye} 使用以 x_{eye} 在人臉影像中取得的眼晴區域紋理資訊 g_{s_eye} 與模型模擬之眼睛區域紋理資訊 g_{m_eye} 的差異向量 $r_{eye}(c_{eye})$ (4.2) 來做預測分析。 c_{mouth} 使用以 x_{mouth} 在人臉影像中取得的嘴巴區域紋理資訊 g_{s_mouth} 與模型模擬之嘴巴區域紋理資訊 g_{m_mouth} 的差異向量 $r_{mouth}(c_{mouth})$ (4.3) 來做預測分析。 c_{other} 使用以 x_{other} 在的人臉影像中取得的眼晴區域紋理資訊 g_{s_other} 與模型模擬之眼睛區域紋理資訊 g_{m_other} 的差異向量 $r_{other}(c_{other})$ (4.4) 來做預測分析。



Re：圖 1.1：區域式主動表現模型定位結果

模型與影像的紋理差異向量 (須做正規化) 表示法如下：

$$\text{整體人臉差異向量： } r_{total}(t) = g_{s_total} - g_{m_total} \quad (4.1)$$

$$\text{眼睛區域差異向量： } r_{eye}(c_{eye}) = g_{s_eye} - g_{m_eye} \quad (4.2)$$

$$\text{嘴巴區域差異向量： } r_{mouth}(c_{mouth}) = g_{s_mouth} - g_{m_mouth} \quad (4.3)$$

$$\text{其他區域差異向量： } r_{other}(c_{other}) = g_{s_other} - g_{m_other} \quad (4.4)$$

錯誤的量測使用平方相加來預測：

$$\text{整體人臉錯誤： } E_{total}(t) = r_{total}^T r_{total} \quad (4.5)$$

$$\text{眼睛區域錯誤： } E_{eye}(c_{eye}) = r_{eye}^T r_{eye} \quad (4.6)$$

$$\text{嘴巴區域錯誤： } E_{mouth}(c_{mouth}) = r_{mouth}^T r_{mouth} \quad (4.7)$$

$$\text{其他區域錯誤： } E_{other}(c_{other}) = r_{other}^T r_{other} \quad (4.8)$$

$$\text{差異向量一般表示： } r(p) = g_s - g_m \quad ; \quad p : t, c_{eye}, c_{mouth}, c_{other} \quad (4.9)$$

$$\text{錯誤一般表示： } E(p) = r^T r \quad p: t, c_{eye}, c_{mouth}, c_{other} \quad (4.10)$$

(4.9)式的一階泰勒展開式為

$$r(p + \delta p) = r(p) + \frac{\partial r}{\partial p} \delta p \quad (4.11)$$

在進行定位時，兩個紋理資訊差為 r ，我們希望能夠選擇適當的 δp 調整模型與位移參數，使得 $|r(p + \delta p)|^2$ 最小(模型紋理資訊與圖片紋理資訊差異最小)，將 $r(p + \delta p)$ 令為零，經推導我們可以得到結果：

$$\delta p = -Rr(p) \text{ Where } R = \left(\frac{\partial r^T}{\partial p} \frac{\partial r}{\partial p} \right)^{-1} \frac{\partial r^T}{\partial p} \quad (4.11)$$

R 中的 $\partial r / \partial p$ 由差分方程式可以計算出來，其中 $\partial r / \partial p$ 為一個 $m \times n$ 的矩陣如式子(4.12 所示)，m 代表區域式主動表現模型的紋理維度，n 代表模型控制參數個數。由式子(4.11)若有模型紋理 g_m 與形狀資訊 x 定義之影像紋理 g_s 的差異向量 $r(p) = g_s - g_m$ ，即可計算出 δp 調整模型與位移參數，使得 $|r(p + \delta p)|$ 最小，也使得模型特徵收斂到正確的位置與正確的紋理資訊。在實際的最佳化過程中，必須每次都去計算它的 $\partial r / \partial p$ ，這會很花費時間，影響到執行的速度。因為紋理上的差異會做正規化的動作。所以我們可以假設 $\partial r / \partial p$ 這個微分式對於追蹤人臉而言趨近一個定值。所以我們可以先利用訓練資料事先計算。

$$\frac{\partial r}{\partial p} = \begin{bmatrix} \frac{\partial r_1}{\partial p_1} & \frac{\partial r_1}{\partial p_2} & \dots & \frac{\partial r_1}{\partial p_n} \\ \frac{\partial r_2}{\partial p_1} & \frac{\partial r_2}{\partial p_2} & \dots & \frac{\partial r_2}{\partial p_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial r_m}{\partial p_1} & \frac{\partial r_m}{\partial p_2} & \dots & \frac{\partial r_m}{\partial p_n} \end{bmatrix} \quad (4.12)$$

$\partial r / \partial p$ 的訓練過程如下：以手工標定的影像為基礎取出影像的紋理資訊 g_s ，計算此張圖所對應的模型參數 c_{eye} 、 c_{mouth} 、 c_{other} 、位移參數 t ， $t = (t_x, t_y, \theta, s)^T$ 並計算對應的模型紋理資訊 g_m 即可計算出差異向量 $r(p) = g_s - g_m$ ，以此數值當作調整的基本值，一次調整一個模型參數或位移參數 p_i ，調整的大小 δp_i 【1】就模型參數而言，先計算所有訓練資料對應的模型參數 c_{eye} 、 c_{mouth} 、 c_{other} 再統計

各別的標準差 σ_{eye} 、 σ_{mouth} 、 σ_{other} ，以標準差 σ_{eye} 、 σ_{mouth} 、 σ_{other} 為參考在極限值 $\pm 0.5 \times \sigma$ 中分割 10 個等分 ($-0.5 \times \sigma$ 、 $-0.4 \times \sigma$ 、 $-0.3 \times \sigma \cdots 0.3 \times \sigma$ 、 $0.4 \times \sigma$ 、 $0.5 \times \sigma$) 以此當作 δp_i 值的改變基準，調整過後可由式(3.16)~(3.18)計算出相對應的模型與影像紋理資訊 g'_m 、 g'_s (在此模型紋理的變異較大) 即可計算出新的差異向量 $r'(p) = g'_s - g'_m$ ，將新的差異向量 $r'(p)$ 與原來的差異向量 $r(p)$ 相減，即可得到 ∂r_i ， δp_i 為我們設定的值， ∂r_i 可計算出來，就可以得到 $\partial r_j / \partial p_i$ $i = 1, 2, 3, \dots, n$ ， $j = 1, 2, 3, \dots, m$ ， m 代表區域式主動表現模型的紋理維度， n 代表模型控制參數個數。 δp_i 有十個數值調整，所以可以計算出十個 $\partial r_j / \partial p_i$ 取其平均為 $\partial r_j / \partial p_i$ 訓練的結果。【2】就位移參數而言定義一個區間也是分割成 10 個等分，此區間的大小為經驗法則，就我們嘗試結果此區間在人臉大小約 60×60 時此區間大小約正負 10 個像素。以此當作 δp_i 的調整基準，調整位移參數 $t = (t_x, t_y, \theta, s)^T$ ，可由原本的形狀資料 x 調整為 x' 。由 x' 定義之新的形狀可以擷取出新的影像紋理資訊 g'_s ，可計算出新的差異向量 $r'(p) = g'_s - g_m$ 將新的差異向量 $r'(p)$ 與原來的差異向量 $r(p)$ 相減，即可得到 ∂r_i ， δp_i 為我們設定的值， ∂r_i 可計算出來，就可以得到 $\partial r_j / \partial p_i$ ，由於 δp_i 也有十個數值調整所以以其平均代表其數值，得到 $\partial r / \partial p$ 即可由式子(4.11)計算 R ， R 就可用在實際定位中用來預測參數的走向。就區域主動表現模型預測因子有 R_t (維度：整體紋理資料數量 \times 位移參數數量)、 R_{eye} (維度：眼睛區域紋理資料數量 \times 眼睛區域模型參數數量)、 R_{mouth} (維度：嘴巴區域紋理資料數量 \times 嘴巴區域模型參數數量)、 R_{other} (維度：其他區域紋理資料數量 \times 其他區域模型參數數量)

$$\delta t = -R_t r_{total}(t) \quad \text{Where } R_t = \left(\frac{\partial r_{total}}{\partial t} \right)^T \frac{\partial r_{total}}{\partial t}^{-1} \frac{\partial r_{total}}{\partial t} \quad (4.13)$$

$$\delta c_{eye} = -R_{eye} r_{eye}(c_{eye}) \quad \text{Where } R_{eye} = \left(\frac{\partial r_{eye}}{\partial c_{eye}} \right)^T \frac{\partial r_{eye}}{\partial c_{eye}}^{-1} \frac{\partial r_{eye}}{\partial c_{eye}} \quad (4.14)$$

$$\delta c_{mouth} = -R_{mouth} r_{mouth}(c_{mouth}) \quad \text{Where } R_{mouth} = \left(\frac{\partial r_{mouth}}{\partial c_{mouth}} \right)^T \frac{\partial r_{mouth}}{\partial c_{mouth}}^{-1} \frac{\partial r_{mouth}}{\partial c_{mouth}} \quad (4.15)$$

$$\delta c_{other} = -R_{other} r_{other}(c_{other}) \quad \text{Where } R_{other} = \left(\frac{\partial r_{other}}{\partial c_{other}} \right)^T \frac{\partial r_{other}}{\partial c_{other}}^{-1} \frac{\partial r_{other}}{\partial c_{other}} \quad (4.16)$$

式子(4.13)為位移參數 $t = (t_x, t_y, \theta, s)^T$ 的預測方法，已知預測因子 R_t 與整體的差異向量 $r_{total}(t)$ 即可求得位移參數的改變值。式子(4.14)為眼睛區域模型參數 c_{eye} 的預測方法，以知預測因子 R_{eye} 與眼睛區域的差異向量 $r_{eye}(c_{eye})$ 即可求得眼睛區域模型參數的改變值。式子(4.15)為嘴巴區域模型參數 c_{mouth} 的預測方法，以知預測因子 R_{mouth} 與眼睛區域的差異向量 $r_{mouth}(c_{mouth})$ 即可求得嘴巴區域模型參數的改變值。式子(4.16)為其他區域模型參數 c_{other} 的預測方法，以知預測因子 R_{other} 與眼睛區域的差異向量 $r_{other}(c_{other})$ 即可求得其他區域模型參數的改變值。

