

# 國立交通大學

## 電控工程研究所

### 碩 士 論 文

應用步態分析於身份鑑別之研究



An Efficient Individual Recognition Using  
Gait Analysis

研 究 生：張婷婷

指 導 教 授：林昇甫 博士

中 華 民 國 一 〇 一 年 七 月

應用步態分析於身份鑑別之研究

An Efficient Individual Recognition Using Gait Analysis

研 究 生：張婷婷

Student: Ting-Ting Chang

指導教授：林昇甫 博士

Advisor: Dr. Sheng-Fuu Lin

國立交通大學

電控工程研究所



Submitted to Institute of Electrical Control Engineering

College of Electrical Engineering

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Master

in

Electrical Control Engineering

July 2012

Hsinchu, Taiwan, R.O.C

中華民國一〇一年七月

# 應用步態分析於身份鑑別之研究

學生：張婷婷

指導教授：林昇甫 博士

國立交通大學電控工程研究所

## 摘要

本論文提出一個自動的身份識別系統，主要目的在於減少人力成本的耗損，並提高監控系統的安全性與可靠性。不同於其他的生物識別系統，例如人臉辨識、指紋辨識、虹膜辨識等等，利用姿態進行身份識別不需與使用者互動、不需近距離拍攝、也不需要太精密的攝影器材即可進行分析，因此步態分析一直是視覺監控領域的熱門研究主題。但由於步態辨識遠距離即可完成，因此相對其他生物識別系統來說受到外在因素的影響也較大，例如：戴帽子、穿著不同等等，所以步態辨識仍然是一個很受挑戰的領域。

在此，本論文的貢獻有兩點：第一，本論文結合多種方法，使用多面向與多角度來分析姿態，不僅平均抽取人體各個部位的特徵，更進一步將特徵分為靜態特徵與動態特徵，以確保本系統面對衣著變因的影響時，還能保持一定的辨識率；第二，本論文提出兩階段式的身份鑑別系統。第一階段主要是利用外輪廓、手部擺動和腳掌離地角度三個特徵來進行候選人篩選，而篩選的目的在於減少候選人，以提高整體辨識率；第二階段則是將輸入的步態影像與第一階段篩選出來的候選人進行匹配，並針對每種特徵不同的重要性對這些候選人進行不同的權重給分，分數最高的即表示與輸入的人物最為相似。從實驗結果來看，本系統確實有不錯的身份識別結果。

# An Efficient Individual Recognition Using Gait Analysis

Student : Ting-Ting Chang

Advisor: Dr. Sheng-Fuu Lin

Institute of Electrical Control Engineering  
National Chiao Tung University

## ABSTRACT

This thesis is a study on individual recognition using gait analysis. In order to reduce the human resource cost on one hand and improve the security and reliability of the surveillance system on the other, we propose an automated individual identification system which can not only recognize identities of individuals but also detect strangers automatically. Different from other biometric systems, such as human face detection, fingerprint identification, iris recognition, etc., the proposed system uses gait to recognize identities of individuals. In addition, the proposed system needs no interaction with people, nor close sensing distance, nor fancy equipment. Because of these advantages, gait analysis has always been a popular research topic.

Furthermore, there are two contributions in this thesis. First, features in this thesis are extracted averagely from head to toe to alleviate the factor of clothing which would cause the distortion of appearance on an individual. Second, we use a two-stage method to recognize identities of individuals. The first-stage is designed to reduce candidates. In the second-stage, we match the individual extracted from the video with candidates selected from the first-stage, and assign weighted scores to the candidates based on their similarities. The candidate who gets the highest score is the most similar one to the individual extracted from the video. According to the results of experiments, the method proposed in this thesis has a good performance.

## 誌 謝

首先，我非常感謝我的指導教授林昇甫博士在這兩年間對我的教導，感謝老師的諄諄教誨與細心指導。同時也要感謝我的家人，一直在背後默默支持我和鼓勵我，使我得以在精神與生活上無後顧之憂，順利完成學業。另外，我要感謝所有幫助我，讓我拍攝做為實驗資料庫的所有人，包含實驗室學長俊偉、裕筆、逸章，一起奮鬥的同學柏宏、俊良、植諺，實驗室學弟妹雅君、明揚、炳晟、兆傑、衍澄，以及朋友們鈞鈞、巧桐、喆夫、甫尊、仲傑、奕君、薇雅、方菱等等，由於你們的幫忙，才能讓我順利建立實驗的資料庫，也由於你們的討論與協助，才能讓我順利完成論文研究。



# 目錄

中文摘要 .....	i
英文摘要 .....	ii
誌謝 .....	iii
目錄 .....	iv
圖目錄 .....	vii
表目錄 .....	xi
第一章 緒論 .....	1
1.1 生物識別方法介紹 .....	1
1.2 相關研究之探討 .....	2
1.3 研究動機 .....	6
1.4 論文主體與貢獻 .....	7
第二章 相關技術與原理 .....	8
2.1 前景背景切割法 .....	8
2.2 靜態特徵 .....	10
2.2.1 步態能量影像 .....	10
2.2.2 傅立葉描述子 .....	12
2.3 動態特徵 .....	14
2.3.1 主成份分析法 .....	14
2.3.2 哈里斯(Harris)角點偵測法 .....	16
2.4 身份鑑別法則 .....	20
2.4.1 最近鄰居法 .....	20

2.4.2	KD tree .....	22
第三章	系統流程 .....	25
3.1	整體系統架構與流程 .....	25
3.2	提取前景 .....	27
3.3	週期估測 .....	28
3.4	靜態特徵 .....	32
3.4.1	外輪廓特徵 .....	33
3.4.2	手部擺動幅度特徵 .....	34
3.5	動態特徵 .....	35
3.5.1	頭部特徵 .....	35
3.5.2	腳掌特徵 .....	38
3.5.3	膝蓋特徵 .....	40
3.5.4	步伐大小特徵 .....	43
3.6	分類法則 .....	45
3.6.1	身份鑑別法則 .....	47
3.6.2	安全監控法則 .....	48
第四章	實驗結果與分析 .....	51
4.1	實驗機制 .....	51
4.1.1	實驗場景與設備 .....	51
4.1.2	實驗資料 .....	53
4.2	實驗結果 .....	54
4.2.1	身份鑑別結果與分析 .....	54
4.2.2	安全監控結果與分析 .....	57
4.2.3	特殊變因情況討論與分析 .....	59
4.3	分析比較 .....	60

4.4	結果討論 .....	64
第五章	結論與未來發展 .....	67
參考文獻	.....	69





# 圖目錄

圖 1.1 文獻[2]性別辨識相關影像。(a)序列側影影像；(b)主成份步態影像(GPCI)。	2
圖 1.2 文獻[4]性別辨識相關影像。(a) 人體步態能量影像；(b)將人體姿態分成 7 區域之示意圖。	3
圖 1.3 文獻[6]身份鑑別相關影像。(a)將人體姿態影像分為七部分之示意圖；(b)利用橢圓去表示此七區域之示意圖。	4
圖 1.4 文獻[8]身份鑑別相關影像。(a)人體六區域示意圖；(b)人體骨架示意圖。	4
圖 1.5 文獻[9]身份鑑別相關影像。(a)長度特徵；(b)寬度特徵；(c)面積特徵；(d)對角角度特徵。	5
圖 1.6 文獻[10]身份鑑別相關影像。(a)人體步態輪廓示意圖；(b)動態輪廓影像(motion contour image, MCI)。	6
圖 2.1 高斯分布示意圖	9
圖 2.2 利用 GMM 擷取前景範例。(a)輸入影像；(b)經過 GMM 前景提取後的結果。	10
圖 2.3 步態能量示意圖。(a) 步態序列影像；(b)步態能量影像；(c) 步態序列影像；(d) 步態能量影像。	11
圖 2.4 Moravec 偵測邊角範例。(a)搜尋視窗在平面時之示意圖；(b)搜尋視窗在邊界時之示意圖；(c)搜尋視窗在轉角時之示意圖。	17
圖 2.5 特徵值分布情形。	20
圖 2.6 KNN 演算法示意圖。	21
圖 2.7 KD-tree 示意圖。	22
圖 2.8 KD-tree 樹狀圖。	24
圖 3.1 本論文系統流程圖。	26
圖 3.2 提取前景範例。(a)灰階背景影像；(b)灰階輸入影像；(c)背景相減後前景影像。	

(d)人體步態影像，即本論文欲處理的步態影像。 .....	27
圖 3.3 B. DeCann 和 A. Ross[16]方法中，抽取靜態特徵之方法示意圖。 .....	28
圖 3.4 文獻[16]中特徵抽取範例。(a)起步時 $v_{\min}$ 、 $v_{\max}$ 和 $v_{\text{ter}}$ 三點的相對位置範例； (b)半週期時 $v_{\min}$ 、 $v_{\max}$ 和 $v_{\text{ter}}$ 三點的相對位置範例；(c) 背背包時 $v_{\min}$ 、 $v_{\max}$ 和 $v_{\text{ter}}$ 三點的相對位置範例。 .....	29
圖 3.5 本論文週期估測方法示意圖。(a)本論文尋找 A、B 和 C 三點之示意圖；(b)本 論文用來估測週期的角度示意圖。 .....	30
圖 3.6 週期估測範例。 .....	31
圖 3.7 一週期內步態影像範例。 .....	32
圖 3.8 本論文靜態特徵中的外輪廓特徵示意圖。(a)步態能量影像示意圖；(b)步態能 量影像二值化後結果；(c)步態能量影像外輪廓特徵；(d)利用 16 維的傅立葉描述 子來描述外輪廓特徵之示意圖。 .....	33
圖 3.9 本論文手部擺動幅度特徵示意圖。(a)步態能量影像手部特徵示意圖；(b)手部 特徵曲線圖。 .....	34
圖 3.10 擷取頭部區域流程示意圖。(a)人體步態影像；(b)擷取人體步態影像最上方七 分之一的區域影像圖；(c)裁剪頭部區域示意圖；(d)最終擷取出來的頭部區域影 像。 .....	36
圖 3.11 頭部特徵抽取結果示意圖。(a)沒戴帽子的原始影像；(b)沒戴帽子的頭部特徵 抽取結果；(c)戴帽子的原始影像；(d) 戴帽子後的頭部特徵抽取結果。 .....	36
圖 3.12 頭部傾斜角度變化曲線範例。(a)第一段步態影像頭部傾斜角抽取結果；(b) 第 二段步態影像頭部傾斜角抽取結果；(c) 第三段步態影像頭部傾斜角抽取結果； (d) 第四段步態影像頭部傾斜角抽取結果。 .....	37
圖 3.13 擷取腳掌區域示意圖。 .....	38
圖 3.14 擷取腳掌特徵範例示意圖。(a)原始輸入步態影像；(b)腳掌區域影像；(c)後 腳掌區域影像；(d)後腳離地角度擷取結果示意圖。 .....	39
圖 3.15 腳掌離地角度一周期變化範例。 .....	40

圖 3.16 擷取腿部區域範例。 .....	40
圖 3.17 擷取膝蓋特徵之流程圖。 .....	41
圖 3.18 前膝蓋特徵抽取範例。(a)原始側影影像；(b)腿部區域影像；(c)腿部區域前 曲線示意圖；(d)前膝蓋特徵點偵測結果。 .....	42
圖 3.19 後膝蓋特徵抽取範例。(a)原始側影影像；(b)腿部區域影像；(c)腿部區域後 曲線示意圖；(d)後膝蓋特徵點偵測結果。 .....	42
圖 3.20 前後腿部膝蓋角度變化。(a)前膝蓋角度變化；(b)後膝蓋角度變化。 .....	43
圖 3.21 偵測步伐距離大小示意圖。 .....	44
圖 3.22 步伐大小特徵範例。(a)一週期內步伐最大之步態影像範例；(b)一週期步伐距 離大小變化圖。 .....	45
圖 3.23 本論文所使用的靜態特徵範例。(a)外輪廓特徵；(b)手部擺動特徵。 .....	46
圖 3.24 本論文所使用的動態特徵範例。(a)頭部特徵；(b)腳掌特徵；(c)前膝蓋特徵； (d)後膝蓋特徵；(e)步伐特徵。 .....	46
圖 3.25 身份鑑別法則流程圖。 .....	47
圖 3.26 閾值選定示意圖。 .....	49
圖 3.27 閾值選定示意圖。 .....	50
圖 3.28 安全監控法則流程圖。 .....	50
圖 4.1 實驗場景圖。 .....	52
圖 4.2 程式執行所使用的個人電腦。 .....	52
圖 4.3 測試資料影像範例。(a)(b)為普通正常行走的影像；(c)(d)為加入變因，穿外套 的行走影像；(e)(f)為加入變因，戴帽子的行走影像。 .....	53
圖 4.4 本系統辨識錯誤範例說明。(a)身份 A 的腳掌離地最大角度步態影像圖；(b) 身 份 B 的腳掌離地最大角度步態影像圖；(c)身份 A 之腳掌離地角度曲線變化圖； (d) 身份 B 之腳掌離地角度曲線變化圖。 .....	55
圖 4.5 身份 A 在其他筆資料中，腳掌離地角度變化範例。(a)身份 A 的第一段步態資 料中腳掌離地最大角度之影像圖；(b)身份 A 的第二段步態資料中腳掌離地最大	

角度之影像圖；(c)利用身份 A 的第一段步態資料分析腳掌離地角度之曲線變化圖；(d) 利用身份 A 的第二段步態資料分析腳掌離地角度之曲線變化圖。..... 56



# 表目錄

表 3-1 本論文所使用的全部特徵、維度、權重大小整理表.....	46
表 4-1 本論文之身份鑑別系統辨識率.....	54
表 4-2 本論文之安全監控系統辨識率.....	58
表 4-3 模擬人穿長裙情況之系統辨識率.....	59
表 4-4 本論文方法與其他文獻方法之身份鑑別辨識率比較.....	61
表 4-5 各方法優缺點比較.....	62
表 4-6 單一特徵身份辨識率.....	63



# 第一章 緒論

安全監控在現代化社會的發展日趨重要，使用鑰匙、認證卡進出公司與家門主要都是為了讓人們安心。若不幸發生公司資料或器材遭到竊盜，甚至家裡遭小偷時，監視設備便成了唯一的線索來源。因此，監控系統已經成為人類不可或缺的基本配備。而本論文主要是將攝影機拍攝到的人進行分析與身份辨識，藉由分析每個人步態之間的差異，提供一個準確性高的自動監控系統。當發現陌生人出現的時候，便可以立刻通知警衛前往關切。此系統不僅可以讓警衛不必二十四小時緊盯著監控攝影機，更增加監控的安全性。

由於步態辨識遠距離即可完成，因此相對其他生物識別系統來說受到外在因素的影響也較大，例如：戴帽子、穿著不同等等。因此，如何擷取有效的特徵將會是本論文的研究重點。本章共分成四節，1.1 先概略介紹目前生物識別上有哪些方法及優缺點，1.2 介紹目前相關研究領域有哪些做法，1.3 闡述本論文的研究動機，1.4 則是本論文之貢獻。

## 1.1 生物識別方法介紹

生物識別方法有很多，包含指紋識別、掌紋識別、虹膜識別、人臉辨識和步態辨識等等。在這麼多生物識別方法中，本論文選擇步態來進行身份鑑別的原因在於指紋、掌紋和虹膜辨識都需要近距離接觸才能進行身份鑑定，但是犯罪者本身當然不會自己去接觸辨識系統。而利用人臉辨識來進行身份鑑別則非常受到影像品質的影響，若欲偵測者在距離攝影機較遠的地方，辨識率就往往無法達到我們所期望的準確。

步態是指人們行走時的方式，這是一種複雜的行為特徵，因此也提供了充足的信息來鑑別每個人的身份。步態分析主要是針對一段行走的影像序列來做處理，但是，由於序列影像的數據量較大，人體外觀也很容易受到外在衣著的影響，因此步態鑑別的計算複雜性比較高，處理起來也比較困難。而利用步態來進行身份鑑別的優點在於

擷取資料的過程中不具有侵入性、不須近距離拍攝，亦可以在不讓受辨識者得知的情況下進行身份辨識進而達到安全監控。除此之外，步態辨識的應用很廣，例如：居家照護，藉由監控老人小孩的行動，分析其姿態與步態，達到自動偵測意外事件與安全監控的功能；阻止犯罪，將公司大樓攝影機拍到的人，藉由分析其步態進行身份鑑別，自動偵測陌生人並通知警衛前往關切，以期降低犯罪率。

## 1.2 相關研究之探討

近年來，步態分析一直是視覺監控的熱門研究領域，步態分析後的資訊可以用來偵測很多事件，例如跌倒偵測、行動不便人士偵測、行為怪異人士偵測、性別辨識和身份鑑別等等。由此可知，基於步態分析的結果可以應用在很多人和很多地方上。

在姿態分析應用在性別鑑別上，文獻[2]提出主成份步態影像法(Gait Principal Component Image, GPCI)，主要就是先算出影像的共變異矩陣，再藉由特徵值、特徵向量與參數的轉換，最後可得 GPCI，即為特徵，如圖 1.1 所示。



(a)



(b)

圖 1.1 文獻[2]性別辨識相關影像。(a)序列側影影像；(b)主成份步態影像(GPCI)。



文獻[2]主要就是使用圖 1.1 (b)中 GPCI 影像的二維矩陣做為分類與身份鑑別的特徵，最後使用近鄰法(k-nearest neighbor, KNN)來進行分類。

文獻[3]和[4]則是先把一周期內的影像全部疊加並取得步態能量影像後，再將人體全部分為七個部分，頭、臂、身體、大腿、前後腿和腳，如圖 1.2 所示：

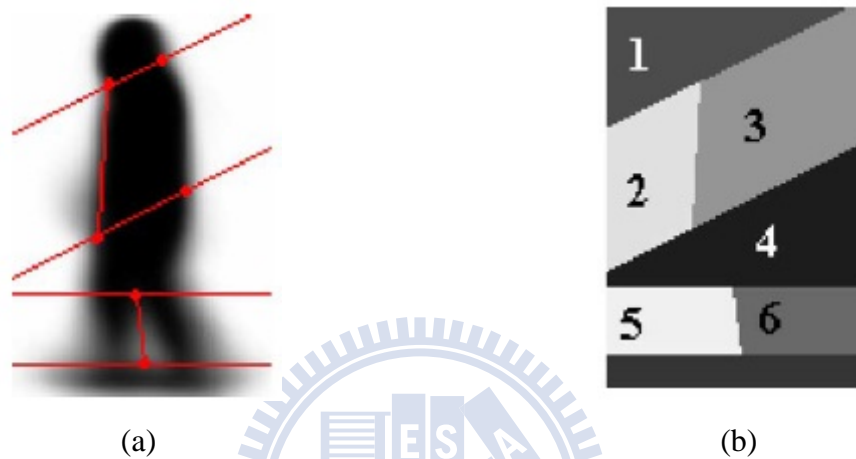


圖 1.2 文獻[4]性別辨識相關影像。(a) 人體步態能量影像；(b)將人體姿態分成 7 區域之示意圖。

文獻[4]主要是針對圖 1.2(a)中每一個區塊的影像進行支持向量基(support vector machine, SVM)分類，最後統計結果以鑑別身份。

在性別識別上，大多數都是採取將一序列影像全部混和、疊加或取平均等等後，再進行特徵抽取。這些特徵在分析性別上確有顯著的效果，但是不同於性別辨識只分為兩大類，身份鑑別要區分的類別更多，所以不僅需要這些靜態特徵，也需要再加入更多動態的、有效的以及具有鑑別力的特徵。

在姿態分析應用在身份鑑別上，早期文獻[5]先找出人體質心位置，再找出整張影像中的主方向和第二大方向，以質心為圓心畫橢圓，質心座標、長短軸長度大小和橢圓角度為特徵來進行身份鑑別。文獻[6]則是文獻[5]的進化版，將人體姿態分為七個橢圓來表示，如圖 1.3 所示：



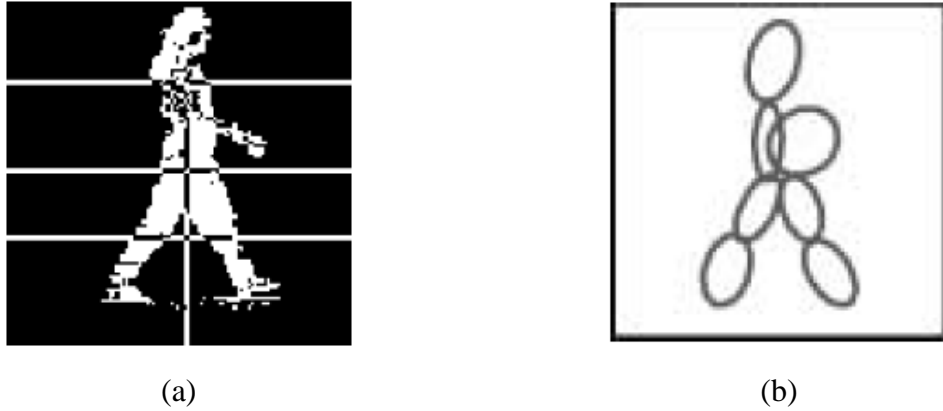


圖 1.3 文獻[6]身份鑑別相關影像。(a)將人體姿態影像分為七部分之示意圖；(b)利用橢圓去表示此七區域之示意圖。

文獻[6] 主要就是針對圖 1.3 (b)中的每個橢圓都使用同樣的特徵抽取法則，即每個橢圓的質心座標、長短軸長度大小和橢圓的偏移角度。對一周期中所有影像皆進行特徵抽取，最後再將這 7 個橢圓的序列特徵利用 SVM 來進行分類與身份鑑別。

文獻[7]和文獻[8]則是利用形態學的方式不斷侵蝕(Erosion)與斷開(Opening)來找出每一張步態影像的骨架圖，再將骨架圖分成六個區域，頭、軀幹、前後腿和前後腳，分別找出各區域的彎曲角度即為特徵，如圖 1.4 所示：

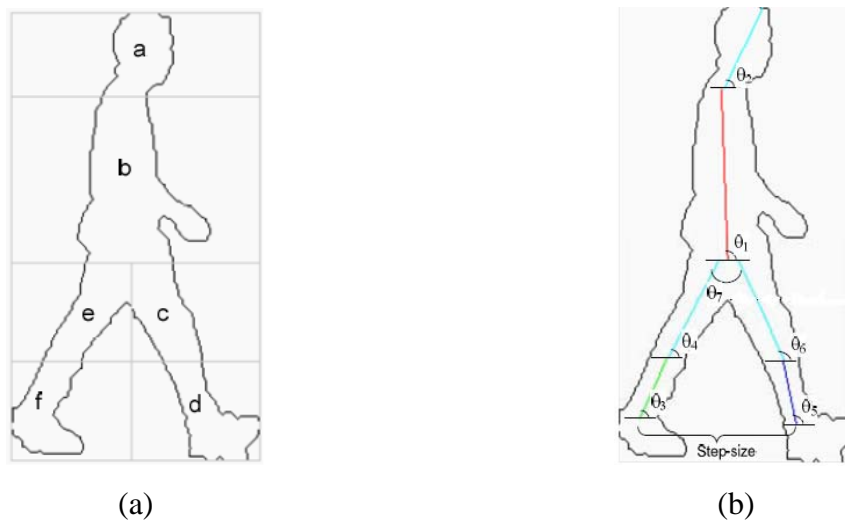


圖 1.4 文獻[8]身份鑑別相關影像。(a)人體六區域示意圖；(b)人體骨架示意圖。

在文獻[8]的方法中，主要就是透過計算圖 1.4(b)中人體姿態骨架圖上每一個關節點的角度，將這些角度當作特徵，最後使用模糊近鄰法(fuzzy k-nearest neighbor)將這些特徵進行分類與身份鑑別。

根據文獻[5]到文獻[8]的方法，都是將人體姿態分為幾個區域來抽取主方向、角度等特徵，但這些特徵都很容易受到變因的影響，例如:穿長裙或外套等等。一旦產生變因之後，人體姿態的部分影像區塊範圍改變，就例如穿長裙子，此時根據文獻[5]到文獻[8]的方法，必須分出下半身的兩隻腳，但此時從人眼只能看到一大片裙子，無法正確得知雙腳的位置，因此下半身腳部區域特徵的準確性就很容易受到影響，而辨識率也必然無法達到預期的成效。

