

機関番号：21201

研究種目：基盤研究(C)

研究期間：2008～2010

課題番号：20500096

研究課題名(和文) 地域FMラジオ放送のオンデマンド放送化および検索技術の研究開発

研究課題名(英文) Research and development of a methodology for on-demand broadcast of local FM radio and a retrieval technology for spoken broadcast data sets

研究代表者

伊藤 慶明 (ITO YOSHIAKI)

岩手県立大学・ソフトウェア情報学部・准教授

研究者番号：90325928

研究成果の概要(和文)：地域FMラジオ放送のオンデマンド放送化および検索技術の研究開発の推進を行った。ラジオ音声を対象と想定し、音声の中の聞きたい区間を容易に検索できる技術の研究を推進し、地域FMでは方言などが含まれるため「語彙外の言葉」でも検索できるように、音素およびさらに細かな系列を開発し、任意の語彙での検索と性能向上を実現・成功した。さらに、様々な音声検索研究手法を公正・公平に評価するため、情報処理学会においてワーキンググループを立ち上げ、共通の音声データ、評価方法の整備を推進し研究者への公開を実施し社会貢献を果たした。

研究成果の概要(英文)：We conducted the research and development of a methodology for on-demand broadcast of local FM radio and a retrieval technology for spoken broadcast data sets. Assuming spoken radio data, we conducted the research and development of a technology for retrieving a section of interest in the data. Because such local FM broadcast includes dialects, we developed the method for retrieving out-of-vocabulary terms by introducing a new unit that is shorter than a phone, and succeeded in open vocabulary retrieval and the improvement of retrieval. Furthermore, we established standard test collections for evaluating retrieval methods that have been recently proposed from many research institutes under the working group in Information Processing Society of Japan, and realized the research basis and the contribution of an academic society activity.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	1,200,000	360,000	1,560,000
2009年度	1,000,000	300,000	1,300,000
2010年度	1,200,000	360,000	1,560,000
年度			
年度			
総計	3,400,000	1,020,000	4,420,000

研究分野：総合領域

科研費の分科・細目：情報学、メディア情報学・データベース

キーワード：インタフェース、音声検索、音楽情報処理、オンデマンド放送

1. 研究開始当初の背景

インターネットの発展に伴って、テキストのみではなく、音声、音楽、静止画像、ビデオ情報などを利用できるようになった。百時間以上の録音・録画が可能となり、扱える情

報量(データ)の増加は著しい。既にインターネット上の検索ニーズ・ビジネスが隆盛になったように、蓄積された音響・映像情報に対しても、聞きたい／見たい部分・場面の検索など、様々な検索ニーズが高まると予想され

る。

岩手県立大学が主体となって地域FMラジオ局を来年度より運用を開始する予定にしている。地域FM局で放送される地域の情報は住民には有益な情報である場合が多い。放送を聞くことができなかつた／重要な部分を聞き逃した際には、後になって実際に放送内容を聞きたいというニーズが高いと考えられ、放送内容を容易に検索できる機能が望まれる。NHKではテレビ番組に対し字幕放送の運用を一部で開始したが、地域ラジオ局や一般のラジオ局では放送内容にテキストを付与して提供することはコストが掛かり現実的ではない。

我々は平成19年度までの科研費テーマで、音声・音楽情報を利用することによりビデオデータの音声・音楽部分の分離と、検索用にビデオの構成要素を抽出する技術の研究を進め、精度良く音声と楽曲ごとのまとまりを抽出できることを確認した。これらの技術を前進させ、地域FMラジオ局の実データを実験フィールドとして、ラジオ放送データの検索／オンデマンド放送の実用化に向けた研究を推進する。これにより、地域に密着した貢献を果たすとともに、音響・映像情報の検索技術の確立にも貢献するものとする。またラジオ放送は視覚障がい者にとって有用な情報獲得手段である一方、視覚障がい者や老人にとって、現状のコンピュータのキーボードを介した検索は非常に困難なことが多い。そこで、視覚障がい者・老人さらには健常者にとっても容易な音声を入力手段とした検索技術についても研究のスコープに入れる。

2. 研究の目的

従来、ラジオ局からの放送はユーザであるリスナーに対して一方通行的な放送形態で、実際に放送内容を後から聞きたいというニーズには応えられなかつた。また、実際のテレビ・ラジオ放送の検索は音声・音楽・生活音・雑音など混在しているため、内容に対して検索することが困難であった。我々は前回の科研費テーマにより音声／音楽の区間を抽出する方法を開発しており、これを発展させ、地域ラジオ局の放送した番組の検索技術の確立を目指す。具体的には以下の実現を図る。

