

建置大型指紋資料庫的架構分析與探討

A Study on the Architecture of Large Scale Fingerprint Database Construction

張正弘

德明財經科技大學資訊科技系

Email: jhchang@takming.edu.tw

摘要

在這個論文中，我們針對大型的指紋建檔需求，設計一種結合指紋本身紋路的基本結構及紋路的分佈數目，能夠延伸基本的指紋 Henry 分類法使具有較多子類別的指紋資料庫系統。本方法不以區域性的特徵來決定分類，而是以指紋的全域特徵為考量的基礎，因此不受指紋的區域特徵限制，在部分指紋紋路喪失的情況下，仍然可以取出紋路分佈的順序來做分類。觀察各類型的指紋，我們發現組成所有指紋的紋路只有十種基本的樣式 (fundamental ridges)，而不同類型的指紋，在不考慮紋路的分支 (bifurcation) 與斷線 (fragment) 的情形下，都只是由這十種不同的指紋紋路樣式依照特定的順序所組合而成的。基於這個原理，我們提出一種語句式 (syntactic) 表達法的新模型，以不同的句子來代表個別的指紋，句子中的字母 (alphabet) 就是那十種指紋紋路的基本形式。並且以此方法為基礎，發展出一個指紋分類系統的新模型。另外，我們也嘗試利用紋路分佈所具備的特質來定義出一個清楚的拒絕分類標準。這個分類法除了能完成 Henry 分類的那七種有效分類以外，也可以使資料庫中的子集合數目大量增加。對於類似全民指紋建檔等的超大型指紋資料庫，或許才可能真正落實。

關鍵詞：指紋資料庫，指紋分類，紋路分佈，奇異點，拒絕標準。

一、研究動機與指紋分類簡介

指紋自動辨識的學術研究已經到達瓶頸的階段，其原因在於日益擴增的指紋資料數量，使得辨識速度

及辨識準確度上皆無法突破。因此，較大的指紋資料庫傾向採取預先分類的措施，以減少指紋辨識所需的比對次數。過去十幾年來有非常多的研究團體致力於指紋分類系統的研究，所使用的方法以 model-based, structure-based, frequency-based 及 syntactic approaches 為主。但是，現今的自動指紋分類系統都無法達到實用的階段，其原因皆為所取得的指紋品質無法符合大部份自動分類系統的需求；指紋自動分類系統的分類正確性過低，及自動分類的處理所需時間過長，使得自動指紋分類系統至今仍甚少實際應用於大資料庫的指紋辨認系統上。

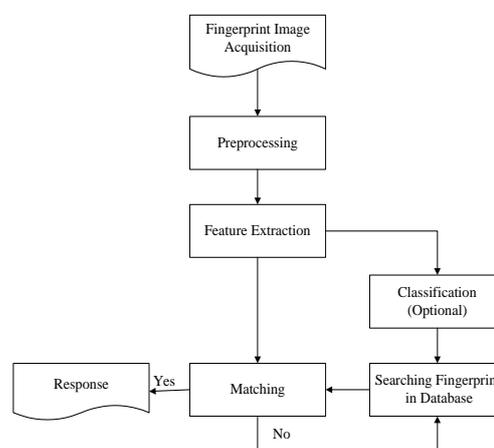


圖 1 具有指紋分類的指紋辨識系統方塊圖

圖 1 中所表示的是一個具有指紋分類的指紋辨識系統方塊圖 [1]。然而，現存指紋自動分類系統的分類正確性過低 (小於 90%)，及自動分類的處理所需時間過長，使得自動指紋分類系統至今仍甚少實際應用於大型資料庫的指紋辨認系統上

[2, 3]。以前面所列的指紋前處理步驟來說，由於過程繁複，因此幾乎有 95% 的指紋比對與分類的時間是花費在前處理上。也因此有必要發展一套能符合實際需求的指紋自動分類系統。

