

音声ドキュメント検索における Web ページを用いたドキュメント拡張の効果

杉本 樹世貴^{†1} 西崎 博光^{†2} 関口 芳廣^{†2}

本稿では、音声ドキュメント検索のための Web ページを利用したドキュメント拡張について述べる。提案手法では、検索対象の音声ドキュメントから音声認識を利用して検索用インデックスを作成する他に、さらにそれらの音声ドキュメントと類似性の高い Web ページからも検索用インデックスを構築する。Web ページの検索は、検索対象の音声ドキュメントを音声認識した結果から Web 検索用のクエリを生成し、それを用いて検索を行う。これら 2 種類のインデックスを適宜組み合わせることで、音声認識誤りや未知語に対して頑健な検索が期待できる。実験の結果、提案手法の有効性を示すことができた。特に未知語を含んだクエリに関しては、補正 11 点平均精度で 4.2%の改善が得られた。

Effect of Document Expansion using Web Pages on Spoken Document Retrieval

KIYOTAKA SUGIMOTO,^{†1} HIROMITSU NISHIZAKI^{†2}
and YOSHIHIRO SEKIGUCHI^{†2}

This paper describes a novel method for spoken document retrieval using Web document expansion. This technique improves document retrieval performance by expanding target spoken documents using Web data. In this research, two types of indexes are built. The one index is made from the transcriptions of the spoken documents, the other is made from Web documents that are highly related to the target spoken documents. The Web documents are retrieved by a search engine that uses Web search queries from transcriptions of the target spoken documents. To utilize both the indexes will achieve robust spoken document retrieval for out-of-vocabulary (OOV) problem and errorful transcriptions by a speech recognizer. The experimental result has shown that the proposed method was very effective for solving the OOV problem. Especially, retrieval performance for only queries including OOV words has achieved improvement of 4.2% comparing with the baseline result.

1. はじめに

1990年代から、アメリカ国立標準技術研究所 (NIST) とアメリカ国防総省が主催する情報検索関連の評価型ワークショップ TREC(Text REtrieval Conference) が開催されている。1990年代後半から2000年までの間、TRECの研究分野(トラック)の1つとして、音声ドキュメント検索トラック (Spoken Document Retrieval Track) が開催され、世界中で音声ドキュメント検索の研究が行われてきた。現在は、ビデオ検索の評価ワークショップ (TRECVID^{*1}) が開催されており、映像中に含まれる音声データを音声認識する等して、ビデオ検索精度の改善が図られている¹⁾。

一方、日本では、2006年に情報処理学会音声言語処理研究会のワークショップである「音声ドキュメント処理ワーキンググループ」が設立された。ワーキンググループは、『日本語話し言葉コーパス』(以下CSJと記す)を用いた音声ドキュメント検索用のクエリと、クエリに対する正解ドキュメントのセットを構築した²⁾。

この評価セットはTREC SDR TrackやTRECVIDが採用しているニュース音声・映像とは異なり、話し言葉音声を対象としている。したがって、音声認識がより難しく、それに比例して検索精度が悪くなる。

この検索テストセットの構築により、日本でも大規模なデータを用いた情報検索技術開発の環境が整った。日本語の音声ドキュメント検索では、これまで、各研究機関が独自のデータベースを用いて研究を行ってきた^{3),4)}。各研究機関が共通して利用できる日本語の評価セットができたことから、音声ドキュメント検索のさらなる発展が期待できる。

将来的に、音声やビデオといったマルチメディアデータが爆発的に増加していくことを考えると、これらの検索技術を開発することは喫緊の課題であることは疑う余地はない。そこで本稿では、この検索テストセットの検索精度改善を図ることを目的に、Webページを用いた新しい検索方法の提案を行う。

^{†1} 山梨大学大学院医学工学総合教育部コンピュータ・メディア工学専攻

Dept. of Computer Science and Media Engineering, Educational Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi

^{†2} 山梨大学大学院医学工学総合研究部

Dept. of Research Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi

*1 <http://www-nlpir.nist.gov/projects/trecvid/>

通常、音声ドキュメントを検索する場合、検索対象の音声データを音声認識することで音声波形をシンボル化し、そこからドキュメントに対するインデックスを作成することで、検索を行うことができる。音声認識処理を用いるため、音声認識誤りや未知語問題（音声認識辞書に必要な単語が登録されていない）が、検索精度に大きく影響する。

