

國立清華大學

碩士論文

以特徵篩選的調昇式學習法
結合支持向量機實作性別和年齡辨識



AdaBoost Multiple Feature Selection
with SVM-based Component Classifiers
for Gender and Age Classification

系別 資訊工程學系 組別 _____

學號姓名 100062553 郭哲綸 (Che-Lun Kuo)

指導教授 張智星 博士 (Jyh-Shing Jang)

中華民國一〇二年七月

摘要

儘管支持向量機 (Support Vector Machine, SVM) 是屬於強分類器，不容易與調昇式學習法 (Adaptive Boosting, AdaBoost) 作結合，但其已被證明為調昇式學習法中有效的組合分類器 (Component Classifier)。本論文即根據此理論提出以特徵篩選的調昇式學習法結合支持向量機來實作性別和年齡辨識。所選用的特徵皆被證實能有效描述人臉，包含局部二元化模版 (Local Binary Pattern, LBP)、局部方向性模版 (Local Directional Pattern, LDP)、局部三元化模版 (Local Ternary Pattern, LTP)、蓋伯濾波器 (Gabor Filter) 以及局部蓋伯二元化模版 (Local Gabor Binary, LGBP)。上述各種方式擷取的特徵，都具有很高的維度，為了加速模型的訓練，以及效能評估的檢測，我們另外使用兩種方式進行降維：主成分分析 (Principal Component Analysis, PCA) 以及線性識別分析 (Linear Discriminant Analysis, LDA)，接著利用支持向量機作為 AdaBoost 的組合分類器，即可在 AdaBoost 每次迭代的過程篩選出最適合的特徵，進而學習並組合出最終的辨識模型。

由辨識結果看來，未進行特徵篩選的 AdaBoostSVM 與 SVM 這兩個分類器有近似的辨識能力，而進行特徵篩選後的 AdaBoostSVM 則有機會得到較上述兩分類器好的性別和年齡辨識結果。

關鍵字：性別辨識、年齡辨識、特徵篩選、主成分分析、線性識別分析、調昇式學習法、支持向量機

Abstract

Although SVM (Support Vector Machine) is a strong classifier that can't be combined with AdaBoost (Adaptive Boosting) easily, it has been proved to be effective component classifier in AdaBoost. In this paper, we propose AdaBoost multiple feature selection with SVM for gender and age classification based on the above-mentioned theory. The chosen features have been proved to be successful in describing facial image. They include LBP (Local Binary Pattern), LDP (Local Directional Pattern), LTP (Local Ternary Pattern), Gabor filter, and LGBP (Local Gabor Binary Pattern). To speed up the training process and get better performance, we reduce dimensionalities by two methods: PCA (Principal Component Analysis) and LDA (Linear Discriminant Analysis). Then, we use SVM as the component classifier in AdaBoost. At each iteration in AdaBoost, we can choose the most suitable feature by evaluating the training error rate. Finally, we can combine these component classifiers into the ultimate recognition model.

From the experimental results, AdaBoostSVM with no feature selection can perform as well as SVM. Moreover, AdaBoostSVM with multiple feature selection has the chance to outperform the other classifiers in both gender and age classification.

Keywords: Gender classification, Age classification, Feature selection, Principal Component Analysis, Linear Discriminant Analysis, AdaBoost, Support Vector Machine

謝誌

首先我要感謝指導老師張智星教授。從大三開始，就是您的專題學生，碩班時期也繼續接受您的指導。這一路四年以來從您身上學到了許多東西，例如：做研究時不要先問別人該怎麼做，而是先自己想想該怎麼做；多參加比賽，不僅能迅速累積自己的實戰經驗，也對未來找工作有幫助；培養運動的習慣，不但能使身體健康，有足夠的體力完成論文這條漫長的路，更有助於思緒的沈澱，讓頭腦清晰，研究的脈絡也才能清楚地貫徹。

接著感謝我的大學好友簡均翰，即使自己忙論文忙得焦頭爛額，也願意在我看不懂論文的算式時，不厭其煩地用簡單的例子讓我了解式子的含意和精髓，讓我的研究能夠步上軌道。

感謝實驗室跟我一起參加由田機器視覺獎的瑋呈。碩一暑假時，我們就有一起參加這項比賽，無奈當時涉獵的論文不夠多，加上彼此課業繁忙，所以僅拿到第五名。希望這次我們的捲土重來能拿到理想的名次。

感謝博班學長阿諾、小龜、葉子和 Davidson。因為你們我才能避免一些研究上方向的錯誤，以及了解做論文研究時，條理清晰和循序漸進的重要性。謝謝同屆碩二的瑋呈、毓哲、高瑋、阿泰、小六、小佩、鄧愷、廷漢、俊源，因為我們彼此的打氣和砥礪，大家的論文進度才能一起穩定地向前。感謝碩一的學弟妹 Lucy、坤優、安子、Cookie、Webber、煜傑、彥章，看見你們讓我彷彿看到了碩一的自己，讓我能夠暫時忘記因趕論文而緊繃的情緒。感謝我的家人和女友在我忙論文時，不斷地支持我，給我加油打氣。

我還不夠好，還有很多需要學習的地方，因此我由衷地感謝一路上願意幫助我、讓我成長的人。如今付梓在即，希望此論文亦能給需要的人幫助和啟發。

目錄

摘要	I
Abstract	II
謝誌	III
目錄	IV
圖目次	VII
表目次	X
第一章 緒論	1
1.1 研究動機	1
1.2 性別與年齡辨識系統	1
1.3 研究方向	2
1.4 章節概要	3
第二章 相關研究簡介	4
第三章 研究方法	6
3.1 影像前處理	6
3.1.1 人臉偵測	6
3.1.2 人臉校準	7
3.1.3 大小調整、灰階化、直方圖均化	8
3.2 特徵擷取	8
3.2.1 灰階強度特徵	8
3.2.2 局部二元化模版	9

3.2.3	局部方向性模版	11
3.2.4	局部三元化模版	13
3.2.5	蓋伯濾波器	14
3.2.6	局部蓋伯二元化模版	15
3.3	降低特徵維度	16
3.3.1	減少取樣點	16
3.3.2	主成分分析	17
3.3.3	線性識別分析	17
3.4	辨識模型訓練	18
3.4.1	支持向量機	18
3.4.2	調昇式學習法	19
3.4.3	調昇式學習法結合支持向量機	21
第四章	研究結果與分析	24
4.1	資料庫簡介	24
4.1.1	PAL 人臉資料庫	24
4.1.2	FERET 人臉資料庫	26
4.1.3	MIR-Google 人臉資料庫	27
4.2	實驗參數設定	29
4.3	效能評估方式	30
4.4	性別辨識實驗	31
4.4.1	PAL 人臉資料庫實驗	31
4.4.2	FERET 人臉資料庫實驗	33

4.4.3	MIR-Google 人臉資料庫實驗	34
4.5	年齡辨識實驗	36
4.5.1	PAL 人臉資料庫實驗	37
4.5.2	FERET 人臉資料庫實驗	39
4.5.3	MIR-Google 人臉資料庫實驗	40
4.6	實驗錯誤分析	43
4.6.1	性別辨識錯誤分析	43
4.6.2	年齡辨識錯誤分析	44
第五章	結論與未來研究方向	46
參考文獻	47	



圖目次

圖 1.2-1 辨識系統流程圖	2
圖 3.1-1 前處理流程圖	6
圖 3.1-2 人臉偵測範例圖	7
圖 3.1-3 人臉五官概略分布圖	7
圖 3.1-4 人臉校準範例圖	8
圖 3.2-1 灰階值特徵範例圖	9
圖 3.2-2 各式模版範例圖	10
圖 3.2-3 擷取 LBP 特徵範例圖	10
圖 3.2-4 LBP 後紋理變化圖	11
圖 3.2-5 對應邊在八個方向反應的 Kirsch 遮罩	12
圖 3.2-6 像素計算 Kirsch 遮罩後的對應位置	12
圖 3.2-7 像素產生 LDP 編碼的範例 ($k = 3$)	13
圖 3.2-8 LDP 後紋理變化圖	13
圖 3.2-9 像素產生 LTP 編碼的範例 ($t = 5$)	14
圖 3.2-10 LTP 後紋理變化圖	14
圖 3.2-11 Gabor 濾波器擷取人臉特徵示意圖	15
圖 3.2-12 LGBP 擷取人臉特徵示意圖	16
圖 3.3-1 人臉影像比例縮小示意圖	17
圖 3.3-2 PCA 與 LDA 的主軸作用示意圖	18
圖 3.4-1 SVM 示意圖	19

圖 3.4-2 傳統的 AdaBoostSVM 與提出的 AdaBoostSVM 之比較圖	22
圖 3.4-3 提出的 AdaBoostSVM 訓練辨識模型之流程圖	22
圖 4.1-1 PAL 年齡直方統計圖	25
圖 4.1-2 PAL 資料庫範例圖	25
圖 4.1-3 FERET 年齡直方統計圖	26
圖 4.1-4 FERET 資料庫範例圖	27
圖 4.1-5 MIR-Google 年齡直方統計圖	28
圖 4.1-6 MIR-Google 資料庫範例圖	29
圖 4.2-1 用 PCA 搭配 LDA 做降維的示意圖	29
圖 4.4-1 PAL 性別訓練辨識率與迭代次數對應圖	31
圖 4.4-2 PAL 性別辨識率直方圖	32
圖 4.4-3 PAL 性別平均準確率均值直方圖	32
圖 4.4-4 FERET 性別訓練辨識率與迭代次數對應圖	33
圖 4.4-5 FERET 性別辨識率直方圖	34
圖 4.4-6 FERET 性別平均準確率均值直方圖	34
圖 4.4-7 MIR-Google 性別訓練辨識率與迭代次數對應圖	35
圖 4.4-8 MIR-Google 性別辨識率直方圖	36
圖 4.4-9 MIR-Google 性別平均準確率均值直方圖	36
圖 4.5-1 PAL 年齡訓練辨識率與迭代次數對應圖	37
圖 4.5-2 PAL 年齡辨識率直方圖	38
圖 4.5-3 PAL UTMVP 年齡評分直方圖	38
圖 4.5-4 FERET 年齡訓練辨識率與迭代次數對應圖	39

圖 4.5-5 FERET 年齡辨識率直方圖.....	40
圖 4.5-6 FERET 年齡平均準確率均值直方圖.....	40
圖 4.5-7 MIR-Google 年齡訓練辨識率與迭代次數對應圖.....	41
圖 4.5-8 MIR-Google 年齡辨識率直方圖.....	42
圖 4.5-9 MIR-Google 年齡平均準確率均值直方圖.....	42
圖 4.6-1 性別辨識之混淆矩陣.....	43
圖 4.6-2 性別辨識常見的錯誤類型.....	44
圖 4.6-3 年齡辨識之混淆矩陣.....	44
圖 4.6-4 年齡辨識常見的錯誤類型.....	45



表目次

表 3.4-1 AdaBoost.M2 演算法之流程	20
表 4.1-1 PAL 男女統計表	25
表 4.1-2 PAL 五類年齡範圍統計表	25
表 4.1-3 FERET 男女統計表	26
表 4.1-4 FERET 五類年齡範圍統計表	27
表 4.1-5 MIR-Google 男女統計表	28
表 4.1-6 MIR-Google 五類年齡範圍統計表	28
表 4.3-1 UTMVP 年齡評分表	30



第一章 緒論

1.1 研究動機

人臉的相關研究，一直是電腦視覺領域中相當重要的一部分。而人臉為主題的辨識，例如性別、年齡、表情、人臉等辨識，近年來，其相關的應用和競賽亦不斷廣泛的發展。例如智慧型手機上的應用程式，可以在拍照後立即估算出使用者的年齡，或是評斷出最像哪個明星；或是針對性別給予不同的廣告；等等都是年齡和性別辨識上的相關應用。此外，今年第八屆的由田機器視覺獎（Utech-zone Machine Vision Prize, UTMVP）[1]亦是國內最著名，以人臉為主題的辨識競賽，辨識的比賽項目分為性別、年齡、表情、人臉等四大辨識。

本論文的研究動機即是為了參加第八屆 UTMVP。研究將專注於性別和年齡辨識，預期能提出合理、有效的演算法設計，並且在實測時能有好的辨識效果



1.2 性別與年齡辨識系統

辨識系統的概略流程如圖 1.2-1 所示。本論文訓練和測試的資料包含一些 Paper 中有名的人臉資料，例如 FERET[2]、PAL[3]等等，也包含 MIR 實驗室從 Google Image Search 收集的資料。訓練和測試的資料同樣都會經過「前處理」、「特徵擷取」、「降低特徵維度」這三個步驟，接著即可透過訓練的人臉資料訓練出性別與年齡的辨識模型，接著再用測試資料評估性別和年齡的辨識結果。

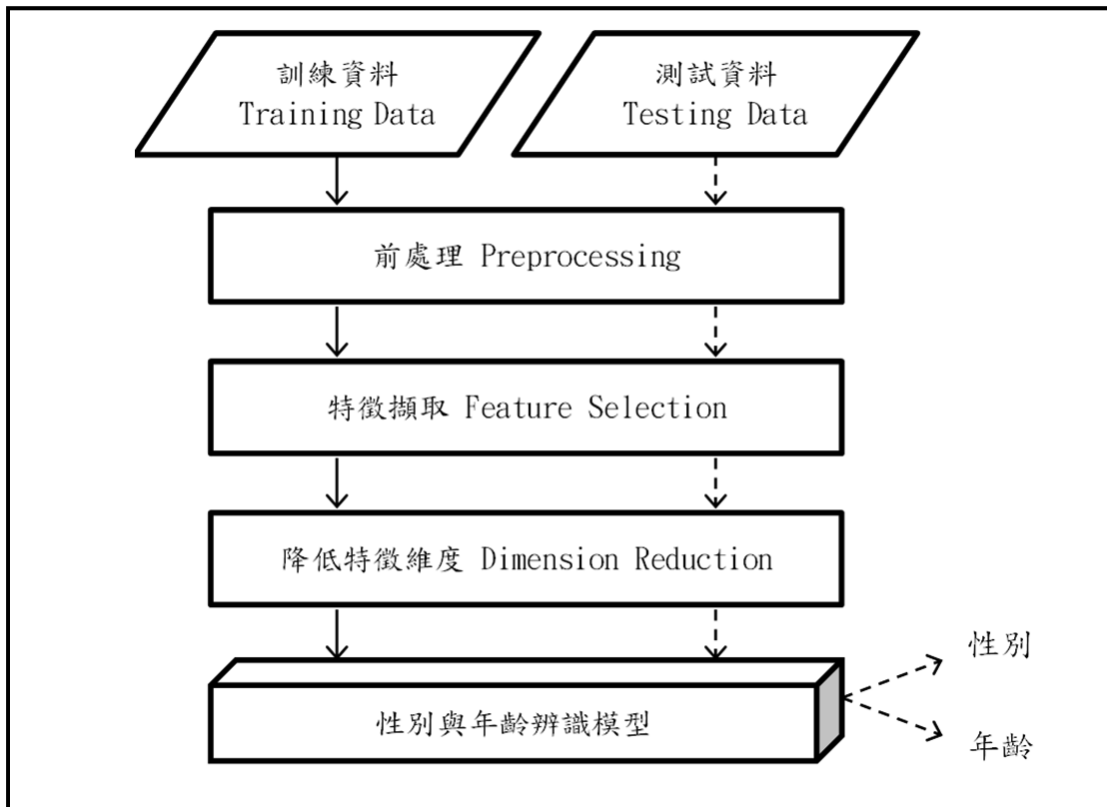


圖 1.2.1 辨識系統流程圖

1.3 研究方向

如 1.1 研究動機所述，此論文是為了參加第八屆 UTMVP，因此研究將會以比賽的規格和可能會碰到的問題為方向來進行。比賽分為程式競賽和論文報告。程式競賽中，主辦單位會隨機挑選人員至攝影機前，各隊必須在規定時間內提報所辨識人員的年齡、表情、人臉、性別。此外，主辦單位將會視情況增加複雜的背景，或是讓辨識人員戴帽子、戴眼鏡等方式增加干擾因素。因此，本論文所設計的演算法必須能克服背景、燈光、外物等干擾因素，且快速辨識出準確的性別和年齡。