LAAM - training parameter Predictor R

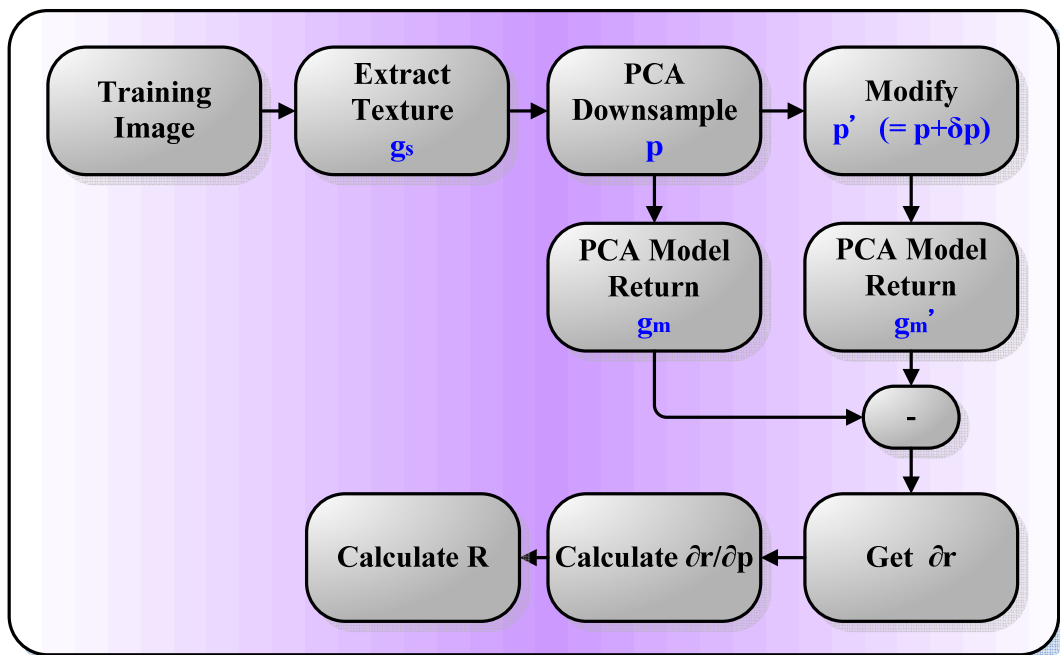


圖 4-1：區域式主動表現模型 - 模型預測因子 R 的訓練流程圖

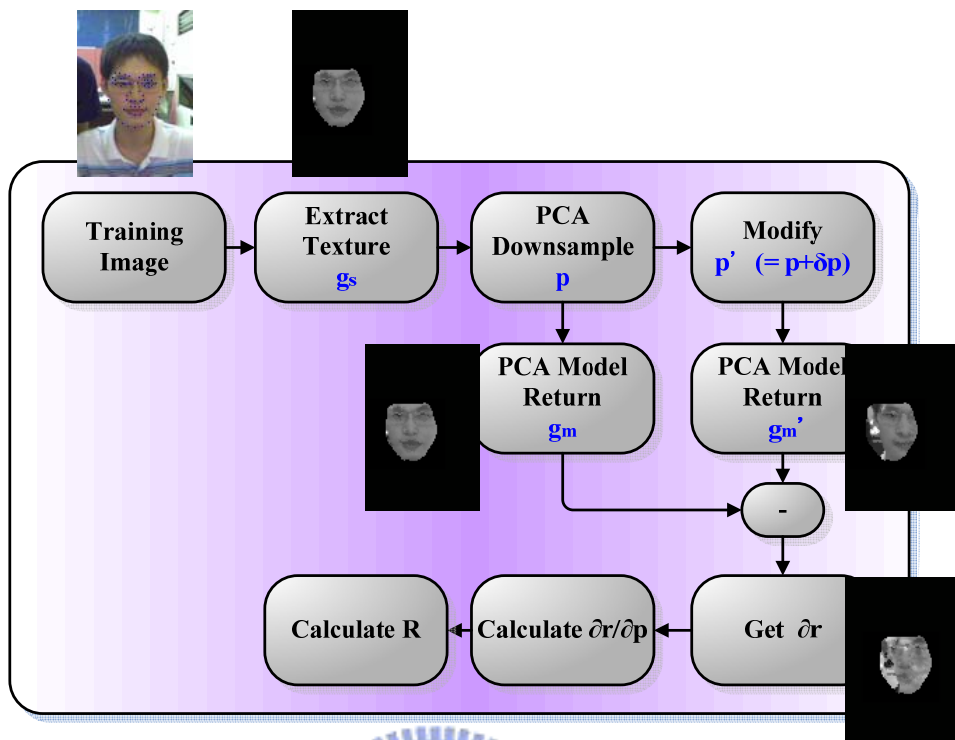


圖 4-2：區域式主動表現模型 - 模型預測因子 R 的訓練流程圖(二)

圖 4.1、4.2 為區域式主動表現模型中預測因子 R 的訓練流程圖，圖 4.2 增加了實際訓練時的影像以便於更能了解整個模型預測因子 R 的訓練流程。圖 4.2 中由手工標定訓練影像擷取紋理資訊 g_s ， g_s 經由區域主動表現模型的主成分分析維度轉換為模型參數 $c_{eye}, c_{mouth}, c_{other}$ ，在此之後分為兩個步驟，第一為利用 $c_{eye}, c_{mouth}, c_{other}$ 計算區域主動表現模型的模型紋理資訊 g_m ，第二為調整模型參數或位移參數 $p : t, c_{eye}, c_{mouth}, c_{other}$ 得到 p' 【1】若 p' 為位移參數 t ($p' = t$)，則根據位移參數的改變而由形狀參數 x 改變為 x' 所對應出新的影像紋理資訊 g'_s 、【2】若 p' 為模型參數 $c_{eye}, c_{mouth}, c_{other}$ ($p : t, c_{eye}, c_{mouth}, c_{other}$)，則根據改變的模型參數 $c'_{eye}, c'_{mouth}, c'_{other}$ 計算區域主動表現模型的模型紋理資訊 g'_m ，將 g'_m 或 g'_s 與 g_m 相減可以得到 r' 與原本的 r (因為為初始為訓練資料轉換出來所以初始的 $r \rightarrow 0$) 相減可以得到 δr 且 δp 已知及可計算出 $\partial r / \partial p$ 在經由式子(4.11)我們就可以訓練出預測因子 R，圖 4.2 中的人形由右到左由上到下依序為：手工標定之訓練影像、由訓練影像擷取出的紋理資訊、模型模擬訓練影像的紋理資訊、模型或位移參數

調整過後的紋理資訊、前兩者相減後的資訊。

4.2 模型更新演算法

得到由 4.1 中所訓練出的參數預測因子 R_t 、 R_{eye} 、 R_{mouth} 、 R_{other} ，我們可以利用式子 (4.13) $\delta t = -R_t r_{total}(t)$ 、(4.14) $\delta c_{eye} = -R_{eye} r_{eye}(c_{eye})$ 、(4.15) $\delta c_{mouth} = -R_{mouth} r_{mouth}(c_{mouth})$ 、(4.16) $\delta c_{other} = -R_{other} r_{other}(c_{other})$ ，建立一個疊代式的演算法作最佳化的收斂來更新參數 $t, c_{eye}, c_{mouth}, c_{other}$ 使其收斂到最佳值。收斂之演算法為：

1. 由 X 的定位擷取影像中紋理資訊 g_{s_eye} 、 g_{s_mouth} 、 g_{s_other}
2. 計算錯誤向量： $r = g_s - g_m$ ： $r(t)$ 、 $r_{eye}(c_{eye})$ 、 $r_{mouth}(c_{mouth})$ 、 $r_{other}(c_{other})$ ，與計算目前的錯誤： $E = |r|^2$ ： E_{total} 、 E_{eye} 、 E_{moth} 、 E_{other}
3. 利用 $\delta t = -R_t r_{total}(t)$ 、 $\delta c_{eye} = -R_{eye} r_{eye}(c_{eye})$ 、 $\delta c_{mouth} = -R_{mouth} r_{mouth}(c_{mouth})$ 、 $\delta c_{other} = -R_{other} r_{other}(c_{other})$ 計算 δt 、 δc_{eye} 、 δc_{mouth} 、 δc_{other}

4. 更新參數

$$t \rightarrow t + k\delta t$$

$$c_{eye} \rightarrow c_{eye} + k_{eye}\delta c_{eye}$$

$$c_{mouth} \rightarrow c_{mouth} + k_{mouth}\delta c_{mouth}$$

$$c_{other} \rightarrow c_{other} + k_{other}\delta c_{other}$$

$$\text{初始值：} k = k_{eye} = k_{mouth} = k_{other} = 1$$

5. 計算新的形狀資訊 X' 與新的紋理資訊 g'_{m_eye} 、 g'_{m_mouth} 、 g'_{m_other}
6. 由 X' 的定位擷取影像中紋理資訊 g'_{s_eye} 、 g'_{s_mouth} 、 g'_{s_other}
7. 計算新的錯誤向量： $r' = g'_s - g'_m$ ： $r'(t)$ 、 $r'_{eye}(c_{eye})'$ 、 $r'_{mouth}(c_{mouth})'$ 、 $r'_{other}(c_{other})'$ ，計算新的紋理錯誤 $E1_{total}$ 、 $E1_{eye}$ 、 $E1_{moth}$ 、 $E1_{other}$
8. If ($E1_{total} > E_{total}$)

{ k 、 k_{eye} 、 k_{mouth} 、 k_{other} 各除以二；

回到第 4 步。}

```

else {
    If      (  $E1_{eye} > E_{eye}$  ) {  $k_{eye}$  除以二並回到第 4 步。}
    else If (  $E1_{mouth} > E_{mouth}$  ) {  $k_{mouth}$  除以二並回到第 4 步。}
    else If ( )  $E1_{other} > E_{other}$  ) {  $k_{other}$  除以二並回到第 4 步。}
    else 接受新的預測。}
直到所有的  $k$ 、 $k_{eye}$ 、 $k_{mouth}$ 、 $k_{other}$  接近零且錯誤率沒有提升。
}

