

整合智慧型手機圖文辨識與圖書館系統之 行動資訊平台

黃純敏	莊倫瓊	徐佩均	鄭仔含	林志成
雲林科技大學	雲林科技大學	雲林科技大學	雲林科技大學	雲林科技大學
資訊管理系	資訊管理系	資訊管理系	資訊管理系	資訊管理系
副教授	大學生	大學生	大學生	大學生
huangcm@	u9623014@	u9623022@	u9623024@	u9623028@
yuntech.edu.tw	yuntech.edu.tw	yuntech.edu.tw	yuntech.edu.tw	yuntech.edu.tw

摘要

利用強大的運算能力協助處理影像資料是當前最重要的應用研究領域之一。許多針對字元辨識方面的技術已被提出，但這些技術往往侷限在單純的字元辨識或是應用於翻譯等功能，並未能在生活方面有更廣泛的應用。由於現今的技術僅可用圖片辨識或二維條碼技術來進行圖書搜尋，而二維條碼技術需下載指定的配套軟體，多了些不方便性。鑑於此，本計畫利用智慧型手機的照相功能、圖文辨識技術、光學字元辨識(OCR)及3D虛擬實境技術，與圖書館系統做結合，並以雲科大圖書館為實作範例，建構成**BooksEasy(不可思議)**此一機制，若使用者在任何地點看到喜愛的書籍，只需以手機拍攝書本封面的標題，就可顯示圖書館是否有此書籍，並隨即導引及展示存放地點，以便使用者快速找尋此書。務期使用者不需其它的配套軟體，以簡單的按鈕即可得到比價資訊，而若無消費意願，系統也提供圖書館資源，讓使用者有更多元的選擇。

關鍵字：圖文辨識技術、光學字元辨識、圖書館利用、智慧型手機、3D虛擬實境

Abstract

With the rapid progress of digital image technology, more and more images are stored and displayed on the Internet. Using powerful computing capability to help image processing has become one of the most important researches in web-based application. There are a lot of techniques which aim at character recognition have been published, but these techniques usually focus on character recognition or the function of translation. They do not widely apply in our daily life. This mobile-based project, "BooksEasy", plans to combine the techniques of photography, pattern recognition technique, Optical Character Recognition, and 3D Virtual Reality with the library of National Yunlin University of Science and Technology in Taiwan as our test case. Due to the limitation of OCR, we focus on image preprocessing in order to enhance image recognition. Under this design concept, the title of a physical book will be captured when a user presses a photograph button and it will be converted to a textual query for searching the nearby library instantly. It is

expected that users will be able to visualize the shelving location of the needed book from the screen of their mobile phone wherever they are. Furthermore, this study will also connect to e-shop for price comparison when the target book is not found.

Keyword : Pattern recognition, Optical Character Recognition (OCR), Library Application, Smart phone, 3D Virtual Reality

1. 前言

在3C產品愈來愈發達的社會，現代人幾乎人手一機，也愈來愈依賴科技產品帶來的便利，造就了手機不離身的現象。而權威網路分析師瑪麗米克(Mary Meeker)也將觀察焦點集中於智慧型手機及行動網路的未來，她預測智慧型手機出貨量將在一年內先超越功能性手機，接著，全球智慧型手機出貨量將在2012年超越個人電腦出貨量，達到5億支("智慧型手機打開資訊革命的入口,"2010; 楊欣霖, 2010)。目前智慧型手機的功能完備化與低價化確已引爆另一波需求，使其年成長率較上年度大幅增加42.5%。說明了智慧型手機已成為現代的掌中主流。

一項針對劍橋大學 (Cambridge University) 與開放大學 (Open University) 的調查(梁鴻翔, 2009)顯示，讀者希望能使用行動電話找到圖書館資訊/服務，而隨著網路通訊科技的快速發展以及行動通訊裝置的日漸普及，讀者對於獲得資訊所願意等待的耐心已愈來愈少，總是希望能立即取得最需要的資料。而在過去，人們只能在電腦平台上搜尋圖書的相關資訊，缺