現存的指紋自動分類系統皆是依據 Henry 分類法[4]將所有的指紋分成八大類（如圖 2 所示），分別是：arch, tented arch, left loop, right loop, plain whorl, central-pocket whorl, double-loop whorl 和 accidental whorl 等；國內外作指紋辨識器的廠商所設計的方法幾乎都無法使用在超過千人以上的場合。這項研究是要設計一個大型的指紋資料庫，所設計的指紋資料庫必須包含萬人以上的指紋資料，分類的正確性必需大於 98%，指紋比對的正確性必需大於 99%，這些規格是目前的技術無法達成的。

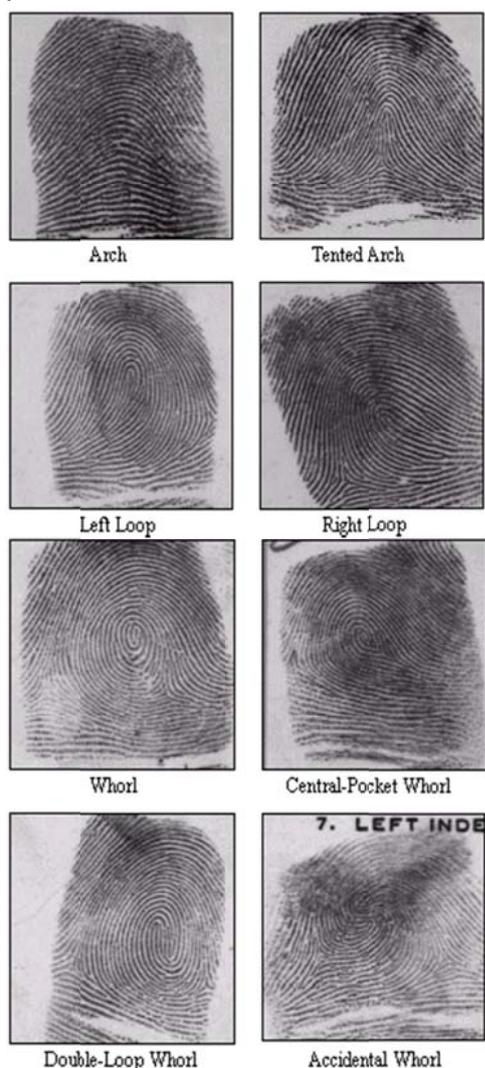


圖 2 Henry 分類法所定義的指紋類型

本論文提出了一個能夠分出七個完整類別的自動指紋分類系統的原創方法。指紋影像是經由各式各樣的管道取得的，這些指紋影像的品質通常並不是相當好。因此若一個指紋分類系統在規劃時沒有將一些意外情況考慮進去，則分類的結果一定不能達到可以實用的標準。基於指紋的變形、不完整及污損的性質，使得指紋分類系統的分類正確性一直無法有效的提昇。一般來說，傳統的指紋分類法都只能分出 Henry's classification 中的三到五類，並且正確率大都低於 90%。本論文所提出的以指紋紋路的分佈順序為前提的自動指紋分類系統，是以紋路的全域的空間特徵為依據，因此能夠避免部分區域由各種原因所造成的雜訊，進而達成了較高的分類準確率。在拒絕率為 5.1% 的情況下，能夠分出完整的七類指紋，並且平均正確率達到 94.8% 以上。對於指紋自動分類系統來說，本論文所提出的方法是能夠分出 Henry's classification 所有類型的唯一方法，以研究的觀點考量，本論文所提出的分類法是一項新的突破。未來若是能夠與指紋辨識的業界結合，將此方法應用於大型的指紋資料庫上（例如全民的身份證指紋資料庫系統），本論文設計的分類法必定能成功的提高該系統的辨識準確度及辨識速度。