- (1) 放送音声の検索技術の確立：放送された番組内の音声区間を同定し、ユーザが聞きたい部分の検索キーワードをキーボードからテキストあるいはマイクから音声で指定することで、その検索キーワードが話されている放送音声区間を高精度に検索する技術を確立する
- (2) 放送音楽の検索技術の確立：放送された番組内の音楽区間を同定し、ユーザが聞き

たい部分の楽曲名をテキスト、あるいはメロディを音声等で指定することで、対応する音楽区間が高精度に検索できる技術を確立する

- (3) 放送番組データのアーカイブ(蓄積)方法および検索のための情報付与方式、ラジオのオンデマンド方式の実現可能性を検討し、その提供方式を研究開発する

3. 研究の方法

本研究では、下記の4件のサブテーマに分けて研究開発を推進した。

- (1) ラジオ放送の音声・音楽区間の識別と分割技術の高精度化の研究開発
- (2) ラジオ放送の音声区間の検索技術の研究開発
- (3) ラジオ放送の音楽区間の検索技術の研究開発
- (4) ラジオ放送のオンデマンド放送に向けた課題整理および技術研究開発

このうち(1)は、平成19年度までの科研費テーマである、ビデオデータの音声・音楽部分の分離技術の研究を前進させ、さらなる高精度化を図った。(2)では、(1)により特定された音声区間を我々の独自の音声検索技術を応用した。具体的には、地域FMでは方言などの音声認識システムの語彙には含まれない様々な言葉が含まれているため、「語彙外の言葉」でも検索できるように、検索時の単位として単語ではなく、音素およびさらに細かな系列(1/2音素、1/3音素、SPS(Sub-phonetic Segment))の研究開発を行い、任意の語彙での検索機能の実現と音声検索性能の向上のための研究を推進した。さらに検索キーワードの入力手段として音声を用いることにより、視覚障がい者や老人さらには健常者にとっても優しいインタフェースの実現を目指し、音声クエリの実現方式の研究も推進した。(3)では、特定された音楽区間やビデオ区間を検索特定するため、新たなビデオ区間の検索方法の研究を推進した。(4)では、上記の技術を利用して、実際のラジオ放送のオンデマンド放送に向けて課題を整理し、放送番組データの収集を試みそのアーカイブ(蓄積)方法の検討を行った。これら4件のサブテーマを有機的に連携しながら研究を推進した。

4. 研究成果

地域FMラジオ放送のオンデマンド放送化および検索技術の研究開発の推進を行った。上記の研究の方法における4つのサブテーマに沿って報告を行う。(1)ラジオ放送の音声・音楽区間の識別と分割技術の高精度化の研究開発については、音楽における繰り返し構造を利用し曲と曲の構造を抽出することによって曲境界を抽出する新しい技術を研

究開発し、国際雑誌論文①に採録されるなど多数の学会発表を行った。(2)ラジオ放送の音声区間の検索技術の研究開発については、地域FMでは方言など様々な音声認識システムの語彙には含まれない言葉も含まれているため、「語彙外の言葉」でも検索できるように、検索時の単位として単語ではなく、音素およびさらに細かな系列(1/2音素、1/3音素、Sub-phonetic Segment等)の研究開発を行い、任意の語彙での検索と音声検索性能の向上を実現・成功した。これらの成果は情報処理学会の雑誌論文③として採録など国際会議等で発表した。(3)ラジオ放送の音楽区間の検索技術の研究開発については楽曲検索を含めビデオの検索技術の研究を推進し、国際会議等(学会発表の⑥等)で発表を行った。(4)ラジオ放送のオンデマンド放送に向けた課題整理においては、ラジオ放送のオンデマンド放送に向けたデータ収集を実施するとともに、データを利用する上で著作権等の課題を整理した。さらに、音声検索研究が盛んに行われるようになり、様々な研究手法を公正・公平に評価するため、情報処理学会の音声言語処理研究会のワーキンググループを立ち上げ、共通の音声データ、評価方法の整備を推進し、既に日本語評価セットの研究者への公開を開始した。これにより研究の基盤構築・学会活動への貢献を果たした。この活動については雑誌論文②で採録された。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計4件)

① Yoshiaki Itoh, Akira Iwabuchi, Kazunori Kojima, Masaaki Ishigame, Kazuyo Tanaka, and Shi-Wook Lee, "Automatic Music Boundary Detection Using Short Segmental Acoustic Similarity in a Music Piece," EURASIP Journal on Audio, Speech, and Music Processing, 査読有, Volume 2008 (2008), Article ID 480786, 10 pages

② Tomoyosi Akiba, Kiyooki Aikawa, Yoshiaki Itoh, Tatsuya Kawahara, Hiroaki Nanjo, Hiromitsu Nishizaki, Norihito Yasuda, Yoichi Yamashita, Katunobu Itou. Construction of a Test Collection for Spoken Document Retrieval from Lecture Audio Data, IPSJ Journal, 査読有, Vol. 50, No. 2, pp. 501-513, 2009/2.

