

私立東海大學資訊工程與科學研究所

碩士論文

指導教授：周忠信

手寫數位簽名的生物特徵研究

**The Study of Biometrics for Digital Handwriting**



研究生：詹焯然

中華民國九十五年三月二十二日

## 摘要

本論文提出一個結合質心向量、速度、軌跡方向改變以及簽名大小-時間比例值等生物資訊的手寫數位簽名相似度計算方法。利用生物特徵相似度，此方法可以用來防止現行數位憑證認證中，數位憑證被竊取或被誤用的情況發生。與其他生物資訊應用相比，手寫簽名生物資訊是一個行為上的生物特徵，因此，即使複製簽名軌跡，但沒有即時簽名發生，也將會被分辨出來。而更重要的是，沒有合法簽名者的同意，這些生物特徵也很難被獲得。在本研究中，當簽名的相似度在 80%以上時，我們即可認為這個擁有數位憑證的使用者就是合法的使用者。除此之外，此系統同時也很容易被實作或建置成網路上的應用系統。

關鍵字：認證、生物特徵、手寫簽名、相似度

## **Abstract**

In this research, by combining the centroid, the speed, the stroke direction change and the size-time ratio of a handwriting trajectory as the biometrics of a person's handwriting signature, an economic method to calculate the biometrics similarity between two handwriting signatures is proposed. The biometrics similarity can be used to prevent the situations such that either the digital certificates are stolen and misused or the users are commandeered illegally. To compare with other biometrics solutions, handwriting signature biometrics is a better candidate to be used in the Internet world since it is a behavior type biometrics and therefore duplicated signature without similar behavior is distinguishable and most importantly it can not be obtained without the person's freewill. The experiments of our solution show that when digital certificate is verified and the returned biometrics similarity is higher than 80% it almost always can guarantee that the person is really own the certificate. The method proposed in this paper can be easily implemented and deployed for the Internet applications.

**Keywords:** Authentication, biometrics, handwriting signature, similarity

# 目錄

摘要.....	i
Abstract.....	ii
目錄.....	iii
表目錄.....	v
圖目錄.....	vi
第一章 緒論.....	7
1.1 研究動機.....	7
1.2 章節架構.....	9
第二章 相關背景技術.....	10
2.1 手寫簽名在實作目的上的分類.....	10
2.2 手寫的生物特徵應用系統.....	11
2.3 簽名辨識的背景技術.....	13
2.4 關鍵點.....	19
第三章 簽名驗證方法.....	22
3.1 資料輸入.....	22
3.2 正規化.....	23
3.3 尋找關鍵點.....	24
3.4 對應兩簽名的關鍵點.....	24
3.5 求得特徵值.....	26
3.6 整體相似度.....	31

第四章 實驗與討論.....	35
4.1 系統之實作.....	35
4.2 實驗.....	38
4.3 總結.....	42
第五章 結論與未來方向.....	43
5.1 結論.....	43
5.2 未來方向.....	43
參考文獻.....	45

## 表目錄

表 3.1 時間-簽名大小比例值之影響系數 .....	32
表 3.2 方向改變向量之影響系數.....	32
表 4.1 真實簽名 1.....	39
表 4.2 抄襲簽名 1.....	39
表 4.3 真實簽名 2.....	40
表 4.4 抄襲簽名 2.....	40
表 4.5 軌跡方向改變所造成之相似度變化.....	40
表 4.6 簽名大小-時間比例值不同所造成之相似度變化 .....	41
表 4.7 簽名形狀不甚相同所造成之相似度變化.....	42

## 圖目錄

圖 2.1 DP-Matching 中每個 grid 的計算 .....	16
圖 2.2 對簽名重新取樣並求其基線.....	18
圖 2.3 判斷各點的領域.....	20
圖 3.1 簽名驗證流程.....	22
圖 3.2 所記錄的資料結構.....	23
圖 3.3 關鍵點對應演算法.....	26
圖 3.4 簽名大小-時間比例演算法 .....	29
圖 3.5 相似的關鍵點，不同的軌跡。 .....	30
圖 3.6 $S_c = 0.5$ 時之調整曲線圖 .....	33
圖 3.7 $S_c < 0.5$ 時之調整曲線圖 .....	34
圖 4.1 系統之簽名輸入畫面.....	35
圖 4.2 求得關鍵點.....	36
圖 4.3 關鍵點對應以及各相似度.....	37
圖 4.4 真實簽名與抄襲簽名 1.....	38
圖 4.5 真實簽名與抄襲簽名 2.....	39
圖 4.6 軌跡方向改變不同之簽名.....	40
圖 4.7 簽名大小-時間比例值不同之簽名 .....	41
圖 4.8 簽名形狀不甚相同之簽名.....	41

# 第一章 緒論

## 1.1 研究動機

近年來，隨著網路的高度發展，網路應用也相對越來越多。這些應用包括電子商務與線上交易[5][6]等。隨著網路應用的增多，如何確保線上交易的安全與正確乃成為了重要的研究方向之一。其中，使用者的身份認證(authentication)更是一項重要的議題[12]。

在網路上確認使用者的身份，最簡單的方法是利用密碼為基礎(password-based)的方法[20]。然而，此種方法卻有容易遺忘、容易混亂、容易被猜測與竊取等重大缺點。除了以密碼為基礎的方法外，現行在網路上最主要的使用者身份認證方法就是利用數位憑證(digital certificate)[41]。數位憑證又稱為電子憑證，是由憑證授權單位發出。一般而言，數位憑證內含個人的公開金鑰(public key)、公開金鑰效期、憑證授權單位、憑證序號以及個人之數位簽章等。數位憑證的主要目的是用以確認此人身分。此外，數位憑證經過編碼(encrypted)，可用在網路交易安全需求上，使網路交易服務達到資料之隱密性、身份確認及不可否認性等安全需求。然而，此種方式雖比以密碼為基礎的方法方便且安全，但亦有其缺點。此缺點即是無法預防數位憑證被竊取或是被誤用[17]。

因此，一個能夠確認數位憑證身份的方法，在網路世界是必要的技術之一。其中，將數位憑證與生物特徵(biometrics)資訊融合，即是方向之一。生物特徵是一種生理上(physiological)或行為上(behavioral)特有的資訊，由於每個人所擁有的生物特徵都不盡相同，因此，生物特徵的應用，是近年來的熱門研究領域之一[7]。

生物特徵大致上可以分成下列兩種分類：

- 生理上：

生理上的生物特徵包括了指紋比對[33]、虹膜與視網膜掃描[13][45]、掌紋比對[36]等。這種生理方面的生物特徵是屬於靜態的，不容易隨著時間而變動。



■ 行為上：

行為上的生物特徵則包括了聲紋分析[11]以及資料輸入風格(input style)[16]等，包括簽名辨識以及鍵盤和滑鼠的輸入習慣等。此類的生物特徵屬於動態行為，對同一個使用者而言，每次輸入後所測量結果可能都不一樣，但是其相似度卻可作為辨識用。而更重要的是，這種行為上的生物特徵很難被事先複製。

要開發一個適合於網路上的生物特徵應用系統，有幾項需求很重要：

- 生物特徵值最好要為動態的，且事先複製沒有用處。
- 收集生物特徵的裝置很容易被取得。
- 此生物特徵系統要能容易實作、容易建置、且容易使用。
- 系統所需要的計算資源越少越好。

綜合以上所述，在眾多的生物特徵中，利用手寫簽名來產生生物特徵值可以說是相當符合上述需求的方法，因為手寫簽名並不需要複雜且昂貴的掃描裝置，只需要一個手寫板即可。因此使用者可以在任何時間、任何地點提供他/她的簽名來進行簽名相似的比較。

本研究開發一個利用分析生物資訊的手寫簽名相似度計算方法來輔助現行的使用者認證方式。使用者利用手寫板輸入自己的簽名，此系統將由簽名中分析獲得所需要的生物資訊，這些生物資訊包括了質心向量(centroid)、速度(speed)、軌跡方向改變之變化(stroke direction change)、以及簽名之大小-時間比例值(size-time ratio)等。藉由這些生物資訊，配合我們所提出生物特徵相似度(biometrics similarity)計算方法，可以得到兩個簽名間的相似度，作為辨識是否為真正合法使用者(valid user)的依據。

利用本研究成果，並配合現行的數位憑證認證，關於數位憑證被竊取或被誤用而導致身份認證出現漏洞的問題也得以解決。因為手寫簽名乃是行為上的動態資訊，其生物特徵很難被模仿且複製的。除此之外，本研究所提出的方法，是一個低成本(low cost)的演算法，即使在配備較低階的環境上，亦可輕易的實作。

## 1.2 章節架構

本論文內容分配如下，第二章介紹關於手寫簽名之相關背景技術；第三章介紹本系統的詳細作法；第四章為本系統相關實驗與討論，第五章為結論以及未來方向。

## 第二章 相關背景技術

本章節將介紹手寫簽名的各種應用分類，以及針對手寫簽名辨識中，相關研究所應用的方法。

### 2.1 手寫簽名在實作目的上的分類

手寫簽名在實作的目的上可以分成幾種分類[32]：

#### (1) 簽名辨識與指認 (Signature verification/identification)：

簽名的辨識與指認著重在對於同樣內容的簽名，如何利用簽名本身的形狀以及其他特徵當作線索去辨識以及指認。其中，簽名辨識與簽名指認的差別在於，簽名辨識乃是一對一 (one-to-one)的，在系統設計上為只需要將一個使用者輸入的輸入簽名(input signature)與儲存在系統資料庫中的樣本簽名 (template signature)比對，判對此輸入簽名是否為真即可。而簽名指認乃是多對一的 (multi-to-one)，系統的設計需求為需要在許多相同內容的輸入簽名中，找出和儲存在系統資料庫中的樣本簽名最像或是和樣本簽名擁有相同的特徵的那些簽名。

#### (2) 作者辨識與指認 (Writer verification/identification)：

每個人在用筆寫文章時，除了字體本身的字跡不同之外，還有個人的書寫慣以及排版方式也都不盡相同，而作者辨識與指認即是利用作者的字跡，書寫習慣和排版方式的不同來對不同內容的文字做辨識以及指認。也就是說，系統會針對不同的手寫文件，來判斷作者為何許人。其中，作者辨識和作者指認的不同處在於，作者辨識為一對一的，系統的設計為將一份文件輸入系統後，經過特徵分析後，與在系統資料庫中的作者文件相比，辨識此份文件是否為該作者所作。而作者指認是多對一的。其中，依功能的需求，又可以分成兩種。第一種的系統設計為在許多不同內容的文件，經過特徵分析，將之和系統資料庫中的作者文件相比較，找出同一個作者所寫的所有文件。第二種的系統設計為，在許多不同的文件中，經過特徵分析，將這些文件依照作者加以分類。在[3][34][46]中，即為相關之研究與探討。

除此之外，手寫簽名的系統，還可因為輸入方式的不同而區分成兩種：

(1) Offline system：

Offline system 指的是使用者輸入自己的簽名，而系統在比對時，只考慮圖形，也就是形狀的部份的特徵。此類系統由於只考慮靜態的特徵，例如簽名中相關的角度、長度等。所以對以辨識為目的的系統而言，其可靠度並非很高，因為模仿者只需要想辦法將簽名的形狀模仿的很像，即使是慢慢的抄襲，都有可能通過系統的判斷。因此，此種 Offline system 在現今網路這種無法見到對方的環境下，所能夠發揮的功能並不大。