乏了方便性及機動性，而手機上網雖然越來越普遍，但多需要使用搜尋器，以便在眾多資訊中找到符合需求的網頁，若可使用手機隨手拍攝擷取圖書封面的書名，即可直接連結圖書館及書店找尋該標的物，對提高圖書館使用及促進網路行銷應有實質的正面效益。然而以目前現階段OCR的技術來說，雖然可以達到不錯的辨識效果，但圖片若過於複雜或不清楚，必然會影響到辨識效果，故本研究著重於圖片的前置處理，利用影像處理技術將圖片做前置處理，降低雜訊及複雜程度，理論上應可提高OCR結果的準確率，進而增加查詢的精準度。如此，當使用者出門在外不管是書店或熱門排行榜的海報，藉由本系統可快速得到書籍資訊，那將會是生活上的一大便利。

由於現今的技術僅可用圖片辨識或二維條碼技術來進行圖書搜尋，而二維條碼技術需下載指定的配套軟體，多了些不方便性。鑑於此，本計畫將手機拍照功能結合圖文辨識及 OCR 技術，建構成 **BooksEasy(不可思議)**此一機制，務期使用者不需其它的配套軟體，以簡單的按鈕即可得到比價資訊，而若無消費意願，系統也提供圖書館資源，讓使用者有更多元的選擇。

本研究將在第二節回顧相關技術。我們在第三節探討本系統的研究方法及步驟、系統實作，於第四節討論評估，並以第五節作為結論及勾勒未來展望。

2 文獻回顧與探討

2.1 圖書館系統應用

隨著技術的發展，傳統的圖書館服務已難以滿足用戶的需求。現代快速的生活

步調與電子資料的普及，再加上難以攜帶的厚重紙本，使得越來越少人會想要造訪圖書館(Yue, 2010)。如何擴充及改善圖書館之功能及如何提供讀者更方便的服務一直是圖書館員追求的目標(Yang, 2004)。而在近幾年，行動裝置的快速發展為擴大圖書館服務提供了一個良好的平台給讀者，因此，使用WAP技術來設計和實現一個功能更豐富的行動圖書館系統便有其可行性。行動圖書館基於 WAP技術將實現動態發布的信息，信息的查詢、更新、虛擬參考服務、電子資源的上傳下載和以搜索權限來存取電子文獻資料庫等。這些功能有助於提高圖書館的服務，以充分地利用圖書館資源(Yue, 2010)。

赫爾辛基科技大學在一次的讀者意見調查中，詢問了讀者擁有行動電話的情形，以及讀者對於透過行動電話收取圖書館資訊的意願，調查的結果顯示，95%的讀者擁有行動電話，而且表示願意透過行動電話獲得圖書館資訊。過去，當我們在圖書館搜尋或存取資訊(例如：書籍、論文或雜誌)，多只能使用館內的電腦來完成上述動作，這帶來了些許的不便。由於硬體設備的不足，讀者必須等待電腦無人使用時，方能查詢所需資訊；當讀者查詢到所需資訊時(例如：索書號、架位號碼)，還必須將它們抄錄於紙本上，這不僅造成讀者的不便，也造成了紙資源的浪費(Rong-Yuh, 2002)。

多年來國內外都有與圖書館數位化相關的研究及發展不斷的被提出，像是美國南阿拉巴馬大學書館 (South Alabama University Library) 的「無屋頂圖書館計畫」(The Library without a Roof Project) (存慮紀念圖書資訊館, 2004)、電子書服務(龐文貞, 2011)等等，都如雨後春筍般，引領圖書館數位化的風潮。由這些例子來看，圖書

館數位化已經是眼前的趨勢了，然而可能因為應用在國內各大學圖書館的機制仍不夠方便，學生使用率並不高。鑑於此，本計畫以突破性觀點，利用方便的行動上網機制，做為吸引學生使用圖書館內資源的誘因，希望幫助圖書館推廣書籍利用，達成本計畫的目標。