```

此演算法首先計算目前紋理錯誤 E_{total} ， E_{eye} ， E_{moth} ， E_{other} ，在算出各個參數的更新參數，再計算出更新過後的紋理錯誤 $E1_{total}$ ， $E1_{eye}$ ， $E1_{moth}$ ， $E1_{other}$ 首先會比較 $E1_{total}$ 與 E_{total} ，若錯誤較大代表我們在位移上的參數預測錯誤，代表模型位置與影像中實際人臉的位置誤差更遠此時我們直接將位移參數與模型參數的移動改為原來的一半，不再繼續半段模型參數的正確性。主要是因為整體上人臉位置相差更遠所以區域模型：眼睛、嘴巴與其他模型位置，沒有與圖像中的眼睛、嘴巴與其他區域對應，所以預測的結果一定不符合我們的預期還有可能越發散越遠所以我們不需經過判斷直接更改模型參數的更新。若錯誤有較小表示我們在位移上的參數預測正確，區域模型：眼睛、嘴巴與其他模型位置，與圖像中的眼睛、嘴巴與其他區域距離相去不遠，所以可以進一步判斷內部模型的預測結果是否正確，若某區域的錯誤率有上升代表我們估計錯誤，則將此區域的參數位移值減半重新估計此區域的模型參數，若所有的區域模型錯誤率都有下降代表我們預測的結果正確，接受此次的預測更新原始模型與位移參數，或當所有參數的位移值都衰減到趨近於 0，則代表目前模型參數與位移參數所在位置是最佳位置。主要先利用整體的紋理資訊預測位移參數，等模型位移接近圖片影像中的人臉後，在進行內部區域模型的個別收斂。

LAAM - Face Feature Location Algorithm

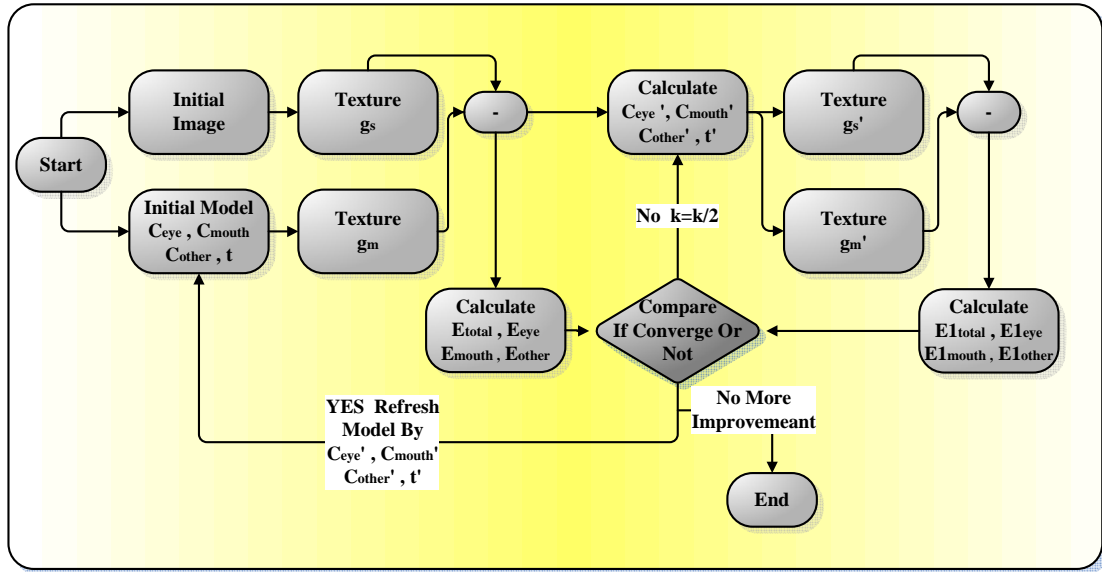


圖 4-3：區域式主動表現模型 - 人臉特徵定位演算法流程圖

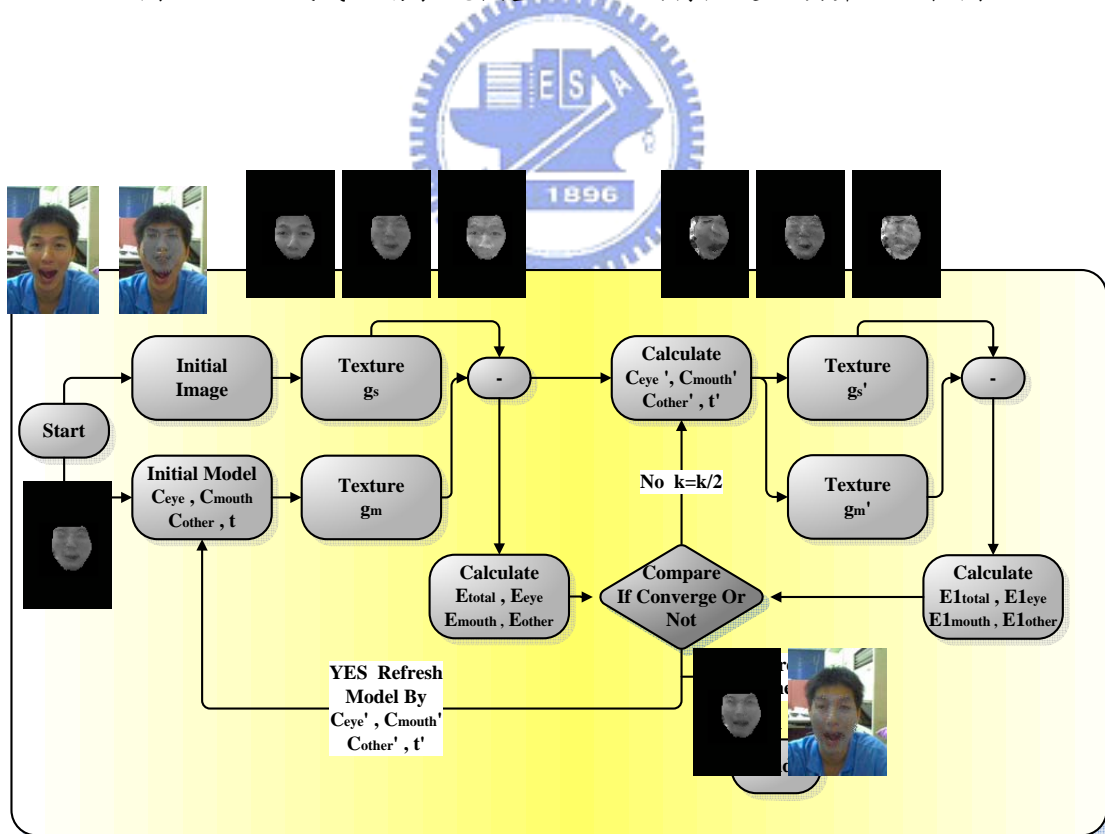


圖 4-4：區域式主動表現模型 - 人臉特徵定位演算法流程圖(二)

圖 4-3、4.4 為區域式主動表現模型人臉特徵定位演算法的流程圖，圖 4.4 在流程圖敘述框的周圍增加了演算法在實際執行時的影像以便於更能了解整個區域式主動表現模型人臉特徵定位演算法流程。先由人臉偵測程式偵測到大概的人臉，初始的位置與大小為初始模型放置的初始點，擷取初始形狀 x 對應的影像紋理資訊 g_s ，將 g_s 與模型初始參數 g_m 相減得到差異向量與誤差先預測模型參數與位移參數所需的調整值，更新參數後可以重新計算出新的模型紋理資訊 g'_m 與新的影像紋理資訊 g'_s 藉由將 g_s 與模型初始參數 g_m 相減得到新的差異向量與新的誤差。比較兩者誤差而調節參數的變化若收斂則更新模型與位移參數重新此流程，若錯誤率沒有下降則結束此收斂程式。



第五章

實驗結果

以手工標定的人臉訓練資料訓練區域式主動表現模型與主動表現模型，比較區域式主動表現模型模擬後形狀資料與紋理資料與原本資料的差異，透過與主動表現模型與原始資料的差異作比較，以實驗得知區域式主動表現模型的模擬能力。另外利用新的人臉影像透過區域式主動表現模型人臉特徵匹配演算法的收斂得到形狀與紋理資料，與原始手工標定的形狀與紋理資料比較，得到收斂演算法的匹配能力。

5.1 實驗設計

5.1.1 人臉表情定義



我們首先定義人臉的表情，經由觀察，人臉變化較大的部份為眼睛與嘴巴區域而且彼此間的變化是獨立的。我們初步定義(如圖 5-1 所示)。

眼部表情：1. 正常、2. 閉眼、3. 閉左眼、4. 閉右眼、5. 張大、6. 皺眉

口部表情：1. 正常、2. 閉嘴微笑、3. 微笑、4. 張嘴、5. 抿嘴、6. o 型嘴



圖 5-1：臉部表情定義

5.1.2 人臉資料特色

準備 230 張手工標定的人臉影像，共五個不同人拍出，一個人臉有 46 張影像。其中每個人各提供 36 張影像(共 180 張)作訓練資料，用來訓練區域式主動表現模型(LAAM)與主動表現模型(AAM)。其餘每個人各提供 10 張影像(共 50 張)用來做測試模型資料用。每個人表情的變化以 5.1.1 定義的人臉表情做排列組合達到所有人臉表情都在訓練資料中。手工標定資料如圖 5.2 所示，每行的變化為同一個人臉不同表情的變化，每列的變化為不同人臉的變化。





圖 5-2：人臉標定資料

5.2 區域表現模型之模型建立測試

使用以上所述 180 張人臉當作模型訓練資料，測試資料 50 張。經由第三章所述之建立模型的步驟(如圖 3.8 所示)，先擷取出形狀與紋理資料，作形狀與紋理上的校正，將資料分為眼睛、嘴巴與其他區域再利用主成分分析建立眼睛、嘴巴與其他區域的統計模型，再將三個模型合成建立出區域式主動表現模型。區域式主動表現模型維度如圖 5-3 所示。