それらの問題を解決するため、検索対象のデータと内容が類似した Web ページを音声ドキュメントのインデキシングに利用する。つまり、提案手法は Web データを用いて音声ドキュメントの内容を拡張する。これによって、本来ならインデックスに登録されて欲しい単語が、音声認識誤りや未知語により、登録されない事態を回避するのが狙いである。

検索実験の結果、Web ページを利用したドキュメント拡張により、検索精度を改善することができた。特に、未知語を含むクエリに対する検索精度が大きく改善したことから、提案手法の有効性を示すことができた。

2. 関連研究

音声認識誤りや未知語問題等を解決するために、90年代後半からこれまで、数多くの検索手法が開発されてきた。これは、前述した TREC の成果の賜物である。

例えば、Wechsler ら⁵⁾ は、音素認識器を使って音声を書き起こし、入力検索語の音素表記とのマッチングを行っている。Ng ら⁶⁾ は、音声の音素表記から作ったサブワードをインデキシングの単位として扱い、DP マッチングにより柔軟なマッチングを行っている。西崎らは³⁾、音声認識結果の誤りに対処するために、複数の音声認識システムを利用したり、音声認識システムに未知語検出処理を組み込むことで、音声認識システムで未知語と判定された区間のみをサブワード単位のインデックスを構築する手法を提案した。岩田らは⁴⁾、語彙に影響されない音声ドキュメント検索として、サブワード単位の認識結果を利用している。また、クエリとドキュメントにおいて、高速にサブワード同士をマッチングさせる手法を提案している。

近年では、音声認識結果のラティス⁷⁾ やコンフュージョンネットワークを利用したり⁸⁾、音声認識結果に信頼度を導入することで⁹⁾、検索性能改善を図っている研究例もある。

日本語でもテストコレクションが整備されたことから、この検索テストコレクションに対する研究も行われている。まず、文献²⁾では、テストコレクションを利用した音声ドキュメント検索の基本的な検索精度が述べられている。また、秋葉らは¹⁰⁾、テキスト翻訳の技術を利用した語彙拡張を音声ドキュメント検索に応用し、テストコレクションでの有効性を示した。さらに、胡らは⁸⁾、音声認識結果のコンフュージョンネットワークを利用すること

で、検索精度が改善できることを示した。前沢らは¹¹⁾、音声ドキュメントを認識する際に、様々な音声認識精度の書き起こしを用意し、認識精度が検索精度に与える影響や、サブワード（音節）をインデックスに用いることが、認識誤りや未知語に有効であることを示した。

音声ドキュメント検索において、本研究と同様に Web を利用する研究も報告されている。例えば、寺尾ら¹²⁾ は、検索に用いる音声クエリと関連性のある Web の情報を利用することで、クエリ拡張を図っている。また宇野ら¹³⁾ は、音声ドキュメントと関連性のある Web 情報の件数と未知語率の関連性を研究している。

音声ドキュメント検索において、クエリ拡張を利用している研究として、Mamou ら¹⁴⁾ が単語の発音に基づく拡張処理を提案している。

文献¹²⁾ に対する、本提案手法の相違点は、検索対象の音声ドキュメントの誤認識や未知語問題を解決する方法として、ドキュメント自体の文書拡張を Web ページによって行うことである。

音声ドキュメント検索において、文書拡張を導入した研究としては、Singhal らの報告がある¹⁵⁾。これは、ニュース音声を検索する際に、その音声と類似したデータをニュースコーパスから選択するという手法である。

今回、我々がターゲットにしているデータは、様々な話題を含んだ学会講演・模擬講演音声であり、これに類似したコーパスは今のところ存在しない。今後、様々な話題の音声（映像）ドキュメントが増加していきだろうが、文書拡張の手法を導入する際に用いる類似コーパスとしては、Web が適していると考えられる。ニュース音声をターゲットとし、同じドメインのニュースコーパスを用いて文書拡張するよりも、Web を使って文書拡張する方が遥かに難しい問題である。