2.2 圖形辨識及影像處理

圖形辨識在多年前就已經存在了。最早應用於數學統計方面，1930年 Fisher 提出了一種判別分析方法，他利用機率分布函數中兩種已知的向量去定義決策規則分類這兩種向量，發展出統計型樣識別 (statistical pattern recognition)。在統計型樣識別中，圖形辨識的主要功能在於以最小的風險選擇模型的類別。1974年，Vapnik 提出“最佳化超平面分類”及支援向量機 (Wang & Chen, 2003)，它是一種監督式學習的方法，可從已知的事物中，經於訓練，預測未知的事物並做分類的動作，也就是透過以往的經驗來對未知的事物進行判斷，這也是圖形辨識的宗旨。

圖形辨識從侷限用在數學統計方面上，逐漸用在實用的地方，包括字元識別、醫學影像分析，軍事情報、語音識別...等 ("Pattern Recognition and Image Processing," 2006)。而圖形辨識的應用雖然廣泛，但精神卻是一致的，其流程大抵可分為：影像處理、特徵表現、特徵選取或特徵萃取、分類以及系統辨識率測試幾個步驟。

影像處理的目的是將任何輸入的圖像形式做信號處理，例如照片或影像畫面，輸出的圖像處理可以是一張圖片或具有特徵性的圖像。從上述圖形辨識的流程便得知，為了讓特徵表現、特徵選取等之後的流程有更精確的結果，完善的影像處理

是必要的(Wu, Pang, & Han, 2011)。Jian Yuan 等人(Jian, Yi, Kok Kiong, & Tong Heng, 2009)提出了影像前處理的三步驟，圖 1 為此三步驟之流程圖。



圖 1 影像前處理三步驟

1. 灰階處理：將彩色圖像轉化為灰階圖像。此處理會降低圖片的複雜性及減少記憶體的需求。而一張彩色影像可展開為一個色彩像素(color pixels)的 $M \times N \times 3$ 陣列，其中每個色彩像素在一特定的空間位置中，可對應到一張 RGB 影像之紅、藍、綠成分的三合一的點。一張 RGB 圖像可視為三個灰階圖像的一個「堆疊」，灰階轉換使用統計式的方法將圖片轉換成數值範圍為[0,255]的 8 位元影像。而將彩色影像轉換成灰階的公式如公式(1)及公式(2)：

$$Y = 0.299R + 0.587G + 0.114B \quad (1)$$

$$Y = (R + G + B) / 3 \quad RGB = (0 \sim 255) \quad (2)$$

2. 圖像平滑化：利用低通濾波來降低圖像雜點，低通濾波能模糊圖片中破碎的字元，並填補輸入圖片中的小縫隙。由於影像中的背景物件，時常出現與文字有類似的紋理特徵，故使用低通濾波對影像產生模糊性，使非文字的紋理特徵與文字的紋理能有較大的特徵區別。
3. 增強對比度：在圖像的背景顏色與文字顏色相近或光線不足的條件下，增強對比度將顯示圖像的特點，有利於之後的辨識。

影像前處理方法甚多，除了上述方法外，Otsu (Mitra, 2002)也提出了一個二值化方法，能自動選取一個最佳的閾值將灰階影像中所有像素點分成兩群：前景資訊與背景資訊，同時使得群內變異最小(群間變異最大)。希望將上述之影像處理結合此二值化方式，來取得最理想的文字影像。下圖(圖 2)為本計畫在 OCR 前之影像處理流程圖，以增加 OCR 準確性。

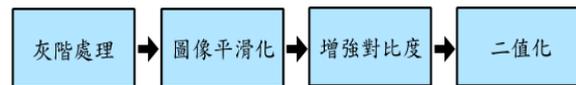


圖 2 二值化之影像處理流程

2.3 光學字元辨識(OCR)