文獻[9]是動態記錄每一張側影影像方框的長度、寬度、面積和方框對角角度來做為特徵，如圖 1.5 所示：

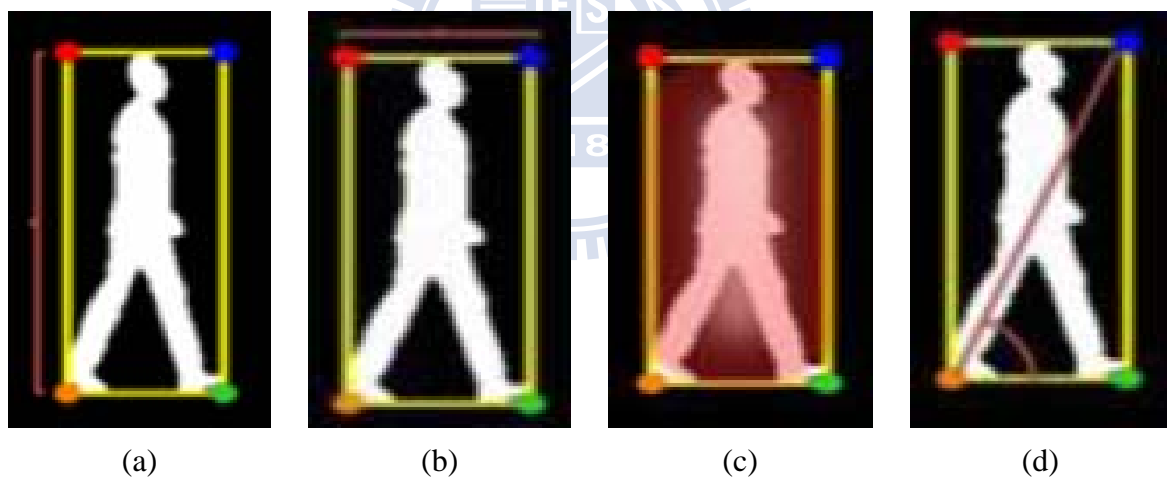
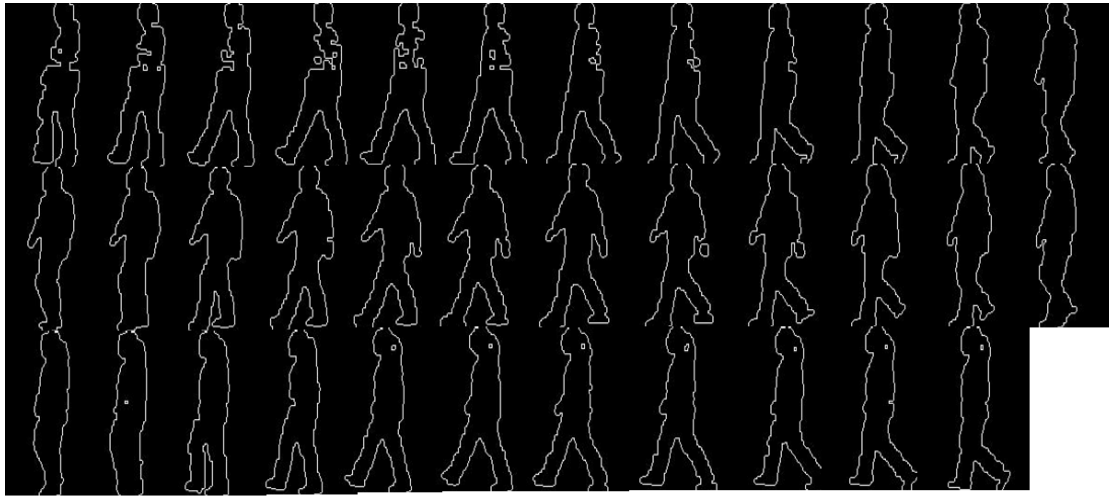


圖 1.5 文獻[9]身份鑑別相關影像。(a)長度特徵；(b)寬度特徵；(c)面積特徵；(d)對角角度特徵。

圖 1.5 為文獻[9]演算法中，抽取特徵的示意圖。這些特徵主要都是依靠在一周期中每一張側影姿態的方框大小，雖然相對其他方法來說，抽取這些特徵的複雜度是較低的，但方框取的好壞便決定了特徵的準確性，而在擷取方框得過程中又很容易受到影像中雜訊的影響，因此就辨識效能來說穩定性還不夠。

文獻[10]先將一周期序列中的輪廓影像相加，得到動態輪廓影像(motion contour

image, MCI)後，如圖 1.6 所示：



(a)



(b)

圖 1.6 文獻[10]身份鑑別相關影像。(a)人體步態輪廓示意圖；(b)動態輪廓影像(motion contour image, MCI)。

經由事先的訓練算出動態輪廓影像的投影矩陣，進入測試階段後，只要將輸入的動態輪廓影像乘上投影矩陣即為特徵。其中投影矩陣的主要目的在於使影像特徵降維，儘管如此，此特徵的維度依然很大，就此文獻的紀錄，其特徵維度大小為  $17 \times 40$ 。

### 1.3 研究動機

在社會新聞中，罪犯在犯罪時常常都會替自己喬裝，不讓自己身上的任何東西甚至是頭髮掉在犯案現場，但有樣東西是很難隱瞞的，那就是走路的姿勢，即步態。因此，本論文希望能藉由分析步態的不同來進行身份鑑別，進而達到自動安全監控的功能。

在相關文獻中步態分析在特徵抽取方面大多都是從同一角度進行分析，例如使用步態能量影像（gait energy image，GEI）分析者，就將一周期的序列影像全部疊加在一起，但此時便失去了分析序列中每一張影像的重要動態變化資訊；相反的，使用人體骨架圖分析者，大部分就只針對動態變化的資訊來進行分析與研究，此時，便失去了靜態特徵的重要資訊。

因此，為了使身份鑑別系統更加強健，本系統不僅結合靜態與動態特徵進行分析與研究，更針對人體姿態的不同區域使用相對應適合的分析方法，平均分散特徵於人體各個部位，主要目的便是希望能更加提高步態辨識之身份鑑別的準確度。

## 1.4 論文主體與貢獻

在進行人體姿態分析的時候，系統的流程主要可以分成四個步驟：前景切割、週期估測、特徵抽取，以及分類辨識。本論文將在第二章介紹相關技術及原理，在第三章則介紹本論文將提出的系統流程，第四章說明本系統的實驗結果與相關領域論文方法之比較與分析，第五章則是結論與未來展望。

在本論文的研究中，將會把重點集中在特徵抽取，以及分類辨識上，本論文所達成的貢獻有兩點，如下所示：

1. 本論文結合多種方法，使用多面向與多角度來分析姿態，針對人體不同區域的特性，提出相對應適合的分析方法。不僅將姿態特徵分為動態與靜態兩大特徵來分別討論，更將特徵平均分散在人體的各個部位，以確保本系統面對衣著變因的影響時，還能保持一定的辨識率。
2. 本論文提出兩階段式的身份鑑別系統。第一階段主要是利用外輪廓、手部擺動和腳掌離地角度三個特徵來進行候選人篩選，而篩選的目的在於減少候選人，以提高整體辨識率；第二階段則是將全部特徵針對第一階段篩選出來的候選人進行權重給分，分數最高的即表示與此人最為相似。從實驗結果來看，本系統確實有不錯的身份識別結果。

## 第二章 相關技術與原理

本章節將會介紹與本論文有關的知識與理論，因此，本章共分為四節，在 2.1 節介紹本論文所使用的前景背景分割法；在 2.2 節介紹靜態特徵的抽取方法；在 2.3 節介紹動態特徵的抽取方法；2.4 節則是介紹本論文用來做身份鑑別的主要演算法。

### 2.1 前景背景切割法

在視覺監控的研究中，前景背景的切割對後續影像處理的結果影響最劇，因此，如何準確地將前景切割出來便成為步態辨識的主要關鍵。而背景前景分割法主要分成三大類：時間差法(temporal differencing)、光流法(optical flow)和背景相減法(background subtraction)。

使用時間差法來提取前景的速度較快，但是擷取出的前景目標容易較破碎與不完整；使用光流法能夠在事先不知道場景的條件下擷取出獨立運動的目標，但是其計算複雜、耗時，不容易達到即時性的要求；而背景相減法只需在事先訓練並建立背景模型，在之後處理中，只要將輸入影像減去背景模型後即可擷取出前景目標影像。因此，在綜合所有因素之後，本論文在這邊是採用高斯混和模型來訓練背景，並利用背景相減法的方式來提取前景，以期盡量保存前景目標的完整性。

在背景相減法中，高斯混和模型(Gaussian mixture model, GMM)是一種較為常用的演算法，且該方法在擷取出前景目標的完整性都有不錯的表現。高斯分布，也就是常態分布。對於隨機變數  $x$ ，其機率密度函數為：

$$p(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2-1)$$

其中  $\mu$  為高斯分布的平均值， $\sigma$  為高斯分布的標準差。如果一組數據符合高斯分布，那麼這些數據大多數會集中在以  $\mu$  為中心的  $-2\sigma$  到  $2\sigma$  範圍內。其分布函數曲線如圖

2.1 所示：

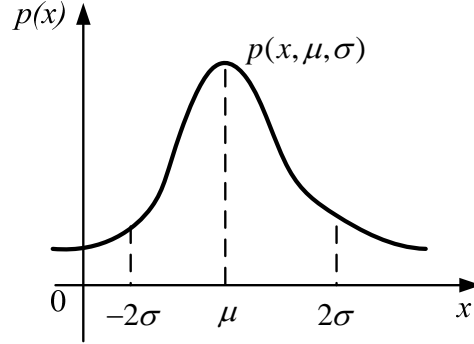


圖 2.1 高斯分布示意圖

對於某一場景的視訊影像，如果其背景相對為靜止狀態，那麼背景中的每個像素點在一段時間序列中的排列就可以用一個高斯分布來描述。但是實際上，對於戶外場景，由於受到各種雜訊的干擾，其背景變化較大，會出現樹枝搖動，光線變化等不同情況。因此用一個高斯分布不能完全描述實際的背景，這種情況下就需要採用多個高斯模型來描述動態的背景，為不同的狀態建立不同的高斯模型。將  $k$  定義為建立高斯模型的數目，那麼當前觀測點像素值的機率密度為：

$$p(x_t) = \sum_{i=1}^k w_{i,t} \cdot p(x_{i,t}, \mu_{i,t-1,k}, \sigma_{i,t-1,k}) \quad (2-2)$$

其中， $w_{i,t}$  是每個單一高斯模型的權重值， $\mu_{i,t-1,k}$  為第  $i$  個單一高斯模型的平均值， $\sigma_{i,t-1,k}$  為第  $i$  個單一高斯模型的標準差。

接著，先對一段背景序列影像進行訓練，透過以上的運算過程計算出平均值、標準差和權重等參數來做為背景影像模型的參數。藉由一段序列影像訓練出來的背景模型可以減低背景中擾動因素的影響，而訓練的序列影像越多，背景模型的精確度也越高，但相對在訓練時間上也會愈長。



最後，利用背景相減法，將輸入影像減去由高斯混合模型訓練出來的背景，即可完成前景提取的動作，如圖 2.2 所示：

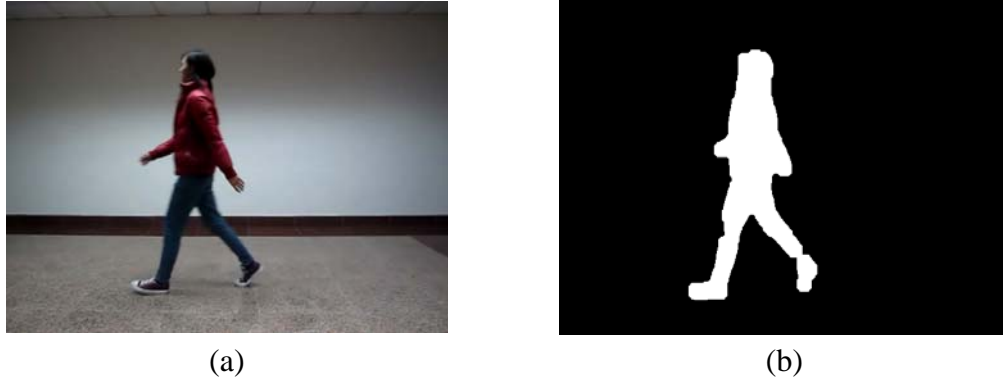


圖 2.2 利用 GMM 擷取前景範例。(a)輸入影像；(b)經過 GMM 前景提取後的結果。

## 2.2 靜態特徵

在這邊的靜態特徵，主要是針對一周期序列影像中的統計平均影像來做分析，所以在抽取靜態特徵之前必須先得到步態能量影像。藉由疊加所有步態序列影像可以得到步態能量影像，接著從步態能量影像中觀察每個人在一周期走路中身體整個外輪廓前後移動的形狀、手部擺動的幅度等等。由於每個人的習慣不同，所以每個人在步態能量影像中的表現也不同。

在本論文中，主要是分析步態能量影像來擷取靜態特徵，其中包含步態能量影像的外輪廓特徵與手部擺動幅度。因此，在本節中，主要又分為兩小節，2.2.1 主要介紹步態能量影像的計算方式與其優缺點；2.2.2 介紹傅立葉描述子的計算過程，在這邊主要是利用傅立葉描述子來描述步態能量影像的外輪廓特徵。

### 2.2.1 步態能量影像

步態能量影像(gait energy image, GEI)，也叫平均步態影像，這是一種屬於統計上的特徵。最初是由 Bobick 和 Davis [11]所提出，Han 和 Bhanu [12]將此方法應用在姿態辨識上，而 X. Li, S. J. Maybank, S. Yan, D. Tao 和 D. Xu[4]更進一步將步態能量影

像應用在性別鑑別上。

步態能量的計算方式如下：

$$G(x, y) = \frac{1}{N} \sum_{t=1}^N B_t(x, y) \quad (2-3)$$

其中  $B_t$  為單張步態影像， $N$  為一週期序列影像中的影像數目， $x$  和  $y$  為二維影像中的座標位置。

在進行影像疊加之前，會先將每一張步態影像  $B_t$  正規化成一樣的高度，在本論文中是將人體步態影像正規化成 140 的高度，且等比例的縮放寬度大小。例如：現在輸入的步態影像大小為  $274 \times 72$ ，將長度正規化成 140 的大小後，接著寬度大小則會被正規化成  $(140/274) \times 72 = 37$ ，所以整張步態影像就會被正規化成  $140 \times 37$  的大小。因為在進行正規化時是等比例的縮放原始影像，所以並不會改變原始輸入影像的體態大小。對每一張步態影像都進行正規化之後，接著以每張步態影像的水平中心線為基準線來進行影像疊加的動作以獲得步態能量影像。

步態能量影像主要是將一周期中的序列影像全部相加再取平均，如圖 2.3 所示：

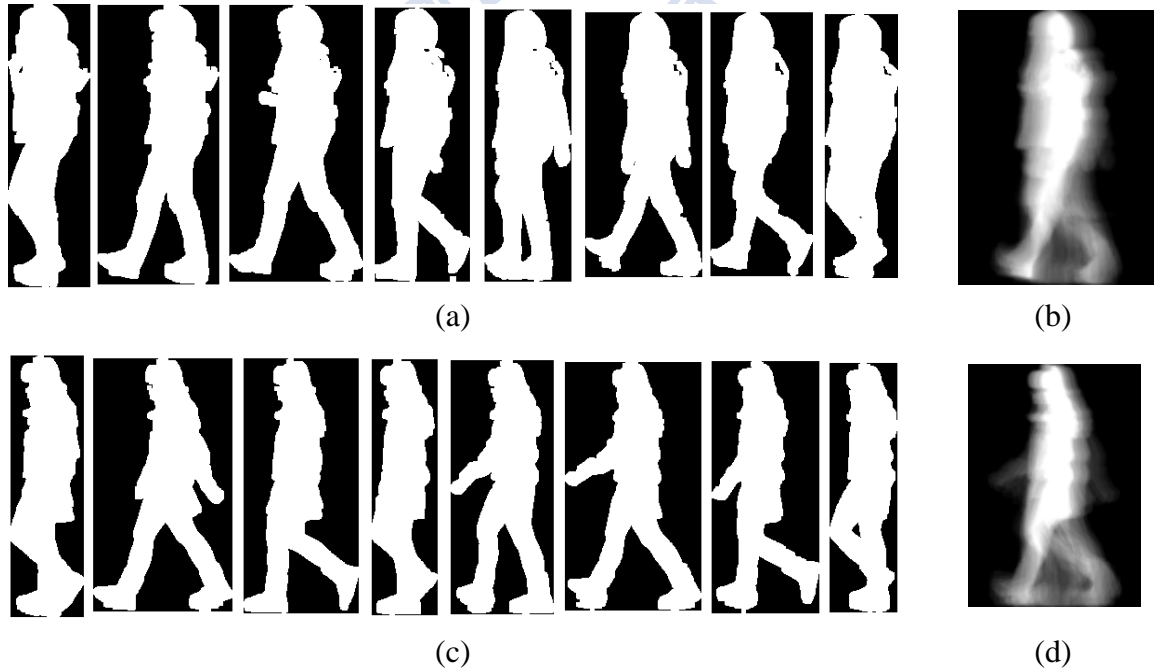


圖 2.3 步態能量示意圖。(a) 步態序列影像；(b)步態能量影像；(c) 步態序列影像；(d) 步態能量影像。



步態能量影像累積了人在行走中的所有姿態，而像素值較高的區域即表示此人在行走中頻繁使用到的位置。圖 2.3 中(a)和(c)為兩個不同的人之步態序列影像，從(b)和(d)的步態能量影像中即可察覺此二人的不同處，例如：(d)在手部擺動的幅度上明顯較(b)來的大、在下半身腳步區域的能量分布圖也不同。由此可知，步態能量的最大優點在於能在單張影像中就表現出人在走路一周期過程中的主要輪廓形狀與變化。

## 2.2.2 傅立葉描述子

在進行影像目標的識別時，首先必須要做的就是抽取目標的特徵，然後用適當的數學表示對目標進行描述。對目標特徵提取的算子稱為目標檢測子(detector)，對目標描述的算子稱為描述子(descriptor)。一般來說，目標描述子需要具備以下四個性質：

1. 完備性：僅僅只有當兩個目標有相同的形狀時，其描述子才會是相同的。
2. 一致性：相似的兩個目標物，其描述子也應該是相似的。
3. 不變性：描述子對目標的幾何變換必須具有不變性，例如平移、旋轉和尺度伸縮等等必須保持不變性。
4. 緊湊性：越好的目標物描述子可以用越少的係數來表示目標的特性。

在所有描述子中，傅立葉描述子(Fourier descriptor, FD)最常用來表示封閉曲線的形狀特徵。其基本的原則[18]就是將目標輪廓曲線轉換成一維序列，對該序列進行一維的傅立葉變換，並從中獲得一系列的傅立葉係數，用這些係數來描述目標物的輪廓。

使用傅立葉描述子的優點在於計算簡單，而計算原理簡單可以讓特徵提取更為穩定，所以在計算過程中，不需要設置大量控制參數就可以獲得結果。此外，由於任何一段序列影像輪廓經過傅立葉變換後，其能量主要都會集中於少數幾個低頻傅立葉係數上，因此採用極少的傅立葉係數就可以描述該序列的特徵。而隨著傅立葉係數的增多，該序列的細節特徵得以更好的被描述。因此，傅立葉描述子對目標物輪廓有非常好的描述能力。

由於人體步態的輪廓線也是一條封閉的曲線，所以在本論文中主要是利用傅立葉

描述子來描述姿態能量影像的外輪廓曲線。首先，把輪廓邊界上每一點都看成一個向量，用複數形式表示如下：

$$z_i = x_i + jy_i, \quad i = 0, 1, \dots, N-1 \quad (2-4)$$

其中  $N$  為輪廓點數，步態輪廓的質心座標為  $(x_c, y_c)$ ：

$$x_c = \frac{1}{N} \sum_{i=0}^{N-1} x_i \quad (2-5)$$

$$y_c = \frac{1}{N} \sum_{i=0}^{N-1} y_i \quad (2-6)$$

選定某個點為起點，本論文中是選擇最右下方的點為起點，沿逆時針方向展開並計算步態輪廓上每一點與質心的距離  $d_i$ ：

$$d_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}, \quad i = 0, 1, \dots, N-1 \quad (2-7)$$

將一個序列的二維輪廓影像經過方程式(2-7)的轉換，對應成一序列的一維距離信號。此時，特徵序列  $\mathbf{D} = \{d_0, d_1, \dots, d_{N-1}\}$ 。接著，對特徵序列  $\mathbf{D}$  進行離散傅立葉變換 (discrete Fourier transform, DFT)，可得到離散傅立葉變換係數  $a_n$ ，如下所示：

$$a_n = \frac{1}{N} \sum_{i=0}^{N-1} d_i e^{-j2\pi ni/N}, \quad n = 0, 1, \dots, N-1 \quad (2-8)$$

方程式(2-8)運算出來的傅立葉係數為複數型態  $a_n = \alpha_n + j\beta_n$ ，接著我們取出這些係數的數值大小  $|a_n|^2 = \alpha_n^2 + \beta_n^2$  來做為特徵，最後再以  $|a_0|^2$  為基準進行正規化處理：

$$a_n^* = \frac{|a_n|^2}{|a_0|^2}, \quad n = 0, 1, \dots, N-1 \quad (2-9)$$

進行正規化的原因在於：正規化後的傅立葉描述子具有尺度、平移和旋轉的不變性。而正規化後的這一序列係數即代表此封閉曲線的傅立葉描述子(Fourier descriptors, FD)如下所示：

$$\mathbf{FD} = [a_0^* \quad a_1^* \quad \dots \quad a_{N-1}^*] \quad (2-10)$$

而此傅立葉描述子即為本論文中用來描述步態影像的外輪廓特徵之方法。

## 2.3 動態特徵

在這邊的動態特徵，主要是對一序列影像中每一張步態影像都去進行抽取特徵的動作。由於每個人習慣上的不同，例如在走路過程中有人習慣低頭走路，有人喜歡抬頭走路、有人走路步伐距離較大、有人走路步伐距離較小等等。由於這些習慣性的小動作造就每個人獨特的走路特色，而本論文就是藉由分析這些小動作來進行身份的鑑別。

在本論文中，動態特徵的部分又分為頭部傾斜角、膝蓋彎曲角度、腳掌離地角度和步伐大小距離，在擷取這些特徵的過程中，主要使用到的方法有主成分分析法(principal components analysis, PCA)和哈里斯角點偵測法(Harris corner detector)。因此，在本節中，又分了兩小節，2.3.1 介紹主成分分析法，主要是用在分析頭部傾斜角和腳掌離地角度；2.3.2 介紹哈里斯角點偵測法，主要是用在分析膝蓋彎曲角度。

### 2.3.1 主成份分析法

主成份分析由 K. Pearson 於 1901 年提出，後來由 Hotelling 在 1933 年再以發展。在統計分析中，主成份分析(principal components analysis, PCA)是一種分析、簡化數據的方法。其主要目的在於希望用較少的變數去解釋原來資料中的大部分變異，也就

是希望能將資料中許多相關性很高的變數轉換成彼此互相獨立的變數。在圖形辨識中，主成份分析算是比較快速又準確的方式之一，他可以對抗圖形平移旋轉的現象，在多個特徵資訊中擷取最主要的幾個來做為特徵依據。

主成份分析法的數學定義是：一個正交化線性變換，把數據變換到一個新的座標系統中，使得所有數據投影的第一大變異數在第一個座標(稱為第一主成分)上，第二大變異數在第二座標(第二主成分)上，依此類推。例如現在有  $N$  組的資料向量  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ ，每個輸入向量都具有  $D$  個維度，利用主成分分析可以將整體輸入向量的維度降低到  $S$ ，其中  $S < D$ 。主要做法流程如下，

1. 將輸入資料向量  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$  組合成一個  $D \times N$  的輸入矩陣：

$$\mathbf{X} = [\mathbf{X}_1 \quad \mathbf{X}_2 \quad \dots \quad \mathbf{X}_N] \quad (2-11)$$