二、指紋自動分類之新模型

這個部份嘗試設計一個新的指紋自動分類系統。本論文中所提出的方法是以指紋紋路的幾何形狀特徵為基礎，以不同的紋路分佈順序來代表不同類別的指紋。觀察各類型的指紋，我們發現組成所有指紋的紋路只有十種基本的樣式，而不同類型的指紋，在不考慮紋路的分支與斷線的情形下，都只是由這十種不同的指紋紋路樣式依照特定的順序所組合而成的（如圖 3 所示）。對於紋路的分支與斷線的情況我們在接下來的處理過程中會有討論。

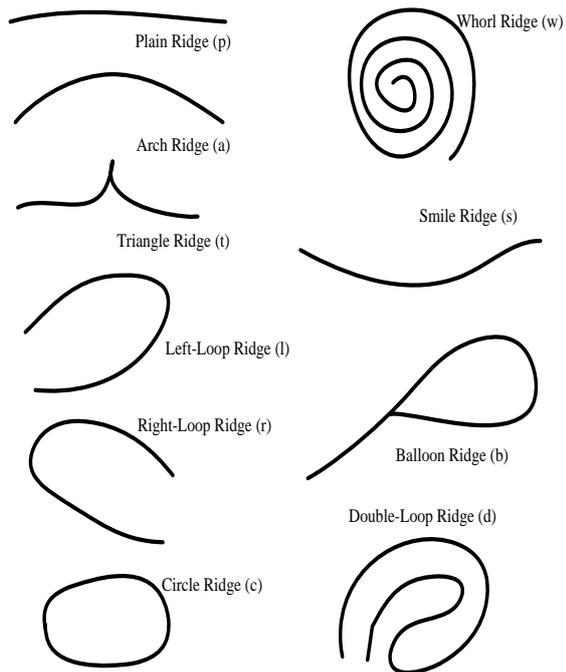


圖 3 組成所有類型指紋的十種基本的紋路（括號內的是代表紋路的縮寫）。

因此，若由指紋影像的下方開始，以各個紋路的中心線為虛擬的搜尋軸由下往上搜尋，則可以得到一個指紋紋路排列順序，不同類型指紋的紋路組合順序皆不相同。圖 4 中所示的是一個指紋的紋路分佈的順序的範例。

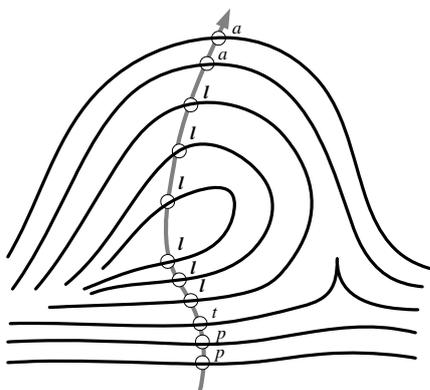
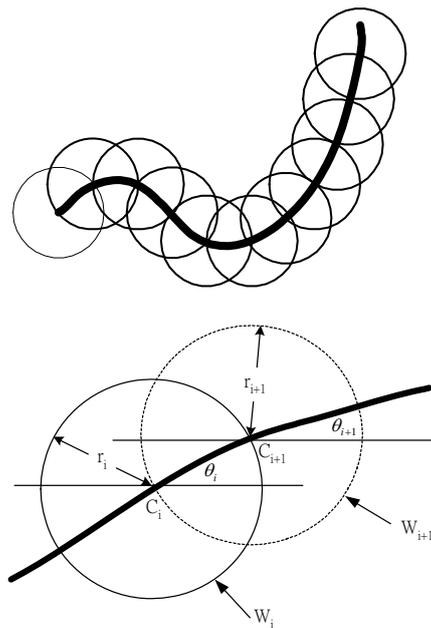


