ƒ ‹ F | Œfia‹ ˙'e›ß• |–M^D‹ ª¤s›p e ƒ¤"G‡ł§i

§Q¥ ¿»y› ¿°^ § ‡N« ¥ ¥i•j·M/§ƒ » ‡ ‹ –– ˙…˘ƒ …v› ,Œ

fi˘fiw (3/3)

›p e^ §¡G §O«‹ ›p e

›p e‰ ¸s„¡GNSC91-2516-S-038-001-

ıƒ · ¶¡¡G91 ƒ~08 /°01 / ƒ 92 ƒ~07 /°31 /Ø

ıƒ ‡ ƒ ¡G» O ¥´ ˙/j ˙ƒ » ‡ ‹

›p e¥ D ø/¡HG§ ı ¸›˜P

ƒ ƒP¥ D ø/¡HGfi}« • ¡A§ı/˝–M

›p e »P/H›ß¡G" L de A

‡ł§i^ «‹¡G§ „ª‡ł§i

‡ B † zŁ¡fG¡¥»›p e¥i/ ‰ ¶ ‹}d,

/ / ¥` Œ 92 ƒ~11 /°1 /Ø

行政院國家科學委員會補助專題研究計畫成果報告

※※※

中 華 民 國 92 年 10 月 1 日

行政院國家科學委員會專題研究計畫成果報告國科會

利用語音辨識技術建立可搜尋之耳鼻喉科教學數位影音資料庫

Building a Searchable Digital Video Database for Otolaryngology Teaching by using Speech Recognition Technology (2-2)

計畫編號：NSC 91-2516-S-038-001-

執行期限：2002 年 08 月 01 日至 2003 年 07 月 30 日

主持人：李飛鵬 臺北醫學大學/臺北醫學大學附設醫院

共同主持人：李友專 臺北醫學大學醫學資訊研究所

共同主持人：徐建業 臺北醫學大學醫學資訊研究所

計畫參與人員：林明錦 臺北醫學院醫學資訊研究所

中文摘要

本計畫主要之目的在於利用語音辨識技術,將耳鼻喉科在教學體影片中的語音轉換成文字. 配合本計劃所開發之軟體,將文字標註時間碼, 利用網路及多媒體資料庫技術,我們可以藉由搜尋特定文字來觀賞影片段落,如此一來學生及醫師即可利用網路快速,不受地點及時間限制,來學習耳鼻喉科教學.

Abstract

The purpose of this project is to create and integrate multimedia medical resource for medical community. In this project, we realize the ° ßpeech to text° ®technology by voice recognition. In this project, the user can access any section of video clip randomly by searching the time code based text. Due to his, the user who using this system can watch the movie clips more efficiently.

Keywords

Otorhinolaryngology, voice recognition , Digital Video, WWW

二、緣由與目的

隨著數位科技的日益成熟,越來越多的影音資料,在我們的生活中扮演了越來越重要的角色. 也因此,人們也嘗試著利用這許多資料,來作各方面的應用. 首先遇到的問題,就是如何做索引及搜尋的工作. 日益龐雜的影音資料,甚至是 3D 立體模型,比起過去的文字檢索,其複雜度更是不可同日而語。而手術錄影帶是專科醫學教育一項重要的資產，傳統的錄影帶都是以類比方式錄影，播放則是線性方式從頭開始到尾。近年來由於數位媒體儲存的進步，影音資料有從類比轉成數位，從線性轉成非線性的趨勢。配合寬頻網路及影音伺服器的架設，可使多人多處同時觀看錄影內容，達到專科醫學遠距再教育的目的。

所謂非線性的播放方式即是可搜尋之影音資料，使用者可依喜好找尋需要之錄影片段，不需從頭到尾把錄影帶看一遍。這種技術在非醫學方面有個很好的例子，就是在柯林頓誹聞案大審判時的錄影內容，Virage公司和AltaVista合作，成功的運用VideoLogger軟體製作成可搜尋之錄影內容，使用者可以文字尋找某一特定的錄影片段，這是第一個在網路可搜尋之影片。

