ƒ ‹ F | Œ fi a ‹ · ' e › ß • | , §U–M^D‹ ª

¡ ¡ ¡

ƒ ‹ F | Œfi a ‹ ' e › ß • ┼M^D‹ª¤s›p eƒ

§ Q ¥ ¸ » y › ¿ ° ^ § ┼ N « ¥ ¥ i • j · M §ƒ » ┼ ‹ —

Building a Searchable Digital Video Database for Otolaryngology Teaching by using Speech Recognition Technology

› p e ‰,N§G 89-2511-S-038-002

ı ƒ ·`¡›G2000 ƒ~08 ⁄°01 ⁄ Ø 2001 ƒ~07 ⁄°30 ⁄Ø

¥ D ø⁄┼HG§ ı ›˜P » O ¥´ ·⁄j ·/» O ¥´ ·⁄j ·™ ┼]´ |

ƒ ŒP¥ D ø⁄┼HG§ı⁄ ˝M » O ¥´ ·⁄j ·´ ;Œ T ‹ ª s'

ƒ ŒP¥ D ø⁄┼HGfi}« •~ » O ¥´ ·⁄j ·´ ;Œ T ‹ ª s'

ƒ ŒP¥ D ø⁄┼HG™ L ┼«n »O˜W⁄j · ƒ »┼┼ ›

› pe »P⁄H›ß¡G¿c« §» » O ¥´ · |´ ;Œ T ‹ ª s'

› pe »P⁄H›ß¡G™ L d• A » O ¥´ · |´ ;Œ T ‹ ª s'

/ / / ”K›n

¥»› pe¥D›n⁄S¥ ™ ƒ‰§ Q ¥ ¸ » y › ¿°,^N§ ┼N „ ¥h™⁄ ƒ fl`,¤ ‰˜´┼« § ‹O⁄ £ ¥ƒ┼Ø ƒ » ┼ ƒ‰b— —·¯Ø..v⁄ ø⁄™» y´ · « ƒ⁄. ƒr ƒ » y ¡ C ƒ⁄...v┼N‰¿O–M‹ ´ · – ¤|⁄@¶ ‡X¥»›p„"' ¶ }o⁄S┼n¯Ø¸– N ⁄ƒ.r— ø fi¶¡‰X › «n™",Œ┼ £ ¡ ¶ ·™¿ ...v– ª‡f‹O¥ Ĥ ⁄æ⁄Ł § ¶¿" , ⁄ƒ┼Ø,CŒ fi ˜ ŵ§¸§┼N ¥i¥ H ´¥ ƒ ¿ ...v¡ A ... ‰æ«Œ‰ ú°⁄Łƒ ¡ – q Y¶›'l • jM⁄l Sw⁄ ƒr¤ [‰..v⁄ø ¸q¤¸ƒ p ƒ⁄@ ·¥˝ § ¡ Œæ¤ƒ¥ ' ... ƒ ·C¯Ø ‡s™"¶ i B¡.Av ⁄¸´ fi §Y¥ i §¥"¸ , § ┼t¸⁄ £ ¤ ƒa⁄ƒi ¶¡›› › Œfiƒ˝ ┼ q^⁄æ´ ƒ ¤ ..ƒ¡ A – q ‰'ƒ¤«D ¤ ¸¤ ·┼ ƒ » ┼ ‹ .—— · ‰ ú°™"^ ˝ ¶ ¡ C t ¥...e¸..W¥ ƒ, ł Ĥ „ ™"‹ [┼] Ą¥ƒłŵ^Hƒh┼ B ƒ P fi ˜[.«v⁄¿" fi e Ą„Ŧ –M‹ ´ · » • ¶ƒ^A"¥—¤ª⁄¡C ' ¿ ŵ‰ú"..°‰æ⁄Łƒ §Y‹O¥i• jM⁄§..v › Œfi¡˝ Ą˜¥¸" ¥ i ¤ ┼ ƒMbS ›n⁄ § ...v⁄ø ‹ q ¡ A ⁄ ‡ ¤«§ Y §¿..v– a ‹ @„ĤMO¡"C§ ┼N ƒb« D ´ ⁄ Ł–ƒ ┼ › « ™ƒ"¤ ł¡·AŊOƒb« _ Ł„y ‰ » D fi j...Ŧ"§Pfi.v⁄ " ‡i¡¥Virage⁄ ‰ ¥¶'M AltaVistaƒX§@¡Aƒ¤¥\™"„B¥¸VideoLogger┼n¯Ø »s§@ƒ¤¥i• jM⁄ § ...v⁄ "fi¢A¥¤¸" ¥¥┼H ƒr·M § ‹Y⁄@fl Sw™"¿ ...v⁄ ø◊q┼Ą« O┼˜⁄ @ ƒb" ,¥i • jM⁄§..v⁄ ◊C ƒ »┼‹ ⁄ ┼ Ĥ ' ł ┼ ĤU¡Ą¤ G ┴–¤˘ ¥¸¯ ª • Ł⁄Ł'"¸ ł⁄Ł¡A– ƒ ¿ ...v┼]┼ ˘ Ø«¥Ki–N⁄ ┼N„L { § ™"‹ ¿ ⁄U¤ ¡G ¤£™"⁄ ┼Nfpf ┼¡™" „ ª·ƒ¤«‹┼N(tympanoplasty)¡A» ┼¡™"¥\fl ' °» ˜u⁄" ł⁄Ł⁄ ┼N(functional endoscopic sinus surgery) ¡A┼ ┼¡™"┼ Y"‰„F^ Ł ª•Ł ┼N (larygomicroscopic surgery)¡ ŒH„ ª·ƒ¤«‹┼N (tympanoplasty)‹ ¤ Ą⁄¥ø⁄ ┼N┼q– ›» ›n¤ ›