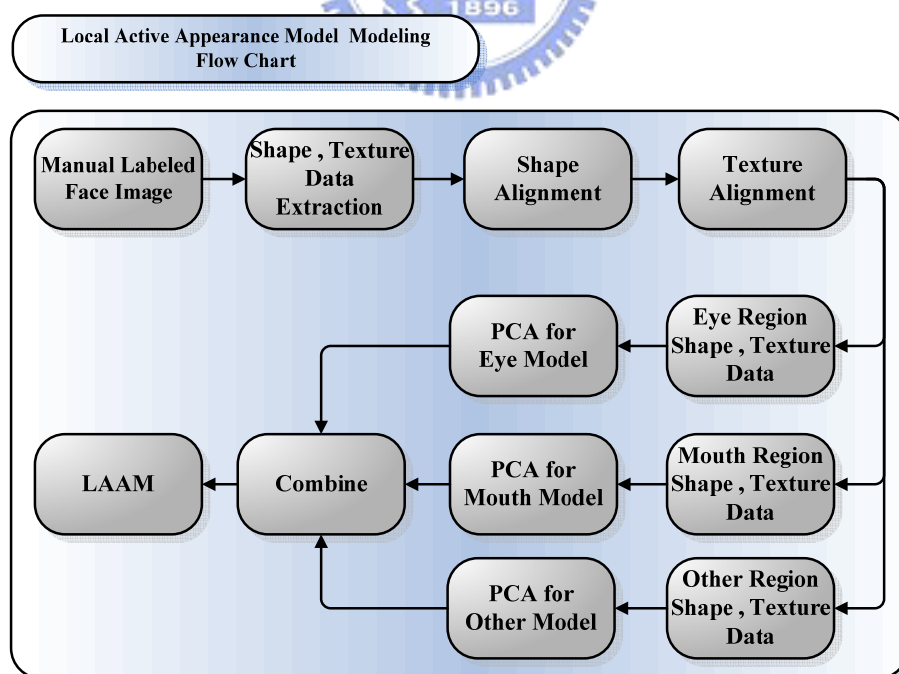


圖 3.8：區域式主動表現模型流程圖

嘴巴部份：形狀維度：1x16，紋理維度：1x109，合成維度：1x125。

眼睛部份：形狀維度：1x52，紋理維度：1x390，合成維度：1x442。

其他部份：形狀維度：1x48，紋理維度：1x1445，合成維度：1x1490。

整體(眼睛、嘴巴、其他)合成維度 1x2060。

主動表現模型的形狀維度：1x116，紋理維度：1x1944，合成維度：1x2060。

兩者的整體維度是一樣的。

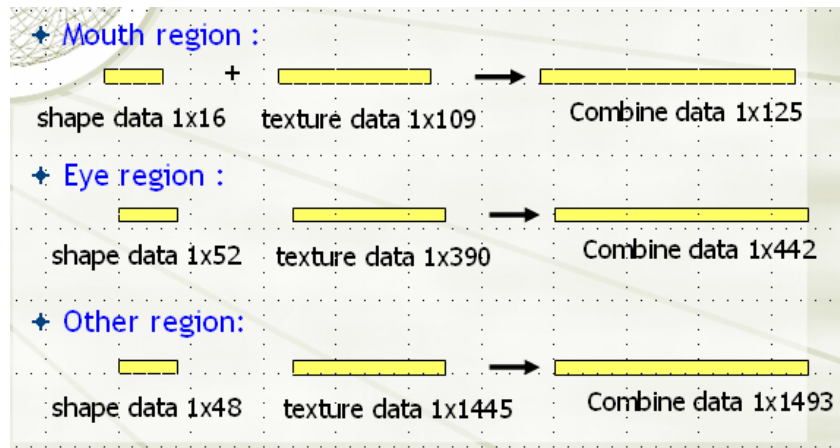


圖 5-3：區域主動表現模型維度

模型選取參數在區域式主動表現模型方面，眼睛部份的特徵向量選取前 179 個(維度 1x442)，嘴巴部份的特徵向量選取前 112 個(維度 1x125)，其他部份的特徵向量選取 179 個(維度 1x1493)。在主動表現模型方面，特徵向量選取前 179 個(維度 1x2060)。資料模擬度都趨近於 100%。雖然表面上而言，區域主動表現模型所需控制參數的數目比主動表現模型多 (179+125+179[LAAM] > 179[AAM])，比較上可能會不公平。但就模型模擬所需的資料儲存，區域主動表現模型需儲存的資料量為 359247(179x442+112x115+179x1493)，主動表現模型需儲存的資料量為 368740(179x2060)，所以在資料量的儲存上區域主動表現模型較主動表現模型少，但是能提供較多的模型參數供系統做變化上的模擬。若主動表現模型要達到相同的參數量，其儲存空間需要原來的三倍。

以測試資料(50 張)進行模型的模擬測試。方法如圖 5.4 所示。先取出測試手工標定影像的形狀與紋理資料，經由 AAM 與 LAAM 的模擬(對 PCA 做投影-降維再做反投影-升維)，模擬過後的 AAM 與 LAAM 形狀與紋理資料在與原始形狀與紋理資料相減比較。

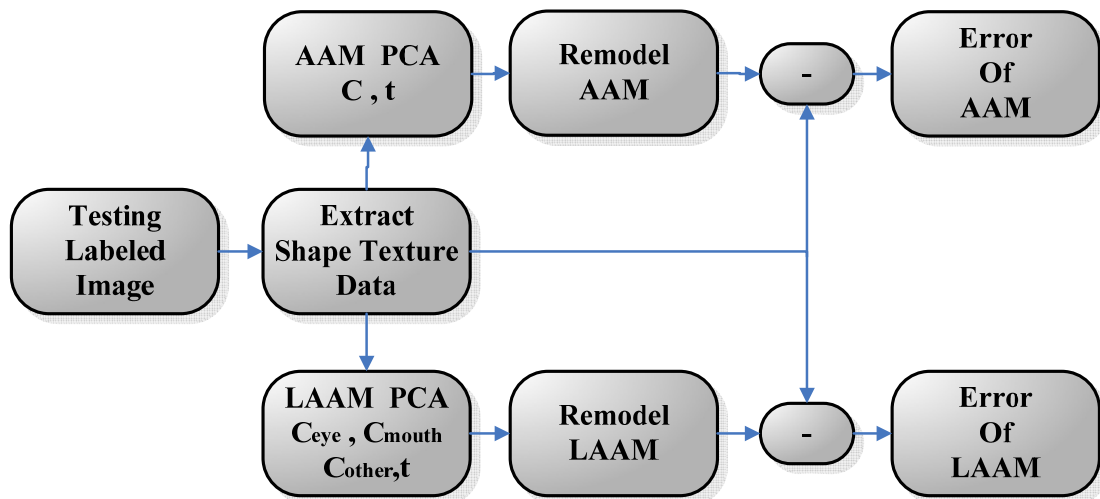


圖 5-4：模型測試流程圖

測試結果如表 5-1 所示，顯示區域式主動表現模型與主動表現模型模擬資料與原始資料的比較結果：

| | Mouth | | Eye | | Other | |
|---------------|-------|------|------|------|-------|------|
| | AAM | LAAM | AAM | LAAM | AAM | LAAM |
| Shape Error | 0.78 | 0.03 | 0.72 | 0.24 | 0.76 | 0.76 |
| Texture Error | 6.43 | 0.47 | 5.76 | 3.08 | 4.51 | 4.07 |

表 1：主動表現模型與區域式主動表現模型比較

經由表一的比對我們會發現，LAAM 與 AAM 表現的比較，在嘴巴與眼睛區域差異明顯(嘴巴區域中，形狀差異 AAM=0.78 而 LAAM=0.03，紋理差異 AAM=6.43 而 LAAM=0.47。眼睛區域中，形狀差異 AAM=0.72 而 LAAM=0.24，紋理差異 AAM=5.76 而 LAAM=3.08)，所以可以得到 LAAM 在變化較大的區域較 AAM 模擬的更精確。在其他區域(鼻子與臉頰)模擬的表現則差不多。由於在特徵值挑選時，模型選取參數在區域式主動表現模型方面，眼睛部份的特徵向量選取前 179 個，嘴巴部分的

特徵向量選取前 112 個，其他部分的特徵向量選取 179 個。在主動表現模型方面，特徵向量選取前 179 個。資料模擬度都趨近於 100%。

在控制的參數上 LAAM 比 AAM 多出兩倍，或許有人會質疑說若將 AAM 的控制參數增加是否會變的更好。所以我們做了另一組實驗。模型選取參數在區域式主動表現模型方面，眼睛部份的特徵向量選取前 179 個，嘴巴部份的特徵向量選取前 112 個，其他部分的特徵向量選取 179 個。在主動表現模型方面，特徵向量選取前 540 個。資料模擬度都趨近於 100%。以測試資料(50 張)進行模型的模擬測試。測試結果如表 5-2 所示，顯示區域式主動表現模型與主動表現模型模擬資料與原始資料的比較結果：

| | Mouth | | Eye | | Other | |
|---------------|-------|------|------|------|-------|------|
| | AAM | LAAM | AAM | LAAM | AAM | LAAM |
| Shape Error | 0.72 | 0.03 | 0.59 | 0.24 | 0.76 | 0.76 |
| Texture Error | 5.75 | 0.47 | 5.07 | 3.08 | 4.10 | 4.07 |

表 2：主動表現模型與區域式主動表現模型比較(二)

由表二可得知，既使主動表現模型 (AAM) 增加控制參數個數，由於資料模擬度的飽合，錯誤率沒有下降太多，結果仍較區域式主動表現模型 (LAAM) 差。而且須儲存資料量也很大。現在來探討區域式主動表現模型 (LAAM) 表現較好的原因：主成份分析法(PCA)主要是做資料的座標軸轉換，若將計算出來的特徵向量依照特徵值做遞減的排列，越大的特徵值所對應的特徵向量其訓練資料在此特徵向量座標軸的分佈標準差越大(資料較分散)，如此挑選少量但特徵值大的特徵向量當作基底座標，由於在這些特徵向量下資料分布都比較分散，即用比較少量的維度表現出一般高維度展現出的資料群。