1.4 章節概要

第二章將介紹論文之相關研究。

第三章將詳述論文使用的方法。將依序介紹前處理、特徵擷取、降低特徵維度所使用的方法，最後則會說明如何使用特徵篩選的調昇式學習法（Adaptive Boosting, AdaBoost）結合支持向量機（Support Vector Machine, SVM）來訓練性別和年齡的辨識模型。

第四章將會說明實驗的相關設定和方法，並分析實驗的結果。

第五章為本論文的結論以及未來研究方向。



第二章 相關研究簡介

性別辨識是人臉辨識領域上很具挑戰性的題目，其對於公司推銷廣告或是禁止異性進入某一區域等等都是很重要的應用。就像很多圖像辨識一樣，如何針對人臉進行特徵擷取以及後續的分類方法是做好性別辨識的關鍵所在。例如 Lian 即提出使用局部二元化模板 (Local Binary Pattern, LBP) 描述人臉特徵，並搭配支持向量機 (Support Vector Machine, SVM) 作為分類器[4]，實作性別辨識；Xia 則是利用局部蓋伯二元化模板 (Local Gabor Binary Pattern, LGBP) [5] 搭配支持向量機在性別辨識上得到好的表現。不過上述特徵擷取的方法，在人臉受到遮蔽物、燈光、表情等環境因素的影響時，容易使得性別辨識不穩定，因此，Jabid 提出了局部方向性模板 (Local Directional Pattern, LDP) [6] 有效的解決此問題。

人臉的年齡辨識亦在近年開始受到重視，各式相關的應用在人機介面、居家照護和安全監控等領域上，展現相當大的發展潛力。不過目前人臉的年齡辨識仍舊是個具有挑戰性的問題，有許多因素會影響到評估答案的準確度，例如欲評估的人臉受到遮蔽物、燈光、表情等環境因素的影響；又或者是一個人人在遺傳上或是臉部保養上的差異，造成有些人長得比實際年齡年輕或是蒼老。因此，年齡的辨識基本上無法精確到一個人的實際年齡，其評估的方式通常是將人的年齡分為不同的範圍，例如分為小孩、壯年人、老人這三個範圍，而需要切分的範圍則是依據應用的種類來決定。如此一來，年齡辨識就可以看作是一項分類的問題，而怎麼描述人臉上有用的特徵，以及分類的方式變成此問題的關鍵所在。

我們通常可以由人臉的五官變化、皺紋紋理或是皮膚性質來判斷一個人大約的年紀，許多方法也從上述這幾點切入來描述人臉上的年齡特徵。Gunay 使用局部二元化模板 (Local Binary Pattern, LBP) 來描述人臉上的皮膚特徵[7]，並在 FERET[2] 資料下以最近鄰居法 (k-nearest neighbor, KNN) 得到了不錯的分類效果；Feng 則是使用蓋伯濾波器 (Gabor Filter) 抽取人臉上的紋理方向特徵[8]，搭配以

線性識別分析 (Linear Discriminant Analysis, LDA) 作為的分類器，在評估消費者年齡的實驗上有好的效果；Caifeng 亦使用 LBP 和 Gabor Filter 作為描述人臉特徵的方法，不同的是他另外先使用調昇式學習法 (Adaptive Boosting, AdaBoost) 挑選有用的特徵，再搭配支持向量機 (Support Vector Machine, SVM) 作為分類器 [9]。

從上述我們可以發現到如何描述描述人臉特徵以及如何對於特徵進行分類對於性別和年齡辨識的重要性。而本論文提出的方法將會以上述各個著名的人臉特徵擷取，搭配調昇式學習法結合支持向量機 (AdaBoostSVM) [10] 作為分類器，針對性別和年齡辨識，訓練出效果較為優良的辨識模型。



第三章 研究方法

以下將依照圖 1.2-1 辨識系統流程圖，逐一介紹「前處理」、「特徵擷取」、「降低特徵維度」、「辨識模型訓練」所使用的方法。

3.1 影像前處理

前處理主要分為五個步驟，分別為人臉偵測 (Face Detection)、人臉校準 (Face Calibration)、大小調整 (Resize)、灰階化 (De-colorization) 及直方圖均化 (Histogram Equalization)。每個步驟都有它的用處存在，而整個流程的主要目的是讓訓練和測試的影像能在同一個基準上。前處理流程如下圖 3.1-1 所示：

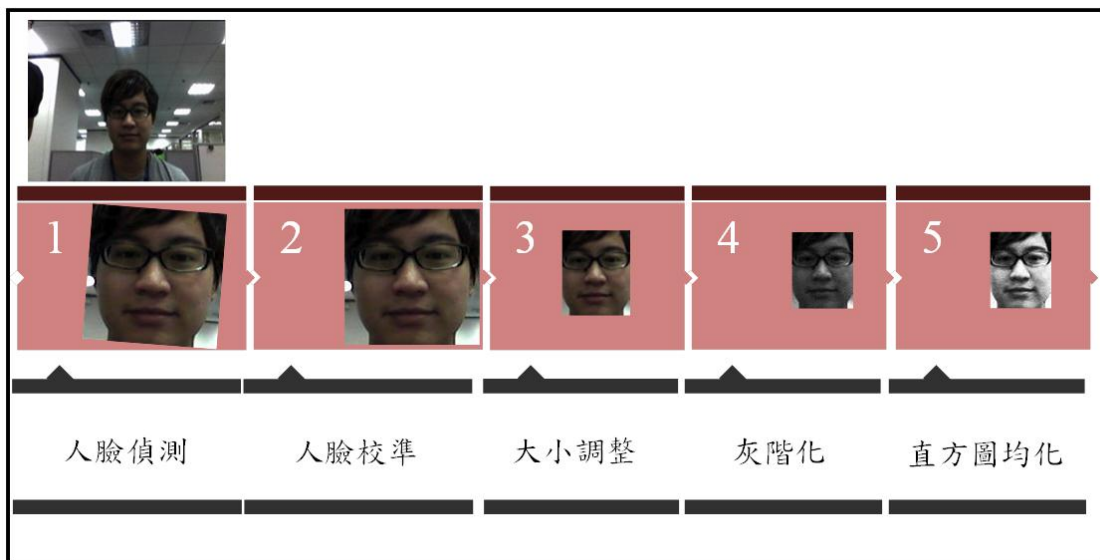


圖 3.1-1 前處理流程圖

3.1.1 人臉偵測

人臉偵測主要是利用 Viola 和 Jones 提出的 AdaBoost 演算法[11]，從 Harr-like 特徵資料庫中分類出最能代表人臉特徵的一組特徵群，再利用此組特徵群找人臉的依據，將人臉從影像中標示出來（如圖 3.1-2 所示）。如此將可以去除不重要的

背景資訊。



圖 3.1-2 人臉偵測範例圖



3.1.2 人臉校準

如圖 3.1-3 所示，人臉校準[12]主要是利用人的雙眼連線幾乎為水平的特性，即可透過雙眼連線的角度，作為校準的依據。如此將可以降低因人臉旋轉造成的誤差。人臉校準範例則如圖 3.1-4。

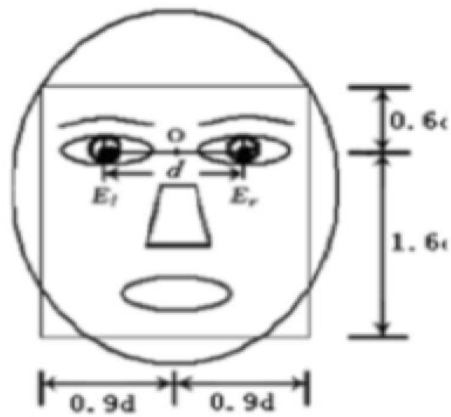


圖 3.1-3 人臉五官概略分布圖



圖 3.1-4 人臉校準範例圖

3.1.3 大小調整、灰階化、直方圖均化

前處理中，我們統一將訓練和測試影像的大小調整為：長寬 76*57，如此不僅能減小影像中的雜訊，也能加快後面步驟的處理。而特徵擷取主要是運作在灰階化的影像上，所以將影像灰階化也是我們的必要步驟。直方圖均化則是加強影像的對比，能減少影像因環境光線的強弱造成的誤差。

3.2 特徵擷取

在人臉資訊辨識的特徵擷取上，多採用局部性的紋理特徵 (Local Texture Feature)，例如表情辨識，高興常表露出牙齒的微笑，和驚嚇則有張開大嘴的特性等等，都是可以透過產生人臉上每一個區塊局部性的紋理特徵，串成一個特徵向量來做分類，進而達到辨識的效果，以下各小節將介紹常用於描述人臉紋理特徵的方法。

3.2.1 灰階強度特徵

由影像的本身灰階強度 (Grayscale) 作為特徵，其擷取的方式是將彩色圖

像的 RGB 轉換到 YUV 的維度上，而 Y (Luminance) 即為灰階值，其值將落在 0 到 255 間。轉換的公式如下：

$$Grayscale = 0.299 * R + 0.587 * G + 0.114 * B$$

灰階值特徵的範例圖 3.2-1 如下：

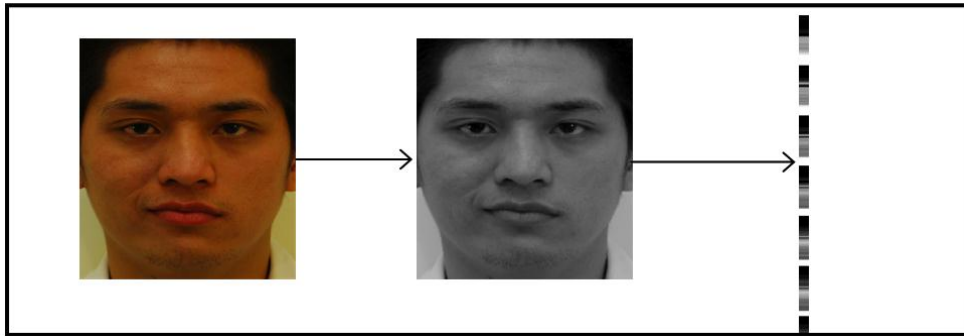


圖 3.2-1 灰階值特徵範例圖

3.2.2 局部二值化模版

局部二值化模版 (Local Binary Pattern, LBP) 是由 T. Ojala 提出[13]，已在多篇論文中，被證實在描述影像的特徵上有良好的表現，後來則常應用於人臉辨識[14]或年齡評估[15]等分類問題。其方法在於對每張影像像素指定一個材質標籤，此標籤是根據影像每個像素點周圍圓半徑 R 上 P 個鄰居像素點所編碼而成，由參考點 (圖 3.2-2 中菱形方塊) 與其所有鄰居點 (圖 3.2-2 中圓形點) 比較像素強度大小取得一個二值化串列，再乘上 2^P 的權值以表示不同的特徵點，公式如下：

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) \cdot 2^p$$

$$s(g_p - g_c) = \begin{cases} 1, & \text{if } g_p - g_c \geq 0 \\ 0, & \text{if } g_p - g_c < 0 \end{cases}$$

g_c 為 LBP 參考點之影像像素值， g_p 為參考點鄰居所有像素值，其中 p 介於 $0 \sim P-1$ 。

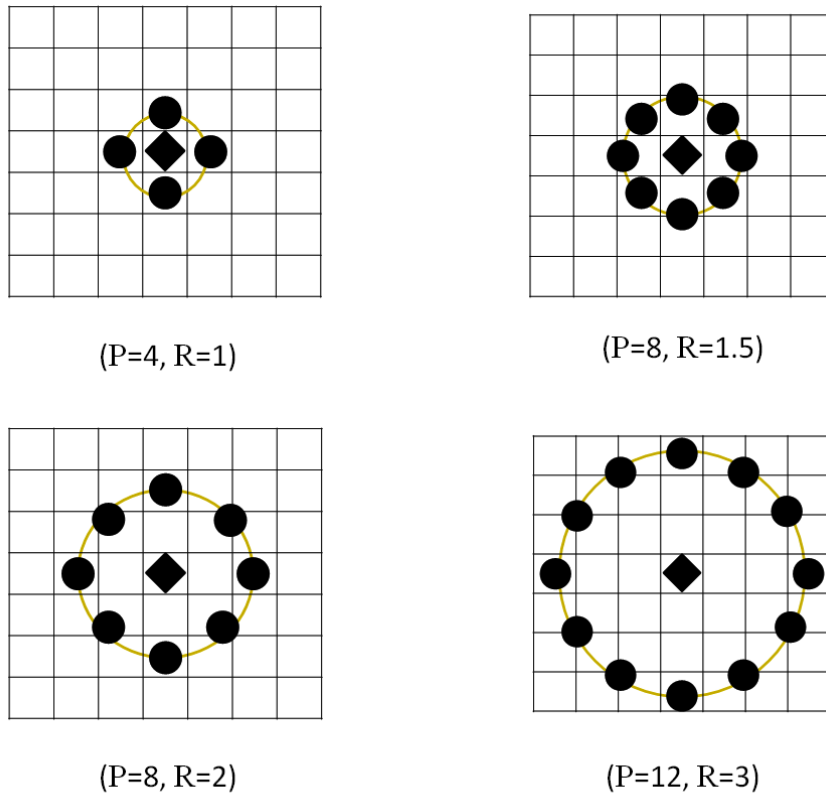


圖 3.2.2 各式模版範例圖

下圖 3.2-3 是以 P=8, R=1.5 的擷取 LBP 特徵的範例，圖 3.2-4 則是說明一張影像經過 LBP 後材質的變化。

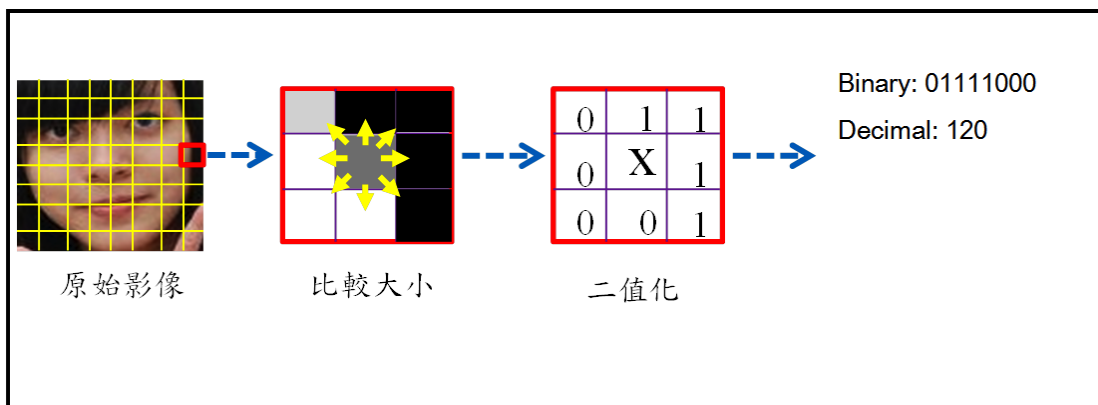


圖 3.2-3 擷取 LBP 特徵範例圖



原始影像



LBP 影像

圖 3.2-4 LBP 後紋理變化圖

3.2.3 局部方向性模版

LBP 提出後，衍生出許多類似概念的局部模版來作特徵擷取。相較於 LBP 只對周圍的梯度大小變化做編碼，沒有將方向性納入作為考量，T. Jabid 等人提出局部方向性模版 (Local Directional Pattern, LDP) [6]，能夠計算各個像素的邊在不同方向的反應值，並將紋理編碼。相較於 LBP，LDP 對於影像在有雜訊或是非單一光源的情形下，能夠擷取出更為穩定的特徵值。