(2) Online system：

在 Online system 中，針對使用者所輸入簽名，系統在比對時，除了考慮靜態特徵，也就是形狀的部份之外，還會考慮到動態的特徵。要考慮到動態的特徵，就要加入時間因素，也就是使用者在輸入簽名時，會同時記錄筆觸所走到的當時時間。從這個時間因素中，系統就可以得到許多的動態特徵，所謂的動態特徵包括了速度、加速度等。利用這些動態特徵，對以辨識為目的的系統而言，將可以更準確的判斷手寫文件的真與假，因為動態特徵是很難模仿的。因此，Online system 是非常適合在現今的網路環境下的。

## 2.2 手寫的生物特徵應用系統

在 2.1 節中我們介紹了手寫簽名的分類，而架構在這樣的分類上，目前也有許多手寫的生物特徵應用系統已經被開發出來，本節即是要介紹這些生物特徵應用系統。

手寫生物特徵應用系統可以分成下面幾種[32]：

■ 文件辨認(Document Recognition)：

文件辨認是一個將手寫文件的內容轉換到電腦上的字元的過程，手寫文件可以是 online 輸入的也可以是 offline 的。此類系統在處理輸入的手寫文件時，會利用一個使用者的特定分類器(user-specified classifier)來幫忙，此分類器可以設

定是要針對字元(character)或是字(word)來做處理。在 offline 的辨認裡，系統利用分類器配合作者的書寫風格(writing style)來辨識文件裡的字。而 online 的辨認則是在 offline 的基礎上，再加上作者書寫時的時間資訊，能更精準的辨認出作者所書寫的字為何。

■ 法庭上的應用(Forensic Applications)：

手寫系統在法庭上的應用多半為判斷某文件是由何人所作，或是從許多文件中找出某人所作的那一份文件。這種為了法律問題而設計的系統，由於來源文件大多是事後才會提出，所以系統型態幾乎都是屬於 offline 的系統。此外，為了要達成辨識的目的，所以此種系統需要針對特定的人士，獲取他們的手寫文件當作範本，在經過特徵比對之後，方可判斷該文件的書寫方式是否與某人相同。但此種系統多半只是為輔助的工具，最後還是需要由人來做最後的決定。

■ 便利性的應用(Convenience)：

手寫系統也可以應用在 user-aware system 上以增加使用者的便利。例如，對一個公用的 tablet PC 而言，使用者可以利用內建的(embedded)手寫辨識系統來當作使用者登入時的依據。根據使用者所輸入的任何手寫文字，系統可以藉由作者指認(writer identification)的功能，判斷使用者是何人，並且自動載入使用者個人化的設定。

■ 存取控制(Access Control)：

存取控制的應用系統，其功能大多都是讓事前有被授權過的合法(valid)人員，能夠在登入時，得到他們應該有的權限。而此類的系統，大多屬於在本地端(local)的簽名辨識，由於是在本地端，所以並不需要特別的機制來保護與處理資料傳送的問題。而系統在使用者簽名後，經過計算與比對，由系統判斷此使用者是否合法。其特色是精準度要高，但是誤判(false reject)的機率要低，而且在必要的時候甚至需要兩次以上的簽名再確認，以增加系統辨識的正確性。

■ 電子簽名(Electronic Signature)：

在以往，簽名即是一個人的合法代表，當某人在一份文件上簽下自己的名字之後，即是表示此人已經看過這份文件，或是會對此份文件上事情負責。然而，

在現今快速發展的社會中，傳統公文往返的方式非常的曠日費時，隨著網路的發展，許多的文件已經改用電子化的方式來傳遞，因此，應用在網路上的電子化簽名也隨之應運而生。除了系統端的簽名辨識的功能之外，要使電子簽名能夠有和傳統簽名一樣的效力，就必須保證此簽名在網路上傳遞的同時，不會遭到攻擊或竄改，因此 PKI 的安全機制多會被搭配應用於這之上。另外，由於簽名者是在網路的另一端，要保證此傳送過來的簽名為真，就必須同時確保簽名者是即時 (real time) 所簽的，所以此種應用系統通常會利用時間資訊來保證簽名的即時性，也正因為如此，此類系統多為 online 的應用系統。

## 2.3 簽名辨識的背景技術

在上一節中，我們介紹了手寫簽名應用系統的分類，而在本節裡，將要介紹的是簽名辨識的背景技術。在手寫簽名辨識的領域裡，許多利用生物特徵的方法已經被開發出來[8][9]，在大部份的研究裡，作者們使用的這些生物特徵包括了簽名的外型、簽名的速度變化、簽名的加速度變化、簽名的相對距離、簽名時的壓力變化以及在簽名時，筆的傾斜角度等。利用了上述的生物特徵，要完成線上手寫簽名的辨識，最常被使用的方法如下[30]：

### ■ Hidden Markov Models(HMM)：

Hidden Markov Models 隱藏式馬可夫模型，是一個利用了機率與統計的原理，來描述訊號所具有的特徵。其最早是在 1907 年，由馬可夫(Markov)提出馬可夫鏈(Markov chain)，之後經由 Baum-Welch 提出馬可夫參數的最佳化後，HMM 成為了一個相當成熟的模式。由於其穩定性，因此它被廣泛的應用在許多的領域上。

HMM 是一個有限狀態機模型(finite state machine)，它主要在表示不論是離散的(discrete)或是連續的(continuous)觀察下，一個流程通過一個有限狀態機時的各種情況。而整個流程即可藉由個種狀態轉移間的機率連結而表示出來。整合這些機率，HMM 的數學模型 可以由下式表示：

$$\lambda = (A, B, \pi) \quad (2.1)$$

其中， $A$  為狀態轉移的機率， $B$  為在某一狀態下產生特定值的機率， $\pi$  為初始狀態的機率。當知道一個 HMM 的參數  $\lambda$ ，對於某個觀察序列  $O$ ，我們可以計算  $P(O|\lambda)$  來表示  $O$  從  $\lambda$  發生的機率。簡單的說，也就是整個流程可以從第一個狀態成功的依序轉移到最後一個狀態，這樣的情況發生的機率。

HMM 的一個重大的問題在於它須要大量的運算。這對狀態數很多的序列而言，其計算非常之沉重。然而即便如此，HMM 依然是個成熟穩定的技術，也有許多的研究是利用 HMM 來做手寫簽名的辨識。在[27]中，作者將手寫簽名重新取樣，並經過篩選以及去除雜訊的步驟後，對於每個新的取樣點都求出它到下一個點的轉折角度，以及區域性的筆劃相對長度，將之當作 HMM 的狀態序列，利用 HMM 來判斷簽名的真實性。在[40][47]裡，作者分別利用簽名的相對角度，相對速度，以及壓力變化，當作其狀態序列，利用 HMM 判斷簽名的真假。

■ Neural Network :

類神經網路，是一種平行計算系統，藉由人工神經元來模擬生物神經網路的能力。類神經網路的架構可以分成三層，分別為輸入層(input layer)、隱藏層(hidden layer)和輸出層(output layer)。其中輸入層為接收資訊的窗口，而隱藏層為主要運算的部份，輸出層則為輸出結果之處。

一個人工神經元包括了節點(node)、臨界值(threshold)、轉換函式(transfer function)以及輸入(input)輸出(output)值。對一個運算神經元而言，輸入值  $X_i$  在輸入神經元之後，藉由各自的權重(weight)進行加總之後和臨界值比對，以決定訊號是要被傳送或是被抑制。透過臨界值的過濾，要被傳送的訊號會被傳送至轉換函式中進行轉換計算，最後產出輸入值。整個類神經網路的神經元，可以用下列數學式表示：

$$net_j = \sum_i^N w_{ji}x_i + b_j \quad (2.2)$$

$$y_j = F (net_j) \quad (2.3)$$

其中,  $y_j$  表示第  $j$  個神經元的輸出值;  $net_j$  為輸入值  $x_i$ 、權重值  $w_{ji}$  與偏移值(bias)  $b_j$  的總和。

類神經網路最重要的就是它的學習功能, 通常網路學習演算法包括正向傳播(forward propagation)與反向傳播(backward propagation)。其基本的流程為, 輸入訊號由輸入層經過相關的權重處理之後, 傳遞至隱藏層。隱藏層透過轉換函式的處理, 得到一個輸出值傳遞到輸出層。但若輸出層未能夠得到滿意的結果, 就進行反向傳播, 並且修正其權重, 使得誤差可以縮至最小, 以此方式達到學習的目的。

和 HMM 一樣, 類神經網路也需要大量的運算, 這同樣也是一個很大的問題。但, 透過類神經網路的學習功能, 手寫簽名辨識依舊可以應用在其中。在[7]裡, 作者使用了簽名中每個點的絕對距離以及每個點到下一個點的角度, 配合了 Ranking Approach 與 Genetic Approach 這兩個演算法所計算出來的值分別當作輸入訊號, 在簽名者輸入數個樣本簽名的同時, 傳入三個不同的類神經網路中學習。當學習完成後, 便可以依照學習結果, 來判斷在驗證過程中所輸入的簽名為真或為假。

#### ■ Dynamic Programming Matching (DP-Matching) :

DP-Matching, 或稱 Dynamic Time Warping (DTW), 是一個用來對兩個不同長度的序列做配對的演算法, 它最原始是被應用於語音辨識(speech recognition)上, 後來才被廣泛的應用在簽名辨識之中。使用此方法的原因在於, 由於每次的簽名都不會一樣, 所以每一次簽名, 所得到的取樣點的數目都可能和樣本簽名的取樣點的數目不同, 因而要比較這些不同長度的序列, DP-Matching 是一個選擇。絕大多數利用 DP-Matching 來做簽名辨識的系統, 大多會有下列三個步驟 [25][26] :

##### (i) 資料取得以及前置處理 :

資料的取得(data acquisition)通常是利用手寫板或 pda 等可以用固定取樣速率來取得使用者簽名軌跡的裝置來得到使用者的簽名。而前置處理



(preprocess)則是需要配合系統後續的辨識工作而有所不同，然而最常見的前置處理工作則是簽名的正規化(normalization)，目的是用來使簽名能在同一個維度下比較。

(ii) 特徵解析：

此步驟即是對簽名中所要參與的所有點，依照系統設計者的需求，分別求出它們的生物特徵。接著利用 DP-Matching 對兩簽名求出最佳配對，並求出兩簽名間的距離(distance)或是利用特定方法求出其相似度。

(iii) 真假判斷：

在求出兩簽名間的相似度之後，便可以依照系統所設計的規則，來判斷此簽名的真假。

藉由 DP-Matching, 對兩個不同長度的向量  $A=\{a_i \mid i = 1 \text{ to } I\}$ ,  $B=\{b_j \mid j = 1 \text{ to } J\}$  而言，要找出此兩向量中所對應(corresponding)的元素組合，我們可以將此兩向量依序排在棋盤的兩端，接著利用公式 2.4，找出在此  $I \times J$  的棋盤中，每一個 grid 裡的配對值，即可求出其路徑，也就是所須要的元素配對組合。圖 2.3 為此公式之圖解意義。

$$d_{i,j} = \max \begin{bmatrix} d_{i-1,j} \\ d_{i-1,j-1} + p \\ d_{i,j-1} \end{bmatrix}, 1 \leq i \leq I, 1 \leq j \leq J \quad (2.4)$$