圖 4 一個 right-loop class 的指紋，紋路的分佈順序由下到上是 *pptllllllaa*

我們是利用 spherical algorithm，以一個任意的起始點開始向左及向右兩個方向來追蹤指紋的紋路。追蹤指紋的同時記錄當時的圓圈中心(C_i)、半徑(r_i)及前進角度(θ_i)，並以此圓與紋路的交會處當作下一個圓圈的中心 (C_{i+1})。



圖五：指紋紋路的追蹤

三、紋路抽取的後處理及型式判別

指紋的紋路經常會有分支的情況，由於我們所預先分析出的紋路類型並不考慮分支(否則可能有數十種紋路類型)，因此分支的紋路在我們的系統中會被分別的追蹤並且形式上分出兩條獨立的紋路出來(如圖 6 所示)。另外，由於指紋在擷取時的不均勻的特性，指紋的紋路或多或少都具有某種程度的斷裂。為了區別出在追蹤時所遇到的是真正的端點或只是紋路的小斷裂，基於指紋紋路是連續的特性，在所追蹤的圓圈遇不到交點的情況，以前一個圓圈的前進角度再追蹤兩個步驟以確認真正的端點(如圖 7 所示)。

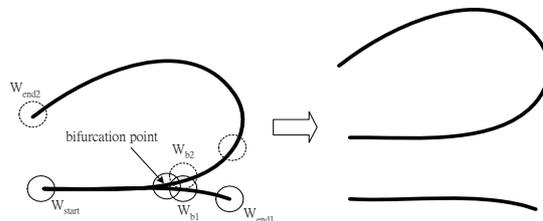


圖 6 具有分支的紋路，在追蹤過後分成兩條不同的紋路 l 及 p 。

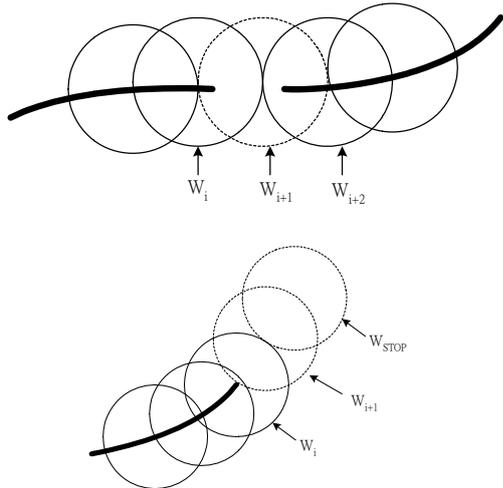


圖 7 紋路的斷裂點與紋路端點的處理。

我們以一個 rule base 的方法來判斷這十種基本的紋路。在追蹤完一條指紋後得到兩個端點，以這兩個端點做出投影軸 (PA)，在投影軸的中心點垂直向上延伸做出一個抽取軸(EA)，抽取軸與紋路的交會點是這條紋路的主要點(principal points)。利用端點所在的位置、主要點的數目、追蹤時的角度及紋路的高度等資訊便可以判斷出紋路的形式(如圖 8 所示)。