LDP 是一個 8-bits 的二進位碼，這個特徵計算的方式主要是比較一個像素中，其對應的邊在不同方向的反應值。其方法為使用 Kirsch 遮罩 (圖 3.2-5)，各遮罩代表八個方向 ($M_0 \sim M_7$) 的描述方式，並能計算出一個像素中以像素為中心計算，代表八個方向的反應值。

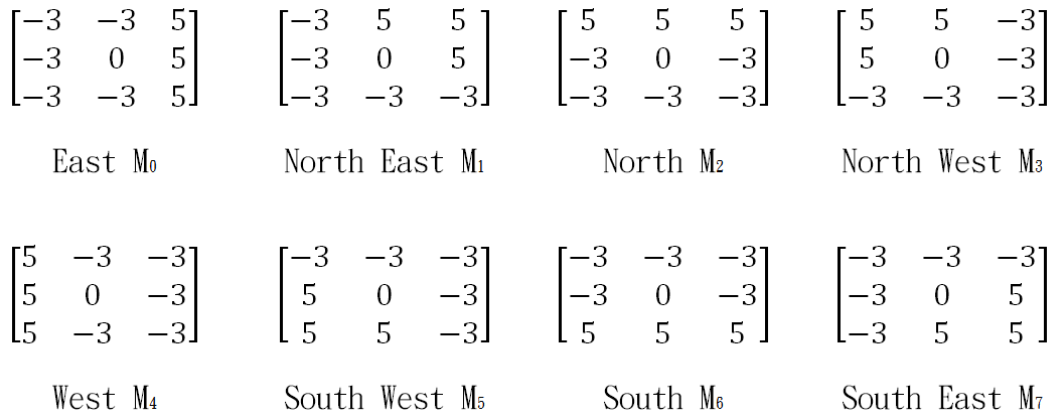


圖 3.2-5 對應邊在八個方向反應的 Kirsch 遮罩

圖 3.2-6 說明了一個像素計算完 $M_0 \sim M_7$ 遮罩後，可以得到相對應值 $m_0 \sim m_7$ ，以及各個值對應的位元位置 $b_0 \sim b_7$ ，每個值代表邊在其對應方向上的重要性。此外，在做 LDP 特徵擷取時，必須決定 k 個主要的方向來產生 LDP 編碼，方法為找出前 k 大的值 $|m_i|$ ，並將其對應的位元位置設為 1，其餘位置則設為 0。如下是 LDP 編碼計算的公式， m_k 為計算完 Kirsch 遮罩的第 k 大值：

$$LDP_k = \sum_{i=0}^7 b_i (m_i - m_k) * 2^i, \quad b_i(a) = \begin{cases} 1, & a \geq 0 \\ 0, & a < 0 \end{cases}$$

如下圖 3.2-7 的範例是以 $k = 3$ ，3 個主要的方向來產生 LDP 編碼。

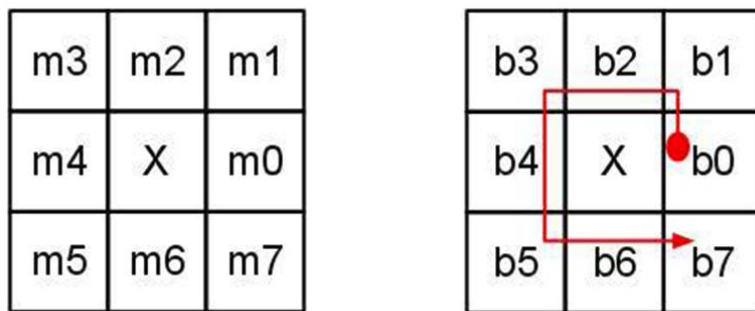


圖 3.2-6 像素計算 Kirsch 遮罩後的對應位置

85	32	26	→	Mask Index	M7	M6	M5	M4	M3	M2	M1	M0
53	50	10		Mask Value	-303	97	161	537	313	97	-503	-399
60	38	45		Rank	5	8	6	1	4	7	2	3
				Code bit	0	0	0	1	0	0	1	1
				LDP code	19							

圖 3.2-7 像素產生 LDP 編碼的範例 ($k = 3$)



原始影像



LDP 影像

圖 3.2-8 LDP 後紋理變化圖

3.2.4 局部三元化模版

針對 LBP 容易受到雜訊影響的問題, Tab 和 Bill 提出了局部三元化模版(Local Ternary Pattern, LTP) [16], Khoa 亦用 LTP 在年齡評估上得到好的效果[17]。其作法主要是將中心的像素質與其鄰近點的像素質比大小, 大小的關係分為三種情況: 大於、小於、接近, 其對應的 LTP 值分別為 1、-1、0。公式如下:

$$LTP(i) = \begin{cases} 1 & p(i) - p(0) > t \\ 0 & \text{if } |p(i) - p(0)| \leq t \\ -1 & p(i) - p(0) < t \end{cases}$$

其中 $p(0)$ 是中心點的像素值, $p(i)$ 為目前中心點鄰近的像素值, t 則是選定的門檻值。下圖 3.2-9 是以 $t = 5$ 的 LTP 編碼範例圖

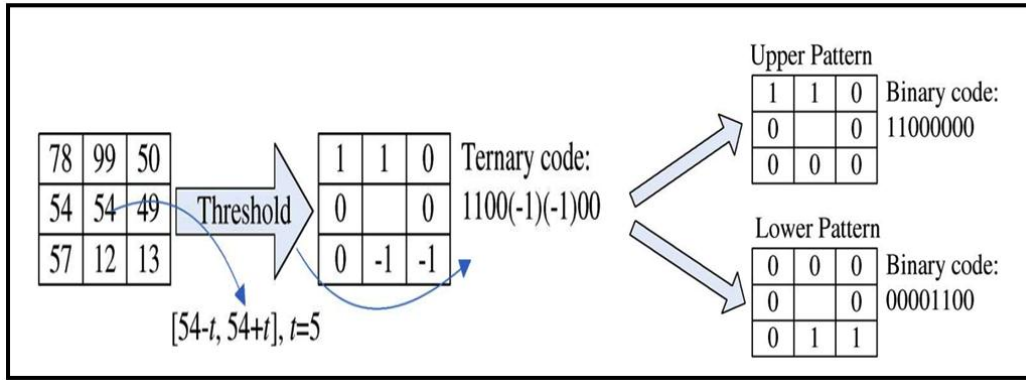


圖 3.2-9 像素產生 LTP 編碼的範例 ($t = 5$)

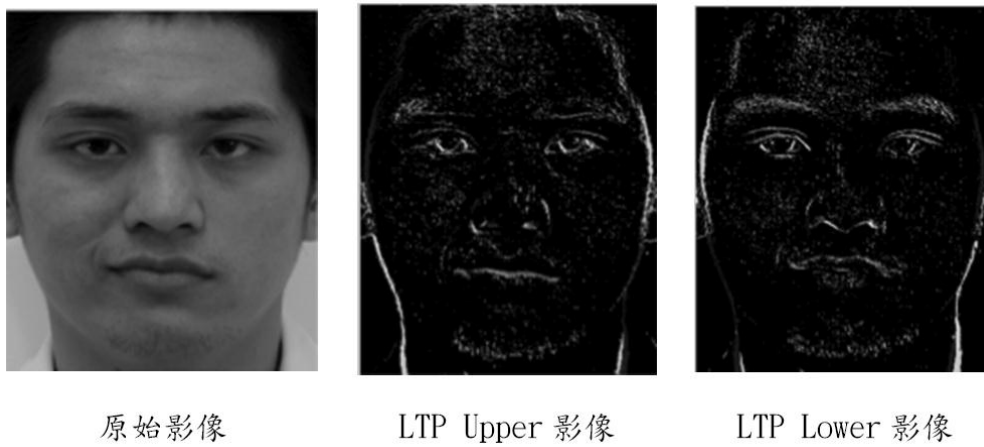


圖 3.2-10 LTP 後紋理變化圖

3.2.5 蓋伯濾波器

蓋伯濾波器 (Gabor Filter) 已經被證實能有效的描述紋理的性質[18]，且對於人臉特徵的描述也有很好的作用。Gabor 可提供影像結構中有關方向性、空間性、頻率性的最大可能解析度，同時還能做到 2-D 影像位置的同步。其濾波器的轉換式可以下公式表示：

$$\psi(x, y, \omega, \theta) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x'^2+y'^2}{2\sigma^2}\right)} \left[e^{i\omega x'} - e^{-\frac{\omega^2\sigma^2}{2}} \right]$$

$$x' = x\cos\theta + y\sin\theta, y' = -x\sin\theta + y\cos\theta$$

其中 (x, y) 是欲計算像素的位置， ω 是徑向中心的頻率， θ 則代表 Gabor 濾波器的

方向。在擷取人臉特徵時，通常我們會選用 5 個頻率和 8 個方向，如下式：

$$\omega_m = \omega_{max} \times \lambda^{-(m-1)}, m = \{1, 2, 3, 4, 5\}, \lambda = \sqrt{2}, \omega_{max} = \pi/2$$

$$\theta_n = (n - 1)\pi/8, n = \{1, \dots, 8\}$$

如下圖 3.2-11 為 Gabor 濾波器選用 5 個頻率和 8 個方向擷取人臉特徵的示意圖：

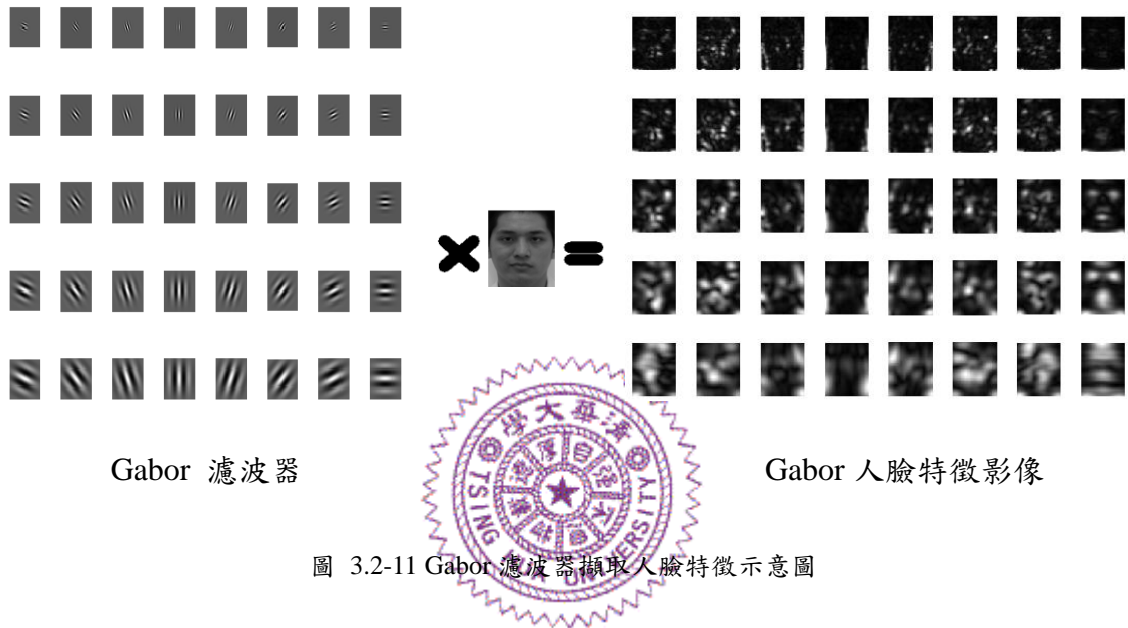


圖 3.2-11 Gabor 濾波器擷取人臉特徵示意圖

3.2.6 局部蓋伯二元化模版

基於 LBP 和 Gabor 濾波器的基礎上，Zhang 提出了局部蓋伯二元化模版(Local Gabor Binary Pattern, LGBP)的概念[19]，對於人臉特徵不僅具有更佳的描述能力，亦能有效減少燈光、遮蔽物、人臉角度等等雜訊的影響。其作法是先將一張影像進行 Gabor 濾波器轉換出多張 Gabor 特徵影像，接著將這些 Gabor 特徵影像套用 LBP 再次擷取特徵，如此即可得到 LGBP 特徵影像。如下圖 3.2-12 為選用 5 個頻率和 8 個方向的 Gabor 濾波器結合 LBP 的示意圖：

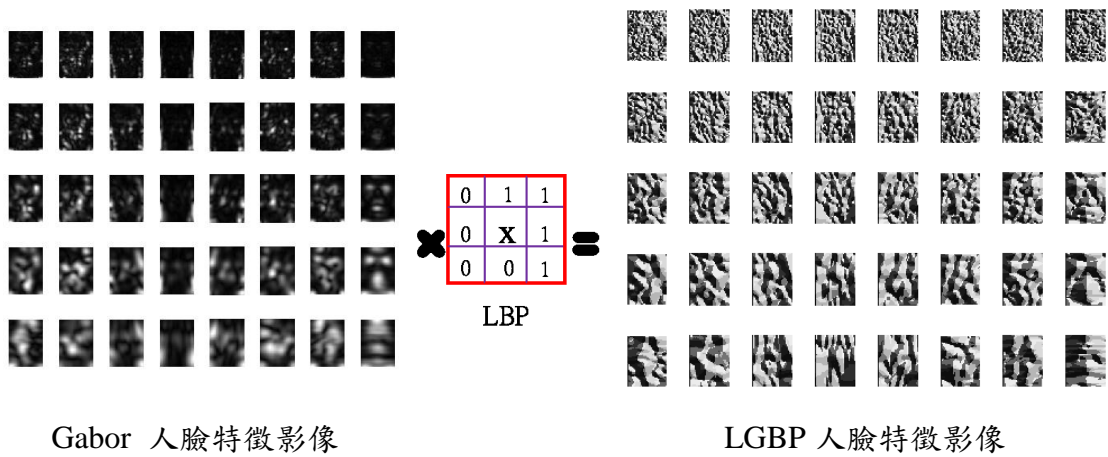


圖 3.2-12 LGBP 擷取人臉特徵示意圖

3.3 降低特徵維度

降低特徵維度不僅能計算更為方便、運算時間加快，甚至還有機會讓辨識的效果變好。目前對影像降維的方法有很多，有最簡單的減少影像取樣點 (Down-sampling)，還有近期主流的主成分分析 (Principal Component Analysis, PCA) 以及線性識別分析 (Linear Discriminant Analysis, LDA)。以下將逐一做介紹。