2. 計算矩陣  $\mathbf{X}$  的平均向量  $\mathbf{u}$ ：

$$\mathbf{u} = \frac{1}{N} \sum_{i=1}^N \mathbf{X}_i \quad (2-12)$$

3. 將每個  $\mathbf{X}_i$  扣除平均向量  $\mathbf{u}$  後，可得到結果  $\mathbf{B}_i$ ：

$$\mathbf{B}_i = \mathbf{X}_i - \mathbf{u} \quad (2-13)$$

4. 計算其共變異矩陣  $\mathbf{C}$ ：

$$\mathbf{C} = \frac{1}{N} \sum_{i=1}^N \mathbf{B}_i \mathbf{B}_i^T \quad (2-14)$$

5. 找出共變異矩陣  $\mathbf{C}$  的特徵值(eigenvalue)  $\lambda_d$ ， $d = 1, 2, \dots, D$ ，和特徵向量

(eigenvector)  $\mathbf{v}_d$ 。接著取出最大的  $S$  個特徵值與相對應的特徵向量  $\mathbf{v}_s$ ,  $s=1,2,\dots,S$ ，將每一個特徵向量除以自己的長度後會形成一組對應的基底向量(basis vector)，而其中每一組基底向量則對應到一個特徵空間(eigenspace)。

6. 將這  $S$  個正規化的特徵向量合成一個  $D \times S$  的轉換矩陣  $\mathbf{W}$ ，即：

$$\mathbf{W} = [\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \quad \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} \quad \dots \quad \frac{\mathbf{v}_S}{\|\mathbf{v}_S\|}] \quad (2-15)$$

7. 最後用  $\mathbf{Y}$  來代表主成分分析的輸出：

$$\mathbf{Y} = \mathbf{W}^T \mathbf{X} \quad (2-16)$$

最後輸出的  $\mathbf{Y}$  為一個  $S \times N$  的矩陣，其中每個行向量就是相對於  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$  做主成分分析的結果。經由這些步驟之後，不僅可以把最具代表性的  $S$  個維度保留下來，同時也可以達到資料降維的目標。

### 2.3.2 哈里斯(Harris)角點偵測法

Harris[13]於 1988 年提出一個區域角點偵測的演算法，此方法是基於 Moravec[14]在 1980 年提出的角點偵測進行衍生後的結果。

Moravec 的方法主要是使用一個方形的搜尋視窗(window)在影像中做小距離的移動，並算出移動前後視窗內影像強度變化的量，即視窗內所有像素值的梯度變化總值，如方程式(2-17)所示：

$$E(u, v) = \sum_{x, y} w(x, y) [I(x+u, y+v) - I(x, y)]^2 \quad (2-17)$$

其中  $I$  為影像， $w$  代表影像搜尋視窗。點  $(x, y)$  為影像座標點位置，點  $(u, v)$  為移動視窗座標點位置。

Moravec 主要是透過  $E(u,v)$  的大小來判斷視窗掃描到的區域為平面(flat)、邊界(edge)或轉角(corner)中的哪一種情況，在這邊參考[19]的說法針對這三種情況會產生的現象來進行分析，如下所述：

- (1) 平面(flat)： 視窗往任何方向移動其  $E(u,v)$  值的結果都很小，表示此區域的梯度變化不明顯，在影像上屬於較平滑的區域。
- (2) 邊界(edge)： 若視窗順著邊界的方向移動，則  $E(u,v)$  值很小；反之，若視窗移動的方向與邊界垂直，則  $E(u,v)$  值很大。
- (3) 轉角(corner)： 不論視窗往哪個方向移動，其  $E(u,v)$  值都很大。根據位移後得到最大的梯度變化值就可以偵測出轉角的位置。

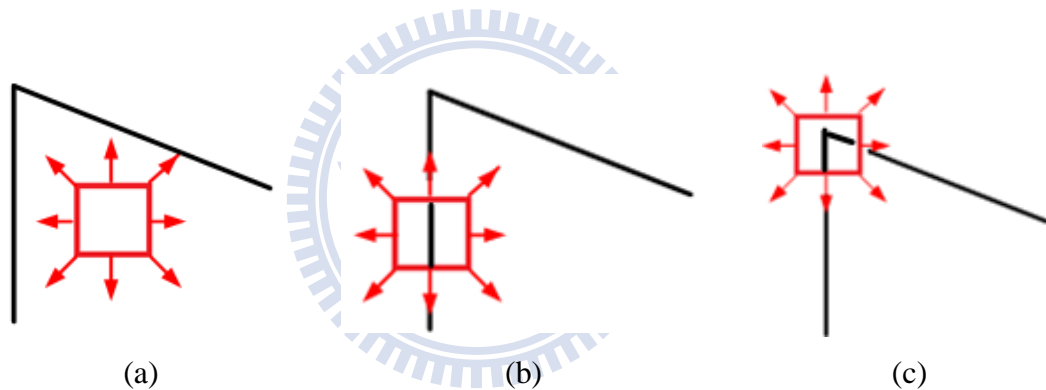


圖 2.4 Moravec 偵測邊角範例。(a)搜尋視窗在平面時之示意圖；(b)搜尋視窗在邊界時之示意圖；(c)搜尋視窗在轉角時之示意圖。

Moravec 希望能找出所有  $E(u,v)$  值都很大的影像座標點  $(x, y)$ ，但是在權衡準確度與運算速度後，Moravec 只做了八個方向的位移，從這八個方向的運算中找出最小的  $E(u,v)$  值，然後產生一張影像，並在這張影像中最小的  $E(u,v)$  找區域最大(local maxima)，而此點就是我們要找的特徵點。但是 Moravec 演算法還是有一些缺點，如下所述：

- (1) 只檢測了視窗函數在八個基本方向上移動的強度變化，不能很準確地找出全部的角點。

- (2) Moravec 沒有對影像進行任何降低雜訊的處理，所以此方法很容易就會受到雜訊的影響。
- (3) Moravec 對於邊界的響應很敏感，相對其  $E(u,v)$  值都很大，因此很難從中辨別是邊界還是角點。

Harris[13]主要就是針對 Moravec 偵測法中的三個缺點加以改良：

- (1) 由於偵測視窗的移動並未兼顧每個方向，所以 Harris 等人利用泰勒展開式(Taylor expansion)，考慮所有最小的變化，忽略方程式(2-17)的高次項後可以表示成方程式(2-18)和(2-19)的樣子：

$$E(u,v) = \sum_{x,y} w(x,y) [I(x+u, y+v) - I(x,y)]^2 \quad (2-18)$$

$$= \sum_{x,y} w(x,y) [xX + yY + I(x^2, y^2)]^2$$

$$E(u,v) = Ax^2 + 2Cxy + By^2 \quad (2-19)$$

其中， $A$  為  $X$  方向的梯度平方值， $B$  為  $Y$  方向的梯度平方值， $C$  為  $XY$  方向梯度，如下所述：

$$A = \sum_{x,y} w(u,v) I_x^2(x,y) \quad (2-20)$$

$$B = \sum_{x,y} w(u,v) I_y^2(x,y) \quad (2-21)$$

$$C = \sum_{x,y} w(u,v) I_x(x,y) I_y(x,y) \quad (2-22)$$

所以根據線性代數的原理，可以將  $E(u,v)$  重新定義為：

$$E(u, v) \cong [u \quad v] M \begin{bmatrix} u \\ v \end{bmatrix} \quad (2-23)$$

$$M = \sum_{x,y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} = \begin{bmatrix} A & C \\ C & B \end{bmatrix} \quad (2-24)$$

(2) 在降低雜訊干擾的方法中，Harris 等人是選擇高斯平滑(Gaussian smooth)來做改善：

$$w(x, y) = e^{-\frac{(u^2+v^2)}{2\sigma^2}} \quad (2-25)$$

(3) 在針對利用  $E(u, v)$  最小值的比較法難以準確辨別出邊界或角點的缺點上，在這邊頂多只能設置一個門檻值(threshold)去定義角點或不是角點。

根據方程式(2-24)的算法， $M$  可由影像本身得到，然後我們針對  $M$  去求出其特徵值(eigenvalue)  $\lambda_1$  和  $\lambda_2$ ，接著將方程式(2-23)改寫為：

$$E(u', v') = [u' \quad v'] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} u' \\ v' \end{bmatrix} \quad (2-26)$$

由方程式(2-26)可以推得，在位移時，若  $\lambda_1$  和  $\lambda_2$  都很大，表示  $E(u, v)$  不管位移角度為何，都會讓  $E(u, v)$  很大，也就表示此區域為角點的區域；相反的，若  $\lambda_1$  和  $\lambda_2$  呈現一大一小的狀態時，表示當位移其中一個方向時會使得  $E(u, v)$  很小，也就表示此區域為邊界的區域；若  $\lambda_1$  和  $\lambda_2$  都很小，表示不管如何位移  $E(u, v)$  都很小，也就表示此區域為平面的區域。詳細的分布情形如圖 2.5 所示：



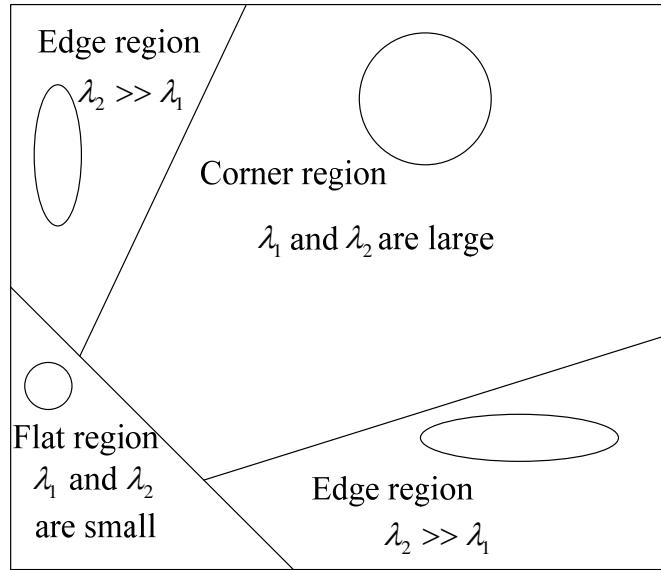


圖 2.5 特徵值分布情形。

## 2.4 身份鑑別法則

本論文的身份鑑別方法主要是先建立資料庫，最後再將輸入資料與資料庫資料做比對，找出最相似的人以完成身份辨識。在這邊最直覺的想法就是選擇使用最近鄰居法( $k$ -th nearest neighbor, KNN)來進行分類，首先算出資料庫中每個點到該點的距離後，再找出距離最近的  $k$  個點。但是利用這樣的方法來做  $k$  個鄰近點搜尋時，會出現一個很大的問題，即速度非常緩慢，必須算出資料庫中每個點到該點的距離後，再搜尋出最近的  $k$  個點，這樣的計算量是相當龐大且沒效率的。

考慮以上的因素，本論文在這邊選擇使用 KD-tree 的方式來搜尋鄰近點，以改善傳統 KNN 做法的缺點，並提升身份辨識的效率。因此，在本節中又分了兩小節來做介紹，2.4.1 主要介紹傳統的最近鄰居法( $k$ -th nearest neighbor, KNN)；2.4.2 則是介紹 KD-tree 的演算法流程。

### 2.4.1 最近鄰居法

最近鄰居法( $k$ -th nearest neighbor, KNN)，意思就是找出  $k$  位最接近的鄰居，是一

種屬於類神經網路的演算法。類神經網路是一種仿生物神經網路的結構與功能的數學模型，是一種能夠學習、能夠總結歸納的系統，而它主要又分成兩種學習方式：監督式學習和非監督式學習。

監督式學習，是指在輸入訓練資料的時候同時也包括了輸出資料，即在訓練過程中，每一筆資料都會有一個對應的正確答案，而類神經網路在學習以後，會依據答案來調整自己的內部參數。非監督式學習，是指在輸入訓練資料的時候並沒有給輸出結果，即每一筆資料並沒有正確答案。所以在輸入新的測試資料時，類神經網路只會依據訓練出來的模型，來判斷和哪一類比較相似。

KNN 在類神經網路中是屬於監督式學習，在經過事先產生資料庫模型後，輸入一筆新的測試資料，如圖 2.6 所示：

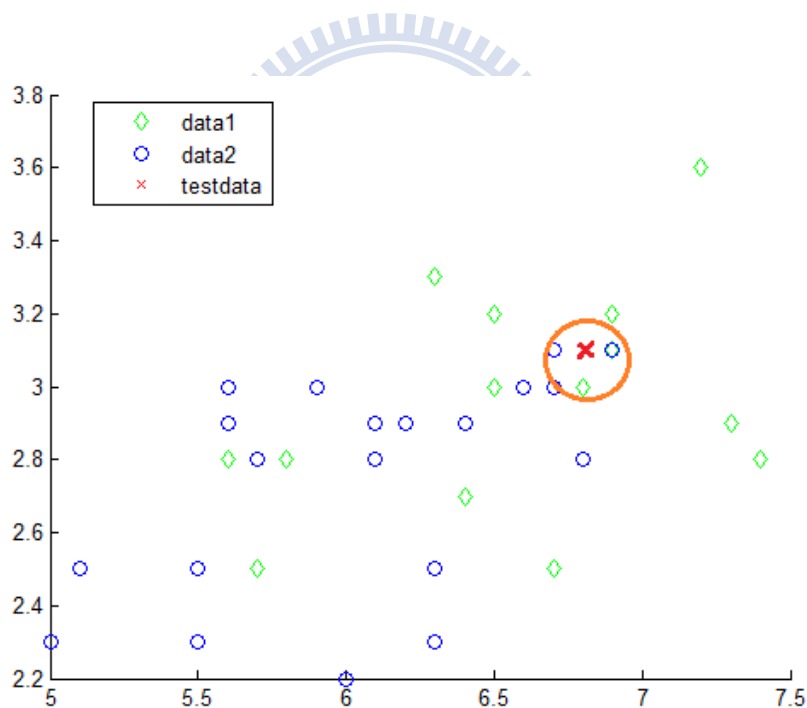


圖 2.6 KNN 演算法示意圖。

圖中紅點為新輸入的測試資料。KNN 演算法中計算新輸入點和資料庫模型中每一點的距離公式主要是使用歐幾里得距離法，如下所示：

$$dis(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2-27)$$

其中  $x$  為新輸入的資料， $y$  為資料庫中的訓練資料點， $n$  為資料點的維度大小。若現在令  $k=3$ ，即找出新輸入資料點的最近 3 個鄰居，搜尋結果如圖 2.6 中橘色圓圈所示，三個鄰居中有兩個鄰居屬於 data2，一個屬於 data1，所以在本範例中此輸入資料將會被判屬於 data2 那一類。若現在是要找最近的  $k$  個鄰居，則以此類推。

## 2.4.2 KD tree

KD tree 最早是由 Jon Louis Bentley 所提出[15]，在本論文中，主要是在資料庫中尋找鄰近點來進行身份鑑別，為了改善 KNN 的缺點，提高分類的效率，我們採用 KD-tree 的方法來解決速度上的問題。KD 一詞即為  $K$  維的意思，是可用於高維度資料的樹狀資料結構。圖 2.7 是一個針對二維資料進行 KD-tree 分類的範例圖，令  $X$  為第一維度， $Y$  為第二維度：

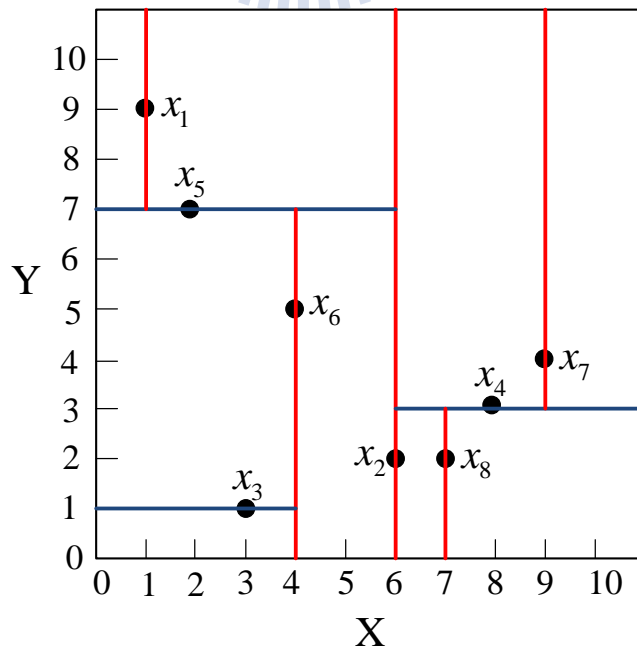


圖 2.7 KD-tree 示意圖。

KD-tree 做法是以第一維度的中位資料為界線，將資料以第一維度一分為二，再以第二維度的中位資料將資料分為兩大類，依序執行到  $K$  維後，再回到第一維，重複以上動作，直到每個小區塊都只有一筆資料為止。其中，中位資料是指將數據按大小順序排列起來，形成一個數列，居於數列中間位置的那個數據，例如：有一筆資料為  $[6, 2, 7, 1, 4]$ ，經排列後得  $[1, 2, 4, 6, 7]$ ，則中位資料等於 4。

在這邊舉例說明如何將二維資料建立成一顆 KD-tree，其法則如下所述：

二維資料點： $x_1 = (1, 9), x_2 = (6, 2), x_3 = (3, 1), x_4 = (8, 3), x_5 = (2, 7), x_6 = (4, 5), x_7 = (9, 4), x_8 = (7, 2)$

第一層：根據資料的第一維度找出中位資料。

第一維度資料有  $[1, 6, 3, 8, 2, 4, 9, 7]$ ，經由排列可以得到  $[1, 2, 3, 4, 6, 7, 8, 9]$ ，在資料為偶數筆時，選擇較大資料的來進行二分法，所以在此範例中選擇  $x_2 = (6, 2)$  來做為第一層。

第二層：以  $x_2$  做為分界，將資料分為第一維度較 6 小的一群  $A : (x_1, x_3, x_5, x_6)$ ，以及較 6 大的一群  $B : (x_4, x_7, x_8)$ 。其中，

- 在  $A$  群中，利用第二維度資料排序後得  $[1, 5, 7, 9]$ 。因為資料為偶數筆，所以選擇較大的資料來進行二分法，而 7 屬於  $x_5$ ，所以選擇  $x_5$  來進行第二層的二分法，因此樹狀圖左半邊第二層得到  $x_5 = (2, 7)$ 。
- 在  $B$  群中，利用第二維度資料排序後得  $[2, 3, 4]$ 。而中位資料 3 屬於  $x_4$ ，選擇  $x_4$  來進行二分法，因此樹狀圖右半邊第二層得到  $x_4 = (8, 3)$ 。

第三層：以  $x_5$  為左半邊的分界，將第二層步驟中  $A$  群剩餘的資料中，找出第一維度較 2 小的一群  $A_a : (x_1)$ ，較 2 大的一群  $A_b : (x_3, x_6)$ ；以  $x_4$  為右半邊的分界，將第二層步驟中  $B$  群剩餘的資料中，找出第一維度較 8 小的一群  $B_a : (x_8)$ ，較 8 大的一群  $B_b : (x_7)$ 。其中，

- 在  $A_a$  群中，因為  $x_1$  的第一維度數值 1 小於  $x_5$  的第一維度數值 2，所以將  $x_1$  放在  $x_5$  左半邊。
- 在  $A_b$  群中，還有兩個資料點  $x_3$  和  $x_6$  還沒二分完成。由於資料點為偶數，所以在這邊選擇較大的數值 4 來進行二分，而 4 屬於  $x_6$ ，所以在這邊選

擇  $x_6$  來進行第三層的二分法。

- 在  $B_a$  群中，因為  $x_8$  的第一維度數值 7 小於  $x_4$  的第一維度數值 8，所以將  $x_8$  放在  $x_4$  左半邊。
- 在  $B_b$  群中，因為  $x_7$  的第一維度數值 9 大於  $x_4$  的第一維度數值 8，所以將  $x_7$  放在  $x_4$  右半邊。

第四層：  $A_{bb}$  中還有一個資料點  $x_3$  還沒二分完成。由於本範例的資料為二維，所以

第四層又回到以第二維度資料來進行二分法。其中，因為  $x_3$  第二維度數值 1 較  $x_6$  的第二維度數值 5 還小，所以  $x_3$  放在  $x_6$  的左邊。

最後得到的結果如圖 2.8 所示：

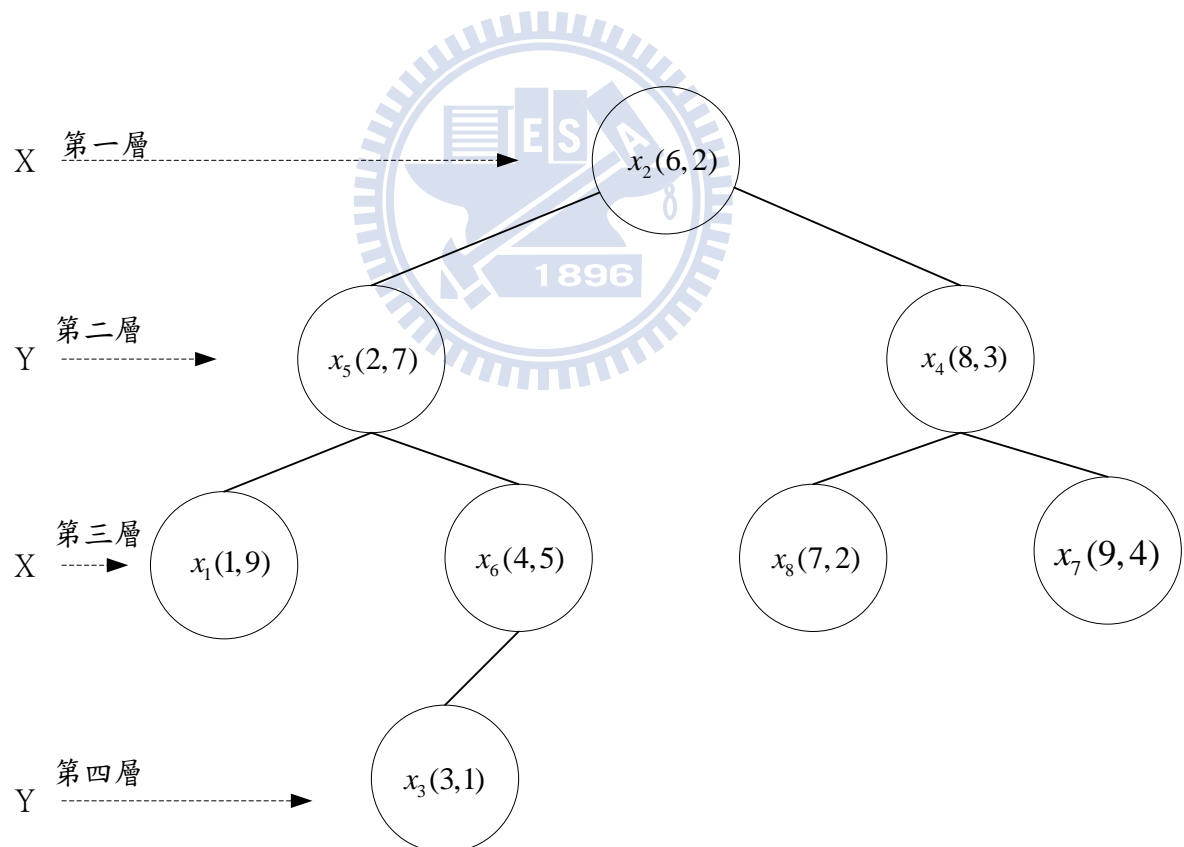


圖 2.8 KD-tree 樹狀圖。

## 第三章 系統流程

考慮姿態變因對於身份鑑別準確度影響之劇，本論文提出一個結合多種方法，使用多面向與多角度來分析姿態，更針對人體不同區域的特性，提出相對應適合的方法來抽取特徵，以期提高單一特徵的辨識率，進而提高整體身份鑑別的辨識率與強健性。