主動表現模型 (AAM) 將一個人臉資料群(1x2060)，利用主成份分析法的座標轉換轉成令一個座標軸。區域式主動表現模型(LAAM)則是將眼睛部份(維度 1x442)、嘴巴部份(1x125)、其他部分(1x1490)三部份利用主成份分析法，分別轉換到三個不同的座標系統。由於人臉的眼睛、嘴巴與其他部分資料群最廣分布的座標軸方向可能不同，三個座標系統可以分別依照他們的資料特性做適當的分布。所以這是區域式主動表現模型 (LAAM) 模擬精確度會優於主動表現模型 (AAM) 的原因。

區域式主動表現模型(LAAM)的模型參數， c_{eye} 、 c_{mouth} 、 c_{other} ，控制著各部份的變化，若要做某個部位的判斷如眼睛表情(眼睛的閉合)或嘴巴表情(嘴巴的張閉)…等可直接結取模擬過後的參數 c_{eye} 、 c_{mouth} 、 c_{other} 去做判斷。

圖 5.5 為區域式主動表現模型與主動表現模型差異度的比較分析，圖 5.5 中第一列為原始資料，第二列為主動表現模型(AAM)根據第一列的原始資料模擬出的最佳結果，第三列為區域式主動表現模型(LAAM)根據第一列的原始資料模擬出的最佳結果。我們可以看出第一行中，原始資料嘴巴呈現微笑的狀態。主動表現模型(AAM)模擬結果嘴巴呈現微張的情形，區域式主動表現模型(LAAM)模擬結果嘴巴呈現微笑的狀態。第二行中，原始資料嘴巴呈現O型嘴的狀態。主動表現模型(AAM)模擬結果嘴巴呈現微張的情形，區域式主動表現模型(LAAM)模擬結果嘴巴呈現O型嘴的情形。第三行中，原始資料眼睛是閉眼的狀態。主動表現模型(AAM)模擬結果眼睛是張開的狀態，區域式主動表現模型(LAAM)模擬結果眼睛則是閉眼的狀態。第四行中，原始資料嘴巴是張大的狀態。主動表現模型(AAM)模擬結果嘴巴是微張的狀態，區域式主動表現模型(LAAM)模擬結果嘴巴則是張大的狀態。由視覺上來說，區域式主動表現模型(LAAM)的模擬結果會較主動表現模型(AAM)的模擬結果好。

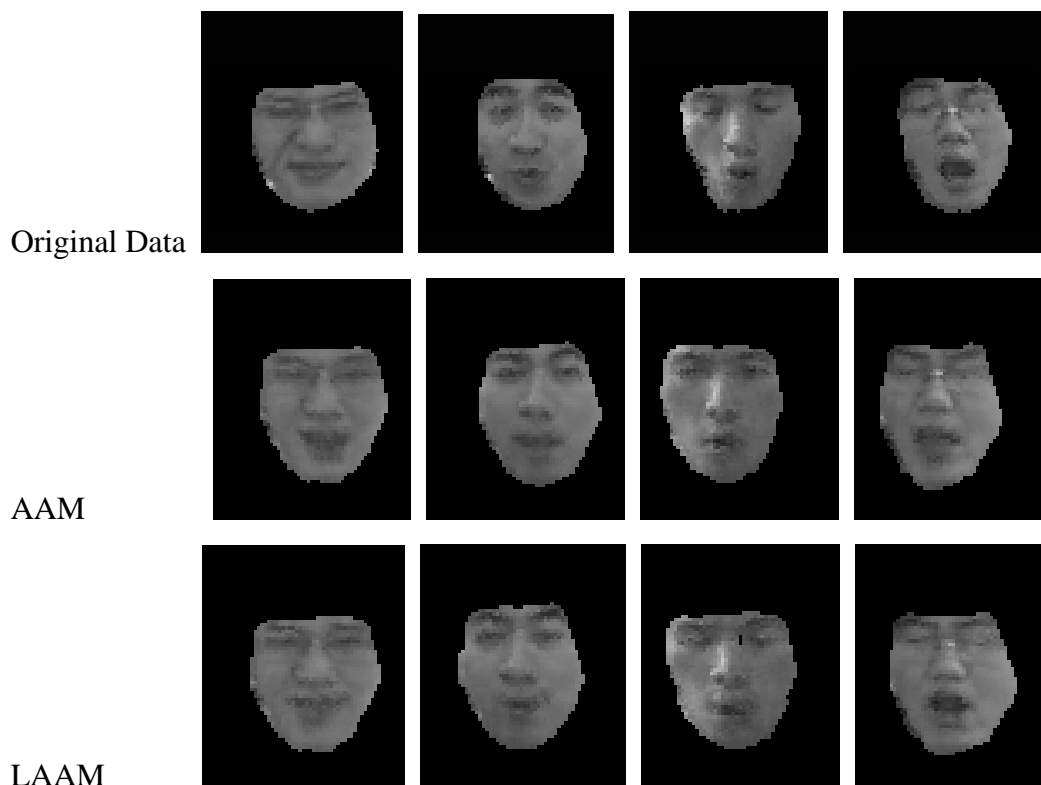


圖 5-5: LAAM AMM 模擬差異

圖 5.6~5.11 顯示主成分分析時，在各種資料模擬度下 AAM 與 LAAM 的比較。5.6 表示嘴巴區域形狀資訊表現情形、5.7 表示嘴巴區域紋理資訊表現情形、5.8 表示眼睛區域形狀資訊表現情形、5.9 表示眼睛區域紋理資訊表現情形、5.10 表示其他區域形狀資訊表現情形、5.11 表示其他區域紋理資訊表現情形、橫軸表示資料模擬度(0%~100%)、縱軸表示模擬與實際的誤差、藍色實線表示 AAM 的分佈情形、紅色虛線表示 LAAM 的分佈情形。可以看到在嘴巴區域形狀與紋理資料上 LAAM 的分佈都較 AAM 低，這代表在嘴巴區域 LAAM 在各個資料模擬度下表現都比 AAM 好。

在眼睛區域除了形狀上低資料模擬度時 AAM 較 LAAM 好外，形狀與紋理資料上 LAAM 的分佈整體都較 AAM 低，這代表在眼睛區域 LAAM 在各個資料模擬度下表現都比 AAM 好。在其他區域形狀與紋理資料模擬誤差不大。由資料模擬程度上來說，不同資料模擬度下，區域式主動表現模型(LAAM)的模擬結果會較主動表現模型(AAM)的模擬結果好。