3.3.1 減少取樣點

本論文減少取樣點的作法主要是將人臉影像等比例縮小。示意圖如圖 3.3-1 所示。縮小的影像解析度會因此降低，但取樣點也會跟著減少。

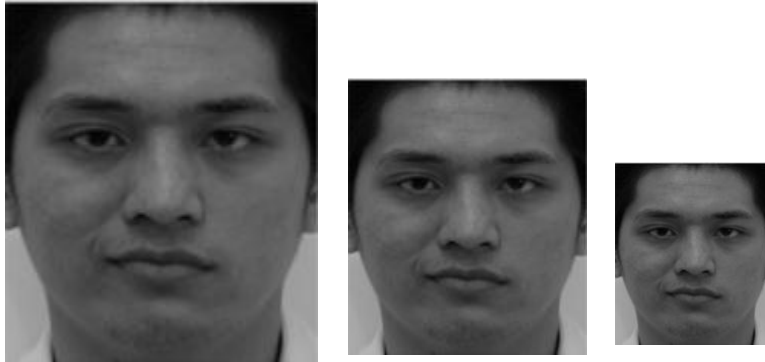


圖 3.3-1 人臉影像比例縮小示意圖

3.3.2 主成分分析

主要向量分析 (Principal components analysis, PCA) [20] 是利用統計的方法，將高維度的資料轉換為少數幾個互相獨立的線性組合。經由線性組合而成的主成分，其變異數最大，且所保留的資訊將最高。許多人臉分類的研究[21][22]，在擷取了高維度的特徵 (例如：Gabor 濾波器特徵) 後，採用了 PCA 作為降維和特徵萃取的方法。

3.3.3 線性識別分析

線性識別分析 (Linear Discriminant Analysis, LDA) [23] 是一種參數型的特徵萃取方法。在 LDA 中，組間分散矩陣以及組內分散矩陣是用來評估類別分散程度的準則。使用時，訓練資料的樣本數必須大於各訓練資料的維度數。LDA 與 PCA 相同，也常在人脸特徵上作為降維和特徵萃取使用，兩者主要的差別和作用可以參考圖 3.3-2。如圖所示，PCA 轉換後的特徵空間主軸，是沿著樣本資料變異最大之方向，轉換後所得的特徵是最具代表性的特徵；而 LDA 則是擁有最小組內差異與最大組間間距的特性，因此透過 LDA 轉換後的特徵空間主軸，相較

PCA 較具有分群的效果，所得的特徵是最具鑑別性的特徵。

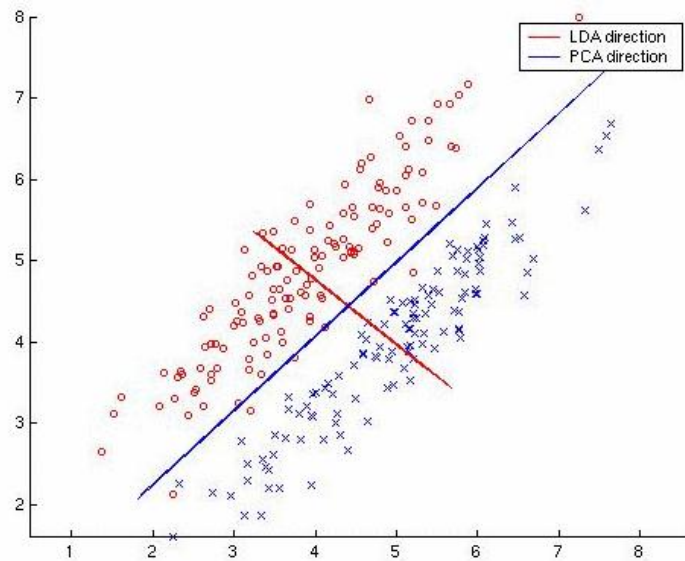


圖 3.3-2 PCA 與 LDA 的主軸作用示意圖

3.4 辨識模型訓練

如第二章文末所敘，本論文提出的新方法將利用 Xuchun 所提出的調昇式學習法結合支持向量機(AdaBoostSVM)[10]作為分類器，並透過 AdaBoost 的特性，在訓練性別與年齡辨識模型的過程中，進行人臉的特徵選取。以下將分別介紹其相關的方法。

3.4.1 支持向量機

支持向量機 (Support Vector Machine, SVM) 是一種由統計學習理論衍生的機器學習理論，由 Vapnik 及其研究團隊所提出[24]。SVM 的概念主要是將所有分類資料樣本視為在一高度平面上的點，期望從中找到一個最能區分資料樣本的決策平面 (Hyperplane)。以圖 3.4-1 舉例來說，若是在二維平面上有兩種不同類別的資料 (方塊與實心圓)，SVM 將會找出一條最能區分兩類的決策邊界 (紅色實線) 或可稱為最佳區分超平面 (Optimal Separation Hyperplane)，並希望這條實

線與兩類資料集的邊界，其之間的距離（Margin）能愈大愈好，如此將能清楚地區分出兩個不同的類別。

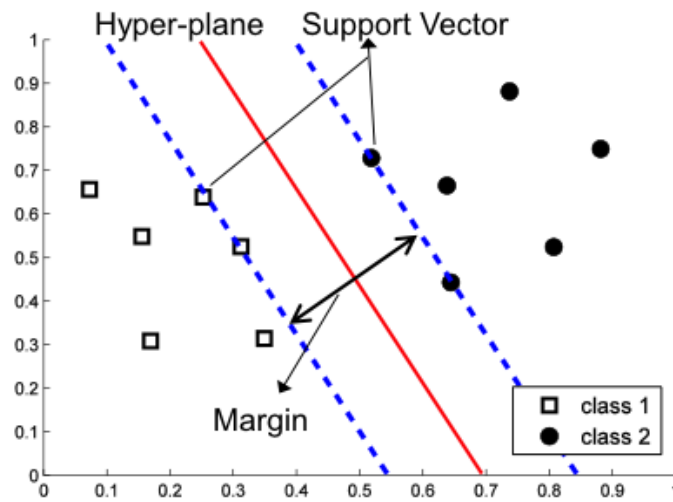


圖 3.4-1 SVM 示意圖

不過在現實世界中並不是所有的資料都可以像上述例子一樣找到線性地最佳區分超平面，此時必須將原始資料透過非線性的映射函數轉換到另外一個較高維度的特徵空間，然後在特徵空間中執行線性分類。這時特徵維度轉換後的 SVM 將被稱為非線性支援向量機，其常用的核心函數有線性（Linear）、多項式（Polynomial）、放射（Radial Basis Function, RBF）和 S 型（Sigmoid）等四種。而本論文只要是使用 RBFSVM，其核心函數如下：

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$$

其中， σ 為一個可調整之參數。

3.4.2 調昇式學習法

調昇式學習法（Adaptive Boosting, AdaBoost）主要是將多個弱學習（Weak Learner）不同的分類假設（Hypothesis），結合出效果較強的分類器（Strong Classifier）[25]。不過原始的 AdaBoost 僅適用於二元的分類上，因此 Freund 和

Schapire 再次提出了 AdaBoost.M2 來處理多元的分類問題，而此法亦是目前 AdaBoost 演算法中，最被廣泛使用的多元分類法。本小節將以 AdaBoost.M2 作介紹，此演算法的詳細流程如表 3.4-1 所示。

表 3.4-1 AdaBoost.M2 演算法之流程

AdaBoost.M2 演算法
<p>輸入：訓練資料 $L = [(x_1, y_1), \dots, (x_N, y_N)]$，其中資料樣本 $x_i \in X$ 與其對應的標籤 $y_i \in Y = \{1, \dots, K\}$。迭代數為 $t = [1, \dots, T]$。</p> <p>初始：樣本權重 $w_i^1 = 1/N$，$i = [1, \dots, N]$，N 為樣本個數。</p> <p>迭代：$t = 1, \dots, T$</p> <ol style="list-style-type: none"> (1) 利用弱分類器演算法訓練得到各個 $h_t: X \times Y \rightarrow [0,1]$ (2) 以弱分類器 h_t 計算各個分類錯誤率： $\varepsilon_t = \frac{1}{2} \sum_{i=1}^N \sum_{y_k (\neq y_i)} w_i^t (1 - h_t(x_i, y_i) + h_t(x_i, y_k))$ (3) 記錄錯誤率最小的 ε_t，以及此時的 h_t (4) 若是 $\varepsilon_t \geq 1/2$，設定 $T = t - 1$ 後，跳離迭代 (5) 設定 $\alpha_t = \ln(\frac{1-\varepsilon_t}{\varepsilon_t})$ (6) 更新樣本權重 w^{t+1} (Z_t 為正規化常數)： $w_i^{t+1} = \frac{w_i^t}{Z_t} \exp(-\alpha_t \frac{1}{2} (1 + h_t(x_i, y_i) - h_t(x_i, y_k)))$ <p>輸出：最後將多個弱分類器組合為強分類器：</p> $H(x) = \operatorname{argmax}_{y \in Y} \sum_{t=1}^T \alpha_t h_t(x, y)$

AdaBoost.M2 演算法需要以附有標籤資訊的訓練資料 $(x_1, y_1), \dots, (x_N, y_N)$ 作

為輸入，其中 $x_i \in X$ 代表大小為 N 的資料樣本集中的第 i 個樣本，而 $y_i \in Y = \{1, \dots, K\}$ 為第 i 個樣本在 K 個類別中所對應的正確類別標籤。對於第一次的迭代訓練而言，此權重將設定為平均分配： $w_i^1 = 1/N$ 。接著，以自行定義的弱分類器演算法來訓練出各個 $h_t: X \times Y \rightarrow [0,1]$ ，其中 $h_t(x_i, y_i)$ 代表 x_i 準確分到其對應標籤 y_i 的信心度大小，其值介於 0 到 1 間，反之， $h_t(x_i, y_k)$ 則代表 x_i 分到其餘錯誤標籤 y_k 的信心度大小，值一樣介於 0 到 1 間。有了弱分類器以後，我們即可計算每次迭代的分類錯誤率 ε_t ，其中 $(1 - h_t(x_i, y_i) + h_t(x_i, y_k))$ 項的值域落於 0 到 2 間，當其值大於 1 時表示 x_i 分到錯誤標籤 y_k 的信心度大於分到正確標籤 y_i ，此時代表某個弱分類器 h_t 的分類錯誤率 ε_t 將因此變大。接著即可搭配權重 w_i^t 選出最小的分類錯誤率 ε_t 及其對應的 h_t ，不過若此時的 ε_t 大於 $1/2$ 則必須離開迭代，並將變數還原為上一次迭代的結果。設定完 α_t 後，即可算出下一階段的樣本權重 w^{t+1} ，透過式子我們可以發現被分錯的樣本將得到較高的權重，反之，分對的樣本權重則較低。如此一來隨著迭代次數的增加，那些不容易被分對的訓練樣本其被取樣的權重值將漸漸增加，使得弱分類器將著重處理這些難分對的訓練樣本。最後，每個迭代階段被選出具有互補能力的弱分類器將結合為最終的強分類器 $H(x)$ 。

3.4.3 調昇式學習法結合支持向量機

調昇式學習法結合支持向量機 (AdaBoostSVM) 是由 Xuchun 所提出[10]，對於處理資料樣本不平均的強況有很好的效果。其主要原理是使用 RBF SVM 分類器作為 AdaBoost 中的組合分類器 (Component Classifier)，而這些組合分類器的差別在於不同的 σ 值設定。而本論文將使用[10]中組合分類器的參數設定方式，加上我們提出的特徵選取步驟，作為訓練性別和年齡辨識模型所使用的方法，兩方法的差異如下圖所示：

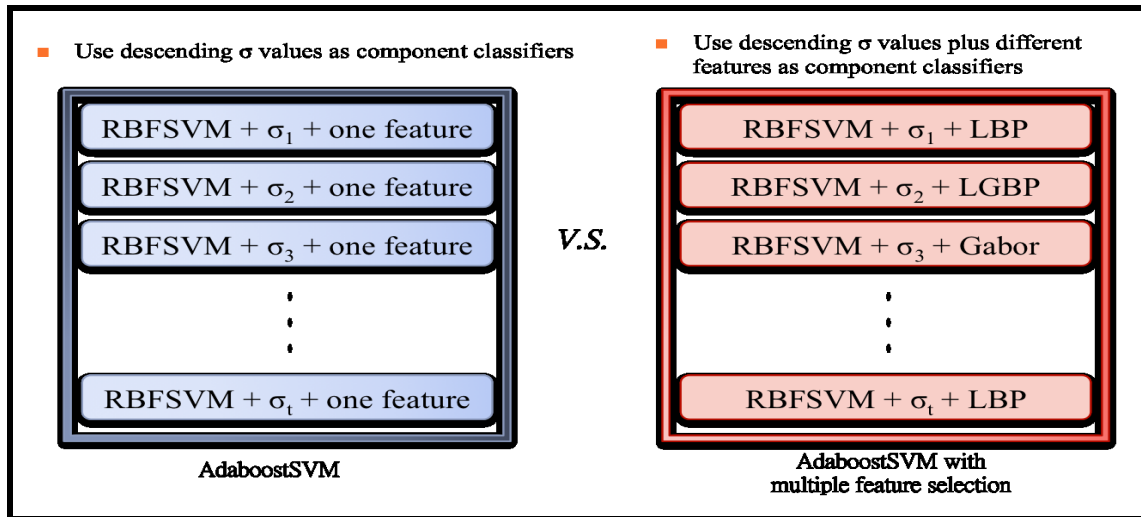


圖 3.4-2 傳統的 AdaBoostSVM 與提出的 AdaBoostSVM 之比較圖

如比較圖所示，傳統 AdaBoostSVM 的辨識模型主要是由很多 RBFSVM 分類器來組成，每個分類器的差別在於不同的 σ ， σ 將依照分類器的順序依序做遞減；而我們的方法跟它最大的不同在於我們各個階段的分類器會挑選出最適合它的人臉的特徵（LBP、LDP、LTP、Gabor、LGBP）。如此，透過 RBFSVM 搭配不同的 σ 和特徵，將可以組合出我們最終的辨識模型。詳細的流程圖如下所示：