Abstract

The purpose of this project is to create and integrate multimedia medical resource for medical community. In this project, we realize the “speech 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 ‹ §"¸¶⁄Øfl¤q┼Ø⁄ƒ"ł..v › Œfi¡ƒ b §›™"¥ ˝ ‹ ¡ ⁄ §Cł¶˘V⁄Ł¶¥« ›n™"¤ ƒ . ⁄] ƒf„H› ⁄]„ ¡ § Q ¥ ┼o┼\ ƒ h Œfi¡¤ § @ ƒ ŬŁ™⁄– ┼¥ ¸ › " ¥ „J¤" ^D¸· N ‹ O ƒ p ƒ fl`™"⁄¤§@j⁄MŒq^e ´ł™"... v › ,Œ¡ki˜ƒ ‹O 3D ¥ ¯ Ø ¸«‹æ_ ƒ » y ¡ C ƒ⁄...v┼N‰¿O–M‹ ´ · – ¤|⁄@¶ › «n™",Œ┼ £ ¡ ¶ ·™¿ ...v– ª‡f‹O¥ Ĥ ⁄æ⁄Ł ƒ ¿ ...v¡ A ... ‰æ«Œ‰ ú°⁄Łƒ ¡ – q Y¶›'l § ¡ Œæ¤ƒ¥ ' ... ƒ ·C¯Ø ‡s™"¶ i B¡.Av › Œfiƒ˝ ┼ q^⁄æ´ ƒ ¤ ..ƒ¡ A – q ‰'ƒ¤«D ‰ ú°™"^ ˝ ¶ ¡ C t ¥...e¸..W¥ ƒ, ł Ĥ „ ™"‹ [┼] Ą¥ƒłŵ^Hƒh┼ B ƒ P fi ˜[.«v⁄¿" fi e Ą„Ŧ –M‹ ´ · » • ¶ƒ^A"¥—¤ª⁄¡C ' ¿ ŵ‰ú"..°‰æ⁄Łƒ §Y‹O¥i• jM⁄§..v › Œfi¡˝ Ą˜¥¸" ¥ i ¤ ┼ ƒMbS ›n⁄ § ...v⁄ø ‹ q ¡ A ⁄ ‡ ¤«§ Y §¿..v– a ‹ @„ĤMO¡"C§ ┼N ƒb« D ´ ⁄ Ł–ƒ ┼ › « ™ƒ"¤ ł¡·AŊOƒb« _ Ł„y ‰ » D fi j...Ŧ"§Pfi.v⁄ " ‡i¡¥Virage⁄ ‰ ¥¶'M AltaVistaƒX§@¡Aƒ¤¥\™"„B¥¸VideoLogger┼n¯Ø »s§@ƒ¤¥i• jM⁄ § ...v⁄ "fi¢A¥¤¸" ¥¥┼H ƒr·M § ‹Y⁄@fl Sw™"¿ ...v⁄ ø◊q┼Ą« O┼˜⁄ @ ƒb" ,¥i • jM⁄§..v⁄ ◊C