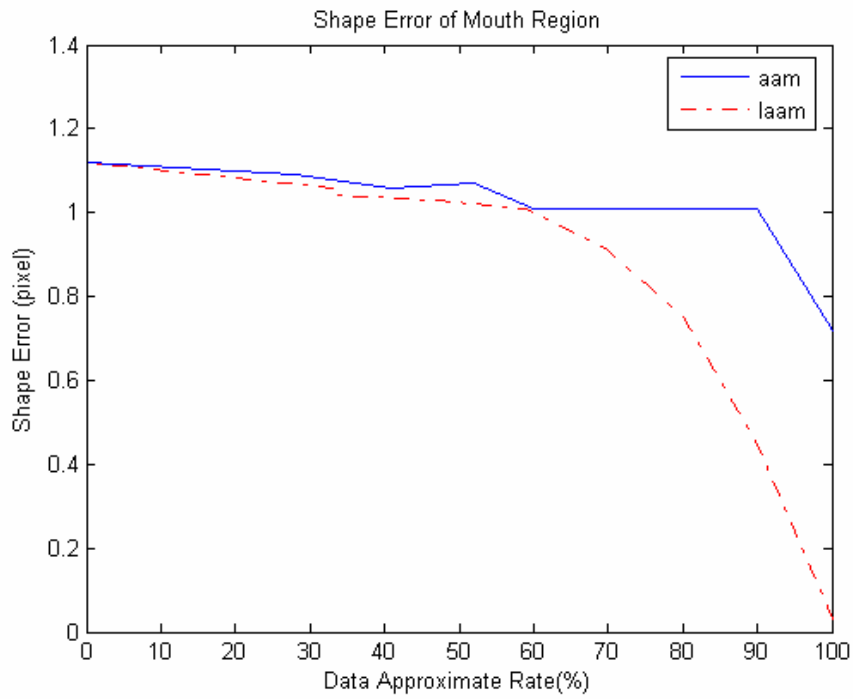


圖 5-6：嘴巴區域形狀模擬誤差

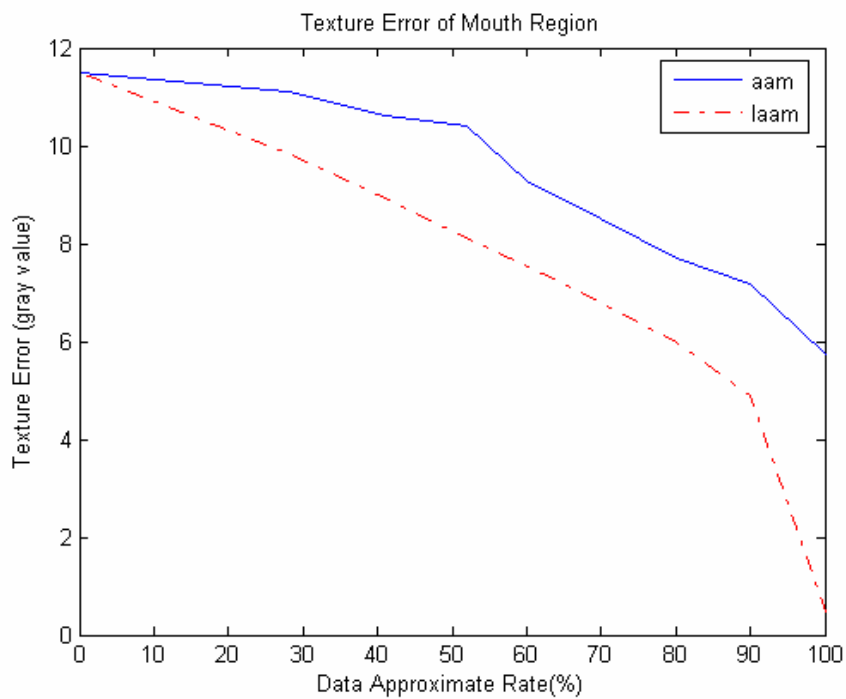


圖 5-7：嘴巴區域紋理模擬誤差

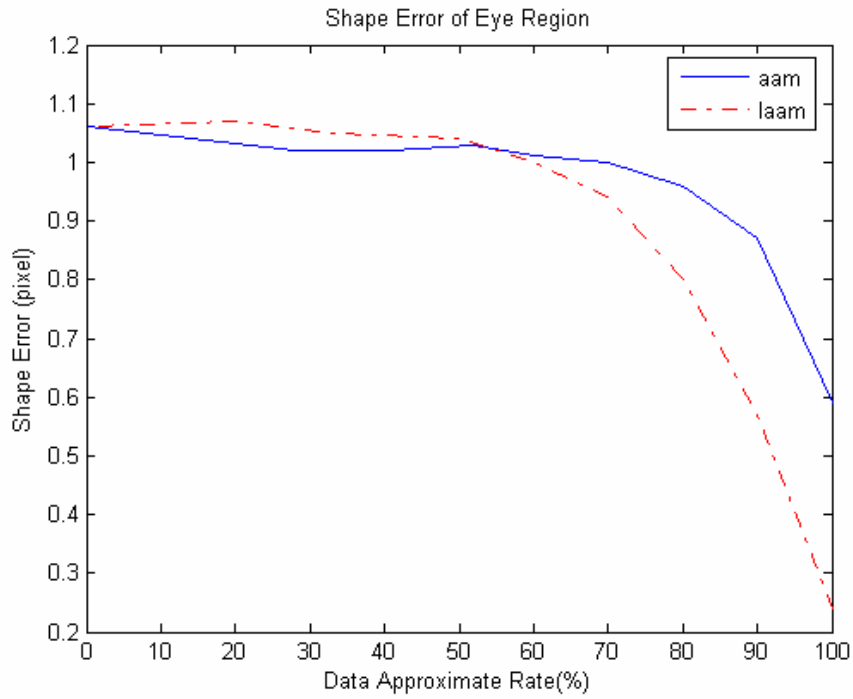


圖 5-8：眼睛區域形狀模擬誤差

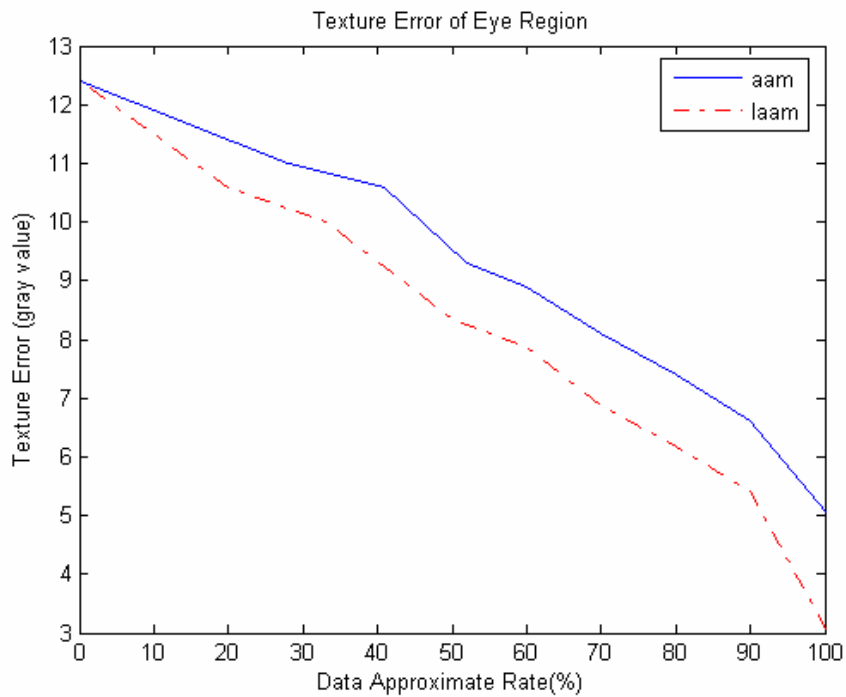


圖 5-9：眼睛區域紋理模擬誤差

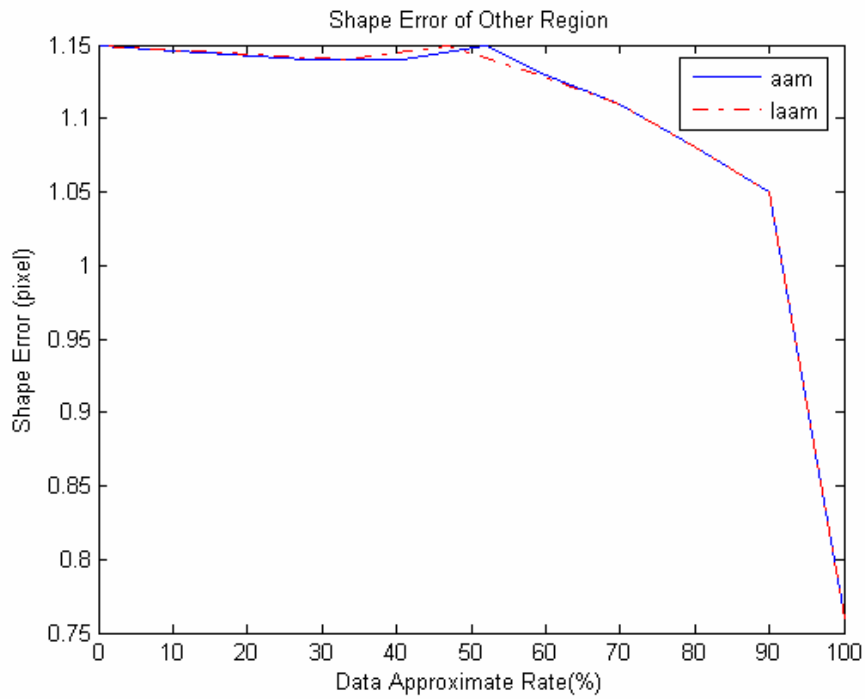


圖 5-10：其他區域形狀模擬誤差

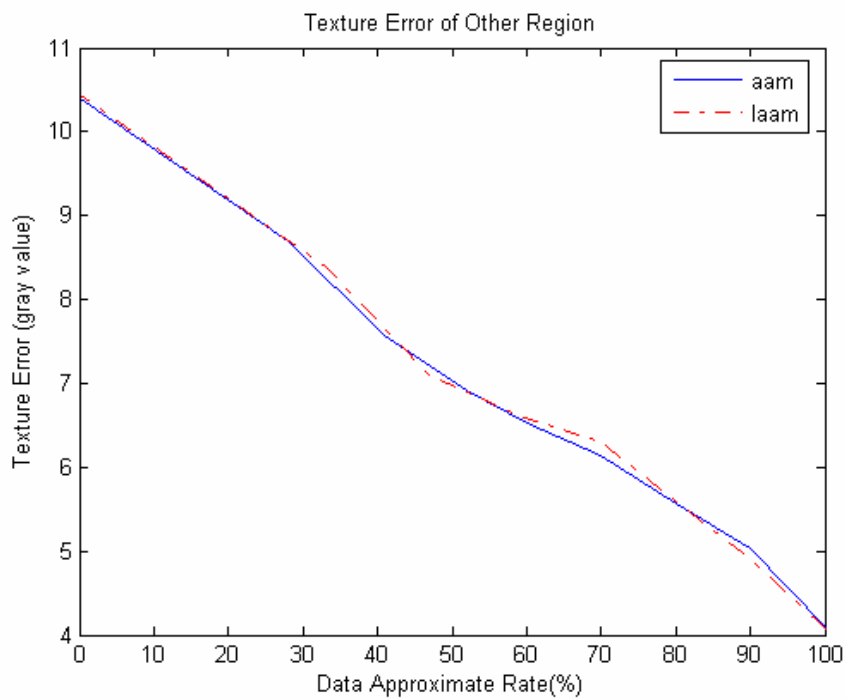


圖 5-11：其他區域紋理模擬誤差

5.3 區域表現模型之人臉特徵定位測試

使用 5.2 建立的區域主動表現模型作人臉特徵匹配。所需的人臉初始位置與大小由 OpenCV 之人臉偵程式提供[19]，主要功能是提供此影像中人臉的大概位置與大小。如圖 5.12 所示，一、三行為人臉影像，二、四行為人臉偵測結果。



圖 5-12：人臉偵測結果

得到人臉大概位置與大小，再利用第四章所述之人臉特徵定位演算法(下圖 4.3)已大概位置與大小為初始，以模型與影像紋理的差異向量預測模型參數與位移參數，去追蹤人臉詳細形狀與紋理資訊。