本章共分為五節，先說明整體系統流程，再進一步說明特徵抽取對象與方法及分類辨識策略；3.1 節介紹整體系統架構與流程；3.2 和 3.3 節說明在進入正式的特徵抽取分析前，所需要進行的預處理，包含前景提取與週期估測；3.4 節介紹系統架構中的靜態特徵分析與抽取方法；3.5 節介紹系統架構中的動態特徵分析與抽取方法；3.6 節則是關於分類決策樹的建立。

### 3.1 整體系統架構與流程

本論文在進行主流程前會先抽取一段背景影像，利用高斯混和模型(Gaussian mixture model, GMM)訓練背景影像模型。之後只要將輸入的影像和背景影像模型相減之後，即可獲得本系統預分析的前景影像。

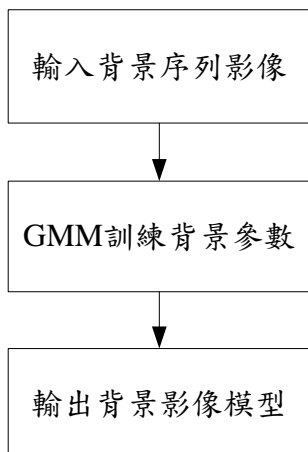
接著，由於人走路是透過左腳和右腳不斷往前行走的循環所構成的，所以其具有週期性，且人在一周期內的走路方向通常是直線的。因此，在獲得一序列的前景影像後，本系統利用頭與雙腳在走路期間角度的變化來估測週期，並只從一序列影像中擷取一周期影像出來進行以下的分析與特徵抽取。

得到一周期影像之後，將進行特徵抽取，所使用的特徵主要又分為兩種，靜態特徵與動態特徵。靜態特徵分別為：外輪廓特徵和手部特徵；而動態特徵分別為：頭部傾斜角、膝蓋特徵和步伐大小。

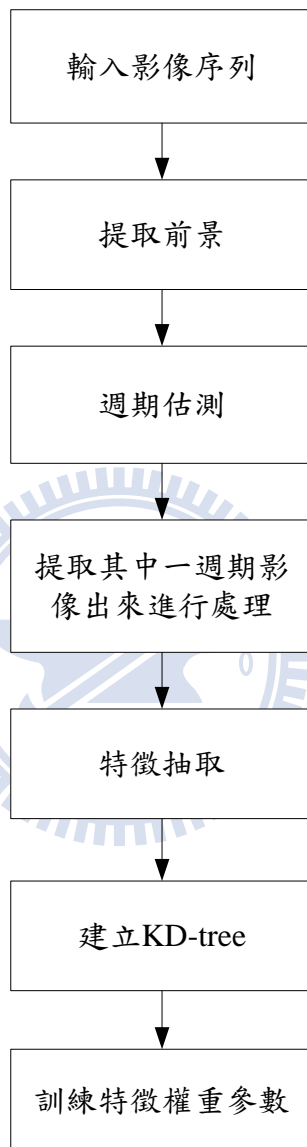
擷取特徵後，便進入身份的分類與辨識階段，本系統是使用 KD-tree 搜尋鄰近點來進行身份鑑別。使用 KD-tree 的目的在於增加系統的效率，系統首先會先建立一個決策樹： $k$  維樹(KD-tree)。之後再將特徵丟進去分類時，系統就會依照 KD-tree 的分

類順序來尋找最相似的人。以下為本系統的架構流程圖：

背景模型訓練流程圖



資料庫訓練流程圖



主系統執行流程圖

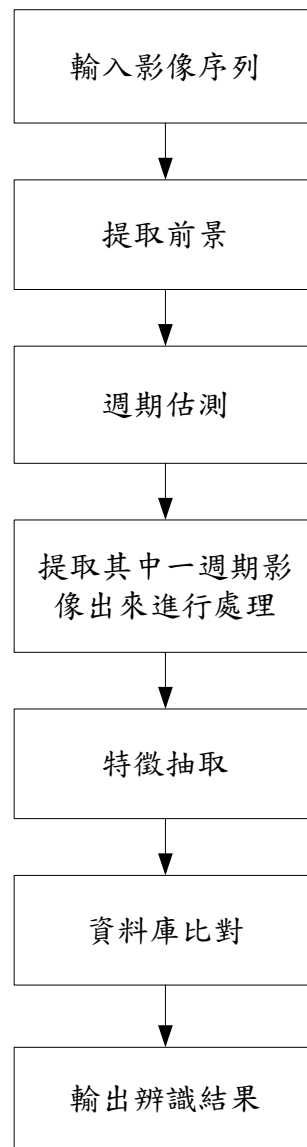


圖 3.1 本論文系統流程圖。

## 3.2 提取前景

在前景背景切割中，本論文是使用高斯混和模型來描述背景狀態。首先輸入一段背景序列影像，接著利用高斯混和模型來訓練背景參數，最後會產生一個背景模型，即為背景影像。詳細的做法流程如 2.1 節所描述的。圖 3.2 為利用高斯混和模型的方法擷取前景之範例。圖 3.2 (c) 為利用前景背景相減之後的二值化前景影像，之後我們對(c)進行一次形態學的膨脹，再將人的步態影像擷取出來如圖 3.2 (d)所示，而(d)即為本論文欲處理的序列影像範例。

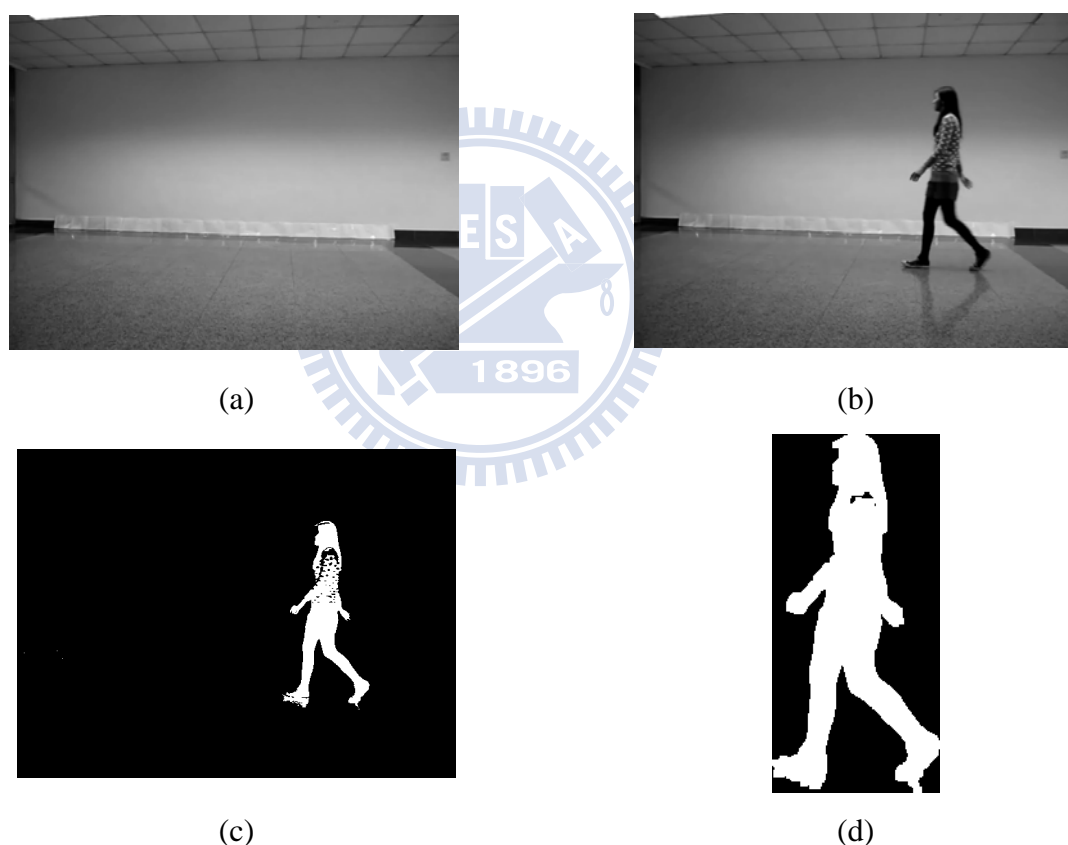


圖 3.2 提取前景範例。(a)灰階背景影像；(b)灰階輸入影像；(c)背景相減後前景影像。  
(d)人體步態影像，即本論文欲處理的步態影像。

本論文之系統即是透過以上的流程，將當前輸入的每一張影像序列與事先訓練好的背景影像模型相減，以得到本系統接下來所要分析的一序列前景影像，即人體步態影像。



### 3.3 週期估測

人的走路行為是由左腳右腳不斷循環的往前走所構成的，所以人的走路是有週期性的。為了降低整體資料量與提升程式執行效率，本論文在得到一序列的前景步態影像後，在這邊只會擷取一周期影像出來分析與抽取特徵。扣除人的異常行為之外，只提取一週期序列影像出來分析還有一個優點，即人在一周期內的走路方向通常是直線前進的，所以在分析過程中不需考慮人在行走中是否稍微偏移了幾度角。

在步態分析中，大家最常使用來估測週期的方法為：利用二值步態影像中前景像素值的總和變化來進行週期估測。但是這方法的風險在於若前景背景在切割時沒有切的很乾淨，像素值的總和變化會忽大忽小，但這忽大忽小的變化是來自雜訊而不是真的步態周期變化，此時就無法估測出真正的週期變化，並會影響到後續特徵抽取與身份鑑別的準確度。

有鑑於上面提到的缺點，本論文在這邊參考文獻[16]的方法，將其抽取特徵的方式加以改良，並將其應在週期估測上，因此，接下來會先就文獻[16]中本論文所參考的部分加以做介紹。在 B. DeCann 和 A. Ross[16]的方法中，抽取靜態特徵的方式首先是先找出三個點，如圖 3.3 所示：



圖 3.3 B. DeCann 和 A. Ross[16]方法中，抽取靜態特徵之方法示意圖。

其中  $\mathbf{v}_{\min} = \{i_{\min}, j_{\min}\}$  ;  $\mathbf{v}_{\max} = \{i_{\max}, j_{\max}\}$  ;  $\mathbf{v}_{\text{ter}} = \{i_{\text{ter}}, j_{\text{ter}}\}$  ,  $i$  為影像中的縱座標位置,  $j$  為影像中的橫座標位置。所以  $\mathbf{v}_{\min}$  所代表的意思為：整張二值影像中像素值為 1 之最左上方的點，約為人頭頂的點； $\mathbf{v}_{\max}$  所代表的意思為：整張二值影像中像素值為 1 之最右下方的點，約為人在走路過程中，前腳腳跟的位置； $\mathbf{v}_{\text{ter}}$  所代表的意思為：整張二值影像中在  $\mathbf{v}_{\min}$  正下方像素值為 1 之最下方的點。找出每一張影像中這三點的位置後，藉由這三點的變化定義出步態曲線，即為特徵。

在本論文中，主要就是利用以上的方法加以做改良來進行週期估測的動作，不同於 B. DeCann 和 A. Ross[16]的方法，本論文在  $\mathbf{v}_{\min}$ 、 $\mathbf{v}_{\max}$  和  $\mathbf{v}_{\text{ter}}$  三點的取法上做了些微的修改，在  $\mathbf{v}_{\min}$  的部分，本論文更改成取頭頂正上方的點，並令為 A 點，即二值影像中最上面不為零那一行的中心點，如圖 3.5 所示。在文獻[16]的方法中， $\mathbf{v}_{\min}$  和  $\mathbf{v}_{\text{ter}}$  是分別屬於二值影像中同一行的最上方與最下方的點，所以  $\mathbf{v}_{\max}$  和  $\mathbf{v}_{\text{ter}}$  的相對位置有時候會因為雜訊的干擾而產生前後互換的現象，如圖 3.4 所示：

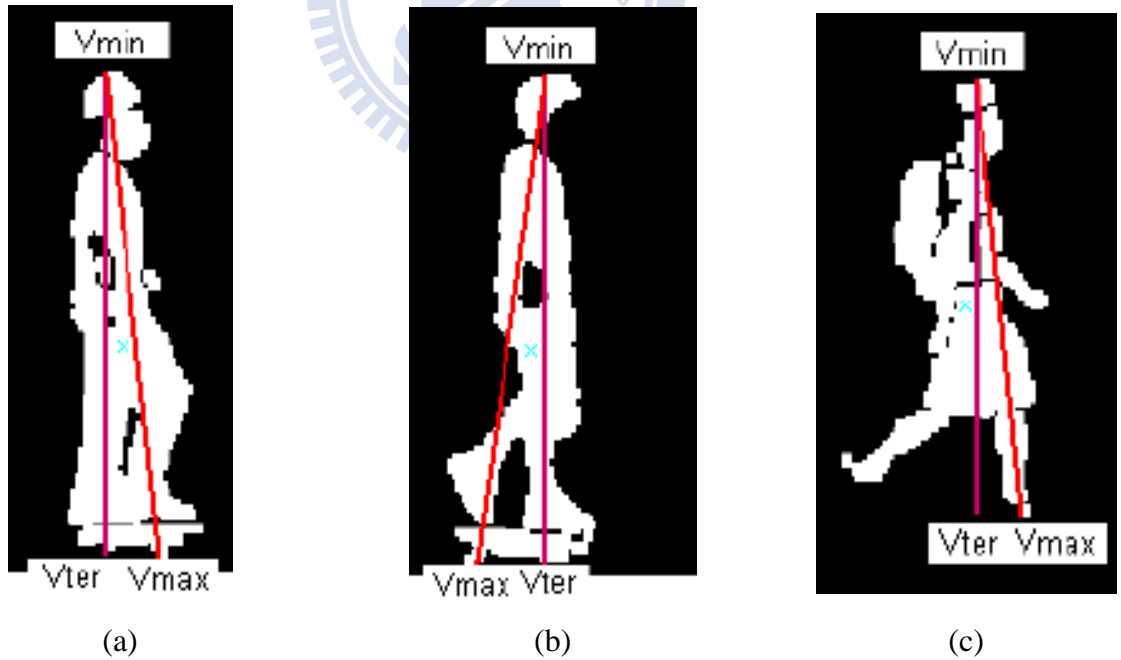


圖 3.4 文獻[16]中特徵抽取範例。(a)起步時  $\mathbf{v}_{\min}$ 、 $\mathbf{v}_{\max}$  和  $\mathbf{v}_{\text{ter}}$  三點的相對位置範例；(b)半週期時  $\mathbf{v}_{\min}$ 、 $\mathbf{v}_{\max}$  和  $\mathbf{v}_{\text{ter}}$  三點的相對位置範例；(c) 背背包時  $\mathbf{v}_{\min}$ 、 $\mathbf{v}_{\max}$  和  $\mathbf{v}_{\text{ter}}$  三點的相對位置範例。

從圖 3.4 中可得知  $v_{\max}$  和  $v_{\text{ter}}$  的位置有可能會一直呈現交換的現象，這就在估測週期這部份來說不是一個好的方法。因此，針對  $v_{\max}$  和  $v_{\text{ter}}$  的部分，本論文改良了文獻[16]的方法，將人體平分為左右兩部份，分別針對左半邊去找出最左下方的點，並令為 **B** 點，針對右半邊去找出最右下方的點，並令為 **C** 點，如圖 3.5 所示：

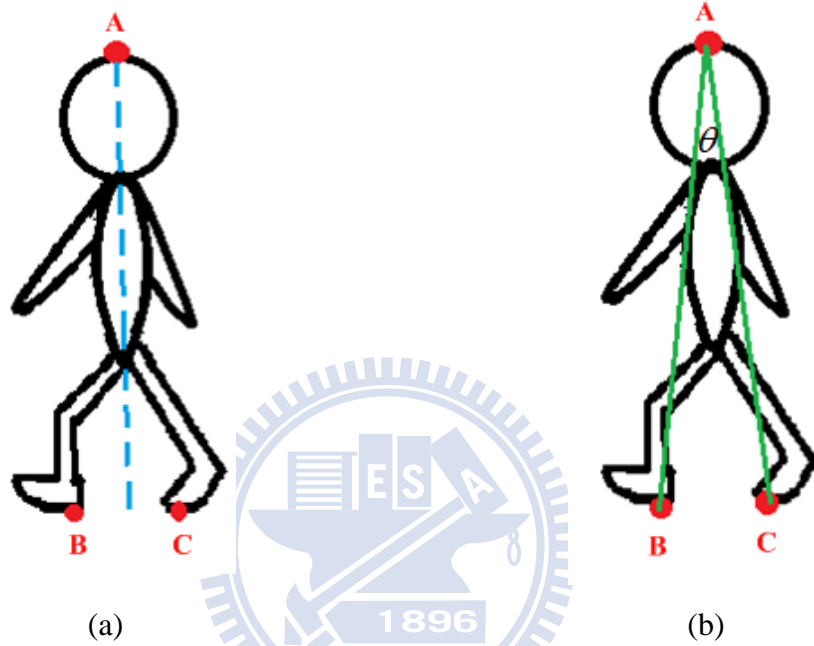


圖 3.5 本論文週期估測方法示意圖。(a)本論文尋找 A、B 和 C 三點之示意圖；(b)本論文用來估測週期的角度示意圖。

找出一序列影像中每張步態影像的 **A**、**B** 和 **C** 三點，如圖 3.5(a)所示，接著算出  $\angle CAB$  的角度大小，如圖 3.5(b)中的  $\theta$  角。其中  $\theta$  角就是人在走路時，兩腳不斷開合所形成的夾角，而本論文就是利用這一系列的  $\theta$  角變化來進行周期的估測。

圖 3.6 為本論文估測週期的範例，經由觀察圖中曲線的變化和其相對應的步態影像，可以得知本論文所使的週期估測方法具有不錯的估測效果。估測出一序列影像的周期變化後，本論文在這邊只擷取出其中一段週期的序列影像出來進行後續的分析。如圖 3.6 範例所示，本論文在這邊只取出第一週期的影像，即序列影像中的第 11 幀到第 42 幀之間的影像出來進行後續的處理，如圖 3.7 所示。

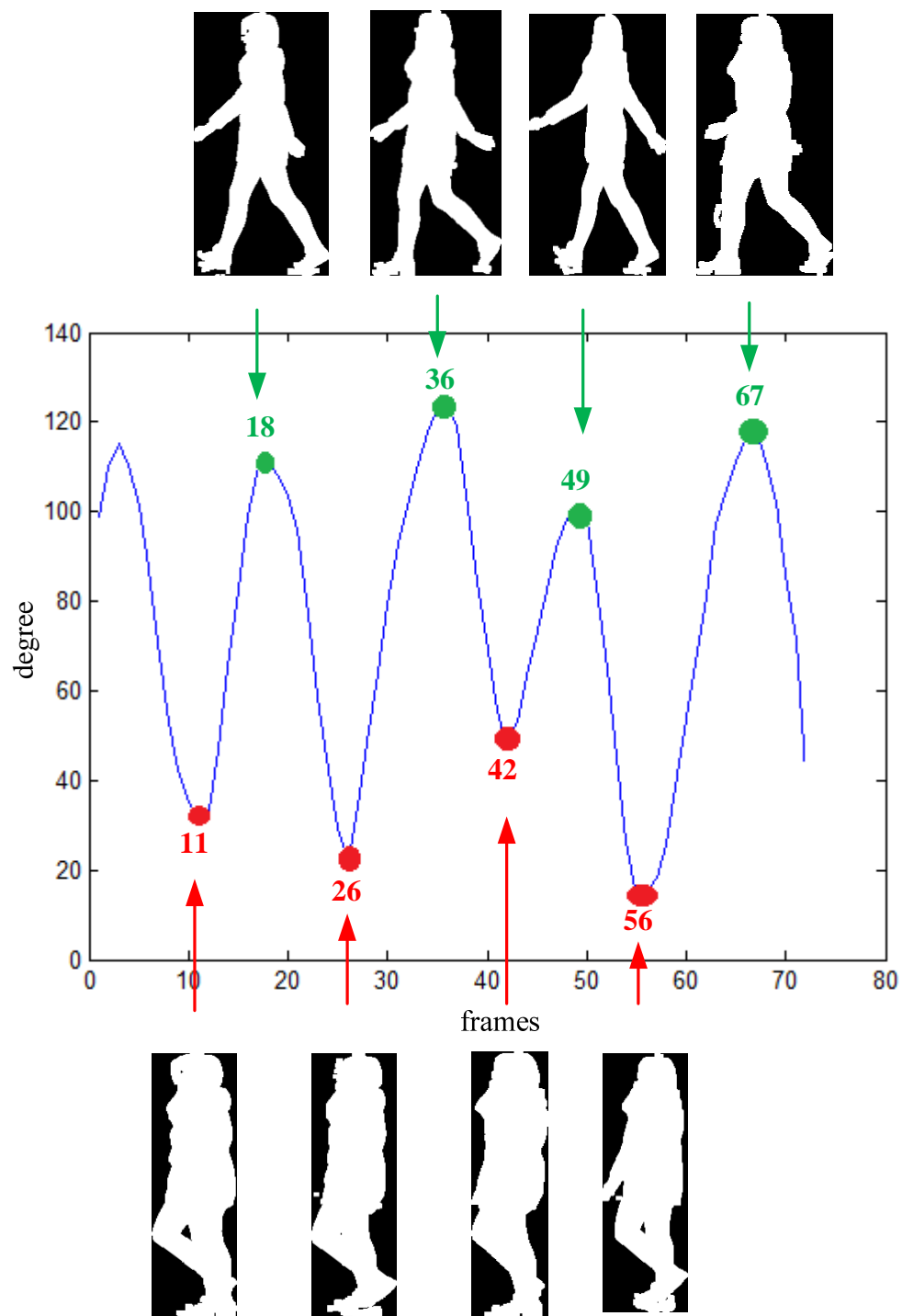


圖 3.6 週期估測範例。



圖 3.7 一週期內步態影像範例。

### 3.4 靜態特徵

本論文在抽取特徵的部分主要又分成了兩大部分，靜態特徵與動態特徵。在本節中，將針對靜態特徵的部分來加以說明與介紹。在靜態特徵的部分，本論文主要的處理對象為步態能量影像(gait energy image, GEI)，並針對人體姿態中的外輪廓與手部擺動幅度來進行分析。因此，在本節中又分了兩小節來進行介紹，3.4.1 介紹本系統的外輪廓特徵；3.4.2 介紹本系統的手部擺動幅度特徵。

### 3.4.1 外輪廓特徵

首先將一週期內的所有步態影像疊加在一起，目的在於得到步態能量影像。接著，取出步態能量影像的外輪廓，利用傅立葉描述子(Fourier descriptor, FD)來進行描述，傅立葉描述子詳細作法流程如 2.2.2 所述，在這邊本論文是統一將影像正規化成 140 像素的高度。在利用傅立葉描述子描述外輪廓的部分，為了盡量降低特徵維度大小並保留特徵的精確性，我們參考人體黃金比例[20]大致為 1:1.618 的說法，推算人大致為七頭身的高度，因此將人立體的前後各取 7 點加上頭頂與腳底共 16 點。本論文之系統在這邊是選擇維度為 16 的傅立葉描述子來進行外輪廓描述。如圖 3.8 所示：

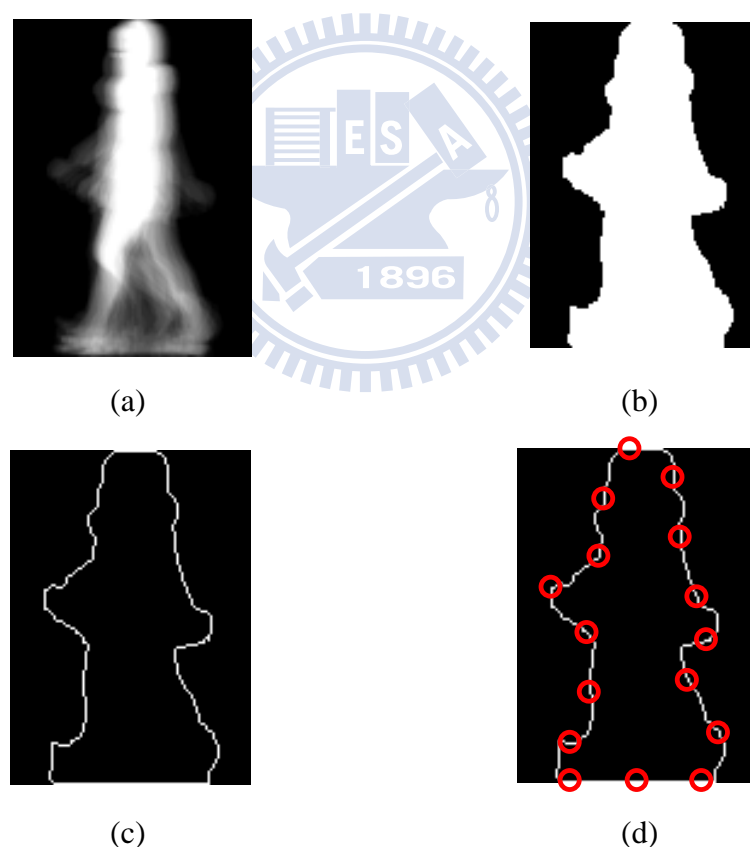


圖 3.8 本論文靜態特徵中的外輪廓特徵示意圖。(a)步態能量影像示意圖；(b)步態能量影像二值化後結果；(c)步態能量影像外輪廓特徵；(d)利用 16 維的傅立葉描述子來描述外輪廓特徵之示意圖。