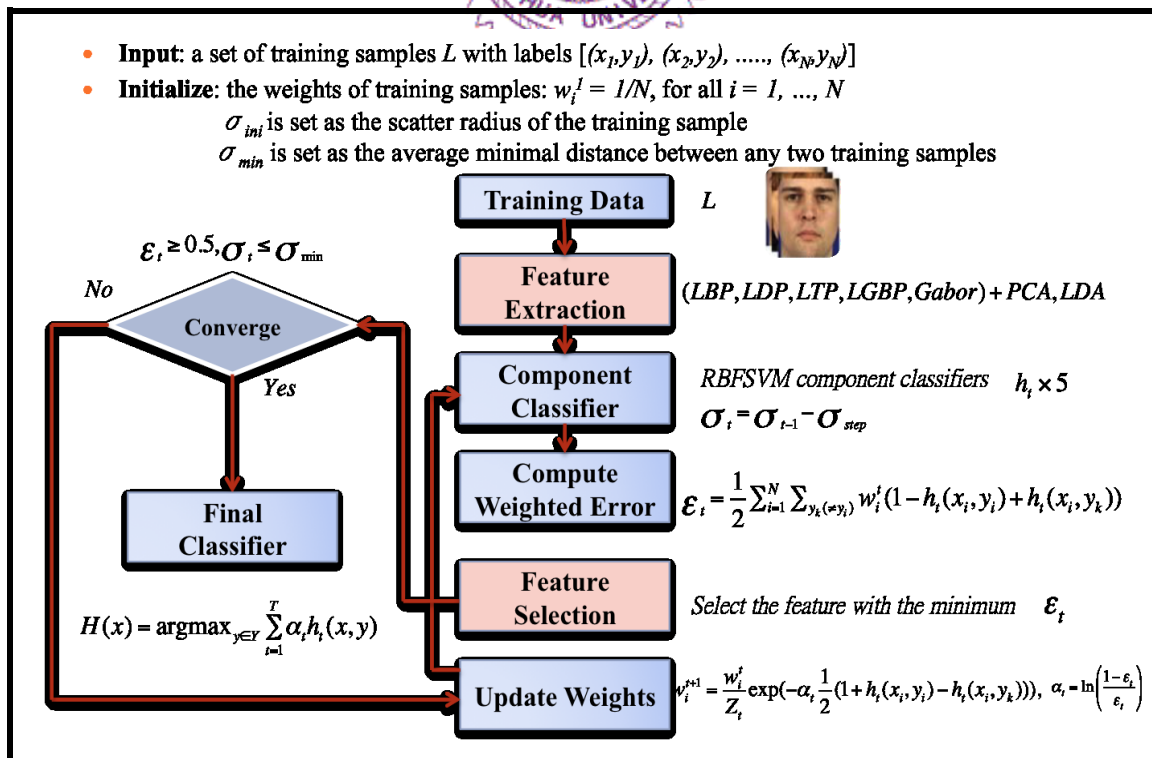


圖 3.4-3 提出的 AdaBoostSVM 訓練辨識模型之流程圖

如流程圖所示，藍色區塊為傳統 AdaBoostSVM 的流程，紅色區塊則是此論文增加的步驟。當我們以附有標籤資訊的 N 個訓練資料 $(x_1, y_1), \dots, (x_N, y_N)$ 作為輸入，其中這些 x 代表人臉資訊，而 y 則是每個人臉所對應的年齡範圍。一開始我們先將每個訓練資料的 w_i^1 初始為 $1/N$ ， σ_{ini} 初始為訓練資料彼此特徵值的最大距離， σ_{min} 則是初始為訓練資料每個兩兩特徵值的平均距離。接著對這些人臉進行特徵選取，特徵分別有 LBP、LDP、LTP、LGBP、Gabor，並以 PCA 搭配 LDA 對特徵進行降維，而每一輪 RBFSVM 的 σ 將逐漸遞減，如此一來每個迭代過程中將有 5 個分類器 h_t 可供比較其這一階段的分類錯誤率 ε_t ，並記錄此時錯誤率較小的 h_t 所對應的特徵，此對應的特徵就是這階段篩選出最好的特徵。若此時 ε_t 小於 $1/2$ 則可繼續下一步驟，否則必須離開迭代。再來即可以此 h_t 更新各個樣本的權重，繼續迭代的進行直到迭代次數終了， ε_t 大於 $1/2$ ，或是 $\sigma_t \leq \sigma_{min}$ 。



第四章 研究結果與分析

本論文將在這章比較提出的 AdaBoostSVM 模型與其他方法的性別和年齡辨識效果。首先我們將介紹使用的資料庫，接著將會說明實驗上的參數設定，以及效能評估的方法，最後則是分析實驗的辨識結果。以下實驗的辨識類別將依照 UTMVP 的規範。性別辨識分為男生和女生共兩類，年齡辨識則分為五類，分別是 0 到 20 歲、21 到 30 歲、31 到 40 歲、41 到 50 歲、50 歲以上。

4.1 資料庫簡介

首先我們先介紹實驗使用的三個資料庫，分別為 PAL 人臉資料庫、FERET 人臉資料庫、MIR-Google 人臉資料庫。

4.1.1 PAL 人臉資料庫



PAL (Productive Aging Laboratory) 資料庫是由 Minear 和 Park 在密西根大學收集而來用為學術使用[3]。這個資料庫具有大量的人臉影像，且年齡涵蓋了 19 到 93 歲的範圍，此外，它不只提供一般的中立 (Neutral) 影像，亦提供不同表情的人臉影像。而在本論文的實驗中，我們將使用其提供的中立影像來做辨識，此中立影像總共包含 578 張不同的人臉，其影像張數以 10 歲為區間的直方統計圖 4.1-1 如下所示，其中 X 軸為年齡區間，Y 軸為張數。而依照 UTMVP 規格分為男女兩類和五類年齡範圍的統計表格如表 4.1-1 和表 4.1-2 所示。圖 4.1-2 則為 PAL 資料庫範例圖。

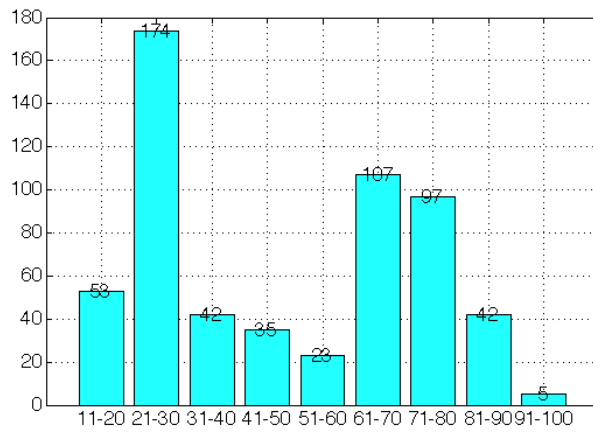


圖 4.1-1 PAL 年齡直方統計圖

表 4.1-1 PAL 男女統計表

性別	數量
男	227
女	351

表 4.1-2 PAL 五類年齡範圍統計表

年齡範圍	數量
0~20	53
21~30	174
31~40	42
41~50	35
51-	274



圖 4.1-2 PAL 資料庫範例圖

4.1.2 FERET 人臉資料庫

FERET(The Face Recognition Technology)人臉資料庫[2]源自 1993 年到 1997 年間，美國國防部的一項以自動影像識別為基礎的智慧型監控系統計畫，而 FERET 則用來提供監控測試以及演算法評估。此資料庫有著相當豐富的人臉影像，參與的拍照人數共有 1208 人，但每人拍攝的條件和張數不盡相同。而在此實驗中，我們使用具有正面人臉的影像 (Regular Frontal Face, fa)，一共 991 張不同人臉的影像，其影像張數以 10 歲為區間的直方統計如下圖 4.1-3 所示，其中 X 軸為年齡區間，Y 軸為張數。而依照 UTMVP 規格分為男女兩類和五類年齡範圍的統計表格如表 4.1-3 和表 4.1-4 所示。圖 4.1-4 則為 FERET 資料庫範例圖。

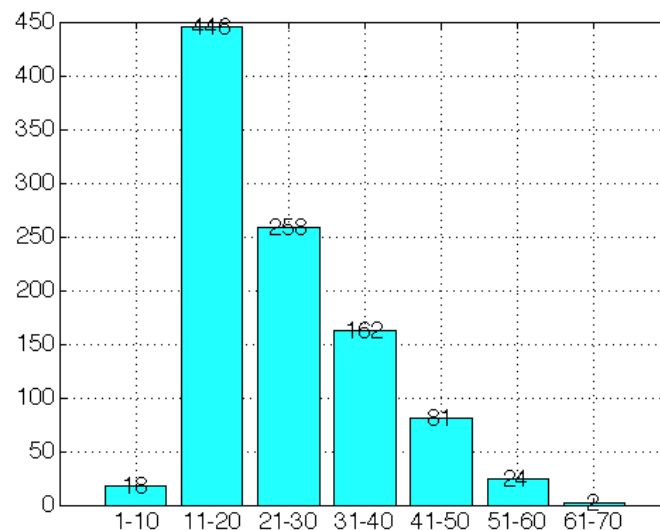


圖 4.1-3 FERET 年齡直方統計圖

表 4.1-3 FERET 男女統計表

性別	數量
男	589
女	402

表 4.1-4 FERET 五類年齡範圍統計表

年齡範圍	數量
0~20	464
21~30	258
31~40	162
41~50	81
51-	26



圖 4.1-4 FERET 資料庫範例圖

4.1.3 MIR-Google 人臉資料庫

MIR-Google 是由 MIR 實驗室透過 Google Image Search 下關鍵字收集而來的人臉資料庫，關鍵字有：小孩大頭照、老人大頭照、明星大頭照、畢業大頭照、履歷大頭照，透過這些關鍵字總共可以收集到 4218 張影像，但這些影像有很多是不能使用的（例如：沒有人臉、光線過暗、人臉角度歪斜），因此，透過人為標記後，我們取得了 1384 張有用的人臉影像，這些影像具有下列特性：皆為亞洲人、人臉端正、年齡和性別為三人投票取眾數的結果。其影像張數以 10 歲為區間的直方統計圖 4.1-5 如下所示，其中 X 軸為年齡區間，Y 軸為張數。而依照 UTMVP 規格分為男女兩類和五類年齡範圍的統計表格如表 4.1-5(未知代表投票結果為不知道性別，此影像通常為小嬰兒或是外表中性的人臉，性別實驗將不考

慮這些影像) 和表 4.1-6 所示。圖 4.1-6 則為 MIR-Google 資料庫範例圖。

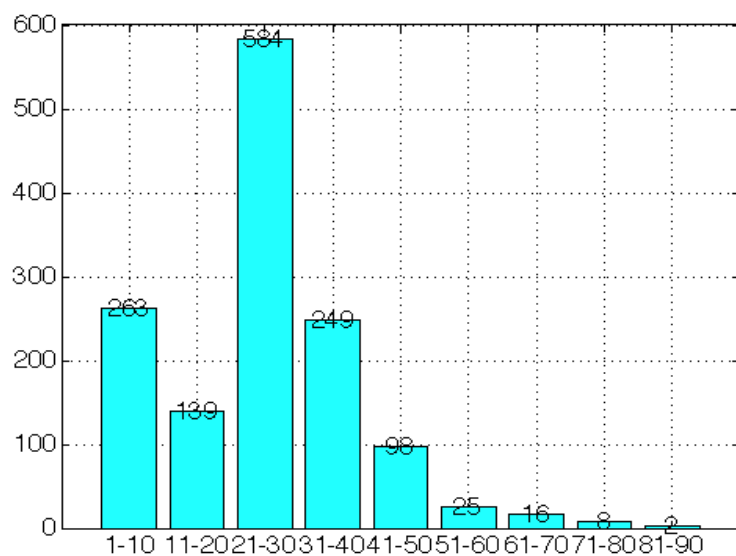


圖 4.1-5 MIR-Google 年齡直方統計圖

表 4.1-5 MIR-Google 男女統計表

性別	數量
男	667
女	652
未知	65

表 4.1-6 MIR-Google 五類年齡範圍統計表

年齡範圍	數量
0~20	402
21~30	584
31~40	249
41~50	98
51-	51



圖 4.1-6 MIR-Google 資料庫範例圖

4.2 實驗參數設定

以下實驗影像的長寬為 76×57 ，各個人臉特徵所使用的參數如下：LBP 預設為半徑 $R = 1.5$ 上 $P = 8$ 個鄰居像素點、LDP 的 $k = 3$ 、LTP 的門檻值 $t = 5$ 、Gabor 濾波器和 LGBP 則是選用 3.2.5 小節提到的 5 個頻率和 8 個方向。

降低特徵維度所使用的參數則是依照實驗測試而來，方法如下圖 4.2-1 所示，我們用 SVM 搭配 LBP 做性別辨識來說明，左圖中我們發現做完交叉驗證後，用 PCA 將特徵降至 75 維效果最好，接著我們將這 75 維的特徵用同樣的方法來測試 LDA，最後發現把 75 維降至 50 維有最好的表現。

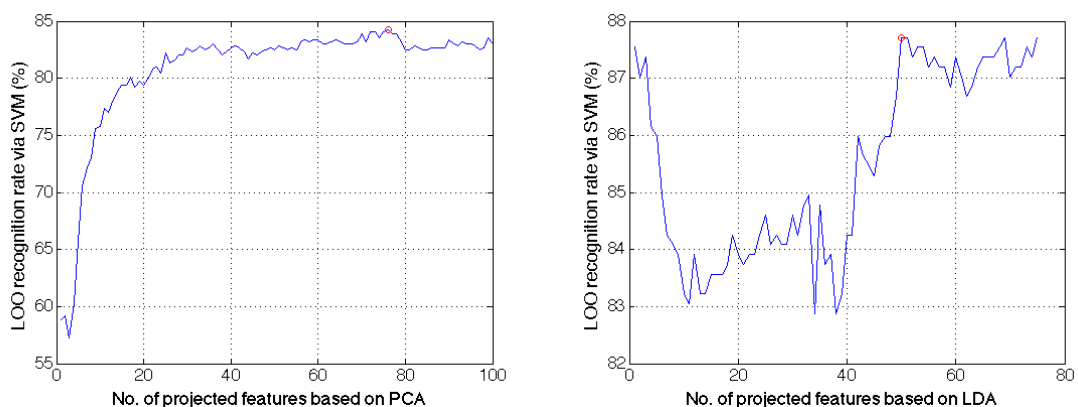


圖 4.2-1 用 PCA 搭配 LDA 做降維的示意圖

辨識模型中所使用的 SVM 為林智仁教授所開發的 LIBSVM[26]，使用的核

心為 RBFSVM，cost 則是設為 1。AdaBoostSVM 中的 σ_{ini} 初始為訓練資料彼此特徵值的最大距離， σ_{min} 則是初始為訓練資料每個兩兩特徵值的平均距離， σ_{step} 則是設為 10。

4.3 效能評估方式

本論文實驗的效能評估將分為三種方式：第一種為辨識率 (Recognition Rate)，值將落在 0 到 100% 之間，愈高代表辨識的愈準確；第二種為平均準確率的均值 (Mean Average Precision, MAP)，算法是將專屬於各類的辨識率做平均，值將落在 0 到 100% 之間，愈高代表對於各個類別的辨識精準度較高；第三種則是 UTMVP 對於年齡辨識所定義的評分方式，如下表 4.3-1 所示，若正確答案為 21~30 歲、31~40 歲、41~50 歲此三區間，答對可得兩分，若答案為 20 歲以下、51 歲以上此兩區間，答對可得三分，且回答答案為正確答案的前後級距可得一分。透過此表，我們即可將測試資料的總分除以測試資料的總數，得到每張影像的平均分數，作為年齡的效能評估方式，此值將落 0~3 分之間 (圖中數字依序為 1: 0-20, 2: 21-30, 3: 31-40, 4: 41-50, 5: 50+，共五個年齡類別)。