本稿では、その問題に取り組み、Web を用いた文書拡張の可能性を示したい。

3. Web ページを利用した音声ドキュメント検索処理

提案手法の処理の流れを図 1 に示す。

まず、検索対象の音声ドキュメントを大語彙連続音声認識システムを用いて音声認識する。このとき、音声認識システムの認識辞書に登録されていない単語が未知語となり、クエリに未知語が含まれている場合、インデックスとマッチングができない。

この認識結果からストップワードを取り除き、インデックスを構築する。本稿では、これを“認識インデックス”と記す。

次に、Web ページから作成するインデックスについて説明する。まず、音声認識結果が

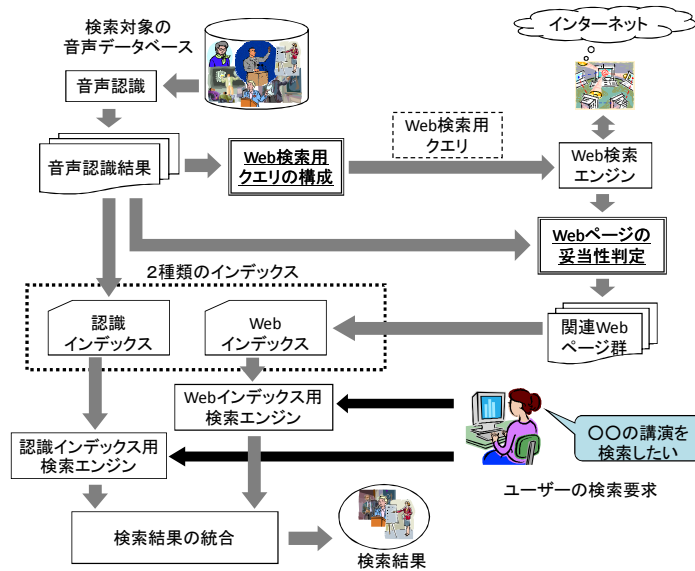


図1 提案手法の処理の概要

Fig. 1 A framework of spoken document retrieval by using Web document expansion

ら、認識されたドキュメントと内容が類似した Web ページを検索するためのクエリを作成する。ここで構築する Web 検索用クエリの“質”によって、どれだけ内容が近い Web ページを収集できるかが決定される。したがって、クエリの生成方法にも技術が必要である。しかし、本研究は Web ページによる音声ドキュメント拡張の有効性を示すことが目的であるため、Web 検索クエリの生成方法は単純な方法を採用した。これについては次節で述べる。

Web 検索クエリによって収集された Web ページの中には、音声ドキュメントの内容と無関係のものも多く含まれている。それらの Web ページは、認識結果との関連性の尺度に基づいて取り除いている。

このようにして集めた Web ページから、認識インデックスを作成した場合と同様にストップワードを取り除き、インデックスを構築する。本稿ではこれを“Web インデックス”と記す。

このように、音声ドキュメントに対して 2 種類のインデックスが構築される。つまり音声認識結果から直接作成される“認識インデックス”と、関連する Web ページから作成され

る“Web インデックス”である。

検索処理では、それぞれのインデックスに対してマッチング処理を行うため、2 つの検索エンジンを用いる。ユーザの検索要求を 2 つの検索エンジンに入力し、それぞれのエンジンから得られた検索スコアを統合することで、最終的な結果を得る。

4. 関連する Web ページの取得

Web 検索を使用するには、キーワード抽出を行い話題に合った検索クエリを構成しなければならない。しかし、音声認識結果が形態素に分割されていること、話題の特定に不向きな単語（機能語や一般語）が多く含まれていることから、話題に合った検索クエリを構成することは困難である。

そこで適切な Web ページを検索するための検索クエリの構成法を検討した。

4.1 Web 検索用クエリの構成法

クエリの構成法として、認識結果で形態素に分割された単語の品詞を判定し適宜連結と削除を行う。

具体的には、以下の手順で行う。

1. 単語の連結：話題を特定する単語（主に名詞）が続く限りそれらを連結し、単語 N-gram を作成する。このとき、Web 日本語 N-gram 第一版^{*1} を使用し、連結した単語が実在するかを確認する。Web 日本語 N-gram に登録されている単語はそのまま連結を行い、登録されていないものに対しては、連結を行わず単体のまま使用する。