因此，在外輪廓特徵的部分其特徵維度大小為  $1 \times 16$  的大小。

### 3.4.2 手部擺動幅度特徵

取得步態能量影像之後，本論文在這邊主要是利用雙手擺動幅度大小距離、手抬起的縱座標高度位置和手與身體之間的距離來做為特徵，如圖 3.9 所示：

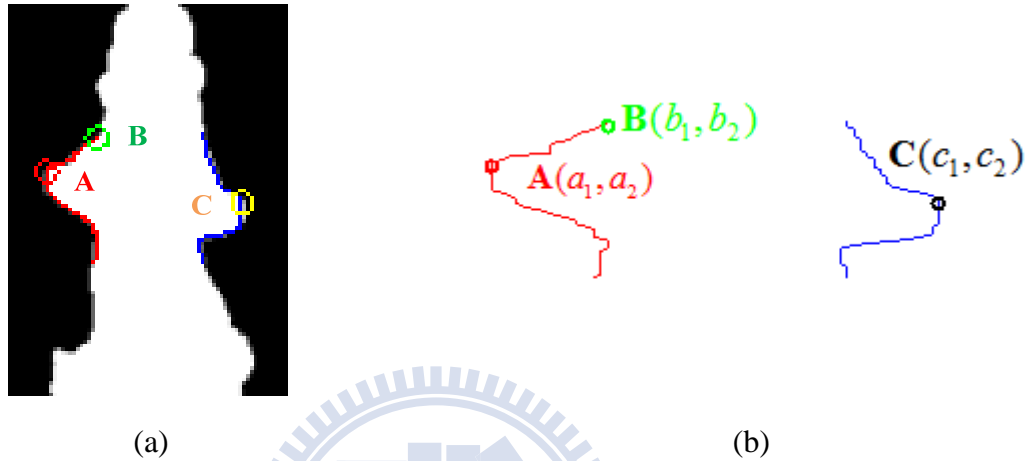


圖 3.9 本論文手部擺動幅度特徵示意圖。(a)步態能量影像手部特徵示意圖；(b)手部特徵曲線圖。

圖 3.9 中的紅色曲線為步態能量影像中手部身體區域的前曲線；藍色曲線為步態能量影像中手部身體區域的後曲線；A 點為前曲線中最左邊的點，即為人在往左前進的一周期影像中手部抬起最高幅度的位置；B 點為前曲線中最右邊的點，即在一周期影像中，幾乎沒有任何擺動位置，即為身體區域；C 點代表的則是在後曲線上最右邊的點，意即手部擺動到身體後方時的位置。

關於手與身體之間的距離、雙手擺動幅度大小距離和手抬起的縱座標高度位置的計算方法如方程式(3-1)到(3-3)所述：

$$hand\_body = |b_2 - a_2| \quad (3-1)$$

$$hands\_distance = |c_2 - a_2| \quad (3-2)$$

$$hand\_high = a_1 \quad (3-3)$$

因此，在手部特徵的部分其特徵為度大小為 1\*3 的大小。

## 3.5 動態特徵

動態特徵與靜態特徵最大的差別在於，靜態特徵是對單一張的步態能量影像來做分析，動態特徵則是針對一周期中每一張影像進行分析，例如：現在有行人 A 的步態序列影像，其週期長度為 32 張影像，靜態特徵就是將這 32 張影像全部疊加在一起，產生步態能量影像後再進行特徵抽取；動態特徵則是對這 32 張中的每一張都做一次特徵抽取，因此動態特徵其實就是紀錄一序列的數值變化。

在本節中，將針對動態特徵的部分來加以說明與介紹。本論文在動態特徵的部分又分成四部分，分別為頭部、膝蓋、腳掌與步伐大小。因此，本節又分了四小節來做介紹，3.5.1 介紹頭部與垂直線的夾角特徵抽取法則；3.5.2 介紹膝蓋彎曲角度的特徵抽取法則；3.5.3 介紹腳掌離地角度的特徵抽取法則；3.5.4 則是說明步伐大小距離的估測方法。

### 3.5.1 頭部特徵

有人在走路的時候習慣低頭看前方地面，有人走路時喜歡抬頭，有人喜歡直視前方前進。走路的樣子有各式各樣不同的變化，但是每個人都會不自覺的傾向於使用自己最舒服、最習慣的姿勢來行走。因此，本論文在這邊也將頭部與垂直線的夾角當作特徵之一，希望藉由多重特徵的搭配組合來提升整體身份的鑑別率。

在本論文中，主要是參考人體黃金比例[20]為 1:1.618，推算人的身材大致為七頭身的說法來將頭部區域單獨抽取出來，意即將人體整張步態影像中的最上面七分之一部份擷取出來，並定義此區域為頭部區域。得到頭部區域影像後，考量所有可能影像頭部影像的變因，歸納其影響最劇的首推「戴帽子」，其中又以鴨舌帽對影像的影響最大。因此，在進行特徵抽取之前，必須先對影像做一些預處理，以降低變因的影響。

在預處理的部分，本論文將已經擷取出來的頭部區域影像再重新做一次修正。將頭部再上下等分為兩份，以此分割線，即鼻子位置附近區域為基準，將頭部影像的左右區域重新裁切，如圖 3.10 所示：





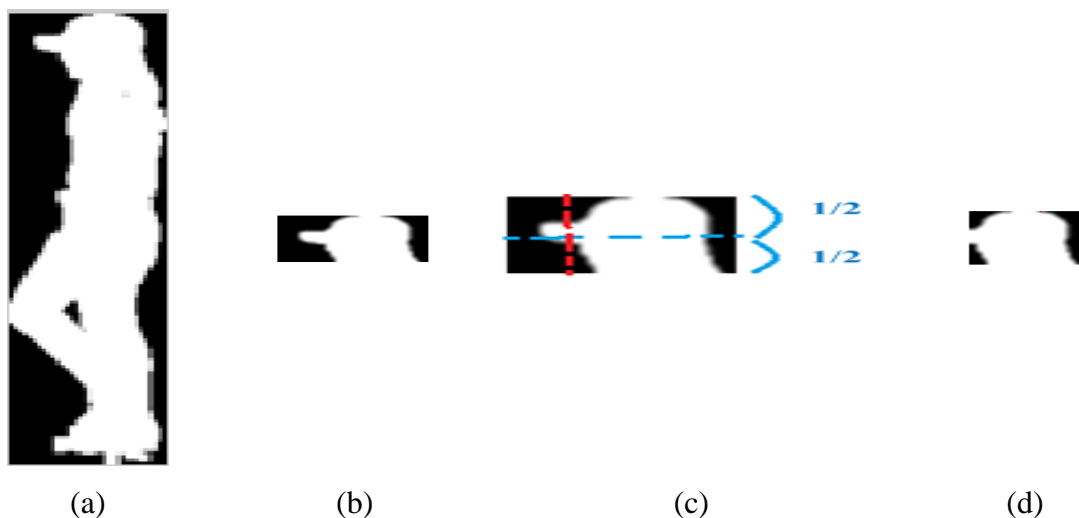


圖 3.10 擷取頭部區域流程示意圖。(a)人體步態影像；(b)擷取人體步態影像最上方七分之一的區域影像圖；(c)裁剪頭部區域示意圖；(d)最終擷取出來的頭部區域影像。

在取得頭部區域的影像後，本論文接下來採用主成分分析法(principal components analysis, PCA)對頭部影像進行分析。主成分分析法的詳細做法原理如 2.3.1 節所述，而本論文在這邊主要是拿來估測頭部二值影像中資料點的主要分布方向，即頭部傾斜的角度與方向。

圖 3.11 中的(b)和(d)是同一個人的頭部特徵抽取結果，差別只在於一個沒戴帽子，一個有戴帽子。圖中綠色的線表示為垂直地面的方向，紅色的線表示頭部傾斜的方向。就圖 3.11 中的範例來看，紅色線在綠色線前方，表示此人為低頭的狀態。

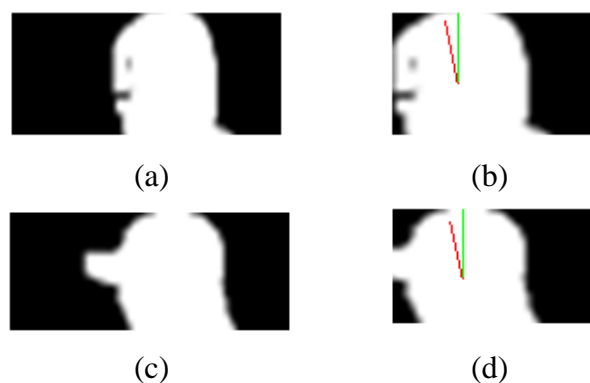


圖 3.11 頭部特徵抽取結果示意圖。(a)沒戴帽子的原始影像；(b)沒戴帽子的頭部特徵抽取結果；(c)戴帽子的原始影像；(d) 戴帽子後的頭部特徵抽取結果。

在頭部特徵的部分，若是往下看的情況時，本論文將角度設為正角度值，若頭部是上仰的情況時，則將角度值設為負角度值，此步驟主要是為了區分屬於往下或往上的情況。考量到人在行走時，頭部可能會忽上忽下的看，因此本論文在這邊是選擇採用人在行走時頭部與垂直線的平均夾角，意即將一週期影像的全部頭部傾斜角取平均做為特徵之一。圖 3.12 為同一人之不同步態序列的一週期頭部傾斜角度變化曲線，由圖中可得知，雖然頭部傾斜角曲線在一週期內的分布不太一樣，但是平均角度卻幾乎分布在同一範圍之中。如圖 3.12 的範例，此人的頭部傾斜角度大致會落在 14 度到 17 度之間。

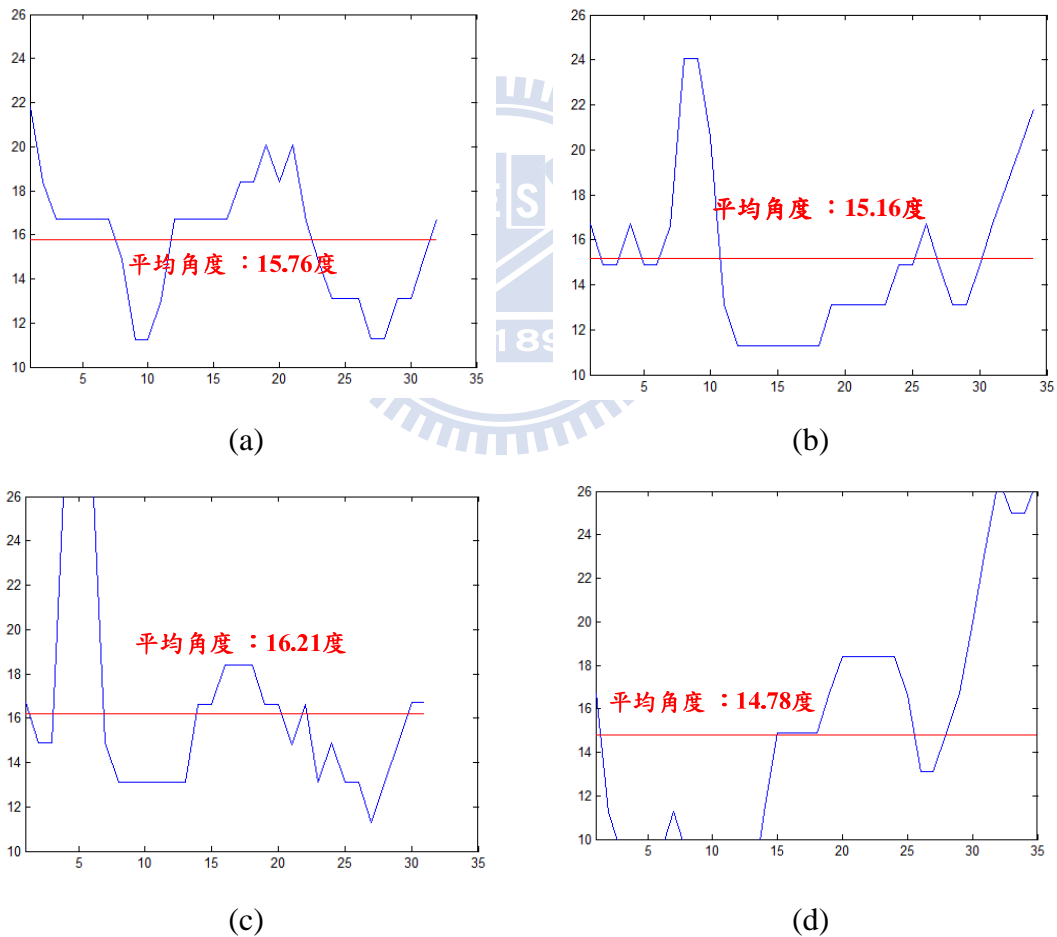


圖 3.12 頭部傾斜角度變化曲線範例。(a)第一段步態影像頭部傾斜角抽取結果；(b) 第二段步態影像頭部傾斜角抽取結果；(c) 第三段步態影像頭部傾斜角抽取結果；(d) 第四段步態影像頭部傾斜角抽取結果。

另外，考量到每個人的頭部長寬不太一樣，所以本論文亦將頭部影像的平均質心座標位置當做特徵之一。因此，頭部特徵包含頭部平均傾斜角和頭部質心座標位置，其維度大小為  $1 \times 3$ 。

### 3.5.2 腳掌特徵

平常在路上，常常不小心就會聽到有人在走路時鞋子拖地所發出的聲音，進而觀察每個人在走路時，有些人習慣性走路不將腳掌抬起，有些人則是一步一步抬起往前走。因此，本論文從此方向下手，偵測每個人在走路時腳掌離地的角度。

本論文主要是先偵測人體步態影像中下半身區域最後一行不為零的位置，以此為基準來擷取腳掌區域影像，如圖 3.13 所示：

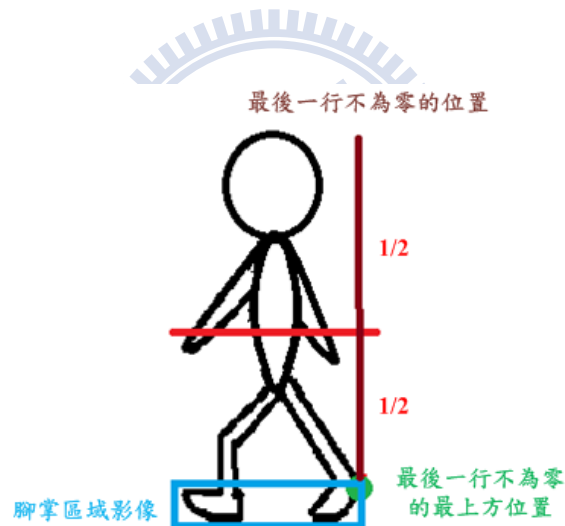


圖 3.13 擷取腳掌區域示意圖。

擷取出腳掌區域影像後，接著會利用分群的方法，進一步判斷此區域內為一個物體或兩個物體。若判斷為一個物體，表示此時雙腳呈現合併的狀態；若判為兩個物體，則表示此時雙腳呈現打開的狀態。在這邊我們只處理雙腳張開的情況，因為當雙腳合併時，除了雙腳原地站立外，就是前腳完全著地，後腳完全離地的情況。而我們主要目的是為了抽取腳掌離地的角度，所以在這邊我們只討論兩個物體，即雙腳張開的情況。

一旦偵測到腳掌區域為兩個物體時，在這邊我們只分析後腳掌離地的角度，所以仿照頭部傾斜角的取法，利用主成份分析法(principal components analysis, PCA)來對後腳掌區域影像進行角度估測。圖 3.14 為實際的腳掌特徵抽取範例，其中(d)中的綠線為與地面平行的水平線，紅線即腳掌的抬起角度，因此，在這邊所說的腳掌與地面之間的角度就是指紅線與綠線之間的夾角。

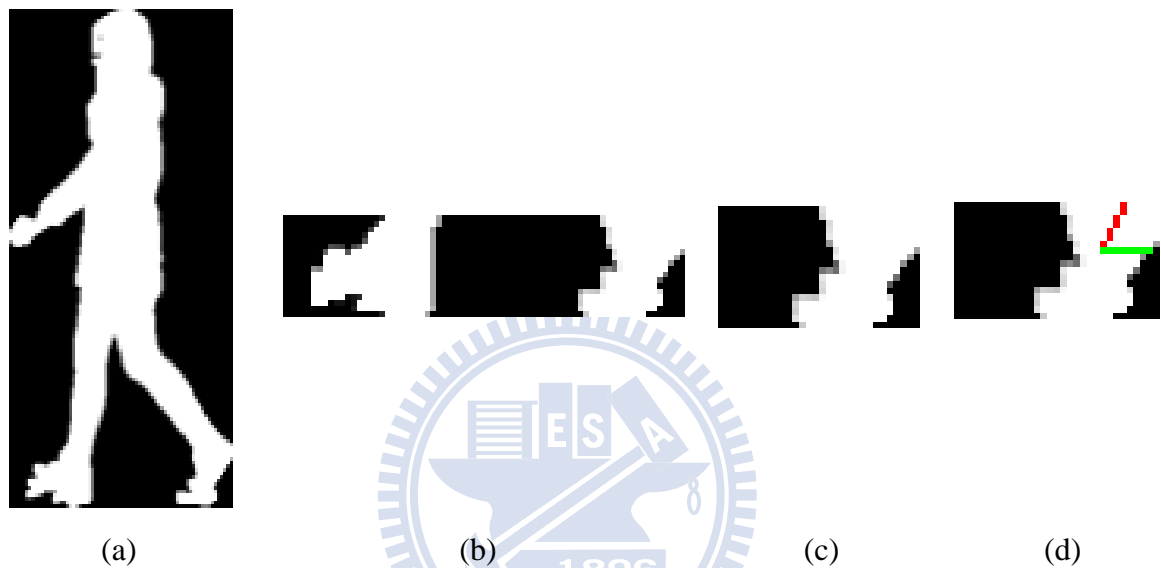


圖 3.14 擷取腳掌特徵範例示意圖。(a)原始輸入步態影像；(b)腳掌區域影像；(c)後腳掌區域影像；(d)後腳離地角度擷取結果示意圖。

由於每個人走路時腳掌抬起的高度都不同，在這邊本論文將一序列腳掌離地角度數據中最大的角度抽取出來當作特徵，如圖 3.15 所示。由於一周期步態中包含左腳右腳各一次的往前走，因此，在圖 3.15 中的腳掌離地角度變化會有兩次的峰值，一次為左腳的離地角度，另一次則為右腳的離地角度。在圖 3.15 範例中，此人的腳掌離地角度為  $84.29^{\circ}$ 。

另外，考量到每個人走路速度都不一樣，因此在這邊我們將速度也加入特徵中。主要是透過計算每個人在一周期影像中腳掌離地的影像張數多寡，即角度不為零的影像張數總和來代表這個人走路時腳移動的速度。在圖 3.15 的範例，腳移動的速度為 16 幀。

因此，腳掌特徵包含腳掌離地之最大角度與雙腳移動速度，其維度大小為  $1 \times 2$ 。

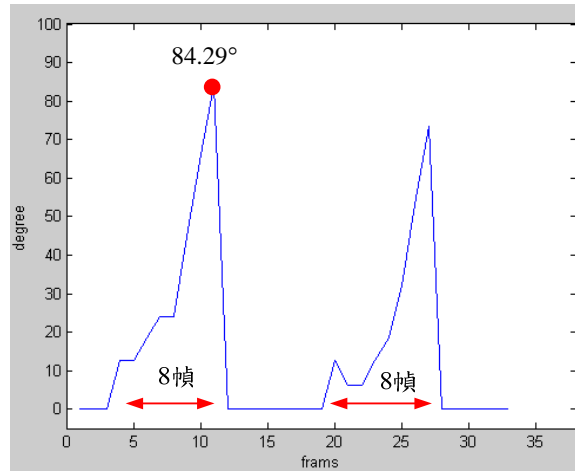


圖 3.15 腳掌離地角度一周期變化範例。

### 3.5.3 膝蓋特徵

有些人在走路的時候會習慣性將前腳蹲一下或後腳往上勾一下，膝蓋的彎曲角度也相對地比較大。由此可知，每個人由於走路習慣上的不同，在膝蓋部分的表現也不太一樣。因此，本論文將膝蓋特徵分成前膝蓋與後膝蓋來進行特徵抽取。

參考人體黃金比例的標準[20]，上下半身的標準比例為  $1:1.618$ ，即下半身佔全身的比例為  $0.618$ ，約為  $3/5$ 。因此，在本論文中在擷取腿部區域影像的標準即是將人體步態影像中的  $3/5$  到腳掌區域影像最上方擷取出來，如圖 3.16 所示：



圖 3.16 擷取腿部區域範例。

觀察人在走路時膝蓋的循環變化，發現前膝蓋的最大彎曲程度會發生在雙腳合併剛要踏出去往前走的時候；後膝蓋的最大彎曲程度則是發生在雙腳打開的時候。因此，為了提高本論文的效率，避免不必要的處理過程，在抽取腳掌特徵的時候，當判斷雙腳為合併時才會偵測前膝蓋特徵，偵測到雙腳打開的時候才會偵測後膝蓋特徵。擷取膝蓋特徵的流程如圖 3.17 所示：

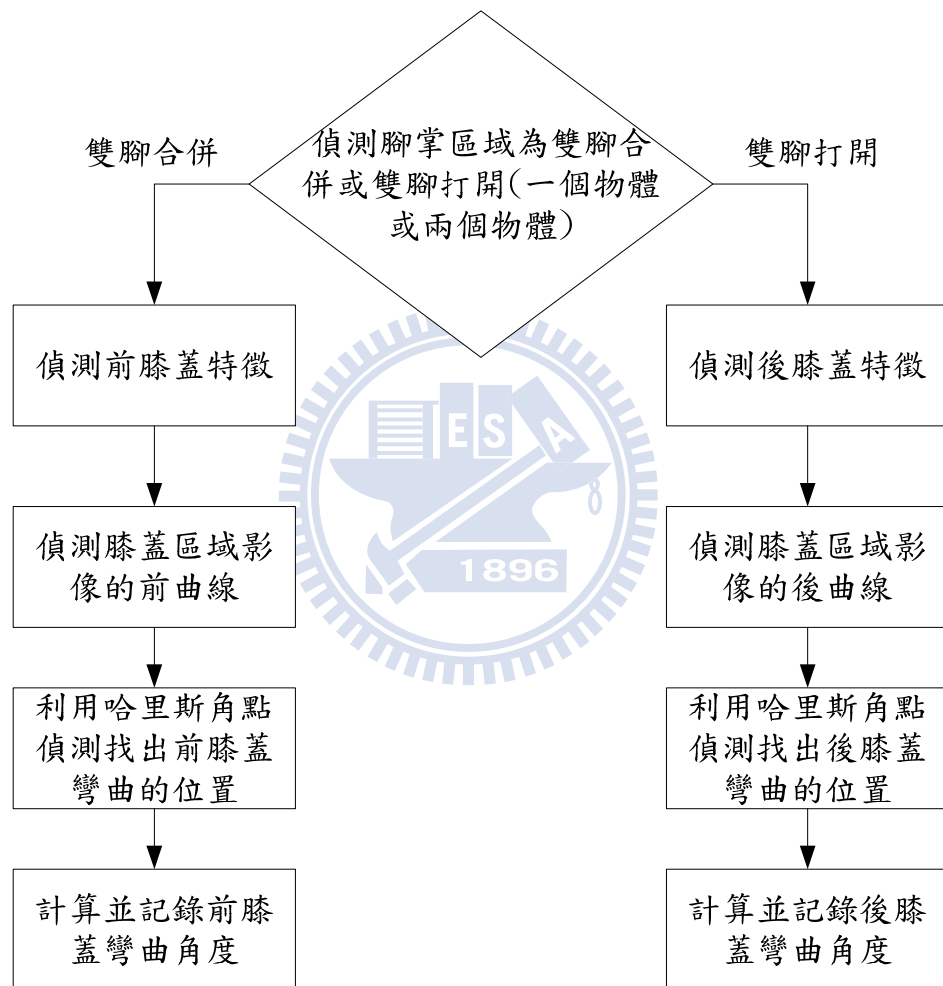


圖 3.17 擷取膝蓋特徵之流程圖。

依據圖 3.17 的流程所述，擷取出腿部區域後，接著我們將偵測此區域的前曲線與後曲線，最後再利用哈里斯角點偵測(Harris corner detection) 分別找出前後曲線的轉角處，即膝蓋彎曲的地方，詳細的哈里斯角點偵測方法流程如 2.3.2 所述。