其中 p 為預設對應到時的加分(bonus)或是沒有對應到時的扣分(penalty)

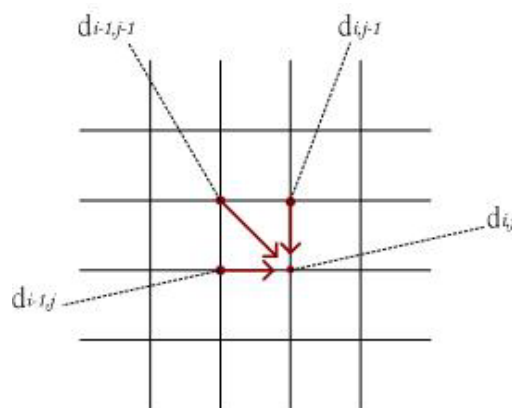


圖 2.1 DP-Matching 中每個 grid 的計算

有許多的研究都是利用 DP-Matching 來幫助配對。而在這些使用 DP-Matching 的手寫辨識研究裡，又可區分成兩種應用，第一種即是傳統的 DP-Matching 做法，也就是在經過前置處理之後，對所有要參與計算的點，皆求出其生物特徵之值，接著利用 DP-Matching 直接求出兩個簽名之間的差異，再配合設計者所訂定之規則來判斷簽名的真假。在[10][14][18][19][38][42]中，系統在簽名的過程裡，作者的前置作業是利用固定的時間區間來對簽名重新取樣，得到新的取樣點，而它所利用的生物特徵為每一個新取樣點的 x 方向與 y 方向的絕對速度  $v_x, v_y$  和絕對加速度  $a_x, a_y$ ，接著利用 DP-Matching 找出其配對並求出兩簽名之間的距離，最後藉由事先訂定的臨界值(threshold value)來判斷簽名的真假。

而在[24]裡，作者的前置作業乃是利用影像處理中的高斯函數對整個簽名影像做平滑化(smooth)的動作，以消除可能的雜訊。

$$g(\lambda) = \frac{\exp(-\lambda^2 / (2\sigma)^2)}{\int_{-L}^L \exp(-\gamma^2 / (2\sigma)^2) d\gamma}, -L \leq \lambda \leq L \quad (2.5)$$

其中， $L$  為 sliding window 大小， $L/2$

在前置作業結束後，作者所利用的是點的 x 方向之距離差異、在 x 方向的相對位置、與前一點的角度差來表示形狀方面的資訊；另外作者求出了每個點在 x 方向的絕對速度  $v_x$ ，x 方向的平均速度  $v_{mean}$ ，利用絕對速度與平均速度的差來表示動態的生物特徵。同樣的，DP-Matching 被用來配對並計算其差異。最後，作者同樣是利用事先設定好的臨界值來區分簽名的真偽。

上述例子表示了典型不利用特殊工具而由簽名中直接找出生物特徵值並配合 DP-Matching 使用的例子。而在[15][21][22][29][35]中，作者則是利用了特殊的裝置來做簽名辨識。作者使用的是可以記錄簽名者握筆資訊以及書寫時壓力變化的裝置來記錄使用者的簽名。在前置過程中，系統對整個簽名的每個點座標皆做正規化，並求出其質心。而在生物特徵方面，這些研究分別用了簽名在每個點的壓力值(pressure)、相對位置、握筆時的傾斜度(inclination)、與質心的角度差、與質心的相對距離 以及在每個點的絕對速度等生物特徵來配合 DP-Matching 的

運算, DP-Matching 針對所有生物特徵  $B=\{B_1, B_2, \dots, B_m\}$  都可以得到差異值  $b=\{b_1, b_2, \dots, b_m\}$  , 接著作者利用了線性組合的方式求出整體差異值  $D$  :

$$D = \alpha_1 b_1 + \alpha_2 b_2 + \dots + \alpha_m b_m, \quad \alpha_1 + \alpha_2 + \dots + \alpha_m = 1 \quad (2.6)$$

相同的, 若是此差異值  $D$  超出了所訂的臨界值範圍, 則輸入的簽名就會被判定為假的簽名。

第二種應用則是 DP-Matching 搭配其他的驗證方法來做簽名辨識。在[23]裡, 作者利用 DP-Matching 搭配類神經網路來完成簽名辨識。其中, 在前置作業中, 作者對簽名重新取樣, 將原本的簽名軌跡, 重新分成等距的 256 個點, 並且找出此簽名的平均基線(base line)。

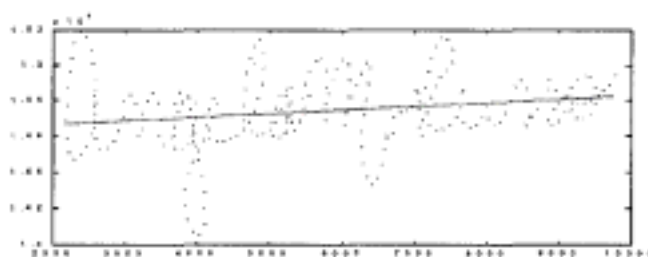


圖 2.2 對簽名重新取樣並求其基線

利用每個新取樣點到平均基線的垂直距離, 配合 DP-Matching, 可以找出兩個簽名之間的點之配對關係。在生物特徵解析的部份, 對於每個點, 作者利用了 AR Model[28]的計算來表示其特徵值。而最後的辨識部份, 作者配合了類神經網路來學習輸入的資料並加以比對, 以得到最後的結果。

DP-Matching 是一個方便的演算法來幫助配對, 然而它有一個缺點就是若是要被配對的那一方, 有兩個非常靠近的選擇, 而且正確選擇就在這兩個其中, 那麼就有可能會有配對錯誤的情況發生。這是因為由於這兩個選擇非常的靠近, 因此皆會小於其臨界值, 那麼最後的結果將有可能是順序在前的那個選擇被配對, 因而若是正確的是後面那個選擇, 則就會發生配對錯誤的情況。

## ■ 其他：

除了以上所提的方法之外，也有許多學者提出了不同的做法。在[31][39]中，作者對每個簽名，求出它的生物特徵後，利用了機率與統計的方式，求出兩個簽名相似的程度。而在[4]中，作者對手寫數位簽名，利用影像處理的方式，去分析其轉折處的相素(pixel)，來判斷此簽名是否是偽造的。在[37][43]裡，作者強化了樣本(template)的建立，由使用者所輸入的簽名中，建立了一個最能適當描述使用者簽名的樣本參考值，使得當需要簽名認證時，可以使誤差達最小。

## 2.4 關鍵點

所謂的關鍵點，指的是一個軌跡中，特別突出的那些點，例如對一個英文字母“v”而言，最突出的地方就是它的尖端處。

對同一個人的同一個簽名而言，此簽名的關鍵點應該在大多數的時間中都是相同的。因此，從一個簽名的關鍵點中，我們可以得到許多有用的資訊，例如簽名軌跡大致上的形狀、簽名的速度分配等。

找尋關鍵點，已有一些演算法被發展出來[1][2][44]，在此我們採用 Jean-Jules Brault 和 Rejean Plamondon 在 1993 年所提出的演算法[2]。此演算法主要的精神在於，一個點要成為關鍵點，以它為中心的鄰點(neighboring points)必定給它許多貢獻(contribution)，使此點得以形成關鍵點，因此，只要求出鄰點所帶給它的貢獻值，即可判斷此點是否為關鍵點。

此方法分成下列幾個步驟來尋找關鍵點：

1. 首先要做前置工作(preprocess)，將原始簽名軌跡  $P = \{p_i(x_i, y_i, t_i, e) \mid 1 \leq i \leq n, n \text{ 為所記錄到的點數}\}$ ，劃分為等距離的線段，其目的在於過濾原始簽名軌跡中的雜訊(noise)。等分後，可以得到新的點集合為  $F = \{f_i(x_i, y_i, t_i, e) \mid 1 \leq i \leq r, r \text{ 是等分後的新點的數目}\}$ 。
2. 接著，對每一個重新等分過的點，找出它們所擁有的領域(domain)。對某點

而言，所謂的領域  $D_i$ ，指的是在此範圍內的其他鄰點會對此點產生貢獻值，若超出此範圍，就對此點無貢獻。每一個點所擁有的領域皆不一樣。

判別領域的方法如下：

- (1) 對於每個點  $f_i$ ，考慮兩個角度  $\theta_f(f_i, f_{i+n})$  和  $\theta_b(f_i, f_{i-n})$ ，其中， $\theta_f(f_i, f_{i+n})$  為第  $f_i$  點與第  $f_{i+n}$  點和第  $f_{i-n}$  點的中點所形成的向量和第  $f_{i+n-1}$  點及第  $f_{i+n}$  點所形成的向量，此兩個向量的夾角。

$\theta_b(f_i, f_{i-n})$  為第  $f_i$  點與第  $f_{i+n}$  點和第  $f_{i-n}$  點的中點所形成的向量和第  $f_{i-n+1}$  點及第  $f_{i-n}$  點所形成的向量，此兩個向量的夾角。

- (2) 找出符合下列條件的最大  $n$  值，則  $n$  即為  $f_i$  的領域：

$$|\theta_f(f_i, f_{i+n})| \text{ 和 } |\theta_b(f_i, f_{i-n})| \text{ 同時小於 } \theta_{max}$$

其中  $\theta_{max}$  為一個預先定義好的臨界值(threshold)， $0 < \theta_{max} < \pi/2$

在本研究中，我們設定的  $\theta_{max} = 3\pi/8$ 。

- (3) 在找出第  $f_i$  點的領域之後，由下面的公式 2.7 計算出此領域中的其他鄰點所帶給此點的貢獻值  $G(i)$ ：

$$G(i) = \sum_{n=1}^{D_i} \text{Cos}(\theta_f(f_i, f_{i+n})) * \text{Cos}(\theta_b(f_i, f_{i-n})) \quad (2.7)$$

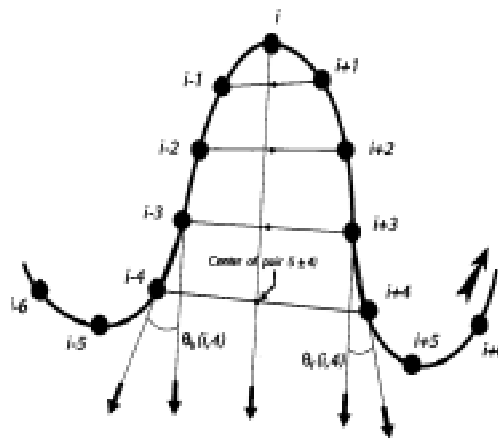


圖 2.3 判斷各點的領域

3. 在算出軌跡中每一個新的等分點的  $G(i)$  值之後，即可開始定出(Locate)關鍵點。由於大部份的  $G(i)$  值都為 0，只有在有轉折處附近的點的  $G(i)$  值才不會為 0，因此，每一群連續不為 0 即是表是一個轉折處，我們只須要在這些不為 0 的值中找出最大的那一個，它就是我們所需要的關鍵點了。