OCR的主要用途是針對既有的文件作文字識別，並將其轉換成電腦所能認識的電子訊號。此技術已被廣泛的應用於書籍與文件的轉換，並將轉換過後的形式發佈到網路上分享。OCR也讓我們能夠對紙本上的文字進行編輯，並能簡易的在文章中搜尋及存取特定的文字或句子，減少我們用肉眼逐字搜尋的時間。

OCR的歷史相當悠久，早在1929年 Gustav Tauschek 即在德國取得OCR的專利，並在1970年代時提出了許多關於OCR的應用。而OCR之軟體則是在1990年代時第一個被圖書館用來辨識歷史悠久的報紙資料(Holley, 2009)。發展至今最成功的應用為圖形辨識、電腦視覺與機器學習，而雖然技術日漸純熟，但仍有無法辨識的部分(Jianli, Nugent, Bowen, & Bowen, 1993)，因此，也有許多學者漸漸的提出與改善OCR相關之論文(Holley, 2009)，其目的都是為了提升OCR的準確率。

OCR 被視為圖形辨識 (Pattern Recognition)的一環，相關論文也一直被提出，使用也相當廣泛，像是辨識音樂

符號(Yakobov, Mash, & Thirer, 2010)、汽機車牌的辨識研究(王振興, 2003), 甚至醫學上也不乏有相關論文, 如骨架擷取(謝明勳, 2004)...等等, 當然在手機上的應用(王炫盛, 2008)更是不在話下。

然而目前由相機所拍攝的影像, 並不能立刻進行 OCR 的辨識, 必須經過幾個影像處理步驟才能將有用的資訊抽取出來。為了提升 OCR 辨識的精準率, 本計畫焦點集中於對影像做前置處理來減少雜質、增加原圖清晰度, 其後再利用 Microsoft Office Document Imaging 所提供之 OCR 元件對書名部分進行文字的辨識, 希透過前置處理的努力, 提升整理辨識效果。

3 研究方法及步驟

3.1 研究方法：

本計畫利用軟體工程原理與技術實作此一資訊系統。基於時間的限制, 本計畫研究將以系統雛形法做為快速開發的方式。系統雛形法優點在於開發快速、成本較低適用於本計畫, 再者因為 Android 系統在技術方面, 整合了作業系統、人性化介面及應用程式, 且軟體可做堆疊 (Software Stack), 所以採用 Android 系統平台配合 OCR 來做為軟體開發的方法。

3.2 研究步驟：

3.2.1 系統架構

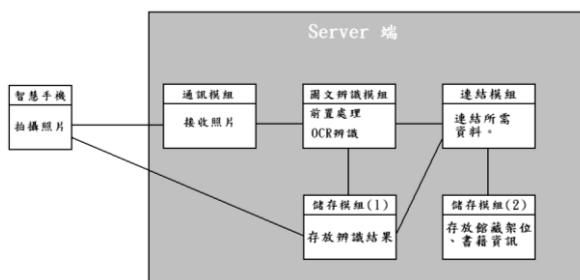


圖3 系統架構圖

在本研究中所規劃的系統架構圖 (圖 3) 顯示, 使用者利用行動上網將智慧型手機所拍攝的書籍封面, 由通信模組將圖片上傳到伺服器。伺服器接收後, 驅動圖文辨識模組、儲存模組和連結模組。圖文辨識模組將照片圖文分離並解析, 利用 OCR 轉換成文字檔, 記錄在儲存模組, 使用連結模組連上資料庫系統進行書籍查詢相關資料。各模組的功能詳述如下：

1. 通信模組：

智慧型手機運行此系統, 進行照片拍攝之後, 會將準備傳送的圖片丟入緩衝備妥, 並利用傳輸介面將資料透過 socket 封包送到 server 的圖文辨識模組中。

2. 圖文辨識模組：

從通信模組接收到的書籍照片, 利用圖文辨識模組, 進行圖文分離動作, 使用 Microsoft Office Document Imaging 的 OCR 元件去分析文字部份的圖片, 是否達到能夠辨識的程度, 如果不能則請使用者重新拍攝, 直到可以成功辨識為止, 再將辨識出來的文字數據丟給儲存模組。