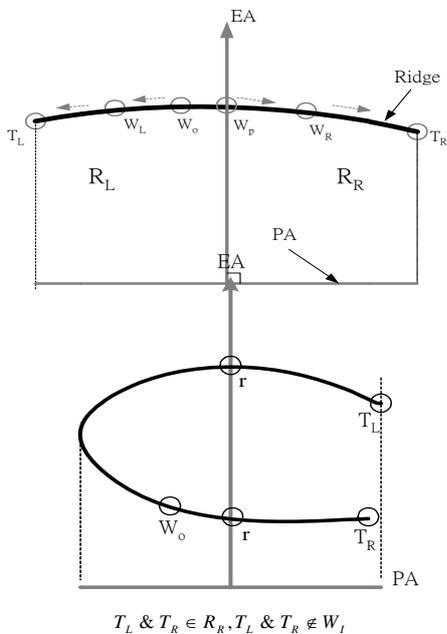


圖 8 判別紋路時所需要的一些參數。

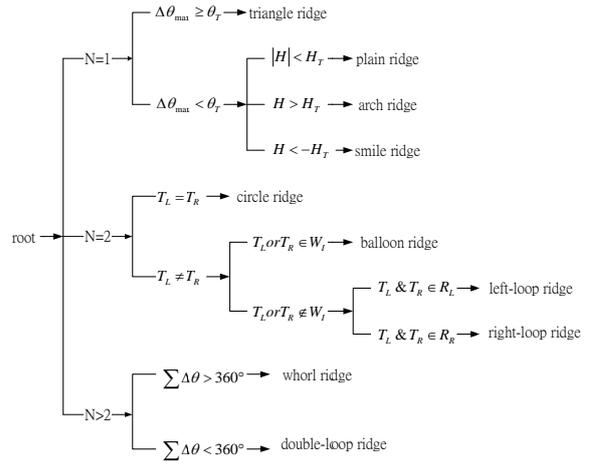


圖 9 十種基本指紋紋路的區別樹(discrimination tree)。

四、指紋紋路的分類模型及實驗結果

觀察各類型的指紋，我們發現組成所有指紋的紋路只有十種基本的樣式，而不同類型的指紋，都只是由這十種不同的指紋紋路樣式依照特定的順序所組合而成的。如表 1 所示，我們以一組 regular expression 來表示此分類的紋路分佈順序，使用 regular expression 來設計模型的優點除了可以簡化系統的設計之外，對於區別紋路時錯誤的指紋，我們設計的系統可以具備容錯的能力。

表 1 代表指紋分類的 regular expression 及經過簡化的 regular expression

Fingerprint Classes	Simplified Regular Expression for Classification
Arch	p^+a^+
Tented Arch	$\alpha^*t^+a^+$
Left Loop	$p^+\alpha^*tl^+a^+$
Right Loop	$p^+\alpha^*tr^+a^+$
Plain Whorl	$\alpha^*s^+\alpha^+(c+w)^+\alpha^+a^+$
Central-Pocket Whorl	$p^+\alpha^*b^+\alpha^+a^+$
Double-Loop Whorl	$\alpha^*s^+\alpha^*dd^+\alpha^+a^+$

我們以一組預先分析出的 NFA 模型為依據(如圖 10 所示)，若此順序能夠符合其中某一個已知類型的紋路排列順序，則接受這個指紋並指定其類型；若依照此輸入順序卻依然無法使 NFA 停留在接受狀態，則將這個指紋歸類為拒絕分類，這種