例「横軸」は「学習」データ「数」縦軸「に」

N-gram 使用 「横軸」は「学習」データ「数」縦軸「に」

N-gram 未使用 「横軸」は「学習」データ「数」縦軸「に」

2. 不要語の削除：話題の特定に向いていない単語（主に助詞、助動詞）を不要語として除去する。

例「横軸」は「学習」データ「数」縦軸「に」

「横軸」学習データ「数」縦軸

これらの方法で整形された単語に対し、出現頻度を求め上位 5 単語の組み合わせを Web 検索用クエリとして使用する。

*1 <http://www.gsk.or.jp/catalog/GSK2007-C/catalog.html>

4.2 Web ページの検索と妥当性判定

収集された Web ページの中には、音声ドキュメントの内容と無関係のものも多く含まれている。

そこで、収集した Web ページと認識結果の類似度を尺度として用いることにより、インデックスとして使用する Web ページの取捨選択を行う。

類似度計算にはコサイン尺度を用いる。これは、各文書をその文書に含まれる単語の出現頻度を要素とした文書ベクトルで表現し、文書ベクトル間のなす角を求めることにより類似度を求めるものである。コサイン尺度の計算式を式 (1) に示す。

$$\cos(d_j, r) = \frac{\sum_{i=1}^m d_{ij} r_i}{\sqrt{\sum_{i=1}^m d_{ij}^2} \sqrt{\sum_{i=1}^m r_i^2}} \quad (1)$$

ここで、 r は認識結果、 r_i は認識結果中の単語 w_i の出現頻度、 d_{ij} は Web 文書 d_j 中の単語 w_i の出現頻度、 m は r_i と w_i に含まれる単語の種類数を表す。

妥当性評価を行った結果、インデックスとして使用する Web ページが 0 件となったドキュメントに対しては、認識インデックスを代用する。

5. 検索システム

5.1 検索手法

音声ドキュメントの検索エンジンには、汎用連想計算エンジン GETA^{*116)} を利用した。GETA を利用することで、大規模なドキュメント-単語集合間の類似度を高速に計算することが可能である。

索引語には、形態素単位を採用する。先行研究¹⁷⁾ では、形態素の他に文字バイグラム、および形態素と文字バイグラムの組み合わせが試されており、その効果が示されている。

認識インデックスおよび Web インデックスを構築する際に、ストップワードを除去して内容語のみをインデキシングする。本研究では、内容語として、名詞、動詞、形容詞、形容動詞を用い、これ以外の品詞をストップワードとする。

ドキュメントとクエリ間の類似度計算には、GETA に用意されている SMART 法^{*218)} を採用した。索引語は、ドキュメント長で正規化された TF・IDF 法により重み付けされてい

る。ベクトルで表現されたドキュメントとクエリから、余弦計算に基づいてドキュメント・クエリ間の類似度を計算する。

5.2 検索結果の統合

認識インデックスと Web インデックス、それぞれのインデックスからの検索結果を統合し、最終的な結果を得る。

2 つのインデックスの統合は、式 (2) に示すように、認識インデックスの検索結果の検索スコアと Web インデックスの検索結果の検索スコアの線形補間により行う。

$$\text{sim}(d) = (1 - \alpha) \times \text{sim}(d|r) + \alpha \times \text{sim}(d|w) \quad (2)$$

ただし、 $\text{sim}(d|r)$ は認識インデックス、 $\text{sim}(d|w)$ は Web インデックスを用いたときの文書 d の検索スコアとなる。 α は、各インデックスに対する重み係数である。この α を、0.0 から 1.0 までの 0.1 刻みで変化させ検索実験を行う。

また、検索クエリに対する未知語率を *OOV* と定義したとき、この *OOV* を重み係数として用いることで (つまり、 $\alpha = \text{OOV}$)、クエリ毎に動的に重みを変化させる手法も実験する。ここで用いる未知語率 (*OOV*) とは、音声ドキュメントを検索するためのクエリに含まれる単語のうち、音声認識辞書に含まれていない単語の割合である^{*3}。