3. 儲存模組：

伺服器端的接受介面, 此模組持續開啟 socket 傾聽網路連線, 當圖文辨識模組辨識出文字後, 即將文字存入資料庫做為個人查詢紀錄, 並將館藏地點詳細地圖存放於資料庫中, 以供後續查詢使用。

4. 連結模組：

伺服器端的程式, 此模組主要功能是連結到資料庫抓取相對映的資料 (EX: 書名、書籍位置、實景圖、帳號) 來查詢所需資訊。

3.2.2 系統實作

在系統設計步驟完成後，依各個模組及子功能將本研究擬開發實作的系統分為三個階段：

(1)第一階段：

本階段開發圖文辨識模組，利用手機照相功能將圖片傳至Server，程式已限制解析度為320*240，再利用灰階、低通濾波及二值化等技術，把圖片背景移除，只存取文字部份的圖片，接著用Microsoft Office Document Imaging的OCR技術，將文字圖片轉換成文字檔，並測試轉換成文字檔資料的正確性，如圖4-圖5。



圖 4 拍攝畫面

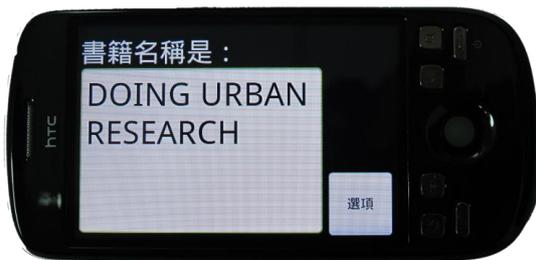


圖 5 確認辨識畫面

(2)第二階段：

本階段將建置圖書館資料庫及個人帳戶資料庫，其中圖書館資料庫需建置平面圖、3D 虛擬實境、實景圖，而 3D 虛擬實境的建模工具為 Google SketchUp。個人帳戶資料庫為我的紀錄功能使用，此功能記錄使用者的查詢，除了增加也可以刪除紀錄，可供使用者重複點選查看，不需要再經由

辨識，可以減少使用者的操作時間。

(3)第三階段：

本階段將擷取出的文字檔，利用智慧型手機的行動上網功能，連結上圖書館資料庫系統，使用書名查詢書籍資料，最後將查詢結果以平面圖、3D 虛擬實境、實景圖顯示於手機上，如圖6-圖8所呈現。

當使用者想購買書籍時，可用本系統的網路比價功能連結到FindBook網站讓使用者進行比價以購買到較便宜的書籍(圖9)。若查無此書，可使用推薦功能，此功能即連結到雲科大圖書館之推薦頁面，若申請成功，將可成為此書的第一位借閱者。

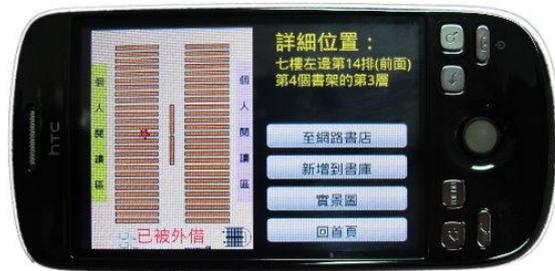


圖 6 查詢架位畫面



圖 7 3D 虛擬實境畫面



圖 8 實景畫面圖



圖 9 網路書店比價畫面

4 系統評估

4.1 圖文辨識模組測試及評估

本研究使用國立雲林科技大學圖書館的館藏當作測試對象，對選擇的書籍進行拍攝，而使用 Microsoft Office Document Imaging 的 OCR 元件來進行辨識，由於 OCR 元件的辨識率受背景的影響極大，甚至造成辨識失敗，鑑於此，本研究對圖片進行前置處理，去除雜訊、簡化背景，來提升 OCR 的準確率，並證明有做前置處理的準確率會比直接拍攝所辨識出的結果來的更加準確。