方法能夠分出完整的七個 Henry 的分類。

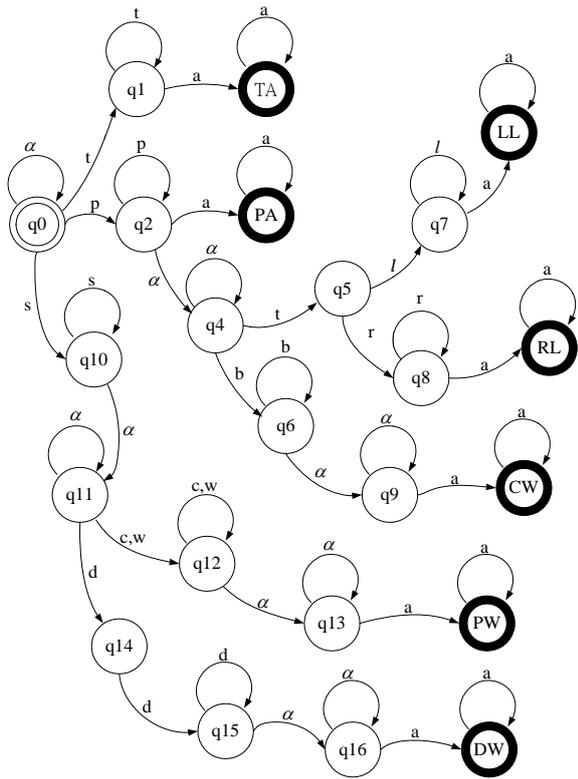


圖 10 以紋路分佈順序為基礎的指紋分類模型

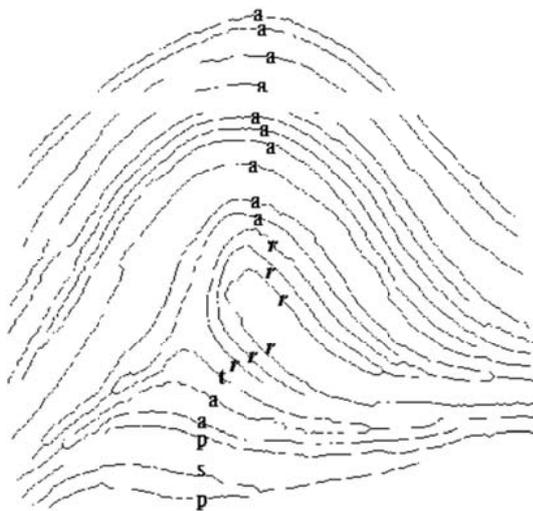


圖 11 指紋紋路的定位、指紋點的連接追蹤及指紋紋路的形式判別。

我們以 NIST 4 指紋資料庫進行實驗，該資料庫具有五種指紋分類，每種指紋各有 400 枚，為了陪和信筒的 7 種分類，我們將其中的 Whorl Type 細分為 PW、CW 及 DW 三類，圖 11 中所示的是指紋紋路判別的部分實驗結果，分類的結果則如表 2

所示。

表 2 分類的實驗結果

True Class		Assigned Class								
Type	Num.	PA	TA	LL	RL	PW	CW	DW	Reject	Accu.
PA	400	375	3	6	4	0	0	0	12	96.6%
TA	400	6	359	11	9	0	0	0	15	93.2%
LL	400	0	0	361	0	6	8	4	21	95.3%
RL	400	0	1	0	354	4	7	8	26	94.7%
PW	312	0	0	4	2	277	3	4	22	95.5%
CW	25	0	0	1	2	0	20	0	2	87.0%
DW	63	0	0	2	2	1	0	55	3	91.7%

五、 結論與未來研究方向

本論文首先提出一個能夠在未知叢集數的多模灰階濃度分佈圖中找出多個閾值的方法。在樣本數夠多的情形下，多模灰階濃度分佈圖中的所有叢集則可以被視為一個混合高斯分佈；我們所發展出來的方法除了能夠預估出準確的閾值以外，也能夠估計出各個叢集的參數。我們捨棄了傳統的統計式重複的調整每個叢集參數的作法，預先求出最佳估計區間，因而降低了計算的複雜度；實驗上利用預先設計的叢集比例，以切割後的各叢集比例來與此預先設計的比例做比較，我們所設計的演算法也具有較高的正確性。

本論文也提出了一個以指紋紋路的分佈順序為前提的自動指紋分類系統。本方法不以區域性的特徵來決定分類，而是以指紋的全域特徵為考量的基礎，因此不受指紋的區域特徵限制，在部分指紋紋路喪失的情況下，仍然可以取出紋路分佈的順序來做分類。對於遭受污損的部分，因其賴以分類的紋路是一個序列，而非單一條特定的紋路，因此紋路所具有的訊息較不容易失去。另外對於不完整的指紋，只要大部分的紋路仍是完整的，仍可以找出代表其類型的順序。對於指紋的變形、旋轉或位移，由於抽取紋路的方法並不依賴其空間的位置，因此並不構成問題。這個分類的方法能成功的分出完整的 Henry 七類指紋，且在拒絕率為 5.1% 的情況下，平均正確率達到 94.8% 以上，對於指紋自動分類系統來說，這是一項突破。