LAAM - Face Feature Location Algorithm

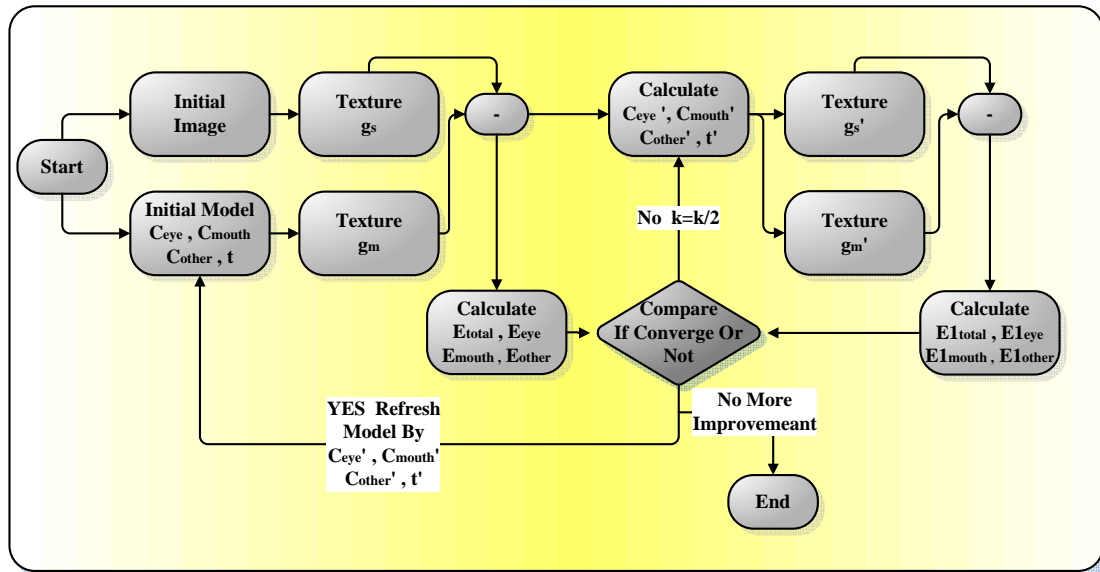
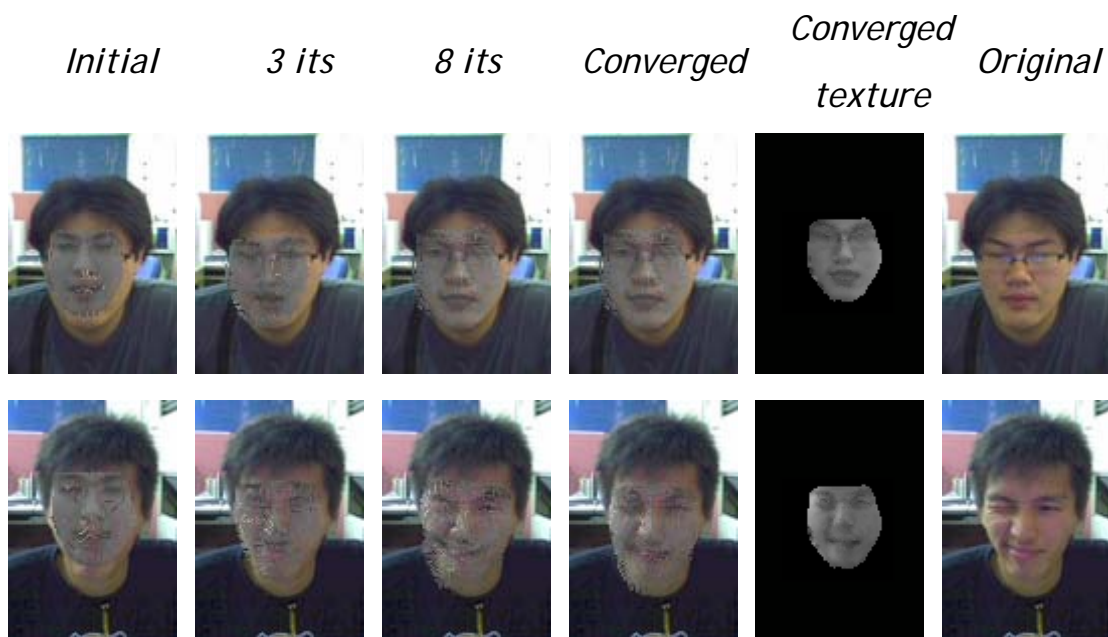


圖 4-3：區域式主動表現模型 - 人臉特徵定位演算法流程圖

人臉特徵定位演算法，追蹤人臉特徵結果如上圖 5-13 所示：第一行為已知初始人臉大概的位置與大小，將區域式主動表現模型(LAAM)初始模型依一定比例放置於影像中，第二行為匹配步驟執行三次結果，第三行為匹配步驟執行八次結果，第四行為匹配步驟執行至收斂的結果，第五行為紋理資訊收斂結果，第六行為初始影像，共有五個結果。可以觀察到人臉的位置與大小越來越接近影像中的



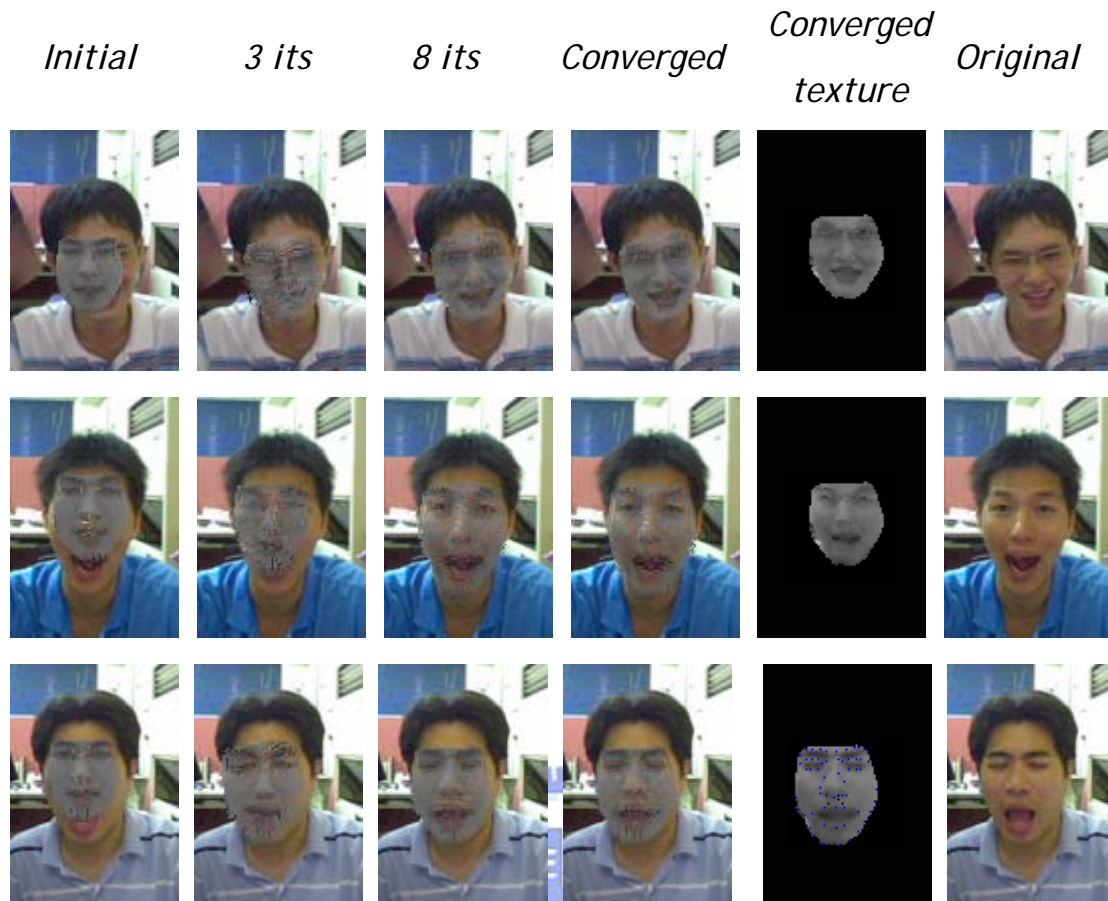


圖 5-13：區域式主動表現模型人臉特徵定位演算法結果

人臉，直到收斂後人臉的形狀於影像中匹配的很好，第五行為收斂的紋理資訊與第六行的原始圖比較可以看到收斂過後臉部的表情與特徵都很符合。由此可以用肉眼觀察出人臉特徵定位演算法的精確度。

下圖 5-14 到 5-18 是使用人臉特徵定位演算法追蹤人臉特徵結果，藍色方框為初始的人臉偵測程式偵測出的大概位置與大小，而綠色圓點標訂出人臉特徵定位演算法追蹤人臉特徵的結果。共有 58 個點。可以看到特徵標定點會隨著人臉表情的不同而跟隨其變化。



圖 5-14：區域式主動表現模型 - 人臉特徵定位演算法結果(人臉一)



圖 5-15：區域式主動表現模型 - 人臉特徵定位演算法結果(人臉二)





圖 5-16：區域式主動表現模型 - 人臉特徵定位演算法結果(人臉三)



圖 5-17：區域式主動表現模型 - 人臉特徵定位演算法結果(人臉四)

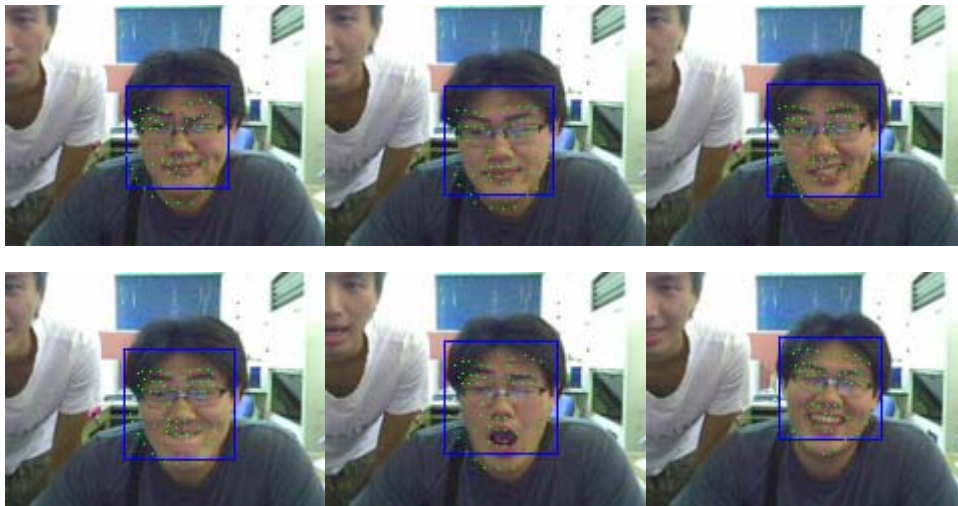


圖 5-18：區域式主動表現模型 - 人臉特徵定位演算法結果(人臉五)

5.4 系統的處理時間

特徵匹配系統處理會隨著人臉不同或初始人臉的位置不同而會有不同的遞迴數目到達收斂如圖 5.13 的收斂情形。經由實際攝影機的測試：系統建構在 Windows 視窗環境下運作，所使用的介面是以 Borland C++ Builder 6.0 作為開發的工具，而開發平台所用的處理器是 Intel Pentium (R) D CPU 2.8GHz，搭載 1.5G 的記憶體，取像使用羅技的一般網路攝影機。平均的特徵匹配系統處理速度約每個 frame 需要 50ms，即為 20fps。