③ 伊藤慶明, 岩田耕平, 石亀昌明, 田中和世, 李時旭, "語彙制限のない音声文書検索における複数サブワードの統合 - 検索語彙に依存した検索性能推定指標の導入 -," 情報処理学会論文誌, 査読有, Vol. 50 No. 2 524

-533 (Feb. 2009)

④ Tomoko Nariai, Kazuyo Tanaka and Yoshiaki Itoh, "Comparative study of focal lengthening in the speech of native speakers and Japanese speakers of English," Journal of Acoustical Science and Technology, The Acoustical Society of the Acoustical Society of Japan, 査読有, Vol. 32, No. 2, pp. 54-61, 2011/3

[学会発表] (計38件)

① Yoshiaki Itoh, Shigenobu Sakaki, Kazunori Kojima and Masaaki Ishigame, "Highlight Scene Extraction of Sports Broadcasts Using Sports News Programs," International Workshop on Multimedia Signal Processing, 査読有, pp.646-649 (2008-10), Cairns

② Go Kuriki, Yoshiaki Itoh, Kazunori Kojima and Masaaki Ishigame, , Kazuyo Tanaka, and Shi-wook Lee, "Open Vocabulary Spoken Document Retrieval by Subword Sequence Obtained from Speech Recognizer," IEEE Workshop on Spoken Language Technology, 査読有, 4 pages (2008-12), Goa

③ Tomoyosi Akiba, Kiyooki Aikawa, Yoshiaki Itoh, Tatsuya Kawahara, Hiroaki Nanjo, Hiromitsu Nishizaki, Norihito Yasuda, Yoichi Yamashita, Katunobu Itou. Test Collections for Spoken Document Retrieval from Lecture Audio Data, In Proceedings of International Conference on Language Resources and Evaluation (LREC 2008), 査読有, (2008), Morocco

④ Tomoyosi Akiba, Kiyooki Aikawa, Yoshiaki Itoh, Tatsuya Kawahara, Hiroaki Nanjo, Hiromitsu Nishizaki, Norihito Yasuda, Yoichi Yamashita, Katunobu Itou, "Developing an SDR test collection from Japanese Lecture Audio Data," APSIPA ASC, 査読有, 2009-10, Hokkaido

⑤ Yuji Onodera, Yoshiaki Itoh, Kazunori Kojima and Masaaki Ishigame, , Kazuyo Tanaka, and Shi-wook Lee, "Spoken Term Detection by Result integration of Plural Subword using Confidence Measure," The 10th Western Pacific Acoustics Conference, 査読有, 8 pages (2009-9), Beijing

⑥ Yoshiaki Itoh, Masahiro Erokumae, Kazunori Kojima, Masaaki Ishigame, Kazuyo Tanaka, "Time-Space Acoustical Feature for Fast Video Copy Detection," 2010 IEEE International Workshop on Multimedia Signal Processing, MMSP 2010, 査読有, Saint-Malo

⑦ Yoshiaki Itoh, Hiromitsu Nishizaki,

Xinhui Hu, Hiroaki Nanjo, Tomoyosi Akiba, Tatsuya Kawahara, Seiichi Nakagawa, Tomoko Matsui, Yoichi Yamashita, Kiyooki Aikawa, "Constructing Japanese Test Collections for Spoken Term Detection," International Conference on Speech Communication and Technology (INTERSPEECH), 査読有, pp. 677-680, 2010, Makuhari

⑧ 栗城吾央, 伊藤慶明, 小嶋和徳, 石亀昌明, 田中和世, 李時旭, "連続単語音声認識結果の読み系列を利用した辞書未登録語の音声文書検索," 電子情報通信学会 研究技術報告, SP2008-53, pp. 61-66 (2008-07) 盛岡.

⑨ 田村哲嗣・桂田浩一・伊藤慶明, "マルチモーダル情報処理の研究動向," 電子情報通信学会 研究技術報告, SP2008-72, pp. 13-18 (2008-11), 岐阜大.