另外，除了由此演算法所找出的關鍵點之外，由於整個簽名中的每一個筆劃(stroke)的第一個點以及最後一個點也是簽名中重要的部份，因此，我們也強制它們為關鍵點。

藉由此方法，我們可以找出在簽名軌跡中的關鍵點，以供我們在計算生物特徵值時使用。

## 第三章 簽名驗證方法

本章節將詳細介紹本系統如何分析生物特徵資訊，以及如何計算兩個簽名之間相似度。

首先，不論是輸入簽名(input signature) 或是樣本簽名(template signature)，資料將一邊輸入且一邊建立起其資料結構。第二步，由於每次簽名的維度大小(dimension)都不同，所以必須對簽名做正規化(normalize)的動作。接著，找出此簽名中的關鍵點(significant point)，並對輸入簽名與樣本簽名做關鍵點的配對。最後，則是針對各組配對好的關鍵點，求出其特徵值(feature data)，並相比較，求出其相似度。

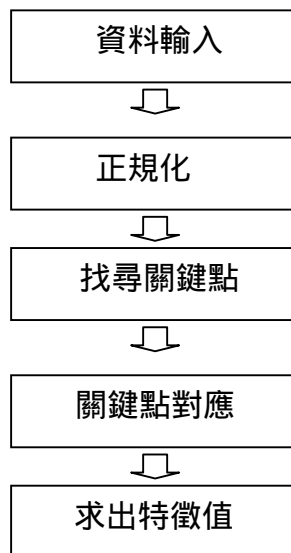


圖 3.1 簽名驗證流程

### 3.1 資料輸入

使用者藉由手寫板來輸入簽名。透過手寫板，將記錄使用者在簽名時所走過的軌跡以及時間  $(x,y; t; e)$ ，其中， $x$  是手寫板所記錄到的點的  $x$  座標， $y$  是手寫板所記錄到的點的  $y$  座標， $t$  是手寫板所記錄到的點的時間戳記(timestamp)。除了所記錄到的點的座標以及時間資訊之外，系統也將同時記錄這個點的型態(type)

$e$ 。所謂的型態，指的是這個點是屬於一個筆劃的端點(起點或終點)，或是屬於一個筆劃內的連續點。在結束資料輸入後，可以得到所記錄到的點集合為

$$P = \{p_i(x_i, y_i, t_i; e_i) \mid 1 \leq i \leq n, n \text{ is the number of recorded points}\}.$$

所記錄的資料結構：

$(x,y)$	Timestamp $t$	Type $e$
---------	---------------	----------

圖 3.2 所記錄的資料結構

### 3.2 正規化

每個人每次的簽名都是不同的，可能長相不同也有可能大小不同，從中得到的生物特徵資訊也就會不同，因此要比較不同簽名的生物資訊，就必須要將它們放在同樣的維度上才能比較，所以，正規化就是必要的步驟。

對於一個簽名，我們要將之針對兩個項目做正規化：座標以及時間。要正規化，需先求得下面幾個值： $x_{max}$ ， $y_{max}$  分別表示為所有 x 座標中的最大值以及所有 y 座標中的最大值； $x_{min}$ ， $y_{min}$  分別表示為所有 x 座標中的最小值以及所有 y 座標中的最小值。對此簽名的所有記錄到的點集合  $P = \{p_i(x_i, y_i; t_i; e_i) \mid 1 \leq i \leq n, n \text{ is the number of recorded points}\}$ 而言，正規化後的點集合  $P^N = \{p_i^N(x_i^N, y_i^N; t_i^N; e_i) \mid 1 \leq i \leq n, n \text{ is the number of recorded points}\}$ 可以藉由下面的公式得到：

座標：

$$x_i^N = \frac{x_i - x_{min}}{|x_{max} - x_{min}|} \times I \quad (3.1)$$

$$y_i^N = \frac{y_i - y_{min}}{|y_{max} - y_{min}|} \times I \quad (3.2)$$

其中  $I$  為一個常數，設定此常數的原因在於我們希望將原本的簽名縮小到一個足夠比較的範圍即可，若是沒有底限的縮小會導致所有的座標值都落在 0 與 1 之間，且差距極小，造成之後計算時的困難。本研究中，所設定的  $I$  值為 100。



時間：

$$t_i^N = \frac{t_i - t_1}{t_n - t_1} \times J \quad (3.3)$$

其中  $J$  為一個常數，設定此常數的原因在於希望原本的時間資訊正規化到足夠比較的範圍即可，若是沒有底限，會造成所有正規化之後的時間值都在 0 與 1 之間，且其中的差距極小，造成往後計算時的困難。在本研究中，所設定的  $J$  值為 5。

### 3.3 尋找關鍵點

要對兩個正規化後的簽名直接比較，對電腦而言並非一件簡單的事。因此在本研究中，我們不對所有的點做比較，而是選擇比較簽名的關鍵點。簽名的關鍵點乃是一個重要的資訊擷取處，從一個簽名的所有關鍵點中，我們可以得到簽名者在簽名時的許多生物資訊來計算各生物特徵的資料值。

我們利用 2.4 節裡所提到的方法來求得簽名中的所有關鍵點。可以得到關鍵點集合為：

$$S = \{s_i(x_i, y_i; t_i; e_i) \mid 1 \leq i \leq m, m \text{ is the number of recorded points}\}$$

### 3.4 對應兩簽名的關鍵點

這一節要介紹的是如何正確的對應兩簽名的每一個關鍵點。要對應的原因是若是沒有正確的對應到關鍵點，則最後求得的相似度將會與正確的相似度有極大的落差。

由於對同一人簽的相同內容的簽名而言，它們的關鍵點並不會百分之百的都出現在完全一樣的位置，但是一定會在類似的時間點出現在極為類似的地方，因此，要對應關鍵點，便可以利用下列幾個資訊來作對應：與第一個點的距離、與第一個點的角度以及與第一個點的時間差。

對一個簽名而言， $S$  為整個簽名的關鍵點集合：

$$S = \{s_i(x_i, y_i; t_i; e_i) \mid 1 \leq i \leq m, m \text{ is the number of significant points}\}$$

可以藉由下面公式求得所需資訊：

距離：

$$d_i = \sqrt{(x_i - x_1)^2 + (y_i - y_1)^2}, \quad 1 \leq i \leq m \quad (3.4)$$

角度：

$$\theta_i = \tan^{-1}\left(\frac{y_i - y_j}{x_i - x_j}\right), \quad 1 \leq i \leq m \quad (3.5)$$

時間：

$$T_i = t_i - t_1, \quad 1 \leq i \leq m \quad (3.6)$$

由上述公式中，對於要比對的兩個簽名，我們可以整理出其各自的資料集  $H_1 = \{h_{1i}(d_{1i}, \theta_{1i}, T_{1i}) \mid 1 \leq i \leq m_1, m_1 \text{ is the number of significant points}\}$ 、 $H_2 = \{h_{2j}(d_{2j}, \theta_{2j}, T_{2j}) \mid 1 \leq j \leq m_2, m_2 \text{ is the number of significant points}\}$ ，藉由這兩個資料集，即可對兩個簽名的關鍵點來做對應。

本研究提出了一個方法，可以利用這兩個資料集，有效的對應兩個簽名間的關鍵點。在開始關鍵點對應前，由於兩個簽名的第一個關鍵點(也就是簽名軌跡中的第一個點)與最後一個關鍵點(也就是簽名軌跡中的最後一個點)，是為簽名中的代表開始與結束的特殊關鍵點，因此我們預設兩個簽名的第一個與最後一個關鍵點已是互相對應的了。

關鍵點對應的方法概述如下：對於  $H_1$  中的每個元素  $h_{1i}$  (也就是每個關鍵點)，都建立一個時間與角度的搜尋區間(searching frame)  $a$  與  $b$  來找尋在  $H_2$  中和自己最為接近的元素  $h_{2j}$ 。本方法利用多次搜尋的方式來達成我們想要的目標。在第一次的搜尋中，對  $h_{1i}$  元素而言，它的時間搜尋區間為  $(T_{1i}+a)$ ，角度搜尋區間為  $(\theta_{1i}+b)$ 。接著，搜尋  $H_2$  裡是否有任何元素  $h_{2j}$ ，其時間與角度的值同時落在此區間中。若恰有一個，則此元素  $h_{2j}$  即為  $h_{1i}$  之對應。若找到多個元素皆落在此區間，則計算這些元素的距離  $d_{2j}$ ，將之和  $d_{1i}$  相比較，其差異最小者即為我們所要選取的對應點。而假若在此搜尋區間中並沒有找到合適的對應點，則邁入下一次的搜尋，此時搜尋區間將會放大一個單位至  $(T_{1i}+2a)$  以及  $(\theta_{1i}+2b)$ ，接著重覆搜

尋的動作。依此類推直到遞迴次數達到預先訂定的最大搜尋次數  $maxSearchTimes$ 。在經過了  $maxSearchTimes$  次的搜尋後，若依然沒有得到合適的對應點，則我們就可以判定此元素  $h_{1i}$  並沒有適當的元素  $h_{2j}$  來對應。而在  $H_1$  中的所有元素都執行過搜尋之後，我們可以得到  $H_1$  與  $H_2$  這兩個資料集的對應組合，也就是對應到關鍵點組合的集合為：

$$M = \{(P_{1i}, P_{2i}) \mid P_{1i} \in S_1, P_{2i} \in S_2, 1 \leq i \leq k, \text{ where } k \text{ is the number of the mapped significant points}\}$$

```

Let  $maxSearchTimes$  the maximum times to search;
Let  $lastIndex$  the last mapped index in  $H_2$ , initial value is 1;
Let  $P$  the set of possible candidates of mapping points;
Let  $a$  and  $b$  be the mapping thresholds for time and angle respectively;
Let  $time(h_{1i}, h_{2j})$  be the function to return the time difference between  $h_{1i}$  and  $h_{2j}$ ;
Let  $angle(h_{1i}, h_{2j})$  be the function to return the angle difference between  $h_{1i}$  and  $h_{2j}$ ;
Let  $dist(h_{1i}, h_{2j})$  be the function to return the distance between  $h_{1i}$  and  $h_{2j}$ ;
Let  $M$  be the set of the mapped significant point pair;

 $M = \{ \}$ 
for  $i = 1$  to  $m_1$  do
  let  $P = \{ \}$ 
  for  $searchedTimes = 1$  to  $maxSearchTimes$  do
    for  $j = lastIndex$  to  $m_2$  do
      if  $(time(h_{1i}, h_{2j}) < searchedTimes \times a)$  and  $(angle(h_{1i}, h_{2j}) < searchedTimes \times b)$  then
         $P = P \cup \{ h_{2j} \}$ 
      if  $|P| \geq 1$  then
        if  $|P| = 1$  then
           $M = M \cup \{(h_{1i}, h_{2x})\}$  where  $h_{2x} \in P$ 
        else if  $dist(h_{1i}, h_{2x})$  is the minimum value of all  $dist(h_{1i}, h_{2y})$ , where  $h_{2y} \in P$  then
           $M = M \cup \{(h_{1i}, h_{2x})\}$ 
         $lastIndex = j$ ;
    break;

```

圖 3.3 關鍵點對應演算法

### 3.5 求得特徵值

在完成了兩簽名的關鍵點對應之後，我們即可從關鍵點中求得我們所需要的特徵值。本研究中所採用的特徵值有三項，分別為質心向量(centroid)、速度(speed)、軌跡方向改變(stroke direction change)、以及簽名大小-時間比例值(size-time ratio)。