在抽取前膝蓋角度特徵的範例如圖 3.18 所示；抽取後膝蓋角度特徵的範例如圖 3.19 所示：

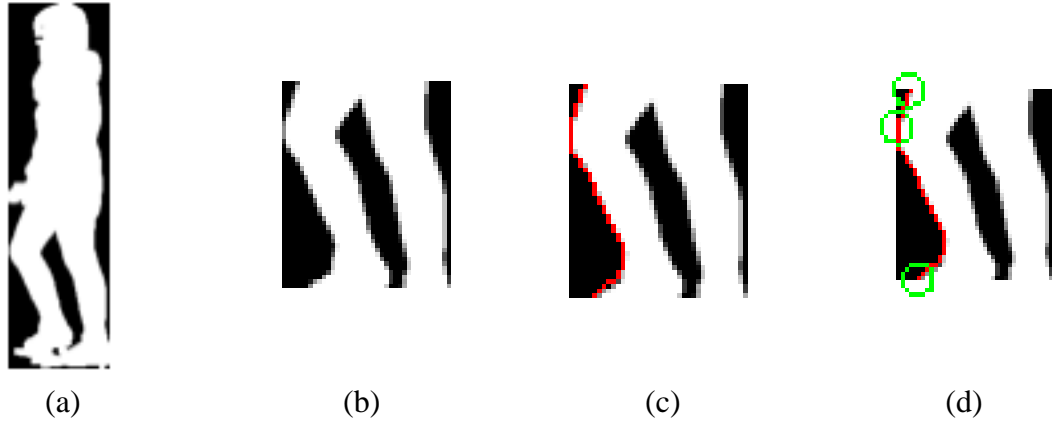


圖 3.18 前膝蓋特徵抽取範例。(a)原始側影影像；(b)腿部區域影像；(c)腿部區域前曲線示意圖；(d)前膝蓋特徵點偵測結果。

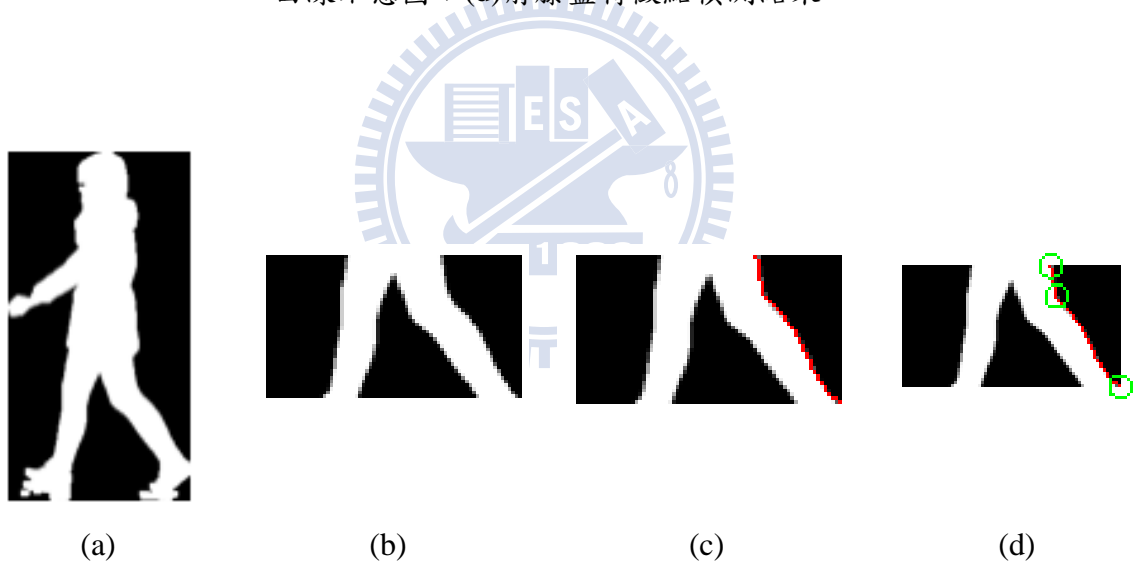


圖 3.19 後膝蓋特徵抽取範例。(a)原始側影影像；(b)腿部區域影像；(c)腿部區域後曲線示意圖；(d)後膝蓋特徵點偵測結果。

經過圖 3.18 和圖 3.19 的流程可得膝蓋的三個特徵點，最上面和最下面的點為曲線上的第一個點與最後一個點，膝蓋區域的彎曲點則是由哈里斯角點偵測所偵測出來的點。藉由這三點的位置，即可算出膝蓋的彎曲角度，如圖 3.20 所示：

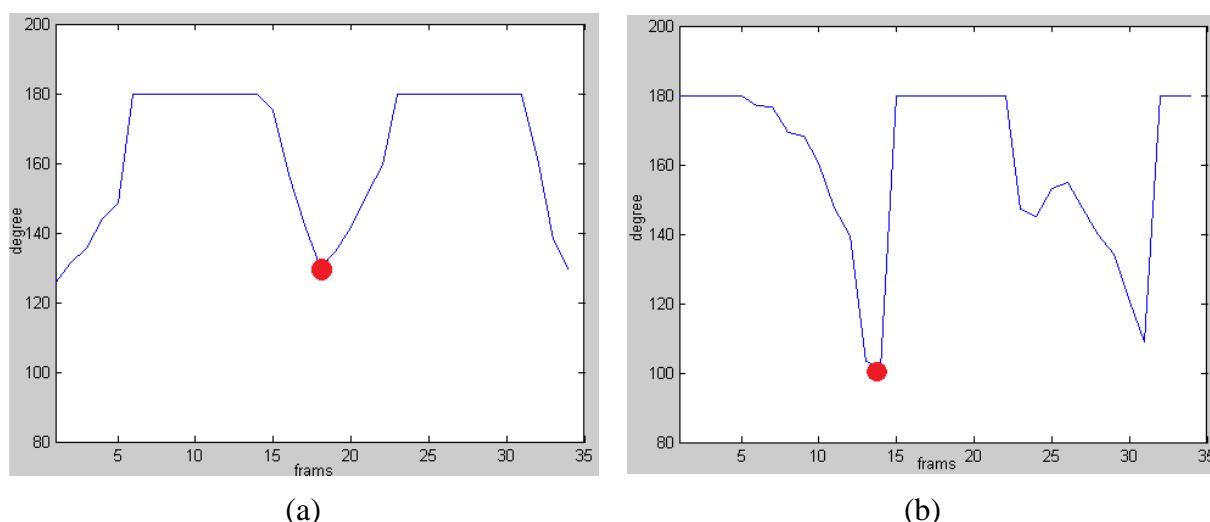


圖 3.20 前後腿部膝蓋角度變化。(a)前膝蓋角度變化；(b)後膝蓋角度變化。

圖 3.20 中 180 度表示腳是直立的狀態，三點呈現一直線，膝蓋沒有彎曲的情況。而本論文在針對膝蓋特徵的部分是選擇抽取前膝蓋和後膝蓋的序列角度數據中最小的角度值來當作特徵，即膝蓋的最大彎曲程度。因此在膝蓋的特徵維度大小為  $1 \times 2$ 。

### 3.5.4 步伐大小特徵

每個人在走路時的步伐大小都不同，最明顯的例子就是男生和女生，由於男生步伐比較大，所以往往女生必須以較快的走路速度才能跟上男生。由此可知，其實每個人的步伐大小都不太一樣，這也就表示步伐大小是具有個人的特色的特徵之一。

在本論文中，步伐大小距離是在估測周期的時候同時就一起記錄下來了。由於在估測周期的時候，必須先找到三個點 **A**、**B** 和 **C**，藉由其  $\angle CAB$  的角度變化來估測週期，而步伐距離大小主要是使用歐幾里得距離法來計算 **B** 點和 **C** 點之間的距離，如圖 3.21 所示：



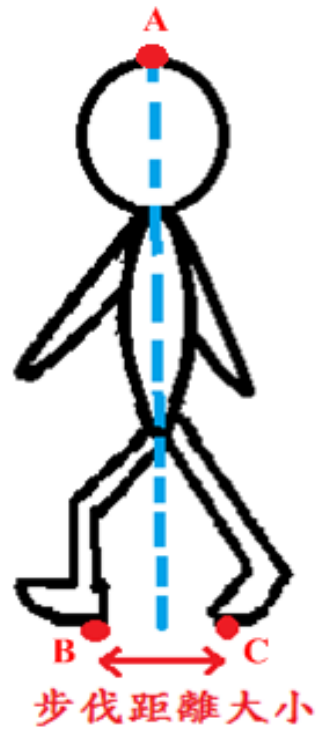


圖 3.21 偵測步伐距離大小示意圖。

其中歐幾里得距離法的計算公式如下：

$$dis(B, C) = \|B - C\| = \sqrt{\sum_{i=1}^n (b_i - c_i)^2} \quad (3-4)$$

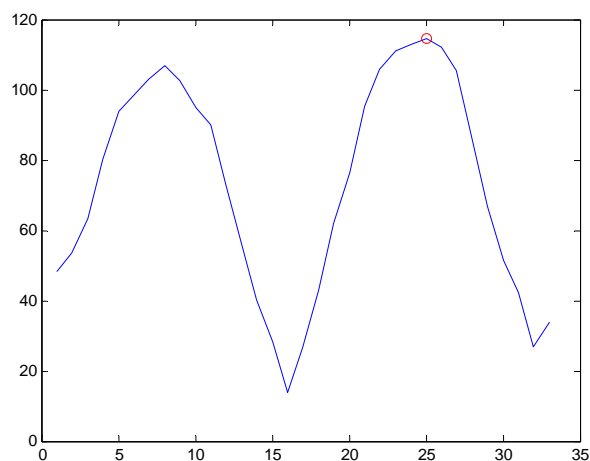
$n$  為資料的維度大小，本論文在這邊的  $n=2$ 。

得到一周期的序列數據後，我們從中找出最大的距離來做為步伐大小的特徵。最大的距離意即此人在這一週期內，雙腳步伐跨最大的距離，如圖 3.22 所示。此外，本論文在這邊一起把每個人走路一周期所花費的時間也放進去，即一週期內影像的總數。

因此，步伐大小特徵包含左右腳平均步伐距離大小和週期時間，其維度為  $1*2$ 。



(a)



(b)

圖 3.22 步伐大小特徵範例。(a)一週期內步伐最大之步態影像範例；(b)一週期步伐距離大小變化圖。

### 3.6 分類法則

在所有特徵中有些是具有明顯的鑑別力，有些則是屬於輔助辨識的角色，所以在進行分類之前，本論文會先將資料庫中的人進行特徵權重大小的訓練，以區分每一個特徵的重要性。而本系統所使用的全部特徵如圖 3.23 和圖 3.24 所示。包含利用 16 維度的傅立葉描述子來描述的外輪廓特徵和手部擺動幅度距離大小為靜態特徵，如圖 3.23 所示；以及頭部平均傾斜角度、膝蓋最小彎曲角度，即膝蓋的最大彎曲程度、腳掌最大離地角度和步伐最大距離為動態特徵，如圖 3.24 所示。全部的特徵與維度大小整理如表 3-1 所示，可知本論文所使用的全部特徵總維度為  $1 \times 28$  的大小。

在訓練每個特徵權重大小的部分主要是先使用 KD-tree 對資料庫內的所有資料建立成一棵樹，以提高後續找最近鄰居時的效率。在資料庫訓練的部分，共包含二十個人，每人四筆資料，共八十筆資料來進行權重大小的訓練，在這邊所說的權重大小其實就是每個特徵其單一的辨識率。訓練的方法就是跟除了當下資料以外的資料庫內所有資料做比較，找出最相似的鄰居資料點，將此判斷的結果與預訂的答案做比較，並從中紀錄每個特徵的單一辨識率為多少。將資料庫內的八十筆資料全部跑過後，即可得到每個特徵單一的辨識率，即每個特徵的權重大小值。

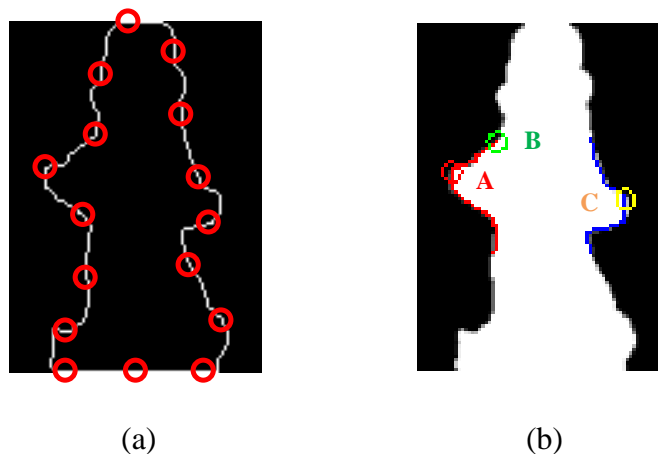


圖 3.23 本論文所使用的靜態特徵範例。(a)外輪廓特徵；(b)手部擺動特徵。

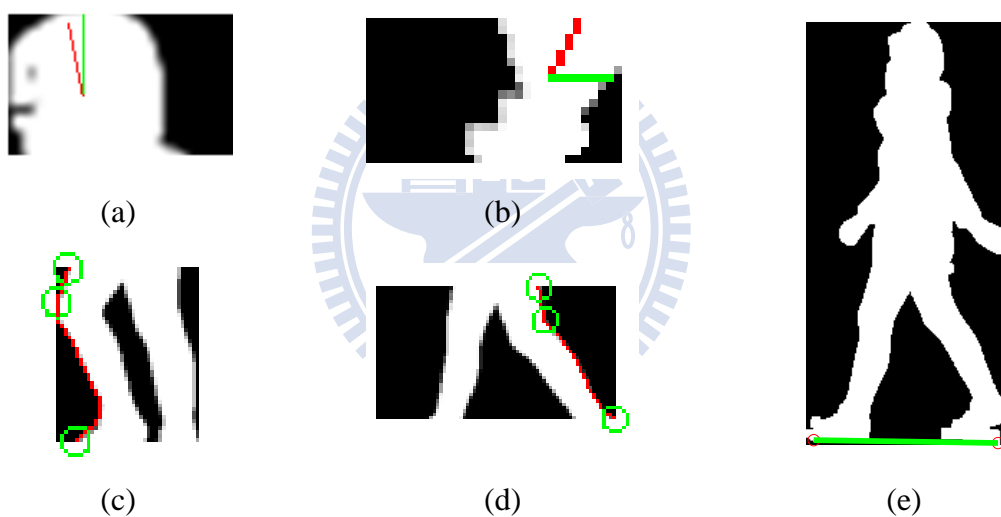


圖 3.24 本論文所使用的動態特徵範例。(a)頭部特徵；(b)腳掌特徵；(c)前膝蓋特徵；  
(d)後膝蓋特徵；(e)步伐特徵。

表 3-1 本論文所使用的全部特徵、維度、權重大小整理表

特徵	外輪廓	手部擺動幅度	頭部	腳掌	膝蓋	步伐大小	總計
維度	1*16	1*3	1*3	1*2	1*2	1*2	1*28
權重	0.85	0.74	0.46	0.59	0.57	0.54	12.99

因此，當現在輸入一筆新的受測者資料時，藉由本系統的流程獲得所有特徵之後，便進入了身份鑑別的階段。而在本論文中又分成兩個系統來進行討論，分別是單純拿來做身份辨識的部分和進一步應用在安全監控上之系統。因此在本節中又分成兩小節來進行介紹，3.6.1 節介紹如何進行身份鑑別；3.6.2 節介紹如何將系統應用在安全監控上。

### 3.6.1 身份鑑別法則

在系統進行身份鑑別之前，必須先利用資料庫訓練每種特徵的權重大小，即區分每種特徵的重要性。接著就是將新輸入的資料利用重要性最大的三個特徵，即權重數值最大的三種特徵先進行第一階段的候選人篩選；最後再將所有特徵針對這些候選人進行第二階段的權重給分，獲得總分最高的候選人，即表示與輸入的人最為相似，亦即為此人的身份。在第一階段和第二階段中，不論是篩選候選人還是進行權重給分，都是選擇最鄰近的三個鄰居點來進行篩選與給分，意即使用 KD-tree 方法來搜尋鄰近點，其中令  $k=3$ 。在這邊選擇  $k=3$  的原因在於，若只找最相似、最近的鄰居點很容易會造成誤判，因為雖然每個人的走路不太一樣，但是差距亦不大，所以在這邊選擇採取類似多票決的方式來進行給分並判斷其身份。圖 3.25 為身份鑑別法則的流程圖：

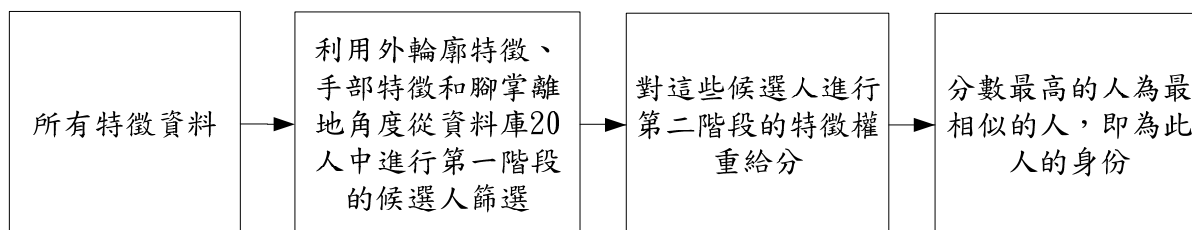


圖 3.25 身份鑑別法則流程圖。

如圖 3.25 所述，本論文在這邊提出兩階段式的身份鑑別系統：

#### 第一階段 候選人篩選

先將外輪廓、手部擺動和腳掌離地角度三個特徵拿來進行候選人篩選，篩選的目的在於減少候選人，以提高整體辨識效能。而為了確保第一階段真的有選到正確答案，所以在這邊是將最近的三個鄰居資料點都列為候選人，意即分別找出這三種特徵與資料庫中最相似的三筆資料。假設都沒有重複的狀況，此時最多的候選人頂多也只有 9 個人，和資料庫中 20 個人比起來，已經減少一半以上的候選人了。而在這邊利用外輪廓、手部擺動和腳掌離地角度三個特徵來進行第一階段的篩選，主要是依據每種特徵重要性的不同，從中選擇最重要的三種特徵來進行第一階段的候選人篩選。

#### 第二階段 相似度判斷

完成候選人篩選後，接著再進行第二階段權重給分。第二階段主要就是將所有特徵，包含外輪廓特徵、手部擺動幅度、頭部特徵、腳掌特徵、膝蓋特徵和步伐大小特徵，共六個特徵與第一階段篩選出來的候選人進行匹配，依照特徵權重比例的不同，為匹配到的候選人進行給分。在所有特徵都完成權重給分之後，接著便統計每個人的分數為多少，分數最高的即為配對到的人，意即此人的身份。

### 3.6.2 安全監控法則

在身份鑑別的部分，因為主要只是要進行身份辨識，所以當現在輸入一筆資料，在身份鑑別系統中不管特徵權重給的總分多低，都一定會在資料庫中找出最相似的人出來，並給予身份。但是在安全監控系統中，主要目的在於安全，因此不僅要能辨識身份，更要能偵測出陌生人，因此如何準確偵測出陌生人變成了安全監控的主要課題。

本論文在區分資料庫人物與陌生人的部分，主要是透過設定閾值來判斷其相似程度是否夠大，足以判斷為資料庫人物並給予身分，否則就判斷為陌生人物。在設定閾值的部分，主要是透過觀察並挑選使整體辨識率最佳的數值來做為閾值大小。本論文所設定的候選人得分總數最高為 12.99，即滿分為 12.99，如表 3-1 所示，所以進行閾值測試的時候，會先從 1 到 13 的閾值大小來進行整體辨識率比較，如圖 3.26 所示，圖中橫軸為閾值大小，縱軸為相對閾值的整體辨識率。

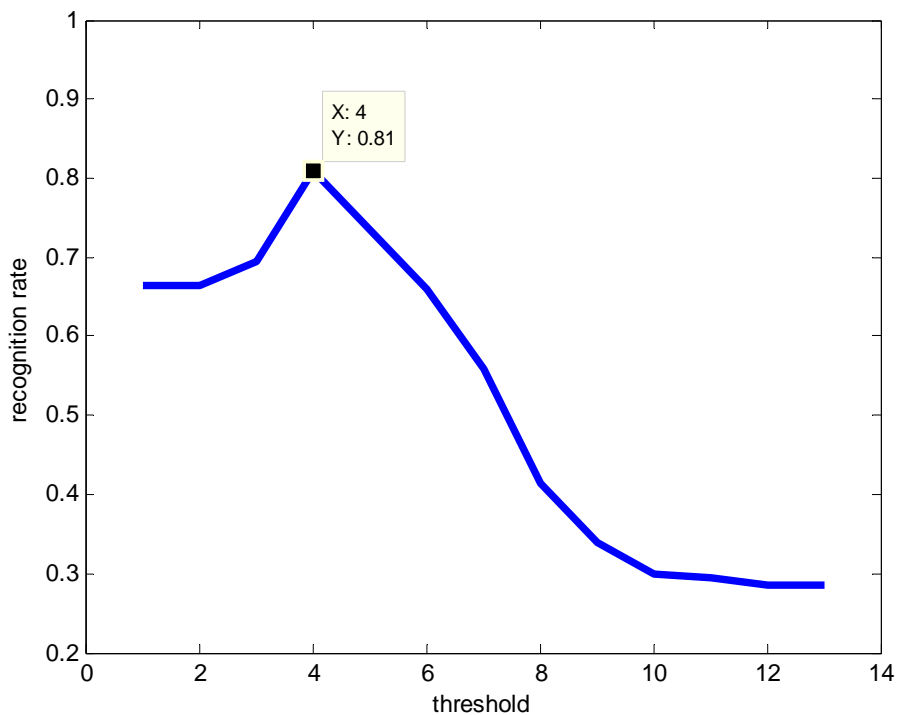


圖 3.26 閾值選定示意圖。

從圖 3.26 可得知，當選定閾值為 4 的時候整體辨識率會有較佳的表現。進一步縮小閾值取樣間隔，再進行一次辨識率統計，如圖 3.27 所示。

從圖 3.27 可得知，令閾值為 4.2 可以使整體的辨識率較高，因此，本論文設定區分資料庫人物與陌生人的閾值為 4.2。若得分最高的候選人總分低於 4.2，表示此輸入的人與資料庫太多人都相似，但是相似程度又不夠大，因此便判定此人為陌生人。

