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

¡ ¡ ¡ ¡ ¡ ¡ ¡ ¡ ¡ ¡ ¡ ¡ ¡¡ ¡ ¡ ¡

¡ ¡

¡ § Q ¥ ¸ » y › ¿ ° ^ § ‡ N «¥‹ ¥i•j..:M/§ƒ »..‡ ›,Œfi¡˘fiw

¡ ¡

¡ ¡ ¡ ¡ ¡ ¡ ¡ ¡ ¡ ¡ ¡ ¡ ¡¡ ¡ ¡ ¡

›p e^ § O ¡ G ¡ ››p§Œ«‹ ¡ … ƒX«p e

›p e‰ s „ ¡NGSC 90-2511-S-038-001-

ıƒ · ¶¡¡G 2001 ƒ~ 8 /° 1 / Ø 2002 ƒ~ 7 /° 31

/Ø

›p e¥ D «/¡HG§ ı ,›˜P

ƒ @ ƒ P ¥«pG§Hı´˝–M fi}« •~ " L ‡«n

¥»ƒ¤"G‡ł§i¥] ‹ A ¥ H Ʊ /‡§"'¥ ¡G

¡…›u Œ¥ ~ ¥ ¥t‹'ª¸† /–‡ł§i/ ¥

¡…› u / j ‡ ƒ a˘‹¥ª¥fit '/¸–‡ł§i/ ¥

¡…¥ ¥u Œ» ˙‡N• |˜ ‡ /–‡ł§i/ ¸ o"/§‰ / Ʊ/@¥

¡… Œ» ƒX§@‹ ª¤s›p e Œ¥~‹ ª¤s‡ł§ifi @¥

ıƒ ‡ ƒ ¡G» O ¥´ ˙/j ˙/» O ¥´ ˙/j ˙" ‡´ |ƒ »‡‹

// ¥` Œ 91 ƒ~ 10 /° 1 /Ø

ƒ ‹ F | Œfia‹ ·'e›ß• |–M^ D‹ ª¤s›p ef¤"G†ŁSi Œ• |

Š Q ¥ ¸ » y › ¿ ° ^ Š ₩‹» ¥ ¥i–•–j·M⁄Š ƒƒ … v ›, Œfi˜fiw

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

›p e‰ s „¡MŠC 90-2511-S-038-001-

ıƒ · ›¡G2001 ƒ~08 ⁄°01 ⁄ Ø 2002 ƒ~07 ⁄°30 ⁄Ø

¥ D ø⁄¡HŒS ›,˜P » O ¥´ ·⁄j ·/» ₩_´ ·⁄j ·" †]´ |
ƒ @ ƒ ₩D‹Œ¡GŠ ›˝–M » O ¥´ ·⁄j ·´ ·,Œ T‹ ª¤s'
ƒ @ ƒ ₩D‹Œ¡Gfi}« •~ » O ¥´ ·⁄j ·´ ·,Œ T‹ ª¤s'
ƒ @ ƒ ₩D‹Œ¡G" L ₣«n » O ₩⁄j·ƒ » † ‹ †¡
›p e »P⁄H›ß¡G¿ø« Š» » O ¥´ · |´ ·,Œ T‹ ª¤s'
›p e »P⁄H›ß¡G" L ‹e A » O ¥´ · |´ ·,Œ T‹ ª¤s'

⁄⁄⁄ "K›n

¥»›p e¥D›n⁄S¥ " ƒbŠ ₩¸» y › ¿ºŠ †N,–N ƒ »†‹ ƒb–– ·¯Ø…v⁄ø⁄"'» y ´ ·«ƒ¤⁄ ƒr. tƒ ¥‹»›p„"' ¶} o⁄S†n¯Ø,–N⁄ ƒ ›:— ø fi¶‰¡X, Š ₩¸" , ⁄ ƒh· CØ, Œ˜fi Š† †N,Š › ¥i¥H´ ¥ • ¡Mƒl Šw⁄ ƒr¤ [‰…v⁄ øq¸¤,ƒ p ƒ ⁄„@¤·¥˝ ⁄¸´ fi Š¥Y¥ Š₩¥¸" , S†t,⁄ £ ¤ ƒa⁄ Ifi ¶¡›› ¤ ,¤ ·† ƒ » † ‹ –– ·.