(1) 質心向量：

想要計算兩個簽名差異的程度，最基本要比較的必定是形狀的相似與否，若是連形狀之相似度都不高，那基本上此簽名幾乎可以說是假的(fake)。而要比較形狀的相似程度，有兩個因素是一定要考慮的，那就是每個點到某基準點之距離和方向。若是兩個簽名的形狀相似，則兩簽名中各自對應到的這些點的距離以及方向的資訊必定都會很接近，其差異值必然不大；反之若是兩簽名中各自對應到的點的距離以及方向之差異值極大，則形狀必定不可能相似。

為了要能成功的判斷兩個簽名形狀的相似程度，選擇一個適當的基準點是很重要的。在本研究中，我們所選擇的基準點即是簽名的質心。質心(centroid)在物理上的意義是一個物體質量的中心。在簽名裡，就是所有座標的平均。我們可以藉由下面公式求得質心座標  $C(x_c, y_c)$ ：

$$x_c = \frac{\sum_{i=1}^n x_i^N}{n} \quad (3.7)$$

$$y_c = \frac{\sum_{i=1}^n y_i^N}{n} \quad (3.8)$$

由於質心是由每個所記錄到的點中所求得的，因此質心可以當作一個簽名在形狀(Shape)方面的總括代表。而要求出形狀的相似程度，我們即可以利用兩簽名中所對應到的關鍵點與其各自質心之間的相對位置和方向之差異來計算。因此我們使用每個關鍵點與質心所形成的向量來計算。使用向量的原因在於，向量即可以表示平面上兩點之間的距離以及方向，而不必再將這兩個特徵分開計算。對於一個簽名中所有已對應完成的關鍵點：

$$M = \{(P_{1i}, P_{2i}) \mid P_{1i} \in S_1, P_{2i} \in S_2, 1 \leq i \leq k, \text{ where } k \text{ is the number of the mapped significant points}\}$$

要求出相似程度，先要求其各自之質心向量。定義  $V_1 = \{v_{1i} \mid (x_{1i} - x_{1c}, y_{1i} - y_{1c}), \text{ where } 1 \leq i \leq k\}$ ,  $V_2 = \{v_{2i} \mid (x_{2i} - x_{2c}, y_{2i} - y_{2c}), \text{ where } 1 \leq i \leq k\}$  為其質心向量，則  $(v_{1i} - v_{2i})$  就為它們兩向量之間的距離。要計算這兩個向量集合的相似程度，我們可利用下列公式來求得質心向量之相似百分比  $S_v$ ：

$$S_v = 1 - \sqrt{\frac{\sum_{i=1}^k (v_{1i} - v_{2i})^2}{k}} \quad (3.9)$$

(2) 速度：

在網路上想要對簽名作認證，只靠判斷簽名的形狀是不夠的，因為模仿簽名的形狀並非一件非常困難的事。若有原始的簽名當參考，入侵者只要稍微練習，即可將簽名的形狀模仿至一定程度的相似。因此要成功的對簽名作認證，我們還必須另外找出一個讓人很難得到原始資訊，也很難模仿成功的生物特徵。其中，簽名的速度就是一種。

每個人在簽名時，都會有自己獨特的節奏(rhythm)，例如有人簽名喜歡一氣呵成的完成他的簽名；有人習慣忽快忽慢；有人會在簽名的某些特定地方稍作停頓等。這些都是節奏的一種。然而，上述這些情況，並非光看原始簽名就可以看得出來的，而且即使模仿者就站在簽名者的身邊看著簽名者簽名，要模仿到一定的相似度也很困難，因此，利用速度這項生物特徵，可以大幅的幫助我們更正確的判斷一個簽名的真假。

對於簽名中的每個對應到的關鍵點而言，它們都有自己的速度。由於簽名是由許多的 pixel 所形成的，因此想要求出每個關鍵點的速度，我們訂定一個 pixel 長度區間  $L$ ，由對應到的關鍵點  $S_i$  開始向前以及向後各走  $L$  個 pixel，並計算它在走這  $2L$  個 pixel 時所花的時間  $t_i$ ，以及總距離  $d_i$ 。利用下列公式求出其速度  $q_i$ ：

$$q_i = \frac{d_i}{t_i} \quad (3.10)$$

定義  $Q_1 = \{q_{1i} \mid q_{1i} \text{ is the speed of } P_{1i}, 1 \leq i \leq k\}$ ,  $Q_2 = \{q_{2i} \mid q_{2i} \text{ is the speed of } P_{2i}, 1 \leq i \leq k\}$  為兩簽名各個對應到的關鍵點之速度。由下列公式求得其相似度  $S_s$ ：

$$S_s = 1 - \sqrt{\frac{\sum_{i=1}^k \left( \frac{q_{1i} - q_{2i}}{\max(q_{1i}, q_{2i})} \right)^2}{k}} \quad (3.11)$$

### (3) 簽名大小-時間比例值(size-time ratio)

在得到速度的相似度後，我們同時可以得到一個時間與簽名大小的比例值。考慮此比例值的原因在於每個關鍵點之間的時間差(duration)，很容易受到所簽的簽名的大小所影響，當所簽的簽名大小改變的時候，關鍵點之間的路徑長(筆劃長度)會相對改變，但是筆劃之間的時間比例卻不一定會隨著改變。

考慮下面的例子：

兩個相同簽名，簽名甲的面積比較小，但是用比較慢的速度簽；

簽名乙的面積比較大，但是用比較快的速度簽。

在這個例子中，此兩個簽名的簽名方式完全不一樣，我們希望能夠突顯此種情況所會造成的影響。

Size of two signatures:  $S_1, S_2$   
Enroll duration:  $T_1, T_2$   
Define  $x, y$   
Size-time ratio:  $S_z$

if  $(S_1 > S_2)$   $x = S_1 / S_2$   
else if  $(S_1 < S_2)$   $x = S_2 / S_1$

if  $(T_1 > T_2)$   $y = T_1 / T_2$   
else if  $(T_1 < T_2)$   $y = T_2 / T_1$

$S_z = \min(x,y) / \max(x,y)$

圖 3.4 簽名大小-時間比例演算法

圖 3.4 表示了我們求簽名大小與時間比例之演算法，此演算法的意義在於求出簽名面積與簽名總時間之間的比例關係，也就是簽名時的速度變化之差異度是否相近。 $S_z$  值越接近 1，表示兩個簽名即使大小有別，但在簽名時的速度變化的差異度其實是差不多的，簡單來說，也就是簽名的風格(style)越像；反之， $S_z$  值離 1 越遠，表示兩個簽名在簽名時的速度變化的差異度越大，換句話說，也就是簽名

的風格越不像。

(4) 軌跡方向改變(Stroke direction change)：

在比較兩個簽名時，很有可能發生一種情況，就是兩個簽名的關鍵點之時間資訊非常接近，因此速度的相似度極高。並且它們的關鍵點也幾乎在同樣的位置，相對於質心的質心向量之相似度也非常高。然而，這兩個簽名的筆劃卻有所不同。



圖 3.5 相似的關鍵點，不同的軌跡。

此圖即表示了這種情況，此圖中的兩個簽名軌跡是由同一個簽名者所簽，它們的關鍵點之時間資訊幾乎一樣，所以速度的相似度很高。另外，關鍵點以及質心的位置也都很接近，因此所求得的質心向量相似度也很高。然而，用肉眼卻可以很明顯的發現，這兩個簽名軌跡其實是有差別的。為了解決這個問題，我們必需應用方向改變向量來幫助我們區分。

軌跡方向改變(Stroke direction change)主要在描述連續的兩個關鍵點之間的軌跡移動趨勢。由上圖的例子來說，我們以肉眼觀察，左圖的第二關鍵點至第三關鍵點之間，x 座標的變化量皆是正向的(由左至右)，y 座標的變化量則皆是負向的(由上至下)；右圖裡同樣的第二關鍵點至第三關鍵點之間，其 y 座標的變化量雖然和左圖一樣，但是 x 座標的變化量卻有所改變，它是先往負的方向移動才再轉為正向。因此，我們可以藉由這種改變來對簽名做區分。

要記錄軌跡移動的趨勢，我們把一個圓分成八個部份來表示軌跡方向的變化。我們可以藉由下面的公式來得到軌跡在經過某個關鍵點  $S_i(x_i, y_i)$  之後的走向：

$$\theta_i = \tan^{-1} \left( \frac{y_j - y_i}{x_j - x_i} \right) \quad (3.12)$$

$$d_i = \theta_i / 45 \quad (3.13)$$

其中， $N_j(x_j, y_j)$ 為關鍵點  $S_i$ 的下一個鄰點。 $\theta_i$ 為  $S_i$ 與  $N_j$ 的角度關係。而  $d_i$ 則是記錄這個角度關係是屬於八個方向中的哪一個。

在求得簽名中每一段軌跡的方向改變之後，我們即可以比較兩個簽名間軌跡方向改變的相似程度。兩個簽名的軌跡方向改變分別為：

$$D_1 = \{d_{1i} \mid 1 \leq i \leq k-1\}$$

$$D_2 = \{d_{2i} \mid 1 \leq i \leq k-1\}$$

在這兩個向量中的每一項都是一個方向的代表，因此，要計算這兩個向量的相似程度，我們比較這兩個向量中每一項的值是否相同，可以得到相同值的個數為  $E$ ，而其相似度即為：

$$S_d = \frac{E}{k-1} \quad (3.14)$$

### 3.6 整體相似度

在 3.5 節中，我們探討了本研究所利用的三種生物特徵資訊並且分別對它們都求出了個別的相似度，以及得到在簽名時簽名大小與時間的關係。而在本節中，我們將討論如何將這些個別的相似度和關係值融合成為一個整體的相似度  $S_f$ ，進而判斷簽名的真假與否。

本研究中，我們認為要判斷一個簽名的相似程度，最重要的就是簽名的速度。因此，在此節的合併相似度之中，我們首先將以時間-簽名大小比例值以及方向變化向量為輔助，來調整速度相似度  $S_s$  的值。

要調整速度相似度，本研究建立了一個規則化的公式(rule-based formula)來做計算。其中，我們訂定了每個特徵值的影響係數(impact factor)。所謂的影響係數，也就是表示某特徵值，對整體相似度的影響到底有多大。這些影響係數的規則是事前所設定的，影響係數會隨著每個特徵的相似度而變化，不論每個特徵值



的個別相似度為多少，只要符合事前設定的規則條件，就可以得到對應的影響係數。

我們針對時間-簽名大小比例值  $S_z$ ，訂定了它的影響係數：

表 3.1 時間-簽名大小比例值之影響係數

$S_z$	
60% ↑	<b>1</b>
40%~60%	<b>0.8</b>
40% ↓	<b>0.5</b>

同時，也對方向改變向量  $S_d$  訂定了它的影響係數：

表 3.2 方向改變向量之影響係數

$S_d$	
25% ↑	<b>1</b>
25% ↓	<b>0.5</b>

由於本研究可以應用在許多地方，每次應用的情況都有所不同，因此管理者可以視所需要的情況，來定義上述的相似度區間值以及相關的影響係數，使得此方法可以更有效率的彈性應用在不同的場合之中。

在建立了相似度的區間範圍以及設定完成其相對應的影響係數之後，我們即可以藉由下面的公式來得到速度的相似度  $S'_s$ ：

$$S'_s = \alpha * S_s * \beta \quad (3.15)$$

然而，簽名的形狀和簽名時的速度是互相有關係的，在簽名驗證的過程中，若沒有考慮到簽名形狀的相似程度，將會有許多的例外情況出現。因此，我們利用質心向量相似度  $S_c$  當參考值，來增加或減少速度相似度  $S'_s$ ，以求得最終的相似度  $S_f$ 。

