



【中文光學辨識系統簡介】

劉秀惠

一、前言：

本館自10月18日開始一連舉辦為期四天的中文光學辨識系統展示，此活動已於10月21日結束。感謝讀者們在活動期間對展示活動的支持和參與。也感謝此次參展的廠商—鼎展、新穎、冠東及力捷（承能）等四家，謝謝他們贊助這次中文光學辨識系統展示活動，帶給大家有關中文光學辨識系統的相關資訊。

二、有關中文光學辨識系統：

中文光學辨識系統（OPTICAL CHARACTER RECOGNITION，簡稱OCR）即利用個人電腦及掃描器自動辨識印刷在紙上的中文文字。就科學技術而言，中文OCR是屬於圖形辨識和人工智慧的結合；在應用上，是以自動輸入方式處理文字資訊，如此可大幅減輕人工輸入的負擔，對於不諳中文輸入法的人及需要大量處理文字的工作也有很大的幫助。

中文OCR包括二個部份，分別為中文光學文字辨識軟體及掃描器。簡單說明如下：

(一)中文光學文字辨識軟體

目前中文OCR軟體共有四家，即此次本館邀請參展的四家廠商，其中有三家係代理或與大陸合作的產品，同時可辨識繁體字及簡體字，只有力捷的丹青OCR是台灣本地廠商研發出來的產品。此版的OCR除了無法辨識簡體字之外，其系統環境也與其他三家不同，因其工作環境是在WINDOWS環境之下，而鼎展與冠東只有DOS版本，新穎則二者皆具。

(二)掃描器

除了中文OCR軟體所能辨識的字形及字數的多寡會影響辨識的正確率之外，掃描器的解析度也會影響到中文OCR辨識的使用效率。理論上掃描器的解析度是愈高愈好，但解析度愈高，所需記憶體也愈大；同時愈高解析度的辨識所花的時間也較長。所以影響辨識能力的主要因素乃在辨識軟體，其他因素則為掃描器、PC環境、紙張、印刷是否清晰及文字是否擺正等。

三、建議：

本文僅就中文OCR作一概略的介紹，關於讀者在展示活動期間所提的軟、硬體等問題，請洽有關廠商，恕無法在此回覆。各位讀者若想再對中文OCR有更深入的了解，您不妨再翻翻以下參考資料，希望對您在觀念上與選擇上有所幫助。

- 1.0&1中文Byte雜誌 1993年4月號
- 2.資訊傳真 1993 Sep.27. pp.52-53.
- 3.資訊傳真 1993 Oct. 4. pp.46-47.