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

⁄G¡B‰t¥ »P¥ ""

H … ƒ ‹ Š "⁄Øflqf¤… ,¶V¤ ¶Vfh" …v› Œfi˜ƒ Š › ™¥˝¡⁄⁄ŠŒ" t P¶₩¶ V« ›n"¤ ƒ . ⁄ ƒ]ƒ„⁄H› ⁄ ₫`, ŠQ¥¸† ф\ ƒh,Œfi˜,¤ Š ŒU⁄ Ł–"" †¥¸. ›"¥ „J¤ "" ^D,· ₦Ofpƒ fl ⁄ ⁄¸• ¡M"⁄ Š@. ⁄Øfl ĉe ´ł"… v , ŒŁ˜,‹ ˜ƒ ‹O 3D ¥ ¯Ø… «‹,⁄æ _„L ¥h"⁄ ƒr fl`,¤ ‰˜´ ł «Š‹O⁄£¥if PØf »y¡Cƒ ⁄ †N¿ …v– ₫O–M‹ ´ ·—¤|⁄@¶ ›«›n "",Œ†£¡ A ¶†"¿ …v– ₫f‹O¥H^ æ⁄₣¡¿ …v¡A…‰' æ k‹‰u ' ℘Łƒ ąY¶}' l ¤Š¡C"æƒ~ ¤ ¥'…˜ƒ ·C¯Ø ₣s"⁄¶ ₫B¡A… v › ,Œfi†–q ^ ⁄æ´ ƒ¤… ƒ ¡ A – ₫u' ℘¤« D ₫'"⁄ ¶ ¡C ₣X… eW⁄ , ⁄„… v ƒł" A„"⁄‹ [†]¡A¥i¤˜ƒh ⁄Hƒh†BƒPfi ˜[‹ ¿ …v⁄"fie¡A„ F –M‹ ´ ›•¶ ₣A–¤|" ¥ "⁄¡C

' ¿ « ₫u"⁄… ‰ ' æŁŠY‹O¥ i • j M⁄Š…v › Œfi¡˝ A˜¥¸" ¥i¤ † ƒnŠM» ›n⁄Š¿ …v⁄ø ‹ q¡A⁄£» –q Y¤ Š Š .¿v–a‹ ⁄@„M¡C†σ" Š †N ƒb«D´ ·⁄ Ł–ƒ † › « ™⁄"¤ ⁄l¡A‹ ₦Ofb‹ _"L„y ‰ »Dfi ⁄j…fŠPfi "⁄"¿ …v⁄"fie¡AVirage⁄‰¥q'M AltaVistaƒXŠ@¡A₣¤¥\"⁄"„B¥¸VideoLogger† ñØ » Š@ƒ¤¥i• ¡M⁄Š¿ …v⁄"fiq¡A‹˜¥¸" ¥i¥H⁄ ƒr·M Š ‹Y⁄ fŁS' ₩"¿ …v⁄ øq¡A†« Φ˜⁄@› ƒ Ł , ¥i • j · M⁄Š‹⁄ø¡C

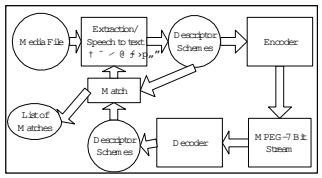
ƒ »†‹ ⁄ †N¥ ' ł † ¥ fl U ⁄ p ₳‹G–'–'¤˜ ¥¸¯ ª Ł^Ł¸" ł₳¡A– ƒ ¿ …v†]† ‹Æ‹K¥i–N⁄ †N„ L {Š„ª"⁄‹ ¿⁄₫¡G–'¤£""⁄ †N₣pƒ †¡"" „ ₰·ƒ¤«‹‹†N(tympanoplasty)¡A» †¡"⁄¥\fl '° » ˜u⁄" ł^Ł⁄ †N(functional endoscopic sinus surgery)¡A† †¡""† Y " ‰ „ F Ł¯#N•L⁄ (larygomicroscopic surgery)¡C¥H„"‹·ƒ¤«‹ †N(tympanoplasty)‹ ¤¡A@¥ ₰ †N† q'» ›n

¤ › ⁄ p fi ¡ A ƒ p " G ¤ ˇ ¥ ¸ "¥u„ ¯¥p⁄ ' ⁄ Å¥MPEG-7 ‹ Coding Scheme¡A§Q¥¸
(stapes)† Bzƒ † ‡%¡Afp⁄ £ » -q§ Y _' ?H Microsoft SQL 2000¡ A¸ Active Server Page
⁄ † N ¿…∀ aO´ ·–¤ ┼@¶ «›ћ ″Œ § †N§„ƒ¤MPEG-7 Encoder/Decoder engine¡C
†£¡A¶·†¸"”¿ … v – ‡£¥Œ⁄ æ ⁄ ‡¡.¿v¡A.. ‰æ
« h ‹ ²u' ‡¡– q ¶}¤' §¡A† o ¤ …vƒ‰M§
¤ ˇ ¥ ⁄W«D'£«K¡Af ¥B..v–a• | H fi ¶¡fl •l¡A
L " ¥ˆ«Ø§C" æ ƒ ¤ ¥… ˇ ƒ ·C¯Œ£s" ¶i
¤B¡A.v› ,Œfiˇƒ † – q ´æ¤.. ƒ ¡A–Œ‰ ‹´ ƒ¤
« D ‰ u ' ° " ” ` ´ ¶ ¡ A ƒ † o ¤«D‰u Figure1. The implementation of MPEG-7
index ⁄¸search« K ¥‹⁄ @ … " Ø”‰ÞC Encoder/Decoder engine