ƒ »┼‹ ⁄ ┼ Ĥ ' ł ┼ ĤU¡Ą¤ G ┴–¤˘ ¥¸¯ ª • Ł⁄Ł'"¸ ł⁄Ł¡A– ƒ ¿ ...v┼]┼ ˘ Ø«¥Ki–N⁄ ┼N„L { § ™"‹ ¿ ⁄U¤ ¡G ¤£™"⁄ ┼Nfpf ┼¡™" „ ª·ƒ¤«‹┼N(tympanoplasty)¡A» ┼¡™"¥\fl ' °» ˜u⁄" ł⁄Ł⁄ ┼N(functional endoscopic sinus surgery) ¡A┼ ┼¡™"┼ Y"‰„F^ Ł ª•Ł ┼N (larygomicroscopic surgery)¡ ŒH„ ª·ƒ¤«‹┼N (tympanoplasty)‹ ¤ Ą⁄¥ø⁄ ┼N┼q– ›» ›n¤ ›

／ p fi ¡ A ƒ p " G ¤ ˇ ¥ ¸ " ¥„(stamp¤) '／"Ψ"／E¤BƒA¥ t› ／ S ƒ »†–M« ´ † B † z ƒ¿†ªApƒ ／£ » – q YS _'O¡H fiv» ƑU› ¥ ÐM" ¤ «Æ¡A' ƒ † v.／ø ／ † N ¿ … v – a ‹ O ´ ·"„¤E／@¶ ¥«ƒÐ／@／H t› ¡A¥HSQ–N¤ »y› ¿ ° ^†n †£¡A¶ †¸"¿ … v – a † £ ‹ O ¥ Ħ ¡Æ¤ƒ¡¿ …v ØS %m¡C t› fi –Mƒ†ƒ W ¥H›^／ ‹ « h ‹ O ‰ u ' ° ／Ł ƒ †qA¥¶}¤l¤…S v øfb·MS ¥D¡Aƒ ／†"¤B˘J／¸· ¥z« ¥H／／ ‹ ¤ ¥¸／W « D ¬／¡ƒA¥K ¥ B … v a•fi¦ ¶H¡ fl•¡A ¥D¡C

L " k ¥ ^／[« O ƒ s ¡ C " æ ƒ ¤ "¥"¶i …˘ƒ ^C'Ø／»ƒ‹s S ¤ł""»y› ¿ ° ／@"‰« O ¤B¡A.. v › , Œ fi ˘ ƒ † – q ¡Æ¤q ƒ¤u˘"ƒ´ ƒ¤ › ^ł^D¡AS › –N–˜¤œ／Uf C ₮"／ ₺k¡A « D ‰'""` ¶ ¡Aƒ † o ¤ «D‰"Ù.'v › , Œ"／ ¥H¶ ¶Q"／S ¤ …v／Ø／"'time stamp": index ／search« ₭¤« ／@› …"Ø／‰ ^D¡C ／@¡B ¥u¿ ° ›^／ –Mƒ†ƒ W "／" t › ¡A¤ ‹ time stamp¡ A／／ – ›z †¡／ «h／'¥H' ł¿／¡A†o¥i¥¸†]'w¿ ° †v"／" cut-off value ¤ „Fƒ ¤C‹G» ¤˘¥ « ¥ ƒ » † ¤^／ –Mƒ† fiw¡A¥H· £"¿ ° †v¡A' ／j»P／／／ ¿° ^ "／"fi ¶Z¡C ／G¡B ƒb t› fi ₫"„GJ¤ ›^／ –Mƒ† ƒ W «K „y(time lag)¡ Apƒ„／]¥i ／¶}／／／ » B^／ ¡Aƒ « K £‹Ð š »y › ¡ Apƒ„•|…W¥[t› "／"Sx^ł« ¡C ／T¡B ／ SO¤ ¥¸›^／ »P／／／ ¿ ° ^†n ¯ Ø B†tz ₮X»y› ¡AƒA–N¤ › †"G¥[¥ H B†¡zA¿ › SOƒ¤¥\"／"†¡／ " ƒX _ ¤¡C »y› ¿ ° ƒ¤¥\«Æ¡A«K¥i¥ H ₫¤»y › ¥ X{" fi ¶ ¤ fl／¡A« ¥ ／ ‹ ° "／"‹ ¸d ／ Ł ƒ¡C／U¤ †˜／T¤ B ₭‹‹O[†]… v ƒ ł A ¡„A¥ ' ／ †N„..v／ø " ₨ "／"›n¤ Ð˘ "¡A' ¥Hƒ ł A "„" x ƒs"¯¶¡／¸† B z†t« ›nfl tƒX¡ Œ ／" … e W"« „†]' |¥..ƒ¤.. ¡Aƒ]' » ¶·¿ Øt« ‹ø ¤C 200Kbits¡A G" · ／·¥H·£¤ ／"†¡" , (intranet)¤ ¥¸ " ‹ ¥D¡C