耳鼻喉科手術由於視野狹小，故常常使用顯微鏡或內視鏡，接至錄影設備後便可將手術過程完整的紀錄下來。常見的手術如耳部的鼓室成型術(tympanoplasty)，鼻部的功能性鼻竇內視鏡手術(functional endoscopic sinus surgery)，喉部的喉頭直達鏡顯微手術(larygomicroscopic surgery)。以鼓室成型術(tympanoplasty)為例，一台手術通常需要兩個小時，如果使用者只對聽小骨中的磴骨(stapes)處理有興趣，如何不需從頭找起呢？

手術錄影帶是醫學教育一項重要的資產，傳統的錄影帶都是以類比方式錄影，播放則是線性方式從頭開始到尾，這些影片在尋找使用上非常不便,而且影帶會隨著時間耗損，無法永久保存。近年來由於數位媒體儲存的進步，影音資料有從類比轉成數位，從線性轉成非線性的趨勢，而這些非線性的影音資料的index 及search便成為一個熱門的課題。

結果與討論

本研究計畫之目的在研究及發展一可檢索搜尋及互動之耳鼻喉科手術數位影音資料庫；並結合網際網路，建立一影音伺服器，具備多媒體互動教學功能，以協助耳鼻喉科醫師再進修或住院醫師訓練，提昇醫療的品質。在第三階段的計劃中，我們希望建立開放的搜尋標準，能使本計畫中的醫學多媒體資料能夠透過標準的架構來交換，本年度計畫結合查詢引擎與後端平台程式透過此一資訊交換標準可以將詮釋加入不同離線多媒體(Off-line Multimedia)更有助於醫學學術研究結果之公開分享。

計畫成果自評

本計劃之目的在於建立可搜尋耳鼻喉科的多媒體教學網站,我們已成功的將語音辨識軟體,應用在耳鼻喉科的手術錄影帶之中,將醫學的專用名詞辨識出來,並配合時間碼的標定,可以藉由 speech to text 的功能,搜尋文字來定位影片.

在今年度計劃中,我們共完成了:

1. 建立開放檢索介面 Request Encoder /Decoder engine :
   在目前資料庫中有標準的 Structure Query Language (SQL) ，我們藉由此計畫建立醫學多媒體的查詢標準，尤其本計畫為之主要內容為手術錄影帶教學，對於醫學有相當重要的地位。
2. 建立網際網路傳輸引擎 HTTP protocol engine:
   我們藉由網際網路的便利性，除了透過我們的網頁查詢資料，也可以將我們的資料透過 HTTP protocol 的傳遞，將影片的詮釋資料透過如 PDA 或不同離線媒體的傳播。

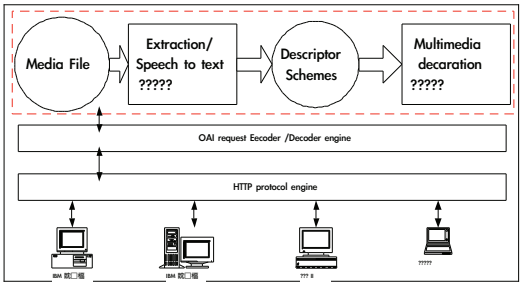

Figure1. Framework of Open Archive Initiative protocol

本計劃團隊完成了第三年所預期的工作目標,並有論文之發表,完全符合計劃書中的預估完成工作成果. 本計劃所開發之軟體與論文也將放置於網站 (http://ENT.TMU.EDU.TW),公開於學術研究使用.

Reference

1. Brazma A, Hingamp P, Quackenbush J, Sherlock Gavin, et al. Minimum information about a micro array experiment – toward standards for microarray data. Nature Genetics (December, 2001):pp.365-371.
2. Digital Imaging and Communications in Medicine (DICOM) Version 3.0. American College of Radiologists / National Electrical Manufacturers Association (1993).
3. Hsu CY, Chen PH, Lee PS, Lin MC, Li YC. Development of a Digital Museum for Human Body. Proceeding of the 3rd International Conference of Asian Digital Library (December, 2000): pp.267-274.
4. Lagoze C, Van de Sompe H. The Open Archives Initiative: Building a low-barrier interoperability framework. Proceeding of the ACM/IEEE Joint Conference on Digital

Libraries (June, 2001): pp.54-62.

5. Li YC. Toward a medical information collective: trends in the development of digital libraries in medicine. Yearbook of Medical Informatics 2001(2001): pp.77-82