為了驗證經過前置處理之圖片辨識出來的準確率會比未做來的高，本研究從圖書找了四十本書籍來實驗，因實驗書籍眾多，舉例十本書籍所實驗出的結果，如圖 10-圖 13：

	書籍圖片	使用 OCR 進行辨識
原圖		 辨識結果：Enhjci1g Quffiy in Ass_ssment
處理後		 辨識結果：Enhancing Quality in Assessment

圖 10 辨識結果一

	書籍圖片	使用 OCR 進行辨識
原圖		 辨識結果：辨識失敗
處理後		 辨識結果：English Literature

圖 11 辨識結果二

	書籍圖片	使用 OCR 進行辨識
原圖		 辨識結果：hi
處理後		 辨識結果：Instrumentation

圖 12 辨識結果三

	處理前辨識結果：Hali.V.V00n62a The
	處理後辨識結果：The GREEN-TRAVEL 501 WCUOOK
	處理前辨識結果：辨識失敗
	處理後辨識結果：msiDe ASla
	處理前辨識結果：辨識失敗
	處理後辨識結果：Inigo Jqnes
	⋮
	處理前辨識結果：RES_D LAYOUT
	處理後辨識結果：RESIDENTIAL LAYOUT
	處理前辨識結果：辨識失敗
	處理後辨識結果：BEYOF FOR
	處理前辨識結果：HE ING OK
	處理後辨識結果：THF DRAWING _OOK4
	處理前辨識結果：Hemtry V4uLARY Games
	處理後辨識結果：Elementary VOCABULARY Gaines

圖 13 辨識結果四

經過影像前置處理後，約有八成左右書籍的辨識準確率比原先未處理來的高，相對的減少需要修正錯字的機會。而有近三成的書籍封面辨識準確率不高，歸因於一些因素，如背景過於雜亂、字與背景過於相似或特殊字體的關係等，使得結果呈現並不甚理想。如圖 14-圖 16 所示：



圖 14 背景過於雜亂的圖片

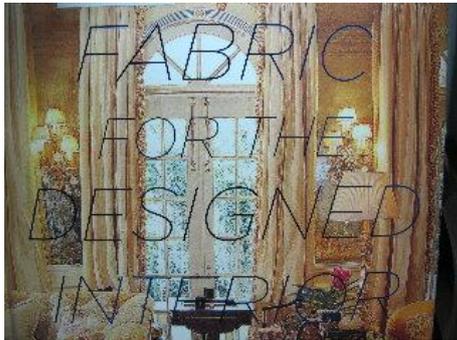


圖 15 字體不明顯的圖片



圖 16 特殊字體的圖片

本研究從這次的實驗中，實證了前置處理確能提升圖片辨識效果的假設，準確率從原先的 32.5% 提升到 84.5% 左右，使得因圖形的雜訊造成辨識不高的機會大幅的下降，讓之後 OCR 辨識的精準度更加可靠。

5 結論與討論

5.1 系統特色

本系統兼具商業價值、創新性及擴充

性三大特色。系統結合圖書館及網路書店，當系統查無此書，可藉由此途徑進行各大網路書店之比價，並可連結至該書店購買書籍，因此極具商業價值；在創新性方面，由於現今的技術主要以圖片辨識(以圖找圖)、二維條碼技術、ISBN 碼查詢，但各自都有不同的缺失，造成許多不便。而本機制只需拍攝書籍名稱，資料來源可從一般書籍或任何海報上的熱門排行，文宣上的書名...等等，藉由本系統所結合之 OCR 及圖文辨識技術，即可查詢館藏架位，也不需搭配任何配套軟體；最後在擴充性的部分，本系統可搭配資訊檢索服務與協定(CNS-13416, Z39.50)，未來若查無此書，可用此查詢全國各地之圖書館藏，且未來若結合網路社群，提供一個平台給使用者，使資訊快速交流，進而營造出愛書的書香社會。