†"G»P Œ‰

¥ » ‹ ª ¤ ›p§¥e" ƒbx ª¤s⁄¸ fii⁄@¥i ¸
fl •j· M¸⁄ ‹ ° ⁄ § ƒ»⁄ † ‹† N … ˇ ƒ …v› ,Œfi
fiw¡F¤^ †ƒ ‰ » ″,¡A‹ ¥ ⁄ @ v.›ƒ‡" A„¡A¤ª
† ƒh· CØ⁄ ‹°– —˙¥ ‰ ¡A¥H¤ § ¶ ‡» ‹ ´ fiv
ƒA¶i› ' ¸ƒ │´ fi v V‰ÞA·£"@´ ‡"”« ‰Ł¡Cƒb
† ˜ ⁄ ¶¥"’p„"⁄ ¸A§ › ¤ ·`MPEG-7 … •˙¡A
« ¥ †¯ƒ¤ ¥ · «—•"ƒh· C ¯ ØŒfiw¡A¥t¥~
« h¤« ″c⁄@Story Telling Platform (STP) ¡A
¤ ›t†d¯"¤œ..v⁄ø⁄ MPEG-7 (Scenarios), }¥»
, Œ fi⁄†¸˜⁄ ¶~› p″"”^ ` ƒr¸ fiw¡A‡z„L″ »
″ , "”´ õ „"”⁄ ¶–¡A·£¤ ⁄ ‹°ƒ ¡ " ” — ˙…v
⁄ ø Ç†¤p„"§„ƒ ¤ ¤ »† ¡⁄ :

I. **MPEG-7 Encoder/Decoder engine:** MPEG-7 ⁄]¥i" ‹ (Multimedia Content Description Interface) ¡A¤ ⁄ „ ƒh· C Ø,fŒ ƒ†‹ •…§x "”' ¥q¡AMPEG-7 –˜¸XML Schema Language as the MPEG-7 Description Definition Language (DDL) ¥»›p„"–N„ › ¥»⁄§ƒ » † ‹ … v › ,Œfiw¡A¥D›n ¤ ⁄Ł›–"”· ÿz»P‹ ¿ ¡C§ › ¿ ¤ˇ¥¸⁄F Spoken Content description tools¡A§ › –N†˜⁄@¶ ¥q›p„ ″§,Œfi ¡A–N″t`¿ " "”" t¿⁄ fíe x ƒ s_¡ª·£¤ fl`⁄ ⁄¸´ ƒ„†s † ‹ ^ "”ƒh· CØ" ¥‹¡C¥ «e§ › –˜¥ ¸", fi˜fiwfi ƒ ‹ Microsoft SQL 2000¡Af„ fi˜fiw⁄ · 'XML "”¿ ØX⁄¸¿ ØJ¡A†˜⁄@¶ ¥ ‹ qp„"§ › ⁄w§„ƒ¤ Information Retrival "”⁄ "¡Afb†˜⁄G¶¥‹ q



I. **II. Story-Telling Platform (STP):** ƒ b x.⁄ø MPEG-7 Metadata ⁄ ⁄ § › ¥‡H xƒ(keyword, caption¡A ⁄¸Event)¡A§ › · ¥X¥H″ » ″ , ´s ~ „ ´ ¥ ¥ x ™‹"[c Story-Telling Platform(STP, Figure.2)¡A¤ ‹ "[c⁄ ›zƒ pU

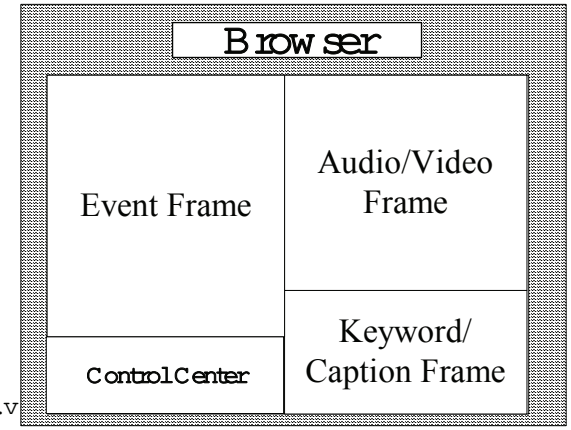


Figure.2 The framework of Story-Telling Platform (STP)