本論文所提出的各種新方法除了能夠使用在指紋系統以外,也能利用於各種不同的應用場合上。列舉如下:

- A. 我們所發展的全域式(global)多模灰階濃度分佈圖的分解技術,可以非常容易被改寫成區域式(local)的演算法。對於亮度不均勻的影像,或是影像中的目標物體並不位於最亮或最暗處時,使用我們所發展的方法可以有有效的切割出有用(或是有興趣)的物體。應用的場合包含:不均勻亮度時的人臉定位、多媒體文件中不同顏色的文字與圖形的分離及 content-base image retrieval 等。
- B. 我們所發展出來的直接灰階影像的指紋定位法,若改變紋路點的連接及紋路追蹤的定義,從另一個角度來看可以直接轉換成從灰階影像做細線化(thinning)的影像前處理動作。根據論文中的分析,直接由灰階影像做細線化不須經過影像二值化的步驟,速度必然較快,也不會有不須要的交互連結或是突刺狀的線段產生。應用的場合包含:物體的外型辨識、曲線(curve)的抽取及辨識等。
- C. 我們所發展的指紋分類模型是基於灰階影像直接做處理的。對於實用上,有些指紋擷取器取進來的影像已經是二值化的影像時,只要改變紋路的定位法,在分類模型完全不須要改變的情況下依然可以完成指紋的自動分類。這代表著我們所設計的模式可以彈性的使用於灰階影像或是黑白影像。

6. 參考文獻

- [1]. Jeng-Horng Chang and Kuo-Chin Fan, "Fingerprint Ridge Allocation in Direct Gray-Scale Domain," Pattern Recognition, Vol. 34, No. 10, pp. 1907-1925, August 2001.
- [2]. N. K. Ratha, K. Karu, S. Chen and A. K. Jain, "A Real-Time Matching System for Large Fingerprint Database," IEEE Trans. Pattern Analysis Mach. Intell., Vol. 18, No. 8, pp. 799-813, 1996.
- [3]. A. K. Jain, S. Prabhaker and L. Hong, "A Multichannel Approach to Fingerprint Classification," IEEE Trans. Pattern Analysis Mach Intell., Vol. 21, No. 4, pp. 348-359, 1999.
- [4]. M. M. S. Chong, T. H. Ngee, L. Jun and R. K. L. Gay, "Geometric Framework for Fingerprint Image Classification," Pattern Recognition, Vol. 30, No. 9, pp. 1475-1488, 1997.
- [5]. J. L. Blue, et al, "Evaluation of Pattern Classifiers For fingerprint and OCR Application," Pattern Recognition, Vol. 27, No. 4, pp. 485-501, 1994.
- [6]. K. Karu and A. K. Jain, "Fingerprint Classification," Pattern Recognition, Vol. 29, No. 3, pp. 389-404, 1996.
- [7]. D. Miao and D. Maltoni, "Direct Gray-Scale Minutiae Detection In Fingerprints," IEEE Trans. Pattern Analysis Mach. Intell., Vol. 19, No. 1, pp. 27-39, 1997.
- [8]. X. Zhuang, T. Wang and P. Zhang, "A Highly Robust Estimator through Partially Likelihood Function Modeling and Its Application in Computer Vision," IEEE Trans. Pattern Analysis Mach. Intel., Vol. 14, No. 1, pp. 19-35, Jan. 1992.
- [9]. Z. M. Kovacs-Vajna, "A Fingerprint Verification System Based on Triangular Matching and Dynamic Time Warping," IEEE Trans. Pattern Analysis Mach Intell., Vol. 22, No. 11, pp. 1266-1276, 2000.
- [10]. X. Zhuang, Y. Huang, K. Palaniappan and Y. Zhao, "Gaussian Mixture Density Modeling, Decomposition, and Applications," IEEE Trans. On Image Processing, Vol. 5, No. 9, pp. 1293-1302, Sep. 1996.