6. 検索実験

6.1 テストコレクション検索条件

検索対象のデータには、CSJ を対象に構築された検索評価用テストコレクションを用いる。

6.1.1 テストコレクションについて

テストコレクションには、CSJ に収録されている「学会講演」「模擬講演」の 2702 講演 (約 600 時間以上の音声データ) に対し、それら講演、または講演の一部を検索するための検索クエリと正解の講演データの組が収録されている。正解リストには、クエリに対して適合している講演 (Relevant)、部分的に適合している講演 (Partially Relevant) が含まれている。クエリの数は全部で 39 個である。

6.1.2 音声認識

テストコレクションには、自動音声認識により得られた 2702 講演の書き起こしも付属している。今回の実験ではそれを用いずに、2702 講演を再度音声認識し直した。その理由は次の 2 つである。

*1 <http://geta.ex.nii.ac.jp/>

*2 <http://geta.ex.nii.ac.jp/getaN2002/gdoc/geta/tutorial/libae/section5.html>

*3 音声ドキュメントに対する未知語率でないことに注意されたい。

1つ目の理由は、付属の書き起こしは、クローズドな音響モデル、言語モデルを用いて音声認識されているため、比較的音声認識率が良いことである。また、全39個のクエリにおいて、未知語を含んでいるものは少ない。今回の研究の目的は、提案手法が未知語や認識誤りに頑健であるかどうかを確かめることであるので、未知語が少ない状況では得られる効果が薄いかもしれない。したがって、講演を再度認識する際に辞書のサイズを少なくすることによって未知語を増やし、未知語が含むクエリが増えるようにした。

2つ目の理由は、言語モデル学習データを作成する際に利用する形態素解析器とWebページを解析する形態素解析器を同一のものにしたかったからである。

現状では、2702講演を音声認識するために必要なモデルを完全にオープンな学習データから学習することは難しい*1。今回の実験では、できるだけオープンになるように、文献¹⁷⁾で用いられている音響・言語モデルよりも、学習データを少なくした。

本研究で、音声認識に用いた音響モデルは、CSJ学会講演から認識評価用セットを除いた学会講演970講演から学習したトライフォンである。使用する特徴量は、12次元のMFCC、 Δ MFCC、 $\Delta\Delta$ MFCCと、 Δ パワー、 $\Delta\Delta$ パワーの全38次元である。

言語モデルは、音響モデルの学習データと同じ学会講演970講演から、トライグラムを学習した。語彙のサイズは17kである。これにより、39個のクエリの約3分の1にあたる11個のクエリに未知語が含まれることになった(クエリ全体に対する未知語率は11.8%)。

音声認識には、Julius rev.4.1を用いた。2702講演に対する音声認識実験の結果、単語正解率で76.9%、正解精度で71.6%となった。

6.1.3 検索単位

本研究では、検索の単位は講演単位とする。すなわち1ドキュメント1講演である。

検索されたドキュメントの適合性の判定は、“適合”はもちろんのこと“部分適合”しているドキュメントも正解とする。

6.2 評価尺度

検索精度を計る尺度には、補間11点平均精度(Interpolated 11-points Average Precision, “11-pt AP”もしくは単にAPと記す)を用いる。

これは、再現率(Recall)が0.0, 0.1, 0.2, ..., 1.0のときの補間精度(Precision)を平均したものである。再現率は0~1.0までの11点であるので、11点の補間精度を平均する。11-pt APの計算式を式(3)に示す¹⁷⁾。