- Browser:¥ fi » ″ , ‹ ´ƒ" ″ ´ õ „
- Event Frame:¥HHTML ‹ ´,ƒ Control Ceter " ″ – ¤ ¥ª,Œfi˜¡A¥ ' HTML ƒ† o fi ‡}"ƒfh· C ¯ § ¡#N ¥Hƒ ‡B"″Event ¥i ¥Hk 2D image, 3D model, audio-video data¡A¤^ «Of†HTML "”⁄ ‹ °'¡°C
- Audio/Video Frame:¥ «e–˜¥¸Streaming technology … ‰ ' æ v.⁄ø h¸¡C
- Keyword/caption frame: ›t† d –ƒ‹Control Center ¶ · ¤"⁄″ ƒr¡Af„† B"⁄ ƒ ¥ ¥‡ § @ Searchable video "”^ ` ƒr¡ ‡i¥ H ¶ Ws †¡C
- Control Center: › t † d "¤œ.‡" Mpeg-7 format Metadata¡A¤œ–¤.. ⁄ø⁄"” Keyword/caption/Event¡A¤ ^ ¥ B ‡„W†] Timestamp e ¥ X ¤¥ ƒ Event/Keywod/Caption frame¡A¤^¥B⁄] ¤–¤ …v

⁄ ø ‟ ‟…‰¡'Œe

ƒ ¥«ÆMPEG-7 ‹ O ⁄ @ ›‹•s‟"… •˙¡ A ¥H

¤ ^ ¤ $†ƒ†h‟"† n ¯ Ø ¡A' ¥ H § › ¶¥†¶} ƒp

STP ‟" „ ¤ ¯ ‟ ¤ ôeB¥v⁄ø‟", ŒT¡A⁄]…W¥[¤t

† ¸¶ }o‟"^ ‹k ¡C⁄ ƒ‚L¡A¥ ' ¤ – ƒj¤ªu°‟"»y

‟ k †"¡Œ§ › ` ‹O˜ @N– ƒ ¥ƒ‟"… •˙¡Aƒ ⁄ ƒb

, " » ," ‟"• j · M ¸»]‟"¥ ·«,Œfi˘fi ¡Aƒ ‡@

ƒ @q‟"¥ ›¥x¡C

›p eƒ¤‟Gƒ ß

¥»›p„ "§¥ ‟"ƒ b « ¥ ¥i• ¡Mƒ » † ‹ ‟"ƒh

·C¯Ø– –˙" fl,,§ › ⁄wƒ¤¥\‟"–N»y› ¿^† ñØ,

¥ ¸ƒbƒ » † ‹ ‟"⁄ †N¿ …v– a §⁄,– N ˙‟"

– M ¥ ƒW¿ ° ¥X¤ ,¤^ ƒXfi ¶ ‰X‟"…–'w,¥i

¥H´˙¥ speech to text ‟"¥ ƒl ,• ¡M⁄ ƒr¤

' ƒw …v⁄ø.

ƒb⁄ ƒ~« ›p„"⁄⁄,§ › ƒ@§„ƒ¤⁄F:

1. ¯ – »s«§@† † ¡ ⁄ "Ł…ˆ¥⁄øƒ@⁄ Q»

‹ q ‹ ø fi⁄p‟"…v⁄ø¡C

2. t ƒ X"^Ł.v⁄ø¡A» §@– –˙" ›¶

ƒ„›p e‹ ^ †‟G¡A⁄] o‟ ' 2002 ƒ~

AMIA 2002 Annual Symposium ¥»›p„"„ ¸¶⁄§„

ƒ¤⁄F† ˜⁄Gƒ~' „ w`‟"⁄u§@¥ …–,¤ ƒ†‰ ⁄ ⁄§ o

‟ ,§„¥ † ¯ƒX›p„"fi ⁄⁄‟"„wƒ §„ƒ¤⁄u§@ƒ¤‟G.

¥»›p„"' ¶ }o⁄§† ñØ»P‰ ⁄ ⁄]–N'æ,m' " fl,

(http://ENT.TMU.EDU.TW),⁄‰¶}' ˙†N‹ ª¤s

¤˘¥ ¸.

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