⑩ 秋葉 友良, 相川 清明, 伊藤 慶明, 河原達也, 南條 浩輝, 西崎 博光, 安田 宜仁, 山下 洋一, 胡 新輝, 中川 聖一, 伊藤 克亘, "SLP 音声ドキュメント処理ワーキンググループ活動報告," 情報処理学会研究報告, 2008-SLP-74-20, 2008, 早稲田大学

⑪ 栗城吾央, 伊藤慶明, 小嶋和徳, 石亀昌明, 田中和世, 李時旭, "Web上の語彙を利用した音声クエリ拡張による検索語検出" SP, SLP, NL, 音声言語シンポジウム, NLC2009-20, SP2009-84, p. 99-104, (2009-12), 東京大学

⑫ 江六前政宏・伊藤慶明・小嶋和徳・石亀昌明, "ビデオ内一致区間検索のための音響情報の時空間配置照合方式" CQ2009-63, PRMU2009-162, SP2009-103, MVE2009-85, p. 45-50, (2010-1), 京都大学.

⑬ 伊藤慶明, 西崎博光, 胡新輝, 南條浩輝, 秋葉友良, 相川清明, 河原達也, 中川聖一, 松井知子, 山下洋一, "音声中の検索語検出のためのテストコレクション構築-中間報告-" 情報処理学会研究報告, 2009-SLP-78 No. 4, (2009-10), 早稲田大学

⑭ 小野寺悠二, 伊藤慶明, 小嶋和徳, 石亀昌明, 田中和世, 李時旭, "複数のサブワード・言語モデルを用いた音声中の検索語検出の高精度化" 第4回音声ドキュメント処理ワークショップ, 6 pages (2010-2), 豊橋科学技術大学

⑮ 西崎博光, 胡新輝, 南條浩輝, 伊藤慶明, 秋葉友良, 河原達也, 中川聖一, 松井知子, 山下洋一, 相川清明, "Spoken Term Detectionのためのテストコレクション構築とベースライン評価," 情報処理学会研究報告, Vol. 2010-SLP-81 No. 13, (2010-7), 東京工業大学

⑯ 秋葉友良, 西崎博光, 相川清明, 河原達也, 松井知子, 伊藤慶明, 胡新輝, 中川聖

一, 南條浩輝, 山下洋一, "NTCIR-9 Spoken Doc: 音声検索語検出と音声ドキュメント検索の評価枠組みの設計", 情報処理学会研究報告, Vol2010-SLP-84 No. 18, (2010-12), 代々木

⑰ 中野拓也, 伊藤慶明, 小嶋和徳, 石亀昌明, 田中和世, 李時旭, "音声中の検索語検出における triphone モデル集約方式の検討," 第5回音声ドキュメント処理ワークショップ, 6 pages (2011-3), 豊橋科学技術大学

⑱ 栗城吾央, 伊藤慶明, 小嶋和徳, 石亀昌明, 田中和世, 李時旭: "音声認識結果の読み系列を用いた音声文書検索の辞書未登録語における高精度化手法の検討," 音響学会論文集, 3-Q-33, pp. 227-228 (2008-09), 九州大学

⑲ 小野寺悠二, 伊藤慶明, 小嶋和徳, 石亀昌明, 田中和世, 李時旭: "複数のサブワードを用いた音声文章検索におけるサブワード信頼度の導入," 音響学会論文集, 3-Q-32, pp. 279-281 (2009-03), 東京工業大学

⑳ 李時旭, 児島宏明, 田中和世, 伊藤慶明: "サブワード連鎖モデルの効率的なクラスタリングに基づく音声語検出," 音響学会論文集, 3-Q-32, pp. 283-284 (2009-03), 東京工業大学

(21) 栗城吾央, 伊藤慶明, 小嶋和徳, 石亀昌明, 田中和世, 李時旭: "音声認識結果の読み系列を用いた音声文書検索の辞書未登録語における高精度化手法の検討," 音響学会論文集, 3-Q-33, pp. 227-228 (2009-09),

(22) 小野寺悠二, 伊藤慶明, 小嶋和徳, 石亀昌明, 田中和世, 李時旭: "複数のサブワード検索結果統合による音声文書検索-検索語毎・候補区間毎の結合加重推定法-", 音響学会論文集, 3-Q-32, pp. 279-281 (2009-09) 日本大学郡山

(23) 栗城 吾央, 伊藤慶明, 小嶋和徳, 石亀昌明, 田中和世, 李時旭: "未知語音声クエリにおける音声中の検索語検出-Webを利用した拡張辞書とサブワードの認識結果の統合-", 音響学会論文集, 1-Q-19, pp. 197-200 (2010-03), 電気通信大学