¥D›n／S» „†]† S¡–N¥¸' ¿ …v–a … ƒ ／˘†B z／»y› ¿ ° ¡A¤ ／／¥ ł A … ˘ ƒ …v ›¯– ／ u Sfl¡A ", ₨«..v „ ƒ^¤œ¥d／¸†n¯Ø»y› ／ "R／u¤ª¡A S

† " GP Q‰

1. ƒ ¶ 10／ƒ¡¤ ª M"' °／ƒl S fi ／$ » † ‹ ／ † Nv¡Aa " ¥–N¤·‹ ¿ … v a…／ƒ¡A ´ ƒ¤¢ ¢ ¢ ł¢ ¡C
2. † ／G¤ M ¥ ÐM´ fiv„ ／ †N„ L { Ł ¡ A˘ ¥¸ƒ V ¶ q †S↺¸MD ¿ ／U¡A‹S«Æ¥H»y› ¿ ° †n¯Ø／¸time code {ƒ †£¥˝ƒ†fi ¶¡„ †／S／ ƒr ¡C
3. t› fi –Mƒ†ƒ W ／¸ N¤˜BJ"˝· ¥z¥H／／ ／‹ ¥D¡A‹ · £"¿ ° †v¡A» ¤˘¥ « ¥ ƒ »† ‹ –Mƒ† fiw¡A¤^„ ¿ ° ^†n¤Ø／¸[¥H %m¡C
4. »y› ¿ ° ƒ¤¥\«Æ¡ A ¤¥‹¸ „ "˝ƒ†fi ¶¡´I ／S／ ƒr fl／¡A«K¥i¥¸search program ／¸ mpeg control program flS'w.. v ø／«œq"˝...‰ 'æ¡A« ¥ ／ ‹ °"˝‹ ¸d ／ Ł ƒ¡C
5. †¸›p／£ƒP t› " "˝»y› ¿ ° ƒ¤¥\†v¡C ¥ » ‹ ª ¤ ›ƒp ¶e(ƒ ¡B» ¡B†)¤ ª M" '／ƒl S fi ／$ ƒ»／ † † Nv¡A« •‰ ¥ ¥ ¤j´ |／¸¥ ¥_´ · |ƒ » † ‹ ,Œ† ' fiv·£¤ ¡C› " ¥–N¤·‹ ¶Q‹ ¿ … ¤aS ¥ … ƒ ／˘¡A´ ƒ¤¢ ˘ ¢ ¢ ¡A, " ₨ ‹ ¢ † ¢ – ¢ fl ¢ ¢¡A«†£²¶´I¢ † byte ƒ –m¡A¤C／@‹ «…¢¸ ¢fl–i¡ Apƒ„ ／@／ p fi"„" ¤ø¥ … £Y»«e›n¢ ¢†¢ ¢–"" ¤s"¯¶¡¡A £Yƒ¤¢ ¢ ¢ ¥£u» ¢·¢fl／ S／@‹ ø¢ †¢•¢fl¢ ¢–¡AƒA´ ƒ¤ Realvideo ¥u» ¢•¢fl¢ ¢— x ¤¡C