表 4.3-1 UTMVP 年齡評分表

Answer Predict	1	2	3	4	5
1	3	1	0	0	0
2	1	2	1	0	0
3	0	1	2	1	0
4	0	0	1	2	1
5	0	0	0	1	3

4.4 性別辨識實驗

以下將依照各個 4.1 小節所述的資料庫分別進行實驗，實驗主要分為兩個項目：第一項實驗為比較尚未進行特徵篩選的 AdaBoostSVM 和 SVM 的效能；第二項則是比較進行特徵篩選後的 AdaBoostSVM。

4.4.1 PAL 人臉資料庫實驗

首先，我們針對各個特徵以及組合過的特徵訓練 AdaBoostSVM 的性別辨識模型，如下圖 4.4-1 所示，x 軸為訓練模型的迭代次數，y 軸為訓練資料做交叉驗證 (Cross Validation) 的訓練辨識率，從圖中我們可以發現單一特徵中，LDP 在訓練錯誤率上表現最差，LBP 表現最好。而將所有特徵組合起來 (All Feature: LBP+LDP+LTP+LGBP+Gabor) 做篩選的辨識率則明顯高於其它線段。此外，我們將組合特徵做篩選的線段挑出來做觀察，可以看到此線段在各個迭代過程所篩選出最適合的特徵 (圖中數字依序為 1: LBP, 2: LDP, 3: LTP, 4: LGBP, 5: Gabor)。

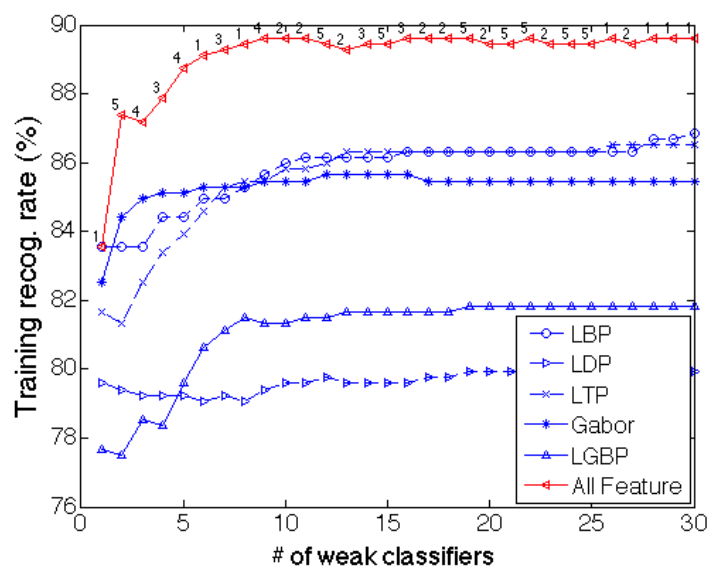


圖 4.4-1 PAL 性別訓練辨識率與迭代次數對應圖

接著我們將用 PAL 人臉資料庫做 5 份交叉驗證 (5-fold Cross Validation)，並以辨識率和 MAP 做效能評估。辨識率的結果如下圖 4.4-2 所示。MAP 的結果如下圖 4.4-3。從中我們可以發現在單一特徵的情況下，AdaBoostSVM 和 SVM 的表現差不多，但在特徵組合後進行篩選的 AdaBoostSVM 則有較好的表現。

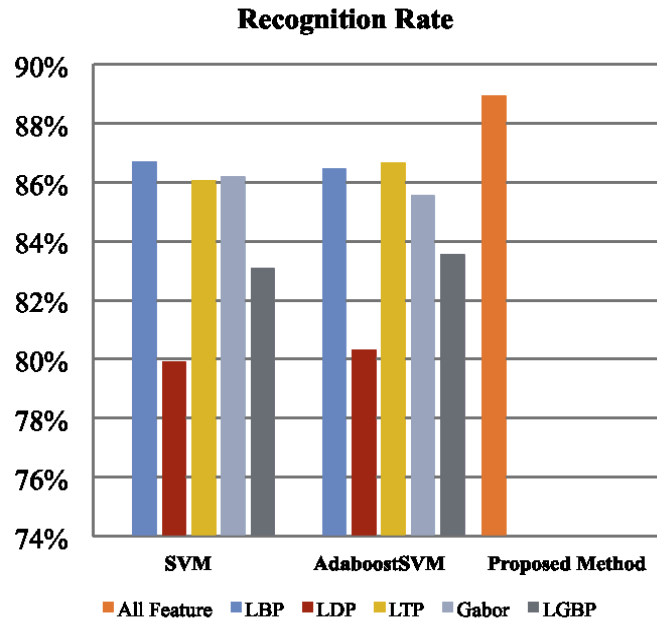


圖 4.4-2 PAL 性別辨識率直方圖

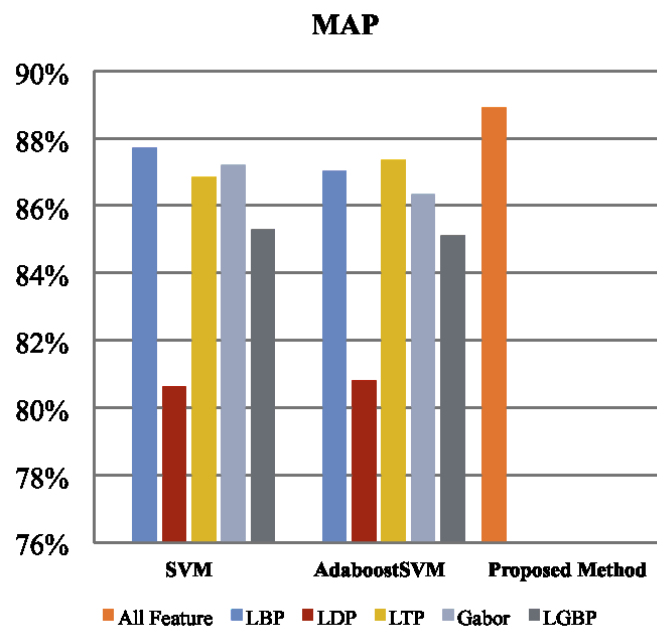


圖 4.4-3 PAL 性別平均準確率均值直方圖

4.4.2 FERET 人臉資料庫實驗

此節同 4.4.1 的實驗方式。如下圖 4.4-4 所示，從圖中我們可以發現單一特徵中 LDP 在訓練辨識率上表現最糟，而 LGBP 則有較好的表現。而將所有特徵組合起來（All Feature: LBP+LDP+LTP+LGBP+Gabor）做篩選的辨識率則明顯高於其它線段。此外，我們將組合特徵做篩選的線段挑出來做觀察，可以看到此線段在各個迭代過程所篩選出最適合的特徵(圖中數字依序為 1: LBP, 2: LDP, 3: LTP, 4: LGBP, 5: Gabor)。

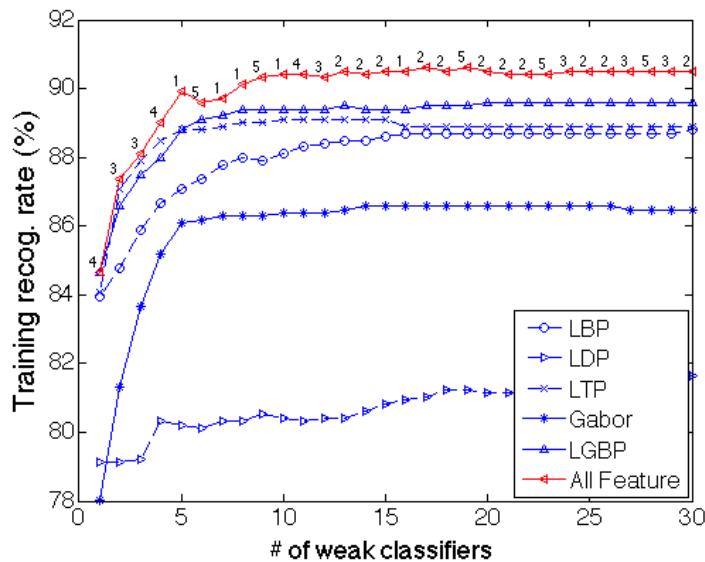


圖 4.4-4 FERET 性別訓練辨識率與迭代次數對應圖

接著我們將用 FERET 人臉資料庫做 5 份交叉驗證 (5-fold Cross Validation)，並以辨識率和 MAP 做效能評估，辨識率的結果如下圖 4.4-5 所示。MAP 的結果如下圖 4.4-6。從中我們可以發現在單一特徵的情況下，AdaBoostSVM 和 SVM 的表現差不多，但在特徵組合後進行篩選的 AdaBoostSVM 則有較好的表現。

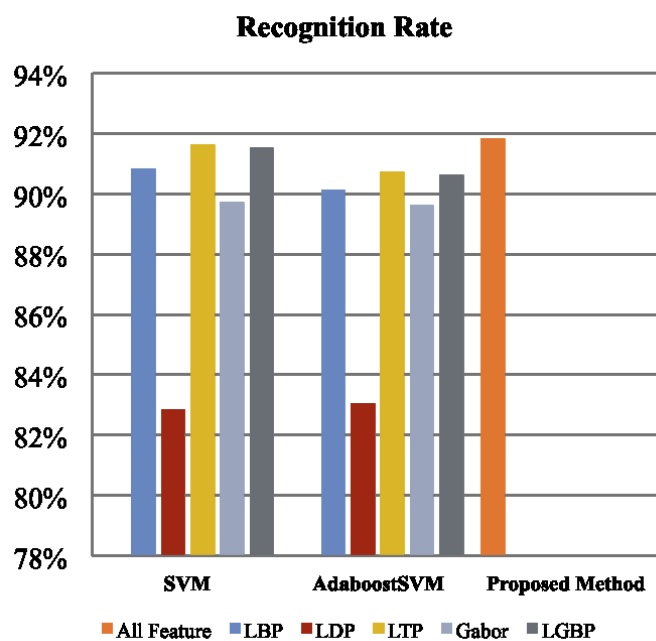


圖 4.4-5 FERET 性別辨識率直方圖

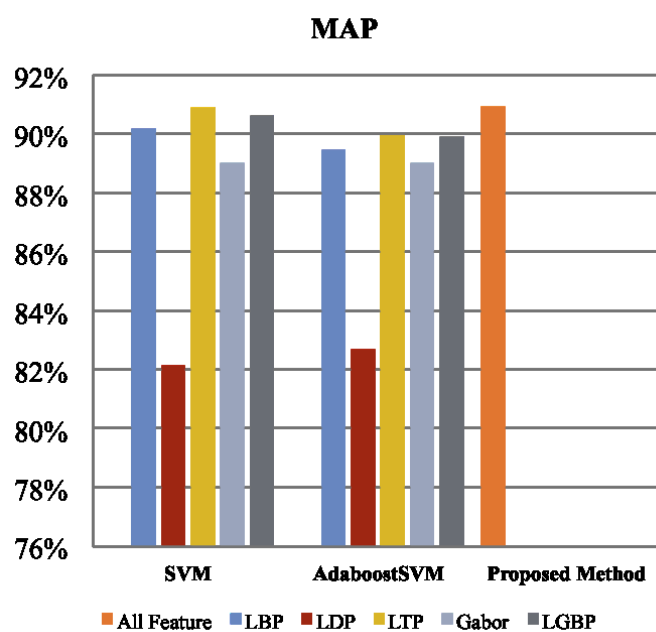


圖 4.4-6 FERET 性別平均準確率均值直方圖

4.4.3 MIR-Google 人臉資料庫實驗

此節同 4.4.1 的實驗方式。如下圖 4.4-7 所示，從圖中我們可以發現單一特徵

中 LDP 在訓練辨識率上表現最糟，而 LBP 則有較好的表現。而將所有特徵組合起來（All Feature: LBP+LDP+LTP+LGBP+Gabor）做篩選的辨識率則明顯高於其它線段。此外，我們將組合特徵做篩選的線段挑出來做觀察，可以看到此線段在各個迭代過程所篩選出最適合的特徵（圖中數字依序為 1: LBP, 2: LDP, 3: LTP, 4: LGBP, 5: Gabor）。

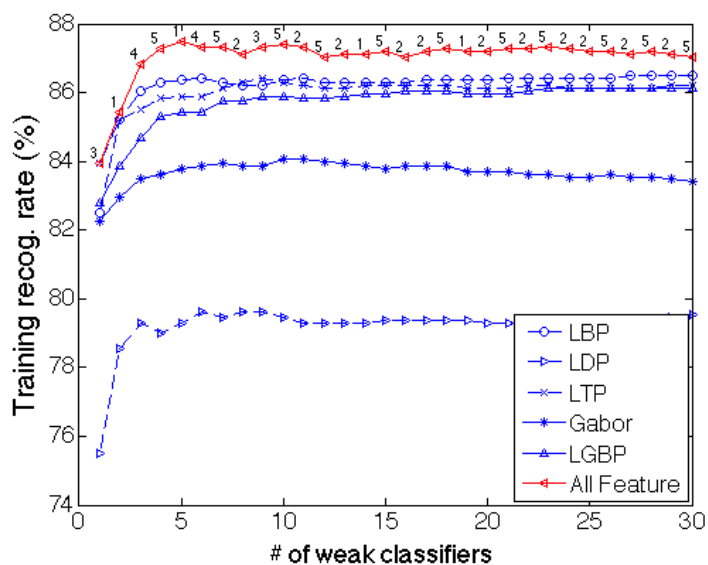


圖 4.4-7 MIR-Google 性別訓練辨識率與迭代次數對應圖

接著我們用 MIR-Google 人臉資料庫做 5 份交叉驗證(5-fold Cross Validation)，並以辨識率和 MAP 做效能評估，辨識率的結果如下圖 4.4-8 所示。MAP 的結果如下圖 4.4-9。從中我們可以發現在單一特徵的情況下，AdaBoostSVM 和 SVM 的表現差不多，而特徵組合後進行篩選的 AdaBoostSVM 則在 MAP 有好一點點的表現。