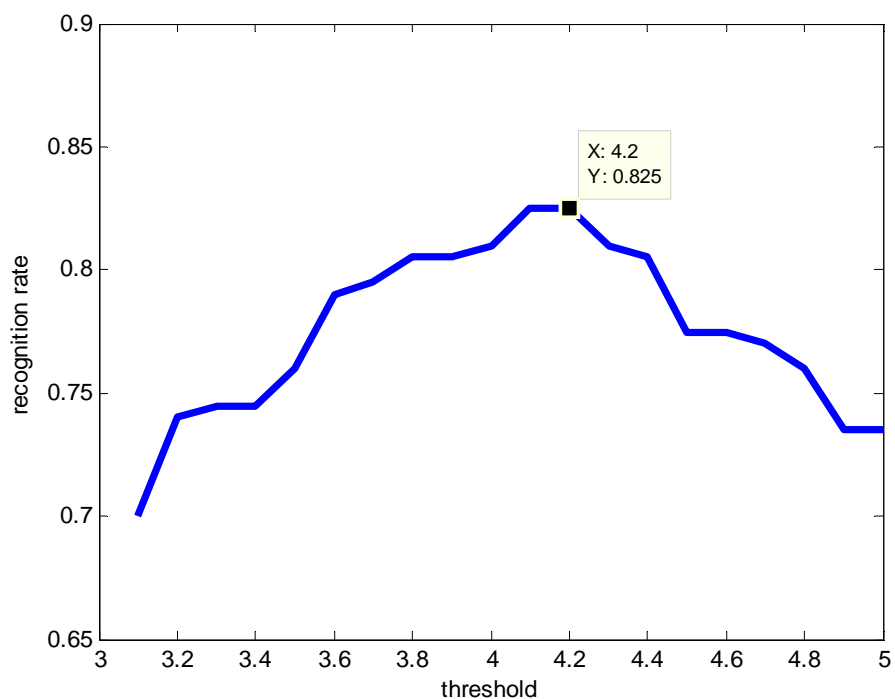


圖 3.27 閾值選定示意圖。

圖 3.28 為安全監控法則的流程圖：

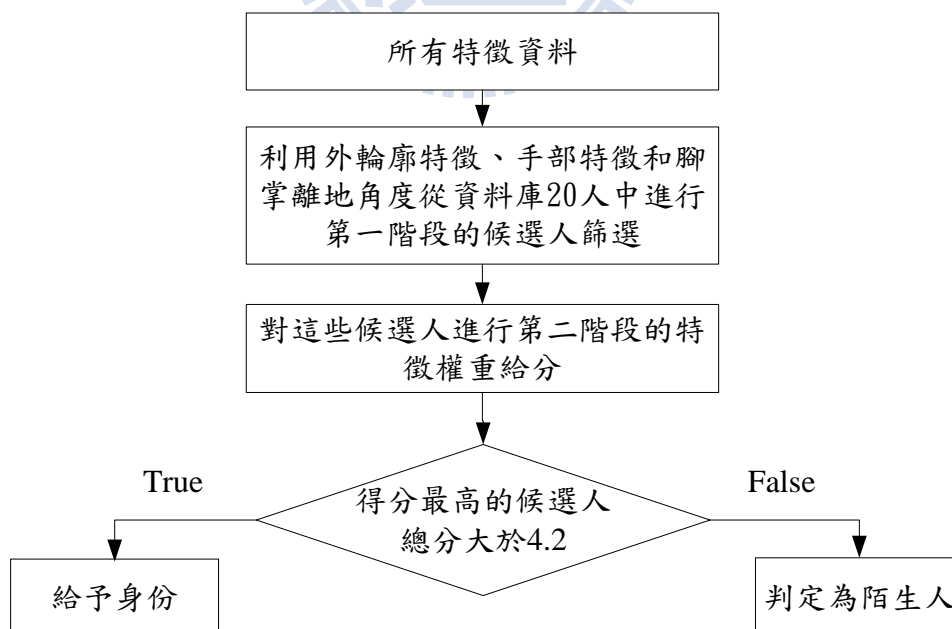


圖 3.28 安全監控法則流程圖。

## 第四章 實驗結果與分析

在本章中主要討論本論文提出的身份鑑別方法之辨識成效，並進一步將此方法應用在安全監控的層面上，即偵測到非資料庫的人時，將其判斷為陌生人，以提供警示通知警衛加強注意或前往關切等等。因此，在本章中又分為四小節來討論，4.1 介紹整個實驗的環境與設備；4.2 介紹利用本論文所提出的方法來進行身份鑑別的實驗結果；4.3 分析本論文方法與其他方法的辨識結果比較；4.4 則是就本論文的實驗結果來進行討論。

### 4.1 實驗機制

本節中主要是介紹在本論文中的實驗機制，包含實驗時所使用的器材、實驗的環境和本論文所使用的資料庫。因此又分了兩小節來介紹，4.1.1 節中介紹我們實驗用的一些硬體架構及設備規格；在 4.1.2 節中會說明本實驗所使用的資料庫大小及包含了哪些形態的資料。

#### 4.1.1 實驗場景與設備

實驗用的設備是一台 canon IXUS 110IS 1210 萬像素的數位相機，利用其攝影的功能來進行拍攝與錄製。將攝影機架設在距離人行走位置約 1.5 公尺遠的地方，如圖 4.1 所示。

在實驗中拍攝人的走路姿態時，相機是固定的且焦距也是固定的。將相機所拍攝的資訊，利用個人電腦將拍攝好的影片經過影像擷取卡儲存成影像序列形式，用這些影像來當作輸入，而這些輸入影像的解析度為 640\*480 像素。至於系統的執行過程則是在中央處理器為 intel i5 3.33G Hz，4GB 記憶體，32 位元作業系統的个人電腦上，如圖 4.2 所示



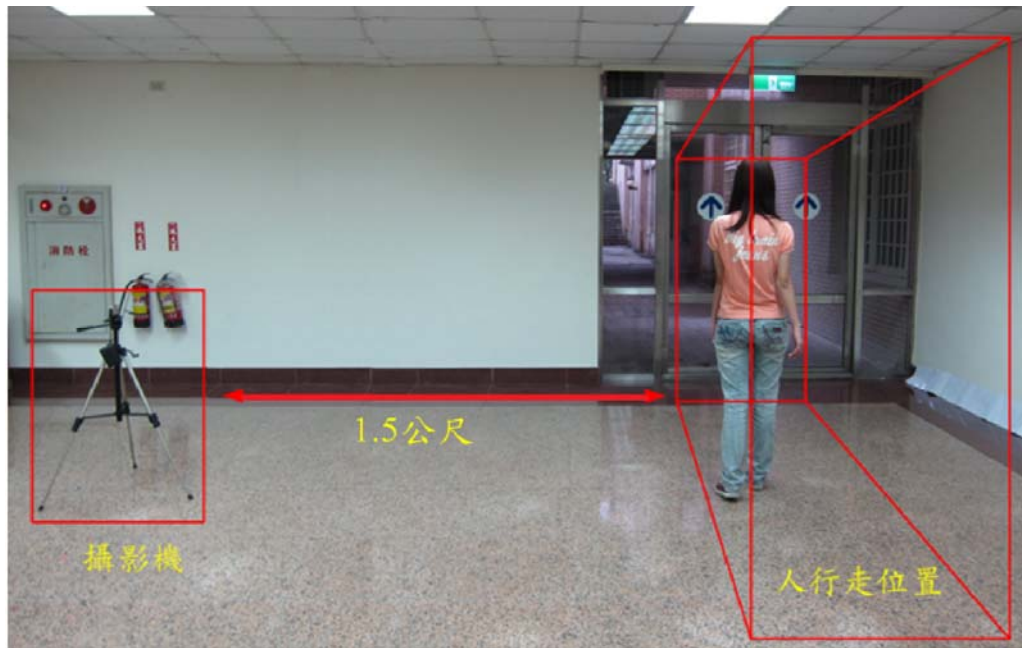


圖 4.1 實驗場景圖。

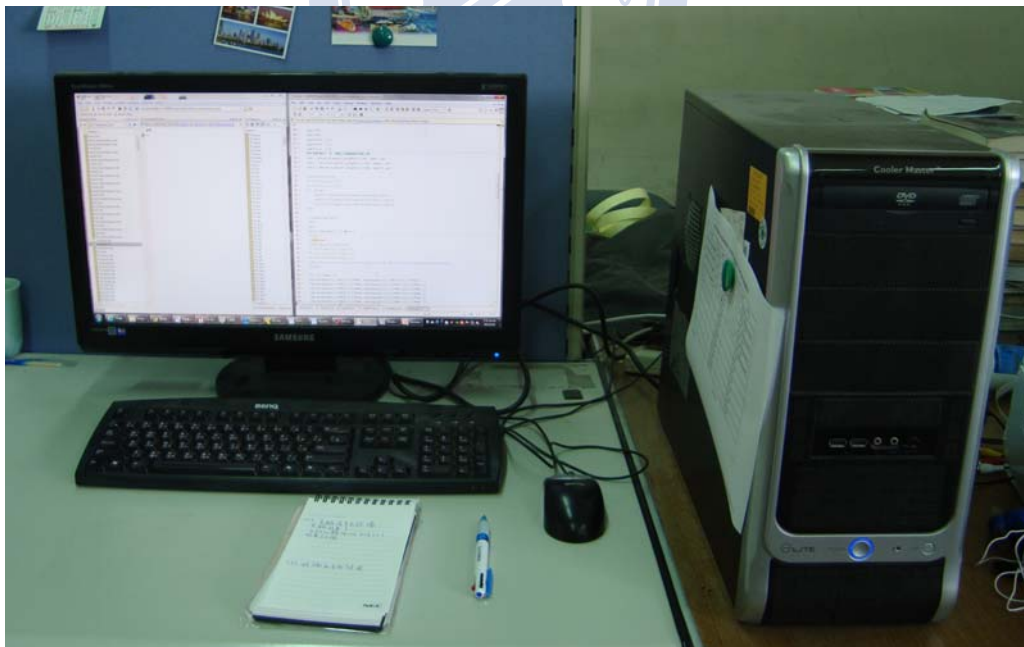


圖 4.2 程式執行所使用的個人電腦。

### 4.1.2 實驗資料

本論文的資料庫內包含 8 女 12 男，共 20 人，每人 4 筆資料，共計有 80 筆資料。在資料庫中的這 80 筆資料單純為原始資料，沒有額外加入變因。

在測試資料中，共包含 200 筆資料。143 筆資料為資料庫內的人物，其中也包含一些具有變因的資料，例如戴帽子等等，如圖 4.3 所示；另外的 57 筆資料為非資料庫的人物，即為陌生人的資料。



圖 4.3 測試資料影像範例。(a)(b)為普通正常行走的影像；(c)(d)為加入變因，穿外套的行走影像；(e)(f)為加入變因，戴帽子的行走影像。

## 4.2 實驗結果

在本實驗中，主要是針對 100 筆沒有格外添加任何變因的原始資料；43 筆加入變因後的資料，例如戴帽子、穿外套，與 57 筆非資料庫內的人物；總共 200 筆的資料來進行本系統的辨識率測試。

在本實驗中，主要又分成三個系統來進行討論，4.2.1 節主要討論就單純的身份鑑別系統來說，本系統的辨識結果；4.2.2 節說明本系統應用在安全監控上的辨識結果；4.2.3 節則是介紹模擬加入特殊變因後的辨識結果。

### 4.2.1 身份鑑別結果與分析

在身份鑑別的部分，主要就是將輸入的人與資料庫做比對，按照圖 3.25 所描述的步驟流程來執行。首先，利用外輪廓特徵、手部特徵和腳掌離地角度三種特徵來進行候選人的第一階段篩選；接著，找出與每一個特徵最相近的三個鄰居點並給與對應的權重值，最後統計候選人中每一個人的總分，分數最高的候選人即為與輸入資料最相似的人，而本系統就是藉由此來判斷其身份。表 4-1 為本論文的系統辨識率，表格左行為輸入測試的資料型態，分別討論 100 筆的原始資料與 143 筆所有測試資料的辨識率。原始資料中皆為未刻意加入變因的步態影像，所有資料中包含原始的步態影像及刻意加入變因的步態影像，例如戴帽子和衣著不同等等。本論文的身份鑑別辨識率定義如下：

$$\text{身份鑑別辨識率} = \frac{\text{所有輸入資料身份匹配正確的總筆數}}{\text{所有輸入資料的總筆數}} \times 100\% \quad (4-1)$$

表 4-1 本論文之身份鑑別系統辨識率

資料(筆數)	辨識率(%)
原始資料(100)	96
原始資料(100) + 加入變因後的資料(43)	93

由表 4-1 的辨識結果來看，可得知本系統在原始資料，意即在未刻意加入變因的情況下進行身份辨識其準確度可高達 96%。而辨識率無法達到百分百的原因在於，雖然人在走路時會有一定的特性，但若發生突發狀況時，例如踢到地面的小東西、與人打招呼或閃避障礙物等等，以上都有可能讓步態產生些微的不同，使系統造成誤判。下列為本論文在腳掌離地角度誤判之範例，如圖 4.4 所示：

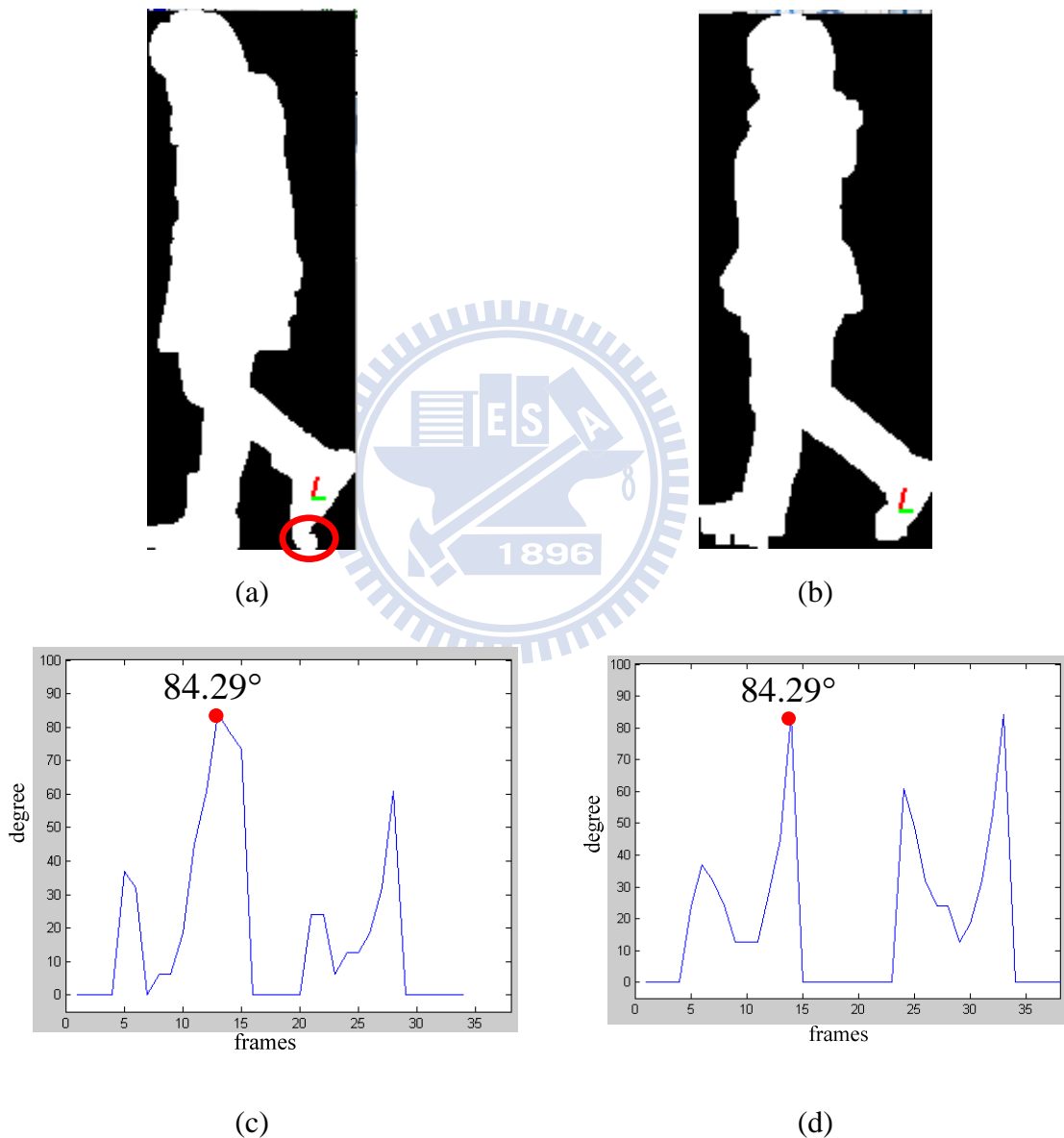


圖 4.4 本系統辨識錯誤範例說明。(a)身份 A 的腳掌離地最大角度步態影像圖；(b) 身份 B 的腳掌離地最大角度步態影像圖；(c)身份 A 之腳掌離地角度曲線變化圖；(d) 身份 B 之腳掌離地角度曲線變化圖。

圖 4.4(a)與(b)中人物之真實身份分別為 A 與 B，但本系統卻在輸入此筆 A 的資料時將其誤判為身份 B。探討產生誤判的原因，首先觀察身份 A 的其他筆步態資料，其腳掌離地角度的抽取結果如圖 4.5 所示。身份 A 的腳掌離地角度特徵大致為 78.69 度，但是在圖 4.4 的範例中，腳掌離地角度抽取結果卻為 84.29 度。進一步觀察身份 A 被判斷為 84.29 度的步態影像圖，如圖 4.4(a)所示，圖中 A 的後腳掌在最下面地面與腳掌的連接處出現一個不屬於腳的白點，在此稱為雜訊。雜訊的來源也許是因為地面出現小石子，或是本系統在一開始前景背景切割時就殘留下來的雜訊點，因此才會將身份 A 的腳掌特徵誤判為身份 B 的 84.29 度。但將雜訊撇除後，從圖 4.5 中可清楚得知，本系統在估測腳掌離地角度的確有一定的準確性。

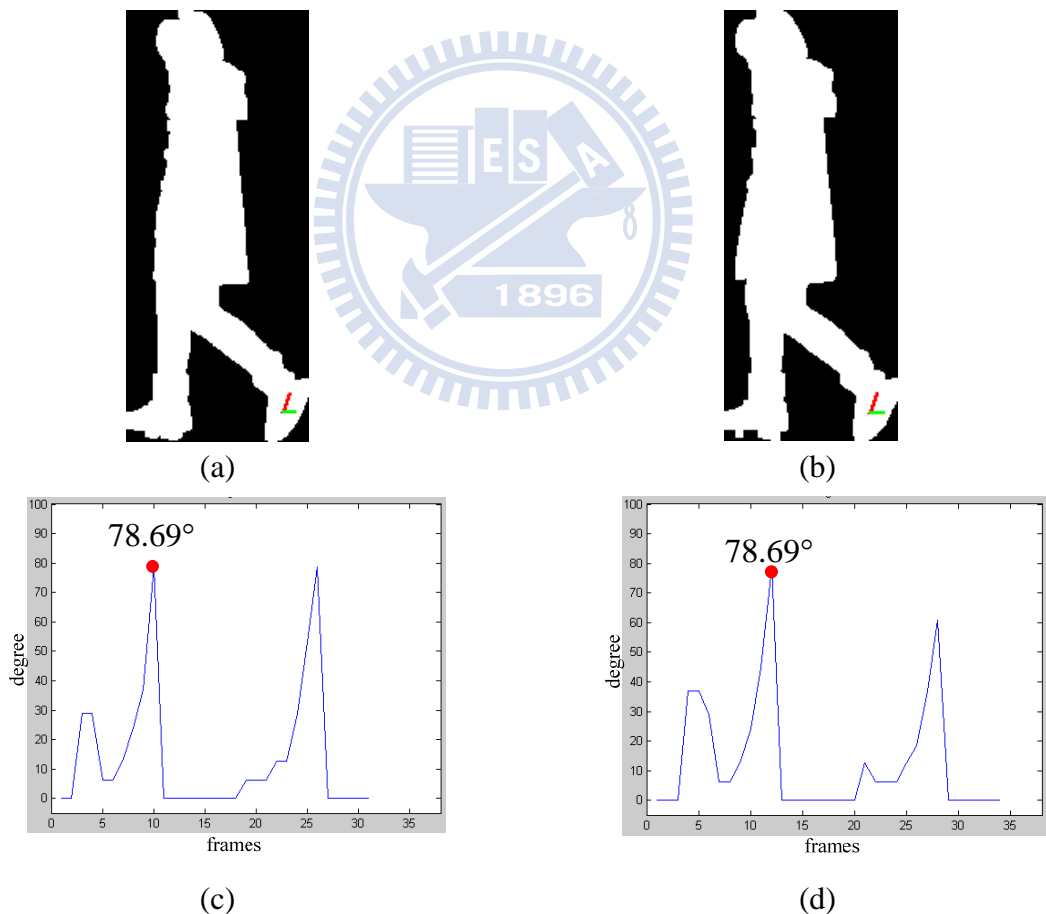


圖 4.5 身份 A 在其他筆資料中，腳掌離地角度變化範例。(a)身份 A 的第一段步態資料中腳掌離地最大角度之影像圖；(b)身份 A 的第二段步態資料中腳掌離地最大角度之影像圖；(c)利用身份 A 的第一段步態資料分析腳掌離地角度之曲線變化圖；(d) 利用身份 A 的第二段步態資料分析腳掌離地角度之曲線變化圖。



因為每個人之間的單一特徵差異不大，所以當受測者在行走時因為外在因素或雜訊的影響，出現不同以往的走路姿態，如圖 4.4 中的範例所示，此特徵就很容易造成誤判。而本論文的優勢在於，本系統是平均分散所有特徵於人體姿態的各個區域，所以就算有單一特徵辨識錯誤，只要外輪廓、手部擺動幅度、頭部傾斜角、膝蓋特徵、腳掌離地角度特徵或步伐大小這六個特徵中有三個以上判斷正確，整體的身份辨識就會是正確的。因此，雖然本系統無法達到百分百的辨識率，但是仍有高達 96% 的辨識準確性。而在加入變因後，本系統依然具有 93% 的高辨識率，由此可知，本系統所提出的方法在對抗變因的時候具有穩定的辨識準確度。

## 4.2.2 安全監控結果與分析

在安全監控的部分，主要就是按照圖 3.28 所描述的步驟流程來執行。首先，利用外輪廓特徵、手部特徵和腳掌離地角度的特徵來進行候選人的第一階段篩選；接著，找出與每一個特徵最相近的三個鄰居點並給與權重值，最後統計候選人中每一個人的總分，若得分最高的候選人總分小於閾值 4.2，即表示此時輸入的人與資料庫太多人都相似，但是相似程度又不夠大，因此在這邊判定此人為陌生人。

表 4-2 為本論文在安全監控部分的系統辨識率，表格左行為輸入測試的資料型態，分別討論 100 筆資料庫人物的原始資料、143 筆所有資料庫人物資料，包含加入變因後的測試資料、57 筆陌生人資料和全部測試資料的辨識率。原始資料中皆為資料庫人物且未刻意加入變因的步態影像，所有資料庫人物資料包含原始的步態影像及刻意加入變因的步態影像，例如戴帽子和衣著不同等等。陌生人資料中皆非資料庫內人物的步態影像，全部測試資料則是將上述的所有資料全部加入並進行辨識率討論。本論文在安全監控部分的相關辨識率定義如方程式 4-2 和 4-3 所示，而本論文就是利用方程式 4-2 和 4-3 來計算安全監控部分的系統辨識率，結果如表 4-2 所示。

$$\text{安全監控辨識率} = \frac{\left( \begin{array}{c} \text{輸入資料為資料庫人物並身份判斷正確的總筆數} \\ + \\ \text{輸入資料為陌生人並判斷為陌生人的總筆數} \end{array} \right)}{\text{所有輸入資料的總筆數}} \times 100\% \quad (4-2)$$

$$\text{陌生人辨識率} = \frac{\text{輸入為陌生人並判斷為陌生人的筆數}}{\text{所有輸入的陌生人總筆數}} \times 100\% \quad (4-3)$$