此調整的過程可以分成兩種情況，我們以  $S_c = 0.5$  當成一個分水嶺：

(1) 當  $S_c > 0.5$ ，我們希望若  $S'_s > S_c$ ，則調整過後的相似度  $S_f$  會比原始的  $S'_s$  更高。而當  $S'_s < S_c$ ，我們希望調整過後的相似度  $S_f$  會比原始的  $S'_s$  更低。圖 3.6 即為此種情況之調整曲線圖。而要達成此目的，我們可以藉由下列公式得到：

$$S_f = S_c - \sqrt{S_c^2 - S_s'^2}, \quad \text{if } S'_s \leq S_c \quad (3.16)$$

$$S_f = S_c + \sqrt{(1 - S_c)^2 - (S'_s - 1)^2}, \quad \text{if } S'_s > S_c \quad (3.17)$$

然而，當  $S_c < 0.5$  且  $S'_s > 0.5$  時，調整曲線需要利用(3.18)的公式來求得  $S_f$ 。

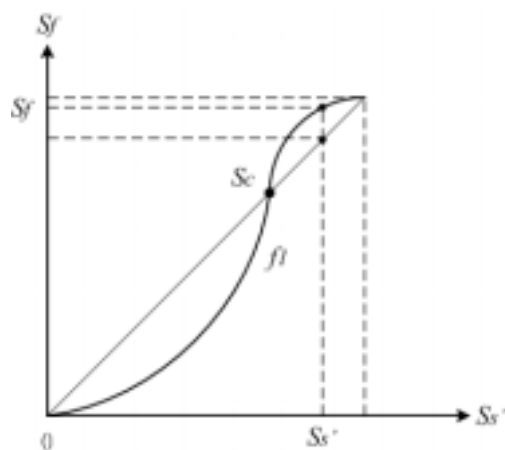


圖 3.6  $S_c = 0.5$  時之調整曲線圖

(2) 當  $S_c < 0.5$ ，在此情況中，若  $S'_s > S_c$ ，則調整過後的相似度  $S_f$  應該要比原始的更低。而  $S'_s < S_c$ ，我們希望調整過後的相似度  $S_f$  會比原始的  $S'_s$  更高，但不應該提高太多。圖 3.7 即為此種情況下之調整曲線圖。而要達成此目的，我們可以藉由下列公式得到：

$$S_f = \sqrt{S_c^2 - (S'_s - S_c)^2}, \quad \text{if } S'_s \leq S_c \quad (3.18)$$

$$S_f = 1 - \sqrt{(1 - S_c)^2 - (S'_s - S_c)^2}, \quad \text{if } S'_s > S_c \quad (3.19)$$

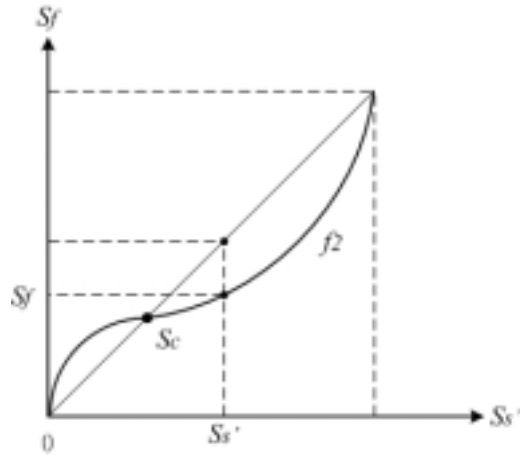


圖 3.7  $S_c < 0.5$  時之調整曲線圖

藉由上述的相似度合併以及相似度調整，我們可以求得最終的相似度  $S_f$ ，利用這個相似度，我們即可以判斷這兩個簽名是否為同一個人所簽。

在下一個章節中，我們將利用幾個案例，來說明在本章中所利用的生物特徵，對簽名的相似度所造成的影響。

## 第四章 實驗與討論

根據第三章的介紹，本章節將以實際的例子以及利用一些簽名輸入來說明本系統之運作。

### 4.1 系統之實作

本系統的實作乃是以 Java 為開發工具，Compaq iPAQ Pocket PC H3850 當作手寫簽名之輸入裝置，並透過 COM 連接埠，將整個簽名軌跡傳至系統，由系統進行運算。

圖 4.1 為簽名輸入系統之畫面，左邊之簽名乃是樣本簽名(template)，右邊則為使用者輸入要進行驗證之簽名。

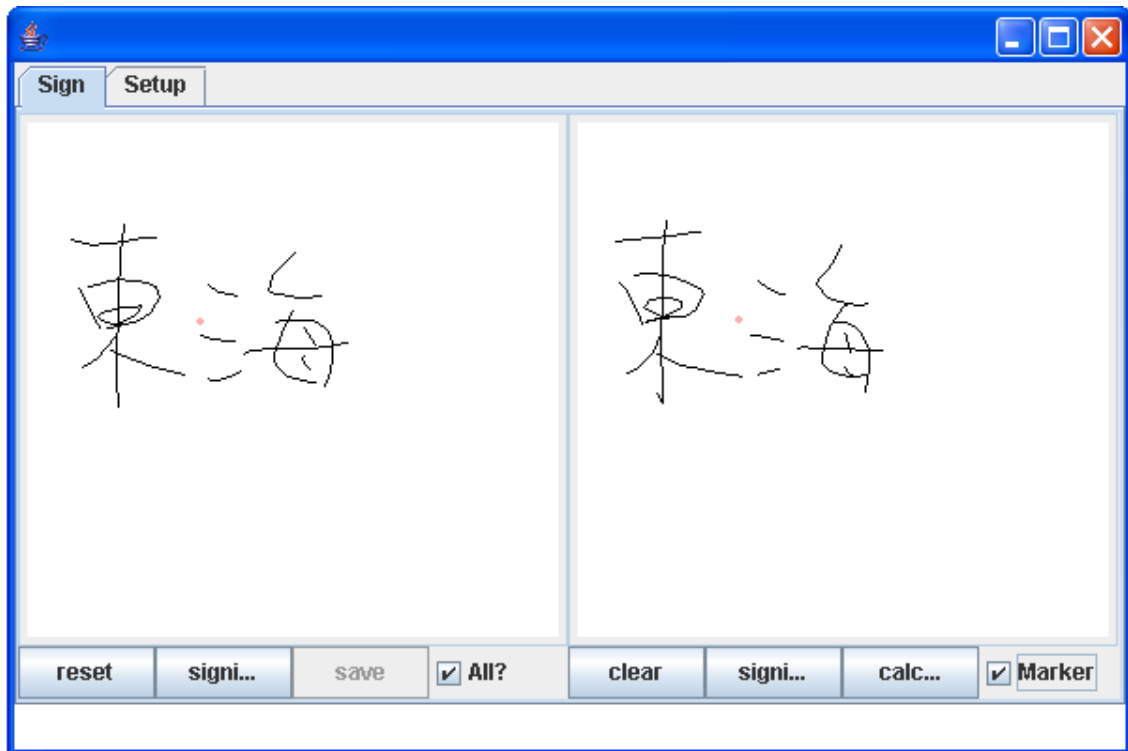


圖 4.1 系統之簽名輸入畫面

圖 4.2 為當簽名輸入之後，由系統尋找出之關鍵的畫面。圖中簽名上之各個圓點即為由系統所找尋到之關鍵點。

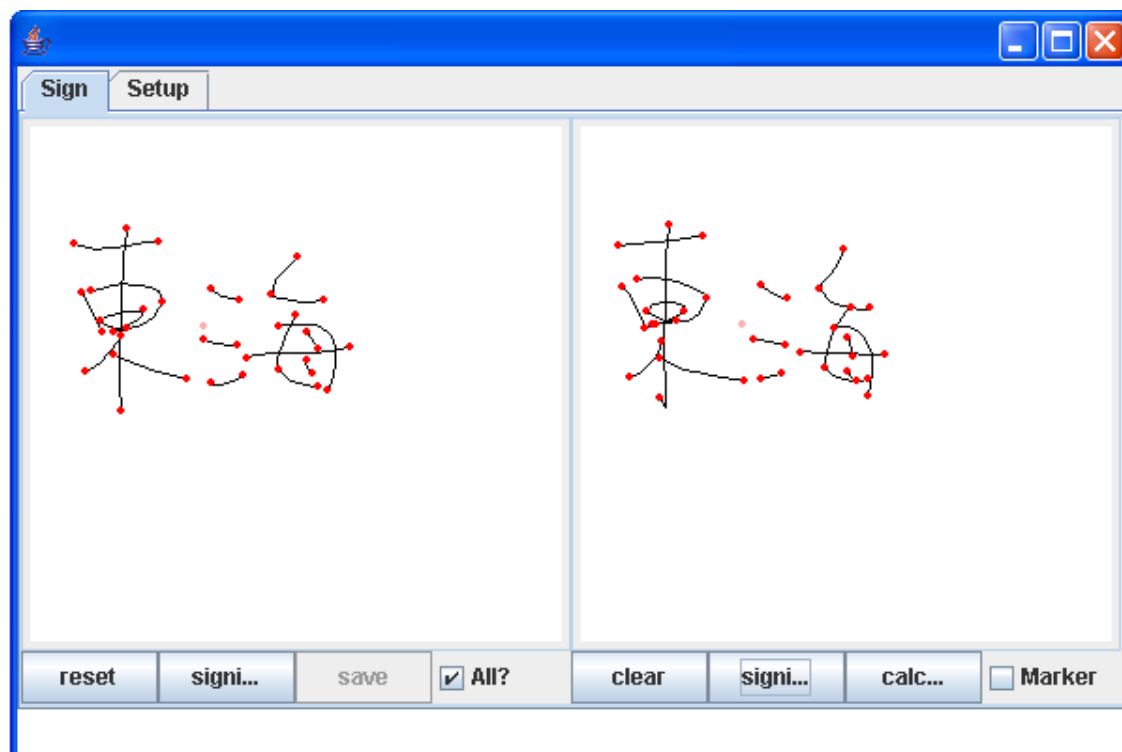


圖 4.2 求得關鍵點

圖 4.3 表示了兩個簽名中，關鍵點的對應以及所計算出之各個項目之相似度。圖中有被方框所框住之關鍵點即是有找到對應之關鍵點。在本系統中，我們為了要對應關鍵點而所應用的時間臨界值  $a$  為 0.02，而角度臨界值  $b$  為 15 度。