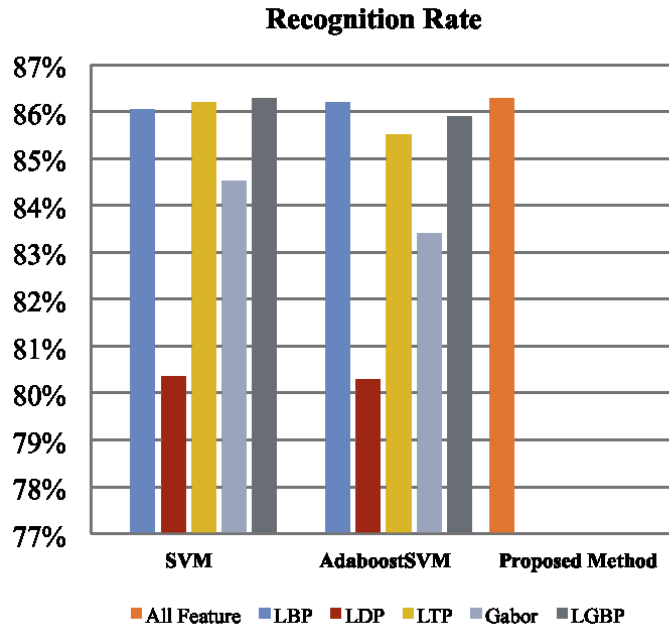


圖 4.4-8 MIR-Google 性別辨識率直方圖

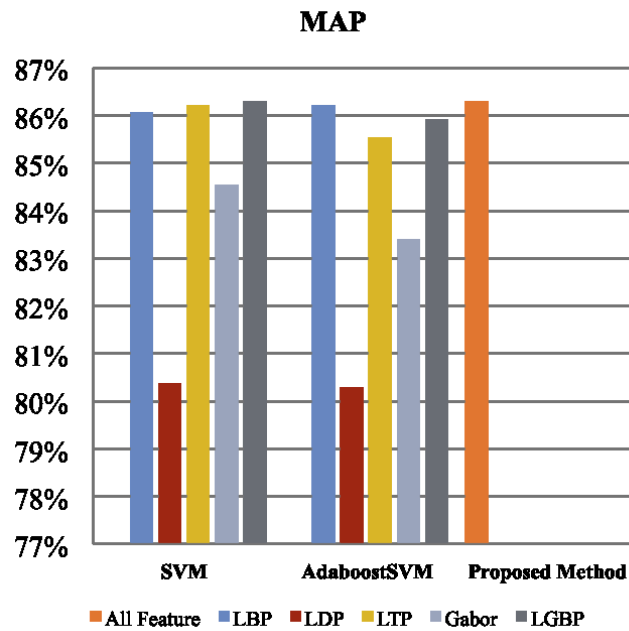


圖 4.4-9 MIR-Google 性別平均準確率均值直方圖

4.5 年齡辨識實驗

以下將依照各個 4.1 小節所述的資料庫分別進行實驗，實驗主要分為兩個項目：第一項實驗為比較尚未進行特徵篩選的 AdaBoostSVM 和 SVM 的效能；第二

項則是比較進行特徵篩選後的 AdaBoostSVM。與性別辨識的差別在於這節將會多探討依照 UTMVP 定義的效能評估方式。

4.5.1 PAL 人臉資料庫實驗

首先，我們針對各個特徵以及組合過的特徵訓練 AdaBoostSVM 的年齡辨識模型，如下圖 4.5-1 所示，x 軸為訓練模型的迭代次數，y 軸為訓練資料做交叉驗證 (Cross Validation) 的訓練辨識率，從圖中我們可以發現單一特徵中，LDP 在訓練辨識率上表現最差，而 LGBP 表現最好。而將所有特徵組合起來 (All Feature: LBP+LDP+LTP+LGBP+Gabor) 做篩選的辨識率則明顯高於其它線段。此外，我們將組合特徵做篩選的線段挑出來做觀察，可以看到此線段在各個迭代過程所篩選出最適合的特徵 (圖中數字依序為 1: LBP, 2: LDP, 3: LTP, 4: LGBP, 5: Gabor)。

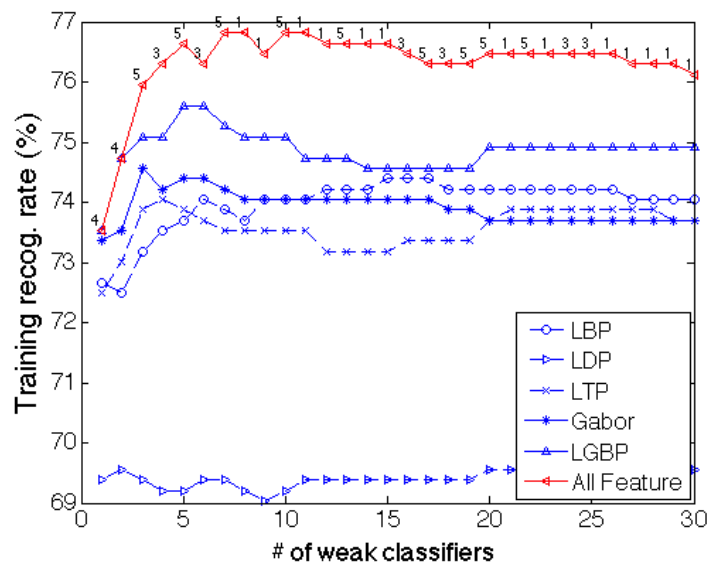


圖 4.5-1 PAL 年齡訓練辨識率與迭代次數對應圖

接著我們將用 PAL 人臉資料庫做 5 份交叉驗證 (5-fold Cross Validation)，並以辨識率和 UTMVP 年齡評分做效能評估。辨識率的結果如下圖 4.5-2 所示。UTMVP 年齡評分如下圖 4.5-3。從中我們可以發現在單一特徵的情況下，AdaBoostSVM 和 SVM 的表現差不多，但在特徵組合後進行篩選的 AdaBoostSVM

則有較好的表現。

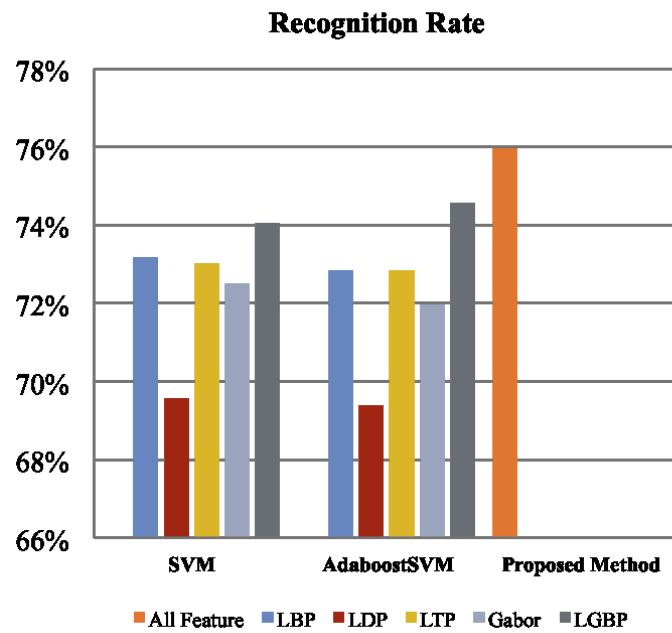


圖 4.5-2.PAL 年齡辨識率直方圖

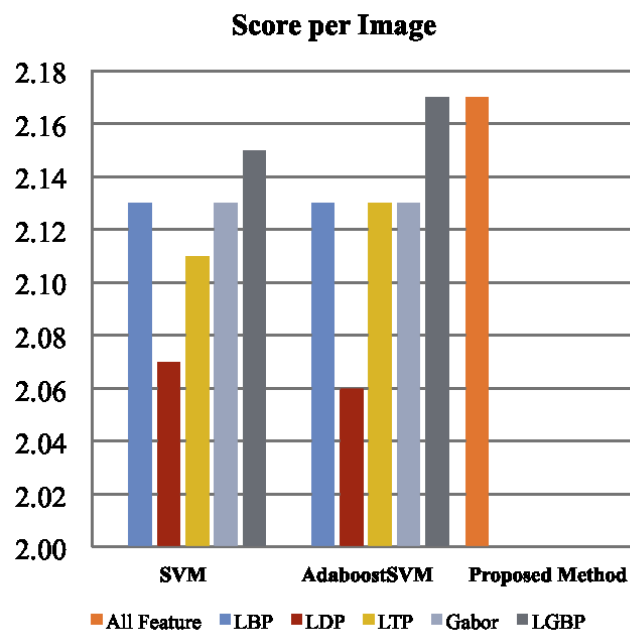


圖 4.5-3 PAL UTMVP 年齡評分直方圖

4.5.2 FERET 人臉資料庫實驗

此節同 4.5.1 的實驗方式。如下圖 4.5-4 所示，從圖中我們可以發現單一特徵中 LDP 在訓練辨識率上表現最糟，而 LGBP 則有較好的表現。而將所有特徵組合起來（All Feature: LBP+LDP+LTP+LGBP+Gabor）做篩選的辨識率則略高於其它線段。此外，我們將組合特徵做篩選的線段挑出來做觀察，可以看到此線段在各個迭代過程所篩選出最適合的特徵（圖中數字依序為 1: LBP, 2: LDP, 3: LTP, 4: LGBP, 5: Gabor）。

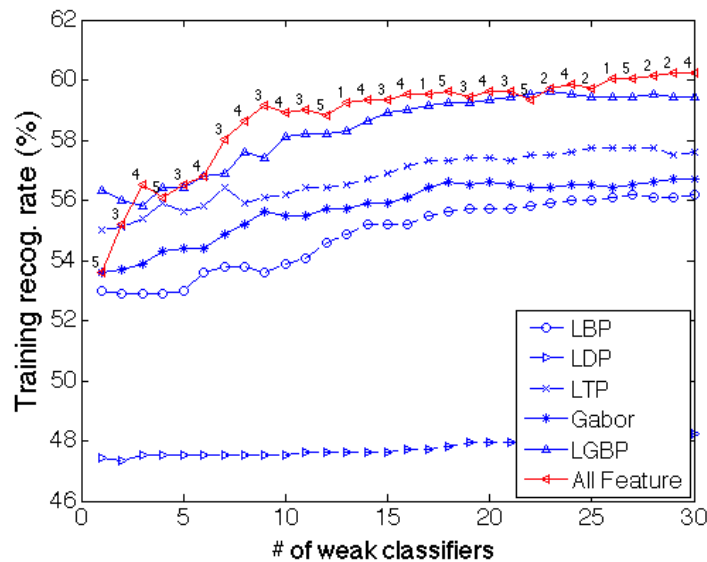


圖 4.5-4 FERET 年齡訓練辨識率與迭代次數對應圖

接著我們將用 FERET 人臉資料庫做 5 份交叉驗證 (5-fold Cross Validation)，並以辨識率和 UTMVP 年齡評分做效能評估，辨識率的結果如下圖 4.5-5 所示。UTMVP 年齡評分如下圖 4.5-6。從中我們可以發現在單一特徵的情況下，AdaBoostSVM 和 SVM 的表現差不多，而這次在特徵組合後進行篩選的 AdaBoostSVM 沒有較好的表現。

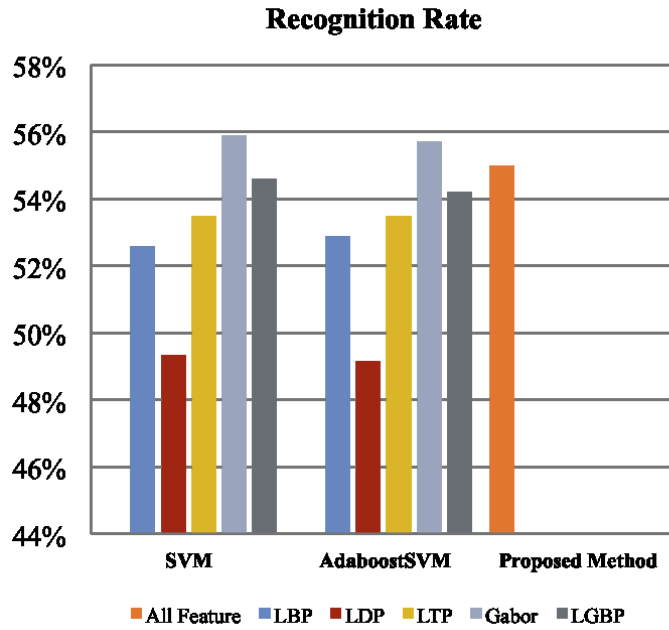


圖 4.5-5 FERET 年齡辨識率直方圖

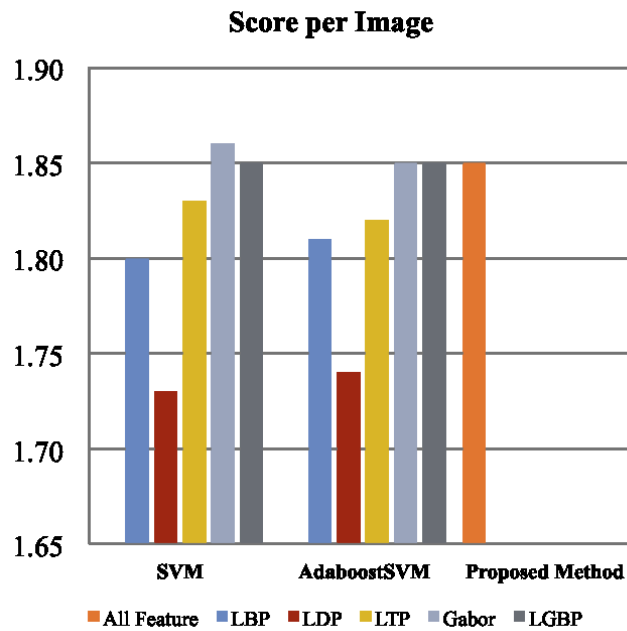


圖 4.5-6 FERET 年齡平均準確率均值直方圖

4.5.3 MIR-Google 人臉資料庫實驗

此節同 4.5.1 的實驗方式。如下圖 4.5-7 所示，從圖中我們可以發現單一特徵中 LDP 在訓練辨識率上表現最糟，而 LBP 則有較好的表現。而將所有特徵組合

起來 (All Feature: LBP+LDP+LTP+LGBP+Gabor) 做篩選的辨識率則明顯高於其它線段。此外，我們將組合特徵做篩選的線段挑出來做觀察，可以看到此線段在各個迭代過程所篩選出最適合的特徵 (圖中數字依序為 1: LBP, 2: LDP, 3: LTP, 4: LGBP, 5: Gabor)。

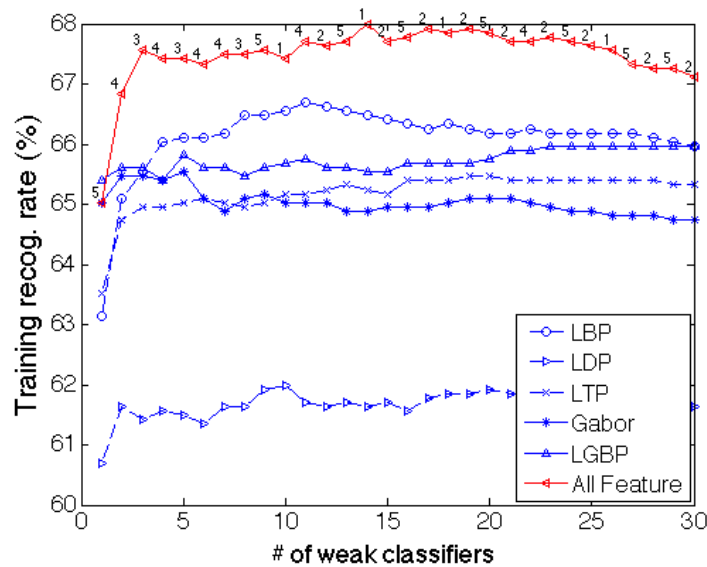


圖 4.5-7 MIR-Google 年齡訓練辨識率與迭代次數對應圖

接著我們將用 FERET 人臉資料庫做 5 份交叉驗證 (5-fold Cross Validation)，並以辨識率和 UTMVP 年齡評分做效能評估，辨識率的結果如下圖 4.5-8 所示。UTMVP 年齡評分如下圖 4.5-9。從中我們可以發現在單一特徵的情況下，AdaBoostSVM 和 SVM 的表現差不多，但在特徵組合後進行篩選的 AdaBoostSVM 則有較好的表現。