5.2 未來展望

由於本系統在實作上，是以大學圖書館做為雛型，未來如有需要也可以應用在實體或是虛擬書店。此外，若加入社群功能之後，藉由眾多的使用者分享心得或舉辦活動，來吸引廣告商購買廣告放置，也可讓網路或實體書店發起團購、周邊商品。期許提供更完善功能，增加更多便利的功能如：連線到電子書庫、書籍片段內容試讀、借閱到期及預約可取書時以簡訊通知、圖書館書籍排行榜...等等，使系統更完整、服務更完善。

6 致謝

本計畫承蒙國科會經費補助與支持，使得研究得以順利進行，謹此致謝，計畫編號: NSC 99-2815-C-224-020-E。

參考文獻

- [1] 王炫盛. 應用於智慧型手機之快速字元辨識技術. 2008.
- [2] 王振興. 多標的汽機車車牌辨識系統之研究. 2003.
- [3] 存廬紀念圖書資訊館. 93 年冬季第六期, 2004.
- [4] 梁鴻翔. 淺談 M-Libraries 行動通訊圖書館服務. 中華民國圖書館學會會訊 17 卷 2 期, 2009.
- [5] 智慧型手機打開資訊革命的入口. 數位時代 *Business Next* 195, 2010, p60.
- [6] 楊欣霖. 抓住 2011 科技產業 10 大趨勢. 數位時代 *Business Next*, 199, 2010, p58.
- [7] 謝明勳. 應用階層式修正型自我組織特徵映射圖網路(HMSOM)於骨架擷取. 2004.
- [8] 龐文貞. ebook business=demand business. 數位時代 *Business Next* 2011 三月號, 202, 2011, p145.
- [9] Holley, R. Analysing and Improving OCR Accuracy in Large Scale Historic Newspaper Digitisation Programs *D-Lib Magazine*, 15, 2009.
- [10] Jian, Y., Yi, Z., Kok Kiong, T., & Tong Heng, L. *Text extraction from images captured via mobile and digital devices*. Paper presented at the Advanced Intelligent Mechatronics, 2009. AIM 2009. IEEE/ASME International Conference on, 2009.
- [11] Jianli, L., Nugent, J. H., Bowen, D. G., & Bowen, J. E. *Intelligent OCR editor*. Paper presented at the Electrical and Computer Engineering, 1993. Canadian Conference on, 1993.
- [12] Mitra, P. Density-Based Multiscale Data Condensation. *IEEE*
- [13] *Pattern Analysis and Machine Intelligence*, 2002.
- [14] Pattern Recognition and Image Processing. *Computers, IEEE Transactions on C-25* (12), 2006, p1336.
- [15] Rong-Yuh, H. *The design and implementation of mobile navigation system for the digital libraries*. Paper presented at the Information Visualisation, 2002. Proceedings. Sixth International Conference on, 2002.
- [16] Wang, S.-J., & Chen, X. *Biomimetic (topological) pattern recognition - a new model of pattern recognition theory and its application*. Paper presented at the Neural Networks, 2003. Proceedings of the International Joint Conference on, 2003.
- [17] Wu, Q.-E., Pang, X.-M., & Han, Z.-Y. Fuzzy automata system with application to target recognition based on image processing. [doi: DOI: 10.1016/j.camwa.2010.08.101]. *Computers & Mathematics with Applications*, 61(5), 2011, p1267-1277.
- [18] Yakobov, V., Mash, L., & Thirer, N. *On chip implementation of an OCR algorithm for musical notation*.

Paper presented at the Electrical and Electronics Engineers in Israel (IEEEI), 2010 IEEE 26th Convention of,2010.

- [19] Yang, Z. B. A. Research of mobile electronic commerce resolution based on wap. *Computer Applications and Software, Oct.*

2004,2004, p103-107.

- [20] Yue, H. Research and development of the mobile library based on WAP technology. *Industrial Mechatronics and Automation (ICIMA), 2010 2nd International Conference, 2, 2010.*