表 4-2 本論文之安全監控系統辨識率

資料(筆數)	辨識率(%)
原始資料(100)	88
原始資料(100) + 加入變因後的資料(43)	84
陌生人資料(57)	82.46
全部資料(143 筆資料庫人物 + 57 筆非資料庫人物)	82.5

觀察表 4-2 的實驗結果，雖然本系統在安全監控方面的辨識率還不夠亮眼，但是也有八成左右的準確度。歸咎辨識率不高的原因可能是因為本系統在同時辨識身份與陌生人的安全監控系統設計中，如圖 3.28 所示，只有簡單地設置一個閾值來區分資料庫人物與陌生人，所以整體辨識率才會無法達到預期中的效果。但是在大多數文獻中，都只研究到利用步態分析來進行身份辨識，較少討論到將其應用在偵測陌生人的系統上，所以在區分資料庫人物與陌生人的判斷方式上，也尚未有一個更有效的區分方法。因此，若是站在輔助系統的角度來說，本系統已經有不錯的表現。

考慮到警衛人員在進行監控時不可能一直緊盯著所有監控螢幕，因此傳統的人為監控系統之安全性將隨著警衛人員工作時間的上升而降低。本系統的主要目的是在於輔助警衛人員、提醒出現陌生人之警示，並減少人力成本的耗損，因此若是作為輔助工具，站在減輕警衛人員的負擔與提醒協助前往現場勘查的立場來說，本系統已經可以為警衛人員提供極大的助益。

### 4.2.3 特殊變因情況討論與分析

人的衣服穿著搭配種類繁多，為了能應對各式各樣的情況，本論文採用的方法並非將特徵完全集中在腳部，也不是將特徵全部集中在靜態特徵上，而是平均分散抽取特徵的位置點，頭、手、前後膝蓋、腳掌、步伐、外輪廓，從上到下從內到外將人體的特徵平均分配在人體的各個部位，以降低衣著、雜訊等變因對辨識率的影響。

考慮特殊情況，若有人穿著長裙出現，從影像中是無法直接區分前後兩隻腳確切的位置。因此本論文與文獻[5]-[8]不同的地方在於，本論文不堅持一定要找到雙腳確切的位置，而是利用膝蓋區域之前後曲線來進行膝蓋特徵的描述。而使用前後曲線的原因在於，若現在出現一個穿長裙的人，在他往前走的時候，裙子的前端必定會被前膝蓋帶起，因此利用膝蓋的前曲線來偵測前腳彎曲角度是完全不受影響的，意即本論文的方法在此情況下亦是適用的。然而在後膝蓋的部分，由於裙子材質的不同，導致有些時候無法準確估測出後膝蓋的彎曲角度，因此後膝蓋的特徵辨識率可能會大幅降低。有鑑於此，本實驗在這邊假設最糟的情況為後膝蓋特徵完全判斷錯誤，因此在本實驗中我們將後膝蓋的特徵拿掉，只利用頭、手、前膝蓋、腳掌、步伐大小和外輪廓特徵來進行辨識，其辨識結果如表 4-3 所示：

表 4-3 模擬人穿長裙情況之系統辨識率

資料(筆數)	身份鑑別(20 類)辨識率(%)	安全監控(21 類)辨識率(%)
原始資料(100)	91	80
資料庫人物資料(143)	84	76
陌生人資料(57)		92.9
全部資料(200)		74



探討並分析造成表 4-3 實驗結果的原因在於，此節中是模擬受測者不管男女都穿上了長裙，並假設最糟的情況為後膝蓋特徵完全判斷錯誤。但是，現實生活中男生不太可能會穿長裙，亦不太可能全部的人同時穿上長裙出現，而後膝蓋特徵的部分也不太可能完全失效。因此，本實驗主要在測試系統在最糟情況下的辨識效能，從實驗結果來看，本系統在最糟情況下仍有七八成的辨識率。

從表 4-3 中的實驗結果可以發現，去除後膝蓋特徵後雖然使整體辨識率下降，但是就單純陌生人辨識的部分準確度反而提高。歸咎造成此現象的原因在於，本系統是利用受測者的所有特徵與候選人特徵做相似度比對，並給予相對應的權重分數，所以拿掉一個特徵會使候選人的整體得分總數下降，一旦低於系統設定的閾值就會被判定為陌生人，所以陌生人的辨識率才會提升至 92.9%，不過就整體辨識率來說，誤判為陌生人的機率也上升，所以整體系統辨識率才會較表 4-1 和表 4-2 來的低。

儘管如此，本系統在模擬最糟情況，即後膝蓋特徵完全辨識錯誤的情況下，系統辨識率還是有一定的水準。雖然少了一些特徵資訊使得整體身份鑑別之辨識率下降，安全監控的部分也只剩七八成，但是若只單單針對偵測陌生人的部分來說，本系統具有高達 92.9% 的辨識率。

### 4.3 分析比較

本論文在本節中將與其他文獻所提出的方法做比較，但是就目前的文獻中，大多都只應用在身份鑑別上，在安全監控的方面較少探討。因此，本論文在本節中將針對身份鑑別的部分來與其他文獻所使用的方法進行辨識率比較，如表 4-4 所示：

表 4-4 本論文方法與其他文獻方法之身份鑑別辨識率比較

方法	原始資料(100 筆)身份鑑別辨識率(%)	原始資料(100 筆) + 加入變因後的資料(43 筆)身份鑑別辨識率(%)
本論文所提出的方法	96	93
B. Lee, S. Hong, H. Lee, and E. Kim[10]	90	88
C. Wang, J. Zhang, L. Wang, J. Pu, and X. Yuan[17]	93	88
H. Ng, H. L. Tong, W. H. Tan, T. V. Yap, and J. Abdullah[8]	82	75
C. Nandini, P. Mukhopadhyay, T. Tanmay, S. K. Ranjan, S. Roy[9]	73	70

H. Ng 等人[8]主張先找出人體姿態的骨架圖後，針對各個關節點來進行角度的特徵抽取，最後使用模糊最近鄰居法(fuzzy-knn)來進行身份分類。C. Nandini 等人[9]所使用的特徵主要都是依靠在一周期中每一張側影姿態的方框大小，包含每一張側影影像方框的長度、寬度、面積和方框對角角度。B. Lee 等人[10]是先找出動態輪廓影像(motion contour image, MCI)後，算出 MCI 的特徵值與特徵向量，藉由參數的設定與轉換，算出投影矩陣，再將特徵向量乘上投影矩陣即為特徵，最後在使用最近鄰居法來進行身份分類。C. Wang 等人[16]主要是針對外輪廓影像來進行特徵抽取，首先先將一周期序列影像等分成四段，對每一段的影像分成 R、G、B 三通道來處理，最後利用主成份分析法(principal component analysis, PCA)和線性識別分析(linear discriminant analysis, LDA)來進行特徵降維，最後利用相似度比對找出最相似的人。詳細分析與探討本方法與其他方法的優缺點，如表 4-5 所示：

表 4-5 各方法優缺點比較

方法	缺點	本論文方法的優點
B. Lee, S. Hong, H. Lee, and E. Kim[10]	特徵維度過大。	根據文獻[10]的描述，其特徵大小為17*40，而本論文提出的方法所使用的特徵維度大小僅為：1*28。
C. Wang, J. Zhang, L. Wang, J. Pu, and X. Yuan[17]	特徵維度過大。	根據文獻[17]的描述，其特徵大小為140*3，而本論文提出的方法所使用的特徵維度大小僅為：1*28。
H. Ng, H. L. Tong, W. H. Tan, T. V. Yap, and J. Abdullah[8]	若下半身穿長裙類的服裝，則雙腳部分的骨架圖之準確性便有待商榷。	針對此方法的缺點，本論文對於膝蓋的角度特徵是偵測前膝蓋區域影像之前曲線和後膝蓋區域影像之後曲線的彎曲角度，而不是使用骨架圖。
C. Nandini, P. Mukhopadhyay, T. Tanmay, S. K. Ranjan, S. Roy[9]	[9]所使用的特徵為：每一張側影影像方框的長度、寬度、面積和方框對角角度。雖然計算複雜度較低，但使用這些特徵卻很容易受到雜訊影像，因此辨識率也較低。	相對此方法的缺點，本論文所抽取的特徵皆平均分散在人體的各個部位，所以即使加入變因後，本論文的辨識率也有不錯的成效。

利用步態分析來進行身份鑑別需要考慮的因素很多，而這也是為什麼步態分析一直是個熱門的研究主題，但是卻還無法將此應用真正運行在現實生活中。因此，藉由研究與分析其他方法的優缺點進而擷取經驗改善缺點，使應用步態分析於身份鑑別之系統能更加完善並增加實現在生活中的可能性。

從表 4-4 可得知，本論文在這邊相較於其他方法來說都有較佳的辨識效果，而表 4-5 則是進一步分析比較各方法的優缺點。雖然文獻[10]和文獻[17]在身份辨識上也有

不錯的表現，辨識率也都有九成以上，但是相對本論文來說，文獻[10]和文獻[17]所使用的特徵維度都很大，對於之後如果要脫離電腦實現在其他硬體上時，特徵維度太大会相對提高硬體的コスト考量。另外，雖然文獻[8]和文獻[9]所使用的方法複雜度較低，但是在面對變因的影響時，辨識率卻下降了不少。而本系統就身份鑑別的部份來說，辨識率高達 96%，加入外套、戴帽子變因後的辨識率也有 93%。由此可知，本系統的身份鑑別方法確實具有不錯的辨識效果。

進一步分析本系統辨識率較高的原因，從表 4-6 可得知，本系統若單純只使用外輪廓一個特徵來進行身份鑑別就已經有 85% 的辨識率，加入更多的特徵確實有提高辨識率的效果，但是能達到的辨識率依然有所限制，並不是一味多加入一些特徵就能達到百分百的辨識率，這亦表示特徵數目或維度的增加與辨識率上升不一定會成正比。而本論文使用這麼多特徵的原因在於，因為若只使用單一個特徵，辨識率很容易就會受到衣著外觀的影響而下降，因此，我們在此平均分散的抽取人體各個部位的特徵，希望能降低衣著變因對本系統的影響。在本系統中，主要進行身份鑑別的其實只有外輪廓、手部擺動幅度和腳掌三個特徵，剩下的頭部、膝蓋和步伐大小三個特徵則是作為輔助的角色，利用這些特徵的搭配組合，以提高整體的辨識率，而這也是為什麼本系統為什麼比別人好的原因。

表 4-6 單一特徵身份辨識率

特徵	外輪廓	手部擺動幅度	頭部	腳掌	膝蓋	步伐大小
辨識率	85%	74%	46%	59%	57%	54%

## 4.4 結果討論

本論文主要是應用步態分析於身份鑑別之研究。針對每個人走路由於個人習慣的不同，產生的不同走路形態來進行分析與特徵抽取。本系統不僅平均抽取人體各個部位的特徵，更進一步又將特徵分為靜態特徵與動態特徵，以期確保本系統面對衣著變因的影響時，還能保持一定的辨識率。最後，本系統在進行身份分類的時候，更提出一套兩階段式的辨識流程，第一階段利用外輪廓特徵、手部特徵和腳掌離地角度特徵來進行候選人篩選，第二階段將全部特徵針對這些候選人進行權重給分，依據總分的高低來進行身份辨識。

除此之外，本系統更將步態分析應用於安全監控上，整個系統的流程和身份鑑別系統一樣，差別只在於最後的分類步驟。在身份鑑別的部分，是選擇候選人中分數最高的人為此人的身份，所以不管總分多低，都一定會匹配到資料庫中某個人的身份。但是不同於身份鑑別系統，在進行安全監控時，主要是希望不僅可以進行身份辨識更可以偵測出陌生人。因此本系統在這邊的判斷原則是設定一個閾值，令閾值的大小為 4.2，並以此來做為判斷是否為陌生人的標準。若此時有人在辨識後所得到的總分不到閾值的 4.2，表示此輸入的人與資料庫很多人都相似，但是相似程度又都不夠高，因此就會被判斷為陌生人。

4.2.1 節主要是針對身份辨識系統的部份來進行分析與探討。從表 4-1 可得知，本系統在正常情況下的身份辨識率高達 96%，在刻意人為加入外套、帶帽子等變因後，雖然辨識率下降了一些，但是還是有 93% 的身份辨識率。探討本系統無法達到百分百的辨識率是因為在現實生活中，擷取步態影像時難免還是會受到雜訊的影響，使得辨識率往往無法達到預期中的好。所以本系統在這邊是透過平均分散地抽取人體姿態中各個部位的變化來做為特徵，不僅可以有效抵抗衣著變因的影響，更能減少雜訊對本系統的影響。平均分散特徵於人體步態各個區域還有一個優點，若現在其中有一個特徵受到雜訊影響，導致辨識錯誤，但是若其他特徵辨識正確的話，此雜訊就不會對本系統的辨識率造成影響。從表 4-4 中與其他方法的身份鑑別率做比較，本系統所



使用的方法之辨識率確實優於其他方法。

4.2.2 節主要是針對應用在安全監控上的分析與探討。由於在相關文獻中，大多都只討論到身份鑑別，還未真正討論到安全監控上，因此在這裡就單純討論本系統在安全監控上的表現。從表 4-2 可得知，本系統在安全監控上大概有八成左右的辨識率，雖然表現還不夠亮眼，但若是站在輔助警衛人員，提供出現陌生人之警示的立場來說，本系統已經能達到不錯的效果。探討辨識率不高的原因推測是在於區分資料庫人物與陌生人的時候，本系統在這邊只單純設定一個閾值來做為區分的標準，令此閾值大小為 4.2，並以此來做為區分資料庫人物與陌生人的標準。單純使用一個閾值來做為區分的標準其實是不夠嚴謹的，但是更有效的區分方法卻還沒有一個大致的方向。因此，就本論文 4.2.2 的實驗結果來討論，站在協助提供警示給警衛人員的立場來說，八成的辨識率已經可以達到不錯的協助效果。

4.2.3 節主要是模擬受測者在穿著長裙的情況下系統的辨識效能。大多數的文獻都是使用人體骨架圖的方式來抽取腳部區域的抽取，但若因為衣著的關係，連人眼都無法準確區分四肢的位置時，又該如何期望系統能利用骨架圖準確找到四肢的位置。因此，本系統所採用的方法不是利用骨架圖來描述雙腳，而是利用前腳的前曲線以及後腳的後曲線來估測雙腳的膝蓋彎曲角度。使用前曲線來描述前膝蓋的彎曲角度好處在於，就算有人穿著長裙行走，但是在人往前走的時候，裙子很自然的就會被前腳所帶起，所以還是可以利用偵測前曲線的方式找出前膝蓋的彎曲角度，而這也是本系統優於其他方法的關鍵之一。後腳膝蓋的部分則是要看裙子的種類才能決定後膝蓋在此情況下是否還具有鑑別力，因此本論文在這邊的模擬實驗中是考慮最糟的情況，意即後膝蓋完全失效沒有鑑別力的情況。表 4-3 為本實驗的模擬結果，因為是模擬最糟的情況，所以整體的辨識率相對也比較低一點，但是就「安全」監控來說，在辨識陌生人的部分有高達 92.9% 的辨識率。

4.3 節主要是將本系統所使用的方法與其他文獻所使用的方法做比較。從表 4-4 和表 4-5 中可明顯得知，本系統的辨識效能確實有較好的表現。進一步分析本論文與其他文獻的不同之處如下所述：

- (1) 本系統是偵測前腳的前曲線和後腳的後曲線來估測膝蓋彎曲角度，而不是使用骨架圖。因此在面臨受測者穿著長裙的情況時，本系統依然有一定的準確性。
- (2) 本系統所使用的特徵平均分散在人體姿態的各個區域，若有單一區域的特徵受到雜訊影響而誤判時，還可以透過其他特徵的互相補足，只要其他特徵辨識正確，就可以減輕此雜訊對本系統的影響。
- (3) 本系統在分類的過程中，必須透過兩階段式的篩選才能確認身份。第一階段主要是利用外輪廓特徵、手部特徵和腳掌離地角度特徵來進行候選人篩選，篩選的目的在於減少候選人，以提高辨識率；第二階段則是將所有特徵針對這些候選人進行給分，總分最高的即為此人的身份。

由於以上的優點，使得本系統在與其他文獻方法做比較的時候，如表 4-4 所示，本系統確實有較佳的辨識成效。



## 第五章 結論與未來發展

本論文主要是應用步態分析於身份鑑別之研究。在本系統中，擷取前景的方式為前景背景相減法。主要是先將一段只有背景的影像透過高斯混合模型來訓練背景模型，之後只要將輸入影像減去背景模型即可獲得前景影像。獲得前景影像後，本系統接著利用人體雙腳與頭部之間的夾角，如圖 3.5(b)所示，來估測每個人走路的週期，並從中擷取出一段週期影像序列出來進行特徵抽取。在特徵的部分，本系統主要又分成了兩大類：靜態特徵與動態特徵。靜態特徵的部分主要是利用傅立葉描述子來描述步態能量影像中的外輪廓特徵和分析步態能量影像中手部區域之前曲線分布情形來抽取手部擺動幅度特徵；動態特徵則是針對週期內每一張影像都進行特徵的抽取，其中包含利用主成份分析法來計算頭部傾斜角和腳掌離地角度、利用哈里斯角點偵測來計算膝蓋彎取角度和雙腳的步伐距離大小。至於最後的身份鑑別法則，本系統提出兩階段式的分類法，第一階段候選人篩選，主要是利用外輪廓特徵、手部特徵和腳掌離地角度三個特徵找出每個特徵與資料庫中最相似的三筆資料來進行候選人的篩選；第二階段相似度判斷，則是將所有特徵針對第一階段篩選出來的候選人進行權重的給分，分數最高的則表示與此人最相似，即此人的身份。

從第四章的實驗結果可得知，本系統在身份鑑別的部分來說辨識率高達 96%，在加入外套、戴帽子變因後的辨識率也有 93%，由此可知，本系統在身份鑑別上確實具有不錯的效果。在安全監控方面，雖然整體的辨識率只有七、八成左右，但若是站在輔助警衛人員、提供出現陌生人之警示的立場來說，已經有不錯的協助效果。在本實驗中，陌生人的辨識率大致也有八、九成的鑑別力。

近一步分析本系統優於其他方法的原因後，可以將本論文的貢獻歸納為以下兩點：

- (1) 本論文結合多種方法，使用多面向與多角度來分析姿態，不僅平均抽取人體各個部位的特徵，更進一步將特徵分為靜態特徵與動態特徵，以期確保本系統面對衣

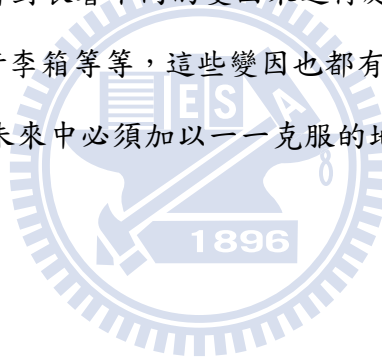


著變因的影響時，還能保持一定的辨識率。

- (2) 本論文提出兩階段式的身份鑑別系統。第一階段主要是利用外輪廓、手部擺動和腳掌離地角度三個特徵來進行候選人篩選，而篩選的目的在於減少候選人，以提高整體辨識率；第二階段則是將全部特徵針對第一階段篩選出來的候選人進行權重給分，分數最高的即表示與此人最為相似。

最後，將本系統進行分析與歸納後，可以將未來發展主要分為兩大方向：

- (1) 因為本系統目前只有針對人體的側影方向來進行分析，但現實情況是人可以往各個方向前進，所以若是能將各個角度的走路姿態都加入一起討論，便可以增加本系統應用步態分析於身份鑑別系統的完整性。
- (2) 在本系統中，目前只針對衣著不同的變因來進行處理，但是現實生活中的變因很多，包含後背包、拖行李箱等等，這些變因也都有極大的可能影響整體系統的辨識率，而這些也是在未來中必須加以一一克服的地方。



## 參考文獻

- [1] M. Hu, Y. Wang, Z. Zhang, and D. Zhang, "Gait-Based Gender Classification Using Mixed Conditional Random Field, " *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 41, no. 5, pp. 1429-1439, Oct. 2011.
- [2] M. Hu and Y. Wang, "A New Approach for Gender Classification Base on Gait Analysis," in *Proceedings of International Conference on Image and Graphics*, Xi'an, Shanxi, pp.869-874, Sept. 2009.
- [3] X. Li, S. Maybank, and D. Tao, "Gender Recognition Based on Local Body Motions," in *Proceedings of International Conference on Systems, Man, and Cybernetics*, Montreal, Quebec, pp. 3881-3886, Oct. 2007.
- [4] X. Li, S. J. Maybank, S. Yan, D. Tao, and D. Xu, "Gait components and Their Application to Gender Recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 38, no. 2, pp.145-155, Mar. 2008.
- [5] J. Little and J. Boyd, "Recognizing People by their Gait: The Shape of Motion," *Quarterly Journal*, the MIT Press, vol. 1, no. 2, Winter 1988.
- [6] L. Lee and W. Grimson, "Gait Analysis for Recognition and Classification," in *Proceedings of International Conference on Automatic Face and Gesture Recognition*, Washington, DC, pp. 148- 155, May 2002.
- [7] H. Ng, W. H. Tan, H. L. Tong, J. Abdullah, and R. Komiya, "Extraction of Human Gait Feature from Enhanced Human Silhouette Images," in *Proceedings of International Conference on Signal and Image Processing Applications*, Kuala Lumpur, pp. 425-430, Nov. 2009.

- [8] H. Ng, H. L. Tong, W. H. Tan, T. V. Yap, and J. Abdullah, "Gait Classification with Different Covariate Factors," in *Proceedings of International Conference on Computer Application and Industrial Electronics*, Kuala Lumpur, pp. 463-467, Dec. 2010.
- [9] C. Nandini, P. Mukhopadhyay, T. Tanmay, S. K. Ranjan, and S. Roy, "An Efficient Human Identification Using Gait Analysis," *International Journal of Research and Reviews in Computing Engineering (IJRRCE)*, vol.1, no. 2, June 2011.
- [10] B. Lee, S. Hong, H. Lee, and E. Kim, "Regularized Eigenspace - Based Gait Recognition System for Human Identification," in *Proceedings of International Conference on Industrial Electronics and Applications (ICIEA)*, Beijing, pp. 1966-1970, June 2011.
- [11] A. F. Bobick and J. W. Davis, "The Recognition of Human Movement Using Temporal Templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 3, pp. 257-267, Mar. 2001.
- [12] J. Han and B. Bhanu, "Individual Recognition Using Gait Energy Image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 2, pp. 316-322, Feb. 2006.
- [13] C. Harris and M. Stephens, "A Combined Corner and Edge Detector," *Alvey Vision Conference*, pages 147-151, 1988.
- [14] H.P. Moravec, "Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover," PhD thesis, Stanford University, Sept. 1980. (published as Robot Rover Visual Navigation. Ann Arbor, MI : UMI Research Press, 1981.)

- [15] J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Communications of the ACM*, vol. 18, no. 9, pp. 509-517, Sep. 1975.
- [16] B. DeCann and A. Ross, "Gait Curves for Human Recognition, Backpack Detection and Silhouette Correction in a Nighttime Environment," in *Proceedings of SPIE Conference on Biometric Technology for Human Identification VII*, Orlando, USA, Apr. 2010.
- [17] C. Wang, J. Zhang, L. Wang, J. Pu, and X. Yuan, "Human Identification Using Temporal Information Preserving Gait Template," *IEEE Transactions on Pattern Analysis and Machine Intelligence*(Early Access Articles), vol. pp, no. 99, pp. 1, Dec. 2011.
- [18] D. Zhang and G. Lu, "A Comparative Study on Shape Retrieval Using Fourier Descriptors with Different Shape Signatures," in *Proceedings of the International Conference on Multimedia and Distance Education*, Fargo, North Dakota, USA, pp. 1-9, 2001.
- [19] 顏智俞(2010)。植基於 *Harris-Laplace* 特徵點檢測且可抵抗幾何攻擊之浮水印技術。碩士論文，朝陽科技大學，台中市。
- [20] 尤善臻、曾鈺娥、程龍君、謝宛陵、陳明玟、吳錦生(2010)。女性美的幾何學原理。美容科技學刊，第 7 卷第 2 期，211 ~ 227 頁。