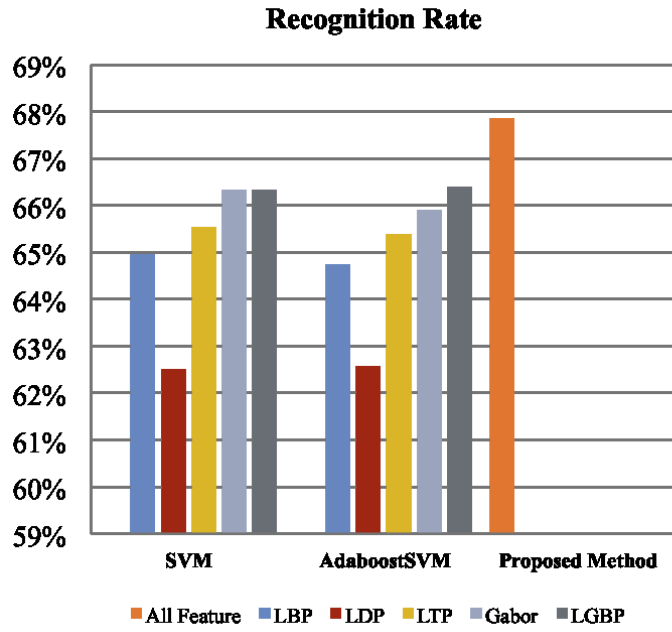


圖 4.5-8 MIR-Google 年齡辨識率直方圖

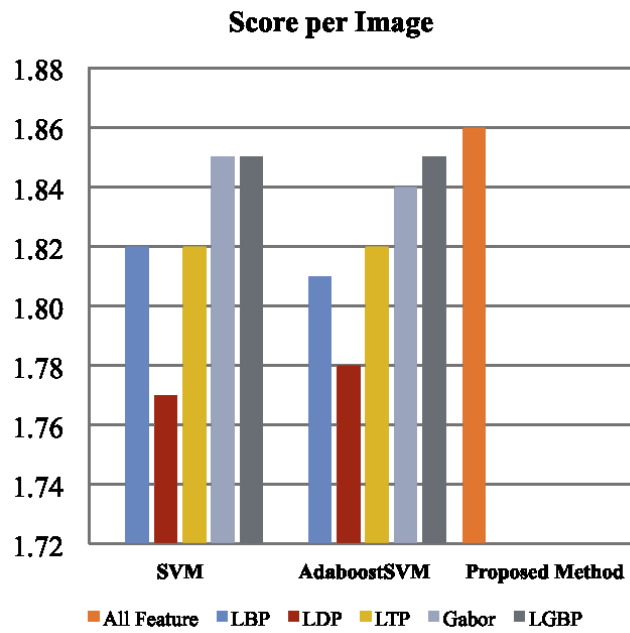


圖 4.5-9 MIR-Google 年齡平均準確率均值直方圖

4.6 實驗錯誤分析

以下將說明本論文提出的方法在性別和年齡辨識上常見的錯誤類型。

4.6.1 性別辨識錯誤分析

我們將以 4.4 小節中 PAL 和 MIR-Google 人臉資料庫的性別辨識結果來做說明。首先，我們先觀察提出方法的混淆矩陣 (Confusion Matrix)，如下圖 4.6-1：

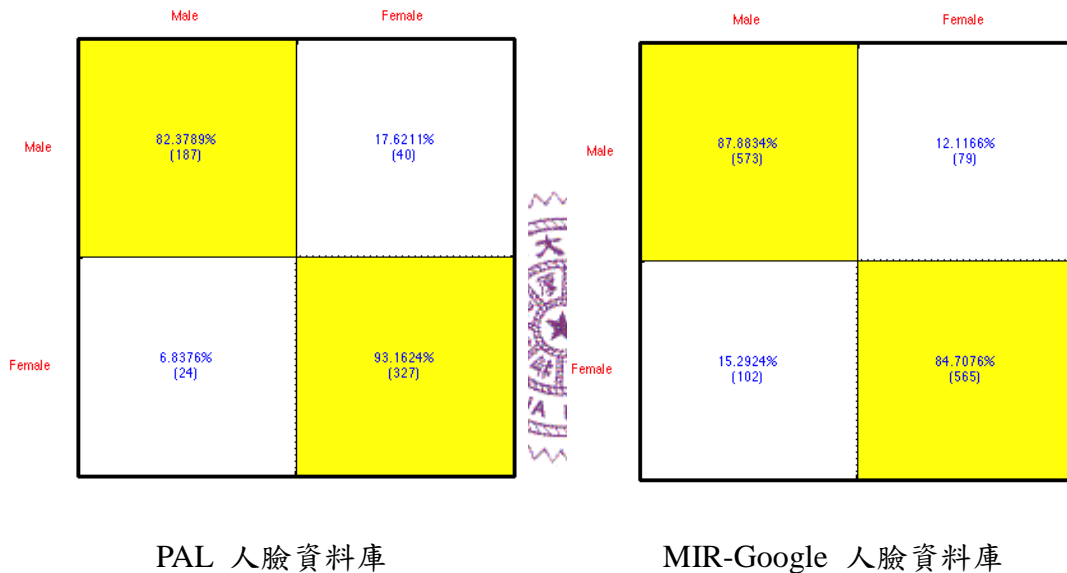


圖 4.6-1 性別辨識之混淆矩陣

從混淆矩陣的分錯區塊 (上圖的白色區塊)，我們挑選了一些常見的錯誤分類影像。如下圖 4.6-2 所示，我們發現到在性別辨識中，我們分錯的原因常常是因為有些人本來就難分辨男女。例如在 PAL 資料庫中，有些老人很難區分男女；而在 MIR-Google 資料庫中，則是小孩較難區分。



圖 4.6-2 性別辨識常見的錯誤類型

4.6.2 年齡辨識錯誤分析



我們將以 4.5 小節中 PAL 和 MIR-Google 人臉資料庫的年齡辨識結果來做說明。首先，我們先觀察提出方法的混淆矩陣 (Confusion Matrix)，如下圖 4.6-3：

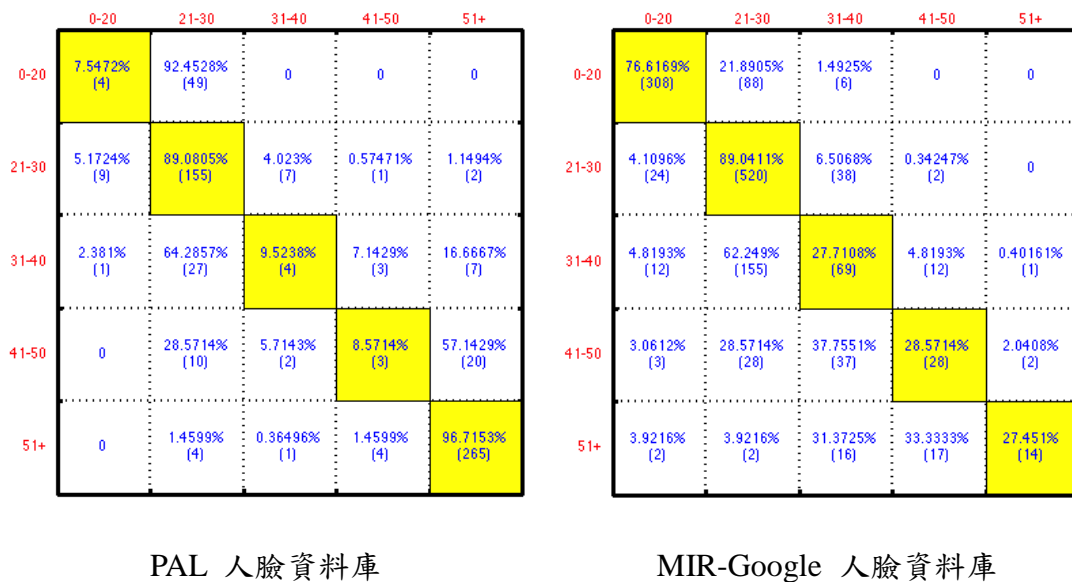


圖 4.6-3 年齡辨識之混淆矩陣

從混淆矩陣的分錯區塊（上圖的白色區塊），我們挑選了一些常見的錯誤分類影像。如下圖 4.6-4 所示，我們發現到在年齡辨識中，我們本來就很難決定一個人落於哪個年齡範圍，而這種情況在 21~50 歲最為明顯。



圖 4.6-4 年齡辨識常見的錯誤類型

第五章 結論與未來研究方向

從實驗結果中我們發現到在單一特徵中，LBP 對於性別辨識有較好的表現，LGBP 則在年齡辨識有較好的表現。而傳統的 AdaBoostSVM 和標準的 SVM 在單一特徵下的辨識結果是不相上下的。此外，本論文所提出的 AdaBoostSVM 結合特徵篩選則在性別和年齡辨識中有較好的辨識結果。

在未來，我們可以試著增加 AdaBoostSVM 可篩選的特徵，進而增加答案的多樣性，如此將有機會提高辨識能力。不過當篩選的特徵變多後，AdaBoostSVM 辨識及訓練所花的時間將因此變多，因此我們可以試著使用維度更小但描述能力不減的人臉特徵，例如使用均勻型局部二元化模版 (Uniform LBP) [27]或是延展示局部三元化模版 (Extended LTP) [28]等等方式來提前降低特徵維度。此外，從年齡的辨識結果我們發現到，21~50 歲這區間三個類別的分類容易分到鄰近的類別，代表這三個類別是不容易區分的，因此，或許我們可以試著先用 AdaBoostSVM 來區分 1~20、21~50、51+ 這三個年齡類別，接著再用年齡迴歸 (Age Regression) 的方式來評估 21~50 歲這三個類別的年齡，如此將有機會減少此類型的錯誤分類。

參考文獻

- [1] 第八屆由田機器視覺獎 (Utechzone Machine Vision Prize, UTMVP)
<http://www.utmvp.com/>
- [2] P. Jonathon Phillips, Harry Wechsler, Jeffrey Huang, Patrick J. Rauss: The FERET database and evaluation procedure for face-recognition algorithms. *Image Vision Comput.* 16(5): 295-306 (1998).
- [3] Minear, Meredith, and Denise C. Park. "A lifespan database of adult facial stimuli." *Behavior Research Methods, Instruments, & Computers* 36.4 (2004): 630-633.
- [4] Lian, Hui-Cheng, and Bao-Liang Lu. "Multi-view gender classification using local binary patterns and support vector machines." *Advances in Neural Networks-ISNN 2006*. Springer Berlin Heidelberg, 2006. 202-209
- [5] Xia, Bin, He Sun, and Bao-Liang Lu. "Multi-view gender classification based on local Gabor binary mapping pattern and support vector machines." *Neural Networks, 2008. IJCNN 2008.(IEEE World Congress on Computational Intelligence)*. IEEE International Joint Conference on. IEEE, 2008
- [6] Jabid, Taskeed, M. Hasanul Kabir, and Oksam Chae. "Gender classification using local directional pattern (LDP)." *Pattern Recognition (ICPR), 2010 20th International Conference on*. IEEE, 2010.
- [7] Gunay, Asuman, and Vasif V. Nابیev. "Automatic age classification with LBP." *Computer and Information Sciences, 2008. ISCIS'08. 23rd International Symposium on*. IEEE, 2008.
- [8] Gao, Feng, and Haizhou Ai. "Face age classification on consumer images with

- gabor feature and fuzzy lda method." *Advances in biometrics*. Springer Berlin Heidelberg, 2009. 132-141.
- [9] Shan, Caifeng. "Learning local features for age estimation on real-life faces." *Proceedings of the 1st ACM international workshop on Multimodal pervasive video analysis*. ACM, 2010.
- [10] Li, Xuchun, Lei Wang, and Eric Sung. "AdaBoost with SVM-based component classifiers." *Engineering Applications of Artificial Intelligence* 21.5 (2008): 785-795.
- [11] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. Vol. 1. IEEE, 2001.
- [12] Shih, Frank Y., and Chao-Fa Chuang. "Automatic extraction of head and face boundaries and facial features." *Information Sciences* 158 (2004): 117-130.
- [13] Ojala, Timo, Matti Pietikäinen, and David Harwood. "A comparative study of texture measures with classification based on featured distributions." *Pattern recognition* 29.1 (1996): 51-59.
- [14] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037-2041, Dec. 2006.
- [15] Shan, Caifeng. "Learning local features for age estimation on real-life faces." *Proceedings of the 1st ACM international workshop on Multimodal pervasive video analysis*. ACM, 2010.
- [16] Tan, Xiaoyang, and Bill Triggs. "Enhanced local texture feature sets for face

- recognition under difficult lighting conditions." *Image Processing, IEEE Transactions on* 19.6 (2010): 1635-1650.
- [17] Luu, Khoa, et al. "Combined local and holistic facial features for age-determination." *Control Automation Robotics & Vision (ICARCV), 2010 11th International Conference on*. IEEE, 2010.
- [18] Lee, Tai Sing. "Image representation using 2D Gabor wavelets." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 18.10 (1996): 959-971.
- [19] Zhang, Wenchao, et al. "Local Gabor binary pattern histogram sequence (LGBPHS): A novel non-statistical model for face representation and recognition." *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*. Vol. 1. IEEE, 2005.
- [20] Jolliffe, Ian T. *Principal component analysis*. Vol. 487. New York: Springer-Verlag, 1986.
- [21] Deng, Hong-Bo, et al. "A new facial expression recognition method based on local gabor filter bank and pca plus lda." *International Journal of Information Technology* 11.11 (2005): 86-96.
- [22] Liu, Chengjun. "Gabor-based kernel PCA with fractional power polynomial models for face recognition." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 26.5 (2004): 572-581.
- [23] Mika, Sebastian, et al. "Fisher discriminant analysis with kernels." *Neural Networks for Signal Processing IX, 1999. Proceedings of the 1999 IEEE Signal Processing Society Workshop.. IEEE, 1999*.
- [24] Cortes, Corinna, and Vladimir Vapnik. "Support-vector networks." *Machine learning* 20.3 (1995): 273-297

- [25] Freund, Yoav, and Robert E. Schapire. "Experiments with a new boosting algorithm." ICML. Vol. 96. 1996.
- [26] Chang, Chih-Chung, and Chih-Jen Lin. "LIBSVM: a library for support vector machines." ACM Transactions on Intelligent Systems and Technology (TIST)2.3 (2011): 27.
- [27] Ojala, Timo, Matti Pietikainen, and Topi Maenpaa. "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns." Pattern Analysis and Machine Intelligence, IEEE Transactions on 24.7 (2002): 971-987.
- [28] Liao, Wen-Hung. "Region description using extended local ternary patterns." Pattern Recognition (ICPR), 2010 20th International Conference on. IEEE, 2010.