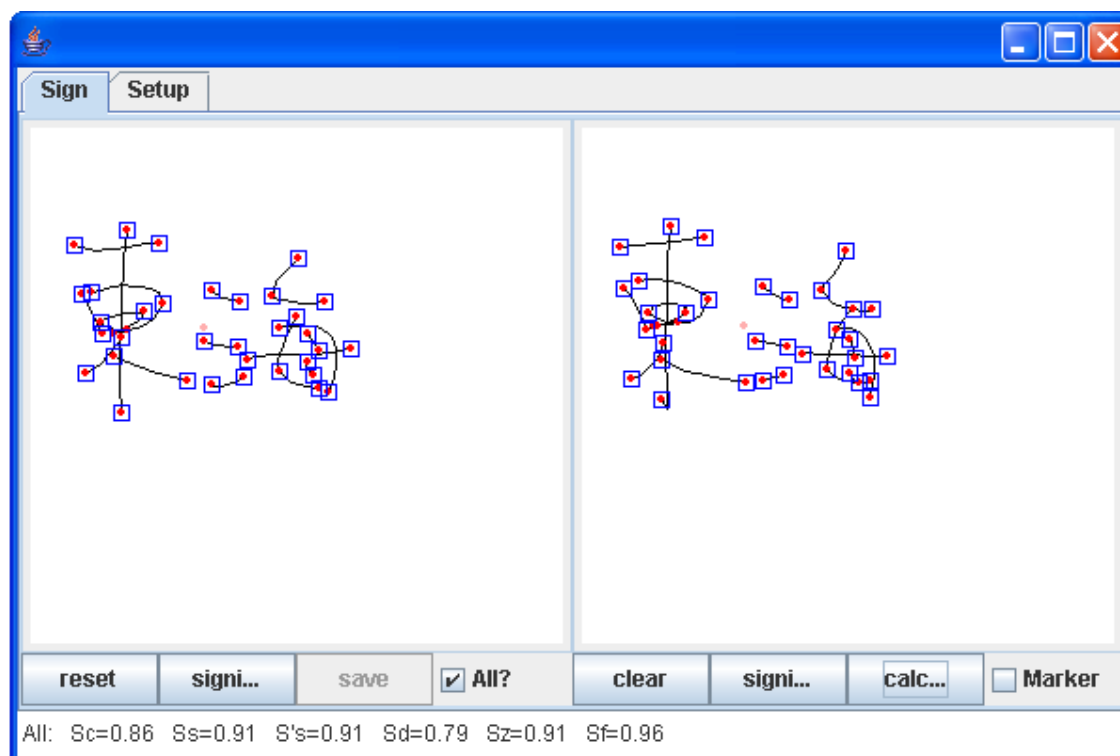


圖 4.3 關鍵點對應以及各相似度

## 4.2 實驗

本節將對各項生物特徵做探討，說明這些生物特徵對整體簽名相似度所造成之影響。

Case 1:

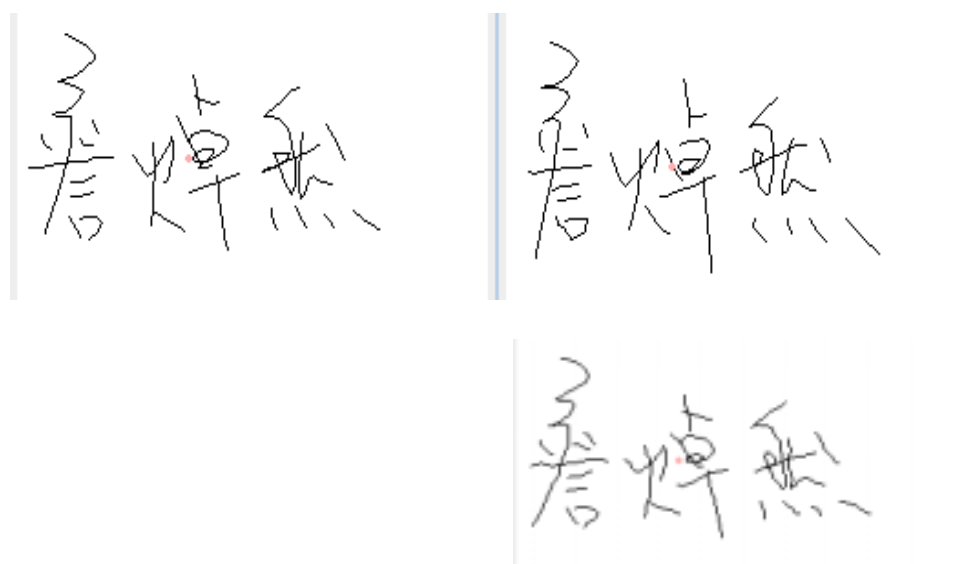


圖 4.4 真實簽名與抄襲簽名 1

在圖 4.4 中，上方右邊之簽名為真正的使用者所簽的簽名，而下方的簽名乃是他人模仿樣本簽名所簽出之簽名。表 4.1 為上方真實簽名之各項生物特徵之相似度數據，我們可以發現每一項數據都非常之高。而表 4.2 為下方他人所抄襲之簽名數據，其中可以發現，跟簽名形狀有關的相似度  $S_c$  與  $S_d$ ，其值都非常之高。然而由於此簽名是抄襲而來，抄襲者必須降低簽名時的速度以維持簽名形狀之完整，因此和速度有關的相似度  $S_s$  與  $S_z$ ，以及  $S'_s$ ，其值皆非常之低。而最後調整後之相似度  $S_f$  也同樣會被拉低。

$S_c$	83%
$S_s$	84%
$S'_s$	84%
$S_d$	76%
$S_z$	93%
$S_f$	89%

表 4.1 真實簽名 1

$S_c$	85%
$S_s$	36%
$S'_s$	18%
$S_d$	79%
$S_z$	37%
$S_f$	12%

表 4.2 抄襲簽名 1

Case 2:

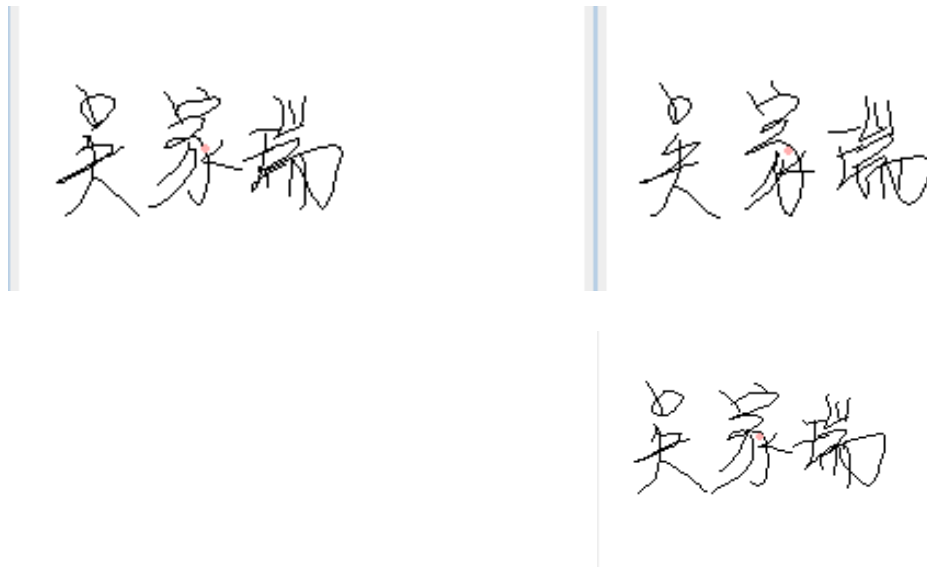


圖 4.5 真實簽名與抄襲簽名 2

圖 4.5 與圖 4.4 為同樣之條件，上方右邊之簽名為真正的使用者所簽的簽名，而下方的簽名乃是他人模仿樣本簽名所簽出之簽名。表 4.3 為上方真實簽名之各項生物特徵之相似度數據，每一項數據都非常之高。而表 4.4 為下方他人所抄襲之簽名數據，同樣的可以發現，跟簽名形狀有關的相似度  $S_c$  與  $S_d$ ，其值都非常高。而和速度有關的相似度  $S_s$  與  $S_z$ ，以及  $S'_s$ ，其值皆非常低。而最後調整後之相似度  $S_f$  也同樣會被拉低。



$S_c$	85%
$S_s$	85%
$S'_s$	85%
$S_d$	71%
$S_z$	93%
$S_f$	85%

表 4.3 真實簽名 2

$S_c$	82%
$S_s$	38%
$S'_s$	19%
$S_d$	73%
$S_z$	36%
$S_f$	2%

表 4.4 抄襲簽名 2

Case 3:



圖 4.6 軌跡方向改變不同之簽名

圖 4.6 表示了同一個簽名者所簽的簽名，其形狀和樣本簽名並沒有差很多，而速度也和樣本簽名非常相似，但是由肉眼可以看出其軌跡還是有所不同。表 4.5 為此簽名的各項生物特徵相似度之值，由其中可以發現，藉由軌跡方向改變值  $S_d$ ，可以表示出此種變化。

$S_c$	76%
$S_s$	86%
$S'_s$	43%
$S_d$	17%
$S_z$	90%
$S_f$	13%

表 4.5 軌跡方向改變所造成之相似度變化

Case 4:



圖 4.7 簽名大小-時間比例值不同之簽名

圖 4.7 表示了一個簽名大小-時間比例值不同之簽名,此種情況在簽名之大小有很大差距或簽名風格和樣本簽名有很大差異時,很容易發生。表 4.6 為此簽名之各項生物特徵之相似度。

$S_c$	89%
$S_s$	90%
$S'_s$	45%
$S_d$	33%
$S_z$	21%
$S_f$	15%

表 4.6 簽名大小-時間比例值不同所造成之相似度變化

Case 5:



圖 4.8 簽名形狀不甚相同之簽名

圖 4.8 表示了一個和速度有關之相似度都很高,然後和形狀有關的相似度卻不高

的簽名。表 4.7 為此簽名之各項生物特徵之相似度。從中可以看出即使簽名之速度很相似，但簽名之形狀不像時，整體相似度依舊會被降低。

$S_c$	48%
$S_s$	84%
$S'_s$	84%
$S_d$	30%
$S_z$	65%
$S_f$	62%

表 4.7 簽名形狀不甚相同所造成之相似度變化

### 4.3 總結

在本章節中，我們由實例說明了系統的運作流程，並且藉由不同的實驗來討論在各種不同的情況下，手寫簽名的相似度因而所造成的改變。這也同時說明了這些生物特徵，對於簽名認證有著非常大的影響力。

## 第五章 結論與未來方向

### 5.1 結論

隨著網際網路的高速發展，越來越多的網路應用系統也被開發出來，提供更便利的服務。在這之中，不乏許多需要高度安全控管的系統，而要對這些系統進行存取或是利用這些服務，就必須要先經過一些驗證身份的程序，方可被準許使用。因此，對服務提供者或系統管理員而言，一個可靠有效的身份驗證程序就顯得格外的重要。

有鑑於傳統應用於身份認證的方法都有其先天上的缺點，因此利用個人生物特徵來幫助判斷合法或非法使用者的方式，逐漸的被人們所利用。本論文即是針對現行數位憑證的認證方式，提出了一個利用手寫簽名來幫助加強認證的方法。本方法只需要一個能夠輸入簽名的配備，不需要太多的計算資源，是個低成本的演算法。除此之外，這個方法利用了質心向量來表示簽名的形狀、利用速度來記錄簽名時重要的生物特徵、並且利用時間-簽名大小比例值來表現簽名時候的風格(style)、而對於有著相同上述資訊，然後軌跡卻不甚相同的簽名，更是用方向改變向量來描述這種情形。最後，配合影響係數以及對於相似度的合併以及放大與縮小，我們可以得到一個可以信任的最終相似度。在第四章中，我們透過實驗更證明了我們所提出的手寫簽名驗證方法的可行性。

利用最終求出的相似度，可以快速有效的判斷目前的簽名者是否為合法的使用者。如此更可進一步的補強了現行數位憑證認證的不足，保證了即使數位憑證被竊取或是被誤用，此非法的使用者也無法完全的通過身份認證，使網路上的交易更有所保障。