O « ‹ „ q , £ ⁄ ¿ " fi fl ‡ ` § J › • ¥ ¿ ' „ J
› « ˙ t › ¡ A ¥ H ⁄ ¿ Ð ¥ D ⁄ M " Q ‰ ..
⁄ " fi ¡ A v › " ƒ ‚ ł " A „ ¥ ƒ ¿ h Ć ¯ Ø

› p e ƒ ¤ G ƒ ß

¥ » › p „ " § ¥ ' ‰ " ƒ b ¥ ¥ i • j · M ƒ h ‡ ‹ " "
· C Ø – — ˙ " fl ‚ § › w ƒ ¤ " ¥ \ Ð y › ¿ ° ^ ‡ n Ø
‡ ¥ ¿ ƒ b ƒ » ‡ ‹ " ¶ ¿ … v – a ⁄ § ⁄ ‚ – N ´ " "
– M ¥ ¿ ƒ W ¿ ° ¥ Ø ‚ ¤ ^ t ƒ X fi ¶ ¡ ‰ X. ⊥ w ‚ ¥ i ¥ H
´ ¥ speech to text " " ¥ \ fl ‚ • j M ⁄ ƒ r ¤ ' w ƒ …v
⁄ ø
ƒ b ⁄ ƒ « › p " ⁄ ‚ § › ƒ § „ ƒ ¤ ⁄ F.

1. ƒ » ‡ ƒ h Ć ¯ » Ø ¶ Z – — ˙ " fl ‚ (http://ent.tmu.edu.tw)
2. Intergrated timecode based annotation system
3. ƒ » ‡ ˆ ` ^ ª fi w

ƒ „ › p e ‹ ^ ¡ ‡ A ¥ G] o " ' 2000 ƒ ~ Œ » ´ ˙ ‚ Œ T ‹ ª Q • | (Medical Informatic syposium in Taiwan 2000) ¤ Ł ¿ ⁄ @ ¡ A ¤ ^ ‡ ¤ ¿ ‰ ⁄ … œ Q (^ D ¥ : A Searchable Digital Video Database for Otolaryngology Teaching By using Speech Recognition Technology § @ " : ¿ c « § » ¡ A fi } « • ~ ¡ A § ı ‚ ˜ ¿ A " L ‡ «) ‚ ¥ » p „ „ " ¶ ⁄ § „ ƒ ¤ ⁄ F † ˜ ⁄ @ ƒ ~ ' w ⁄ • " " ⁄ u § @ ¥ … – ‚ ¤ ^ ƒ ‡ ‡ ¤ ¿ ‰ ⁄ … œ ⁄ § ð ‚ § „ ¥ † ¯ ƒ X › p " fi ⁄ ⁄ " " „ w ƒ § „ ƒ ¤ ⁄ u § @ ƒ ¤ " G ¥ » p „ " ¶ } o ⁄ § ‡ n ¯ Ø Ð ‰ ⁄ ⁄] – N ' æ m ' " fl ‚ (http://ENT.TMU.EDU.TW), ⁄ ‰ ¶ } ' ˙ ‡ N k ª ¤ g ˘ ¥ ¿

Reference

1. Hauptmann, A.G “Speech recognition in the Informedia Digital Video Library: uses and limitations”, Tools with Artificial Intelligence, 1995. Proceedings., Seventh International Conference on , 1995 : 288 -294
2. " L ‡ " « n ¡ B ‡ \ ¯ v fi ¶ ¡ G « Ð x fl ‰ F " C ' ° ⁄ ⁄ ƒ " ¢ ⁄ ` { § – – ˙ ‚ g ¯ ¡ Œ 1995 ƒ ¡ 30 ¡ ø 184-191
3. K. Sparck Johns, G. J. F. Johns, J. T. Foote, and S. J. Young, “Experiments on Spoken Document Retrieval”, Information Processing & Management, vol. 32, no. 4, pp. 399-417, 1996.
4. Young, S.J.; Brown, M.G.; Foote, J.T.; Jones, G.J.F.; Sparck Jones, K., “Acoustic indexing for multimedia retrieval and browsing”, Acoustics, Speech, and Signal Processing , 1997 : 199 -202 vol.1
5. Hauptmann, A.G.; Wactlar, H.D. “Indexing and search of multimodal information” Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on , Volume: 1 , 1997 : 195 -198 vol.1