第六章

結論與未來展望

6.1 結論

我們提出一個對可變形物體有效模擬的模型：區域式主動表現模型(LAAM)。對於物體中趨近於獨立變化部分，如：人臉中的眼睛與嘴巴。將這些區域各自獨立(眼睛區域、嘴巴區域與其他區域)以主成分分析法(PCA)製作區域式統計模型，再做人臉上的結合形成區域式主動表現模型，區域式主動表現模型比主動表現模型更能夠精確的模擬出人臉的形狀與紋理表情。在追蹤特徵形狀與紋理部份，區域式主動表現模型使用區域式的紋理相減差異去預測模型參數，而位置、大小、旋轉參數則用全域紋理相減差異來預測其收斂數值。在預測的步驟中由於速度上的需求，在預測各參數的改變大小的最佳化步驟使用的預測因子R由訓練資料計算得知。若與主動表現模型(AAM)比較，我們提出的區域式主動表現模型(LAAM)有以下這幾個好處。

1. 將區域式主動表現模型(LAAM)模擬結果與主動表現模型(AAM)模擬結果比較，區域主動表現模型(LAAM)平均錯誤率較低(如表一所示)，其模擬度就人眼觀察也較高(如圖 5-5 所示)，在各種資料模擬度上(0%~100%)區域主動表現模型(LAAM)的表現都較主動表現模型(AAM)好。所以由實驗可得，區域式主動表現模型(LAAM)模擬結果較主動表現模型(AAM)好
2. 在主動表現模型(AAM)中須考慮各區域各種變化的排列組合(如人臉中眼睛表情與嘴巴表情的排列組合)才能模擬出較佳的結果，若各表情間的排列組合太少主動表現模型在模擬時會只有某些表情正確而其他的表情錯誤。在區域式主動表現模型(LAAM)中，各區域的模擬是獨立的，所以在建立模型時，只需將此區域的所有變化獨立考慮，不需要考慮到各區域間變化的排列組合。如此就同一個人的

訓練資料可以大幅減少，且隨著追蹤物體中趨近獨立變化的區域越多，兩者所需要的訓練資料量差異越大。所以區域式主動表現模型能夠更有效的準備訓練資料。由於所需訓練資料張數減少，在相同的品質要求下，主成份分析(PCA)後特徵向量的挑選也可以減少，如此可以減少程式需預存的記憶體空間，也由於模型維度較小所以模型模擬執行時間可以縮短，所以可以增加執行時的速度。實驗中，區域主動表現模型需儲存的資料量為 359247 (179x442+112x115+179x1493)，主動表現模型需儲存的資料量為 368740 (179x2060)，所以在資料量的儲存上區域主動表現模型較主動表現模型是比較小的，在參數個數上，區域式主動表現模型個數為 470(179+112+179)，而主動表現模型參數個數為 179，區域主動表現模型與主動表現模型比較，區域主動表現模型維持一定的資料模擬度，在資料儲存上相對較小，它犧牲了模擬參數的增加以換取比主動表現模型更好的模擬度。

3. 區域式主動表現模型(LAAM)的模型參數， c_{eye} 、 c_{mouth} 、 c_{other} ，獨立控制著各部份的變化，在特徵匹配過後，若要做某個部位的表情判斷，如眼睛表情(眼睛的閉合)或嘴巴表情(嘴巴的張閉)…等可直接結取模擬過後的參數 c_{eye} 、 c_{mouth} 、 c_{other} 送入分類器去做訓練判斷。

由於模型已做改變，由三個區域式的獨立模型：眼睛統計模型，嘴巴統計模型與其他區域統計模型組成計算錯誤率時會有四個判斷用的錯誤率 E_{total} 、 E_{eye} 、 E_{moth} 、 E_{other} ，需要預測區域模型參數：眼睛部份模型參數 c_{eye} 、嘴巴部份模型參數 c_{mouth} 、其他部分模型參數 c_{other} 與整體的位移參數 $t = (t_x, t_y, \theta, s)^T$ ，由於區域模型參數紋理差異關係只與此區域有關而位移參數則與整體的紋理差異有關，所以我們修改了主動表現模型(AAM)的人臉特徵定位演算法，使其適合區域主動表現模型(LAAM)的人臉特徵定位，以做人臉特徵定位的動作。

6.2 未來展望

目前我們提出的區域主動表現模型(LAAM)的人臉特徵定位演算法的實驗收斂結果錯誤率仍與主動表現模型(LAAM)的人臉特徵定位演算法的錯誤率差不多。未能有效的發揮區域主動表現模型(LAAM)的優點。未來希望能夠修改其演算法，將區域主動表現模型(LAAM)的優點發揮到最大。



參考文獻

- [1] T.F. Cootes, G.J. Edwards, and C.J. Taylor, “Active Appearance Models”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 6, June 2001
- [2] C. Nastar, B. Moghaddam, and A. Pentland, “Generalized Image Matching: Statistical Learning of Physically-Based Deformations,” Computer Vision and Image Understanding, vol. 65, no. 2, pp. 179-191, 1997.
- [3] A. Lanitis, C.J. Taylor, and T.F. Cootes, “Automatic Interpretation and Coding of Face Images using Flexible Models,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 743-756, July 1997.
- [4] G. Edwards, A. Lanitis, C. Taylor, and T. Cootes, “Statistical Models of Face Images – Improving Specificity,” Image and Vision Computing, vol. 16, pp. 203-211, 1998.
- [5] T. Ezzat and T. Poggio, “Facial Analysis and Synthesis Using Image-Based Models,” Proc. Second Int'l Conf. Automatic Face and Gesture Recognition 1997, pp. 116-121, 1997.
- [6] M.J. Jones and T. Poggio, “Multidimensional Morphable Models: A Framework for Representing and Matching Object Classes,” Int'l J. Computer Vision, vol. 2, no. 29, pp. 107-131, 1998.
- [7] T. Vetter, “Learning Novel Views to a Single Face Image,” Proc. Second Int'l Conf. Automatic Face and Gesture Recognition, pp. 22-27, 1996.
- [8] M. Gleicher, “Projective Registration with Difference Decomposition,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1997.
- [9] G. Hager and P. Belhumeur, “Efficient Region Tracking with Parametric Models of Geometry and Illumination,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 10, pp. 1025-1039, Oct. 1998.
- [10] S. Sclaroff and J. Isidoro, “Active Blobs,” Proc. Sixth Int'l Conf. Computer

- Vision, pp. 1146-1153, 1998.
- [11] M. La Cascia, S. Sclaroff, and V. Athitsos, “Fast, Reliable Head Tracking under Varying Illumination: An Approach Based on Registration of Texture Mapped 3D Models,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 4, pp. 322-336, Apr. 2000.
- [12] T. F. Cootes and C. J. Taylor. “Active shape models.” In David Hogg and Roger Boyle, editors, 3rd British Machine Vision Conference, pages 266–275. Springer-Verlag, September 1992
- [13] N. Dowson and R. Bowden. “Simultaneous modeling and tracking (smat) of feature sets.” In *Computer Vision and Pattern Recognition Conference 2005*, San Diego, USA, 2005.
- [14] P. Felzenszwalb and D. Huttenlocher. “Pictorial structures for object recognition.” *International Journal of Computer Vision*, 61, 2005.
- [15] Hou XW, Li SZ, Zhang HJ, Cheng QS. “Direct appearance models.” In: Kasturi R, Medioni G, eds. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, Vol 1. Kauai: IEEE Computer Society, 2001. 828~833.*
- [16] SONG Gang, AI Hai-Zhou, XU Guang-You, “Texture Constrained Facial Feature Point Tracking” *Conf. of Journal of Software*, vol. 15, no. 11, 2004
- [17] David Cristinacce and Tim Cootes, “Facial Feature Detection and Tracking with Automatic Template Selection.”, in *Proceedings of 7th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 429-434, Southampton, UK, April, 2006.
- [18] I. L. Dryden and K. V. Mardia. “Statistical Shape Analysis.” John Wiley & Sons, 1998.
- [19] P. Viola and M. Jones. “Rapid object detection using a boosted cascade of simple features.” In *Computer Vision and Pattern Recognition Conference 2001*, Hawaii, USA, volume 1, pages 511–518, Kauai, Hawaii, 2001.
- [20] I. M. Scott, T. F. Cootes, and C. J. Taylor. “Improving appearance model

- matching using local image structure.” In Information Processing in Medical Imaging 18th International Conference, pages 258–269, July 2003.
- [21] T. F. Cootes, G. V. Wheeler, K. N. Walker, and C. J. Taylor. “Coupled-view active appearance models.” In *Proc. British*
- [22] G. Edwards, T. Cootes, and C. Taylor. “Face recognition using active appearance models”. In Proceedings of the European Conference on Computer Vision, volume 2, pages 581–695, 1998.
- [23] Yan SC, Liu C, Li SZ, Zhang HJ, Shum H, Cheng QS.” Face alignment using texture-constrained active shape models.” *Image and Vision Computing*, 2003,21(1):69~75.
- [24] Song G, Ai HZ, Xu GY. “Hierarchical direct appearance model for elastic labeled graph localization.” In: Lu HQ, Zhang TX, eds. *Proc. of the SPIE Vol. 5286*. Beijing: SPIE, 2003. 127~132.
- [25] Li SZ, Yan SC, Zhang HJ, Cheng QS. “Multi-View face alignment using direct appearance models.” In: Davis L, ed. *Proc. of the Int’l Conf. on Automatic Face and Gesture Recognition*. Washington: IEEE Computer Society, 2002. 309~314.
- [26] Wiskott L, Fellous JM, Krer N, von der Malsburg C. “Face recognition by elastic bunch graph matching.” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1997,19(7):775~779.
- [27] J. A. Nelder and R. Mead. “A simplex method for function minimization. *Computer Journal*,” 7:308–313, 1965.
- [28] Kwang Ho An, Dong Hyun Yoo, Sung Uk Jung and Myung Jin Chung “Robust Multi- view Face Tracking,” Dept. of Electrical Engineering and Computer Science Korea Advanced Institute of Science and Technology Daejeon 305-701, Republic of Korea
- [29] M. A. Turk and A. P. Pentland. “Face recognition using eigenfaces.”. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 586–591, Hawaii, June 1991.

[30] D. Cristinacce, T. Cootes, and I. Scott. “A multi-stage approach to facial feature detection.” In 15th British Machine Vision Conference, London, England, pages 277–286, 2004.