### 5.2 未來方向

在本研究中，我們提出了幾個演算法來幫助我們對手寫簽名做認證，這些演算法皆應用了一些臨界值來幫助達成我們所要的目的。然而，臨界值的選取對系

統的正確性有著極大的影響。因此，如何更正確的選取適當的臨界值，來增加系統在驗證時的效率以及正確性，將是我們未來另一個研究的重點。

## 參考文獻

- [1] E. Anquetil and G. Lorret, "Perceptual model of handwriting drawing. Application to the handwriting segmentation problem," *Proceedings of the 4<sup>th</sup> IEEE International Conference on Document Analysis and Recognition*, Vol.1, August 1997, pp 112-117.
- [2] J. Brault and R. Plamondon, "Segmenting Handwritten Signatures at Their Perceptually Important Point," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.15, No.9, 1993, pp 953-957.
- [3] S. H. Cha and S. N. Srihari, "Multiple Feature Integration for Writer Verification," *Proceedings of 7<sup>th</sup> International Workshop on Frontiers of Handwriting Recognition*, Amsterdam, September 2000, pp. 333-342.
- [4] S. H. Cha and C. C. Tappert, "Automatic Detection of Handwriting Forgery," *Proceedings of 8<sup>th</sup> International Workshop on Frontiers in Handwriting Recognition*, August 2002, pp. 264-267.
- [5] T.C. Chiang, "Internet and World Wide Web Technologies and Opportunities," *Proceedings of the IEEE International Conference on Industrial Technology*, December 1996, pp 858-862.
- [6] Q. Dai and R. J. Kauffman, "Business Models for Internet-Based E-Procurement Systems and B2B Electronic Markets: An Exploratory Assessment," *Proceedings of the 34<sup>th</sup> Annual Hawaii Internet Conference on System Sciences*, Hawaii, January 2001, pp 10.
- [7] R. A. J. Everitt and P. W. McOwan, "Java-Based Internet Biometric Authentication System," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.25, No.9, 2003, pp 1166-1172.
- [8] M. C. Fairhurst and E. Kaplani, "Perceptual Analysis of Handwritten Signature for Biometric Authentication," *Proceedings of IEE on Vision, Image and Signal Processing*, Vol.150, December 2003, pp. 389-394.
- [9] M.C. Fairhurst, "Signature Verification Revisited: Promoting Practical Exploitation of Biometric Technology," *Electronic & Communication Engineering Journal*, Vol.9, December 1997, pp. 273-280.

- [10] B. Fang, "Tracking of Feature and Stroke Positions for Off-line Signature Verification," *Proceedings of International Conference on Image Processing*, Vol.3, June 2002, pp. 965-968.
- [11] M. George and R. King, "A Robust Speaker Verification Biometric," *Proceedings of IEEE 29<sup>th</sup> Annual Int'l Carnahan Conference on Security Technology*, Sanderstead, 1995, pp 41-46.
- [12] L. Gong, "Increasing Availability and Security of an Authentication Service," *IEEE Journal on Selected Areas in Communications*, June 1993, pp 657-662.
- [13] S. Gordon, "Ocular Biometrics: For Your Eyes Only," *Opto and Laser Europe*, No.84, May 2001.
- [14] C. C. Han, P. C. Chang, and C. C. Hsu, "An Online Signature Verification System Using Multi-Template Matching Approaches," *Proceedings of IEEE 33<sup>rd</sup> Annual International Conference on Security Technology*, October 1999, pp. 477-480.
- [15] S. Hangai, S. Yamanaka, and T. Hamamoto, "On-Line Signature Verification Based on Altitude and Direction of Pen Movement," *IEEE International Conference on Multimedia and Expo*, New York, 2000, pp 489-492.
- [16] B. Herbst, and D. Richards, "On an Automated Signature Verification System," *Proceedings of IEEE Int'l Symposium On Industrial Electronics*, Pretoria, 1998, pp 600-604.
- [17] N. Issa, "Internet security," *Proceedings of 2004 International Conference on Information and Communication Technologies: From Theory to Applications*, April 2004, pp 623.
- [18] A. K. Jain, A. Ross, and S. Prabhakar, "An Introduction to Biometric Recognition," *IEEE transactions on Circuits and Systems for Video Technology*, Vol.14 ,Issue1, January 2004, pp 4-20.
- [19] A. K. Jain, F. Griess and S. Connell, "Online signature verification," <http://citeseer.ist.psu.edu/jain02line.html>.
- [20] K. S. Jung, J. Y. Kim, and T. C. Chung, "Password-Based Independent Authentication and Key Exchange Protocol," *Proceedings of the 2003 Joint Conference of the Fourth International Conference on Information, Communication and Signal Processing*, December 2003, pp 1908-1912.

- [21] Y. Komiya and T. Matsumoto. "On-line Pen Input Signature Verification PPI (pen-Position/pen-Pressure/pen-Inclination)," *Proceedings of 1999 IEEE Int'l Conf. on Systems, Man, and Cybernetics*, Tokyo, October 1999, pp 41-46.
- [22] M. Kondo , D. Muramatsu, M. Sasaki, and T. Matsumoto, "Nonlinear Separation of Signature Trajectories for On-line Personal Authentication," *Proceedings of 2003 IEEE Int'l Conf. on Multimedia and Expo, ICME '03*, Baltimore, April 2003, pp II - 349-52.
- [23] W. S. Lee, N. Mohankrishnan, and M. J. Paulik, "Improved Segmentation Through Dynamic Time Warping for Signature Verification Using a Neural Network Classifier," *Proceedings of 1998 Int'l Conference on Image Processing, ICIP 98*, Chicago, October 1998, pp 929-933.
- [24] B. Li , K. Q. Wang, and D. Zhang, "On-line Signature Verification for E-finance and E-commerce Security System," *Proceedings of 2003 Int'l Conf. on Machine Learning and Cybernetics*, Xi'An, November 2003, pp 3002-3007.
- [25] R. Martens and L. Claesen, "Online Signature Verification by Dynamic Time-Warping," *Proceedings of 13<sup>th</sup> International Conference on Pattern Recognition*, Vol.3, August 1996, pp. 38- 42.
- [26] R. Martens and L. Claesen, "Dynamic Programming Optimization for On-line Signature Verification," *Proceedings of the 4th Int'l Conf. on Document Analysis and Recognition*, Ulm, August 1997, pp 653-656.
- [27] Ithipan Methasate and Sutata Sae-Tang, "Online Thai Handwriting Character Recognition Using Stroke Segmentation with HMM," *Applied Informatics International Symposium on Artificial Intelligence and Applications*, Innsbruck, February 2002, pp. 59-62.
- [28] N. Mohankrishnan, M. J. Paulik, and Mohamad Khalil, "Online Signature Verification Using a Nonstationary Autoregressive Model Representation," *Proceedings of the IEEE International Symposium on Circuits and Systems*, Chicago, May 1993, pp. 2303-2306
- [29] I. Nakanishi, N. Nishiguchi, Y. Itoh and Y. Fukui, "Online Signature Verification Method Utilizing Feature Extraction Based on DWT," *Proceedings of International Symposium on Circuits and Systems*, Vol.4, May 2003, pp. IV-73 – IV-76.
- [30] V. S. Nalwa, "Automatic on-line signature verification," *Proceedings of IEEE*, Vol. 85, pp.215-239, 1997



- [31] J. D Penagos, N. Prabhakaran, and S.V Wunnava, "An efficient Scheme for Dynamic Signature Verification," *Southeastcon '96, IEEE Proceeding, 'Bringing Together Education, Science and Technology'*, April 1996, pp 451-457.
- [32] F. Ramann, C. Vielhauer and R. Steinmetz, "Biometric Applications Based on Handwriting," *IEEE International Conference on Multimedia and Expo*, Lausanne, August 2002, pp 573-576.
- [33] A. Roddy and J. Stosz, "Fingerprint Features-Statistical Analysis and System performance estimates," *Proceedings of the IEEE*, Vol.85, No.9, September 1997, pp 1390-1422.
- [34] H. E. S. Said, T.N. Tan and K. D. Baker, "Writer Identification Based on Handwriting," *IEE 3<sup>rd</sup> European Workshop on Handwriting Analysis and Recognition*, July 1998, pp. 4/1-4/6.
- [35] D. Sakamoto, M. Kondo, H. Morita , D. Muramatsu, M. Sasaki and T. Matsumoto, "Dynamic Biometric Person Authentication Using Pen Signature Trajectories," *Proceedings of 9th Int'l Conf. on Neural Information Processing*, Singapore, November 2002, pp 2078-2082.
- [36] R. Sanchez-Reillo, C. Sanchez-Avila, and A. Gonzalez-Marcos, "Biometric Personal Identification through Hand Geometry Measurement," *IEEE transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, October 2000, pp 1168-1171.
- [37] C. Schmidt and K. F Kraiss, "Establishment of Personalized Templates for Automatic Signature Verification," *Proceedings of 4th Int'l Conf. on Document Analysis and Recognition*, Vol.1, Ulm, August 1997, pp 263-267
- [38] M. Tanaka, Y. Ishino, H. Shmada and T. Inoue, "DP Matching Using Kalman Filter as Pre-Processing in Online Signature Verification," *Proceedings of 8<sup>th</sup> International Workshop on Frontiers in Handwriting Recognition*, August 2002, pp. 502-507.
- [39] C. Vielhauer, R. Steinmetz and A. Mayerhofer, "Biometric Hash Based on Statistical Feature of Online Signatures," *Proceedings of 16<sup>th</sup> International Conference on Pattern Recognition*, Vol.1, August 2002, pp. 123-126.
- [40] T. Wessels and C. W. Omlin, "A Hybrid System for Signature Verification," *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks*, Vol.5, July 2000, pp. 509-514.

- [41] S. William, "Network security essentials: applications and standards," *ISBN 0-13-016093-8*, United States of America, 2000, pp. 72-78.
- [42] B. Wirtz, "Stroke-based Time Warping for Signature Verification," *Proceedings of the 3<sup>rd</sup> Int'l Conf. on Document Analysis and Recognition*, Vol.1, Montreal, August 1995, pp 179-182
- [43] X. H. Xiao and R. W. Dai, "A Hierarchical Online Chinese Signature Verification System," *Proceedings of the 3<sup>rd</sup> International Conference on Document Analysis and Recognition*, Vol.1, August 1995, pp. 202-205.
- [44] K. Yue and W. Wijesome, "Improved Segmentation and Segmentation Association for Online Signature Verification," *Proceedings of IEEE International Conference on System, Man, and Cybernetics*, Vol.4, October 2000, pp. 2752-2756.
- [45] Y. Zhu, T. Tan and Y. Wang, "Biometric Personal Identification Based on Iris Pattern," *Proceedings of Int'l Conf. on Pattern Recognition*, Vol.2, Barcelona, September 2000, pp 805-808.
- [46] Y. Zhu, T. Tan and Y. Wang, "Biometric personal identification based on handwriting," *Proceedings of 15<sup>th</sup> International Conference on Pattern Recognition*, Vol.2, September 2000, pp. 797-800.
- [47] M. Zou, J. Tong, C. P. Liu, and Z. L. Lou, "On-line Signature Verification Using Local Shape Analysis," *Proceedings of 7<sup>th</sup> Int'l Conf. on Document Analysis and Recognition*, Edinburgh, August 2003, pp314-318