" ¿/@

A Searchable Digital Video Database for Otolaryngology Teaching by using Speech Recognition Technology

Chien-Hung Lu, Chien-Yeh Hsu, F.P. Lee, K.N. Lin

(¿ C «§»fi } «•~§ ı ›˜P " L ǂ«n

Graduate Institute of Medical Informatics, Taipei Medical College, Taiwan
Taipei Medical College Hospital
National Taiwan University Hospital

Summery

The purpose of this project is to design and build an interactive and searchable video database for Otolaryngology Teaching. By using this system, doctors can search the otolaryngology surgery video clips they really need. After building a video server and connected to the internet or intranet, we can construct a web based multimedia medical specialist teaching system to improve the quality of medical care.
After digitalization of the video, a single otolaryngology specialist dubs all the video for the operation procedures. Speech recognition technology was applied on these voices and time markers named as "time stamps" are generated to index the video. Incorporated with the database and web system the interactive searches become possible.

Introduction

The main disadvantage of medical video data, especially the operation video, is difficult to search even after digitalization. Many times we have to search from the beginning because lack of interactivity. In this project, we design and build an interactive searchable video database. The doctors can search the otolaryngology surgery video clips. After building a video server, we construct a web based multimedia medical specialist teaching system to improve the quality of medical care.
Many otolaryngology operations need microscope or endoscopes due to small surgery field. There are many videotape records of surgery but the searching and storage is difficult. For example, microscopic surgery of the ear, endoscopic paranasal sinus surgery and laryngomicroscopic surgery are all common ENT operation in National Taiwan University Hospital and Taipei Medical College. So the digitalization of these tapes and building of video server contribute a lot to the preservation and promoting of knowledge on this field.
Currently, speaker-dependent speech recognition software has good recognition rate up to 95%. A single otolaryngology specialist dubs all the video for the operation procedures. Speech recognition technology was applied on these voices and "time stamps" are made to index the video.

Materials and Methods

We collected 60 hours of otolaryngology surgery video. The videos are provided by senior doctors in the National Taiwan University Hospital and Taipei Medical College Hospital. The video are digitized first and transformed to AVI format with 320x240 resolution. Then the video data are compressed to MPEG format and further transfer to Realvideo format.

By using Ulead Video Studio 4.0 (• ¦n •¦ …v), a single otolaryngology specialist dubs all the video mpeg files for the operation procedures. At the same time speech recognition was applied on these voices by using SpeechPro (¶ W fl ƒ ƒ)• and time markers named as "time stamps" are generated to index the video. Chinese vocabulary library for otolaryngology was created first for better speech recognition.

A timer counter programmed by Visual Basic marked the time when key word present. Like "stapes (`E „" )" appears at time 6'15" on the demonstration surgery video. If someone interests only in "stapes ( `E „")" part of the whole surgery, he can look directly from time 6'15". No need to search from the beginning.

Results

An interactive and searchable video database for Otolaryngology surgery will be designed and built. In this project, lots of surgery video for otolaryngology has been collected and rearranged for teaching purpose.

Discussion and Conclusion

The bottleneck of this method is the dubbing. The one who dubbed the video must have the domain knowledge about otolaryngology surgery. And he have to practice many times for better speech recognition rate although the otolaryngology vocabulary library has been made first. Many times when he speak the wrong words he has to redo again. So long surgery video is very times consuming and the repetition of the same key word made the searching difficult. So this method can only applied on short video. More systemic method had to be developed for large amount video database.

Reference

1. Leland, Jon ,"Online Video: The New Medium" ,Videography, vol. 23, no. 11, pp. 48(2), Nov 1998

2. " L ‡ ˝ « n ¡ B ‡ \ ¯ v fi ¶ ¡ G « Dxfl ‰F"C'°⁄⁄ƒ "¢⁄ ` { § – – ˙ , g ¯¡ 1995ƒ30´ 184-191

3.K. Sparck Johns, G. J. F. Johns, J. T. Foote, and S. J. Young, "Experiments on Spoken Document Retrieval", Information Processing & Management, vol. 32, no. 4, pp. 399-417, 1996.

4.Glavitsch, U. and Schauble, P. "A System for Retrieving Speech Documents", ACM SIGIR Conference on R&D in Information Retrieval, pp. 168-176, 1992.