(24) 江六前政宏, 伊藤慶明, 小嶋和徳, 石亀昌明, "ビデオ内音響情報の時空間配置を特徴量とした一致区間検索方式の評価," 4R-5, 情報処理学会全国大会, (2010-3), 東京大学

(25) 小笠原彩夏, 梁良, 石亀昌明, 小嶋和徳, 伊藤慶明, "個人特徴の誇張を利用した似顔絵作成システムの評価," 1D-5, 情報処理学会全国大会, (2010-3), 東京大学

(26) 杉井ひかり, 石亀昌明, 伊藤慶明, 小嶋和徳, "カオスニューラルネットワークにおけるパラメータの検討~自動作曲への応用を目指して~, " 2D-2, 情報処理学会全

国大会, (2010-3), 東京大学
(27) 今康德, 小嶋和徳, 伊藤慶明, 石亀昌明, “並列分散遺伝的アルゴリズムにおける群構造化による探索性能への影響に関する考察,” 4U-6, 情報処理学会全国大会, (2010-3), 東京大学
(28) 前田紘輝, 小嶋和徳, 伊藤慶明, 石亀昌明, “多点局所探索法による画像高解像度化における評価画像選定方法についての検討,” 4X-4, 情報処理学会全国大会, (2010-3), 東京大学
(29) 阿部俊祐, 小嶋和徳, 伊藤慶明, 石亀昌明, “形状・色特徴を利用した毒キノコ識別支援ツールの評価方法に関する研究,” 2Y-5, 情報処理学会全国大会, (2010-3), 東京大学
(30) 小友知己・伊藤慶明・小嶋和徳・石亀昌明, “動画検索のための MPEG-2 データ中の特徴量抽出方法の検討,” 情報処理学会 FIT2010, I-071 (2010-9), 九州大学
(31) 小友知己・伊藤慶明・小嶋和徳・石亀昌明, “MPEG-1/2 中の直流成分を利用した動画検索方式の評価,” 情報処理学会 2010 全国大会 (2011-3), 東京工業大学
(32) 斉藤裕之, 伊藤慶明, 小嶋和徳, 石亀昌明, 田中和世, 李時旭, “複数サブワード認識結果統合による音声中の検索語検出の精度向上 — 複数の音響モデル・言語モデルの利用 —,” 情報処理学会 2010 全国大会 (2011-3), 東京工業大学
(33) 本館修史, 伊藤慶明, 小嶋和徳, 石亀昌明, 田中和世, 李時旭, “音声クエリによる未知語問題の改善と検索精度の向上,” 情報処理学会 2010 全国大会 (2011-3), 東京工業大学
(34) 大原宏太, 小嶋和徳, 伊藤慶明, 石亀昌明, “看板画像からの文字抽出における処理領域の限定と特徴量の補強に関する研究,” 情報処理学会 2010 全国大会 (2011-3), 東京工業大学
(35) 吉田農里, 小嶋和徳, 伊藤慶明, 石亀昌明, “撮影画像データを対象とした画質評価方法の検討,” 情報処理学会 2010 全国大会 (2011-3), 東京工業大学
(36) 増尾拓朗, 小嶋和徳, 伊藤慶明, 石亀昌明, “SIFT 特徴の軸方向左右反転による擬似特徴を利用した教師特徴の増補,” 情報処理学会 2010 全国大会 (2011-3), 東京工業大学
(37) 谷藤史崇, 伊藤慶明, 小嶋和徳, 石亀昌明, 田中和世, 李時旭, “適切なモデル間距離による音声中の検索語検出の精度向上,” 日本音響学会春季研究発表会, 2-P-58, 2 pages (2011-3), 東京工業大学
(38) 中野拓也, 伊藤慶明, 小嶋和徳, 石亀昌明, 田中和世, 李時旭, “音声中の検索語検出における出現頻度情報を用いた

triphone モデル数の検討,” 日本音響学会春季研究発表会, 2-P-62, 2 pages (2011-3), 東京工業大学

6. 研究組織

(1) 研究代表者

伊藤 慶明 (ITOHI YOSHIAKI)
岩手県立大学・ソフトウェア情報学部・
准教授
研究者番号 : 90325928

(2) 研究分担者

田中 和世 (TANAKA KAZUYO)
千葉工業大学・工学部・教授
研究者番号 : 70344207

大川 茂樹 (OOKAWA SHIGEKI)
千葉工業大学・工学部・教授
研究者番号 : 40306395

伊藤 憲三 (ITOHI KENZO)
岩手県立大学・ソフトウェア情報学部・
教授
研究者番号 : 30305297