*1 ただし、クロスバリデーションを行えば、膨大な時間がかかるがオープンなモデルを構築することは可能

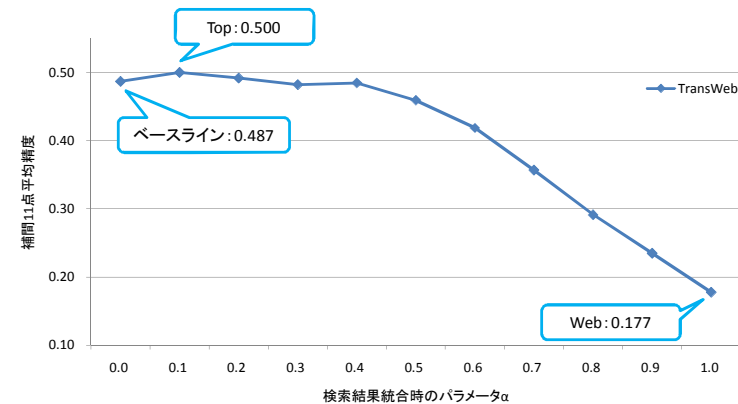


図2 手動書き起こしデータに対する補間11点平均精度の変化
Fig.2 Variation of 11pt-AP for text document retrieval.

$$IP(x) = \max_{x \geq R_i} P_i, \quad AP = \frac{1}{11} \sum_{i=0}^{10} IP\left(\frac{i}{10}\right) \quad (3)$$

ここで、 R_i と P_i は、それぞれ、検索結果の上位*i*番目までに着目したときの再現率と精度である。

検索実験では、上位1000件までの結果を出力し、上記の式により評価する。

6.3 実験結果

Webページを用いたドキュメント拡張の評価実験の結果を図2、図3、図4に示す。

まず、図2は音声ドキュメントを音声認識システムではなく、人手により書き起こしたドキュメント集合を対象にした検索実験結果である。すなわち、音声認識誤りを含んでいないため、純粋なテキスト検索結果となる。検索対象のドキュメントに対する類似Webページは4節で述べた方法で集められている。

図2を見ると、 $\alpha = 0.0$ のとき、すなわちWebによるドキュメント拡張を行っていない結果(ベースライン)に比べ、ドキュメント拡張を行った方が検索結果が良くなっている。 $\alpha = 0.1$ のときに、ベースラインよりも1.3%の改善が得られている。この結果より、音声認識誤りのないテキスト検索においても、Webによるドキュメント拡張の効果があること

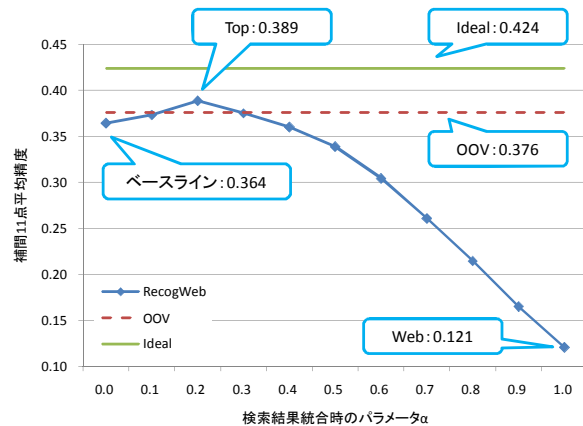


図3 音声ドキュメント集合に対する補間 11 点平均精度の変化 (全 39 クエリ)
Fig.3 Variation of 11pt-AP for spoken document retrieval (all 39 queries).

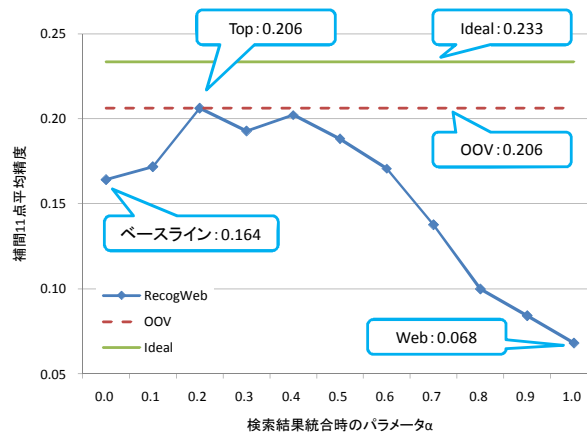


図4 音声ドキュメント集合に対する補間 11 点平均精度の変化 (未知語を含む 11 クエリ)
Fig.4 Variation of 11pt-AP for spoken document retrieval (only 11 queries including OOV).

が示された。

次に、図 3 は、全 39 個のクエリーに対する音声ドキュメント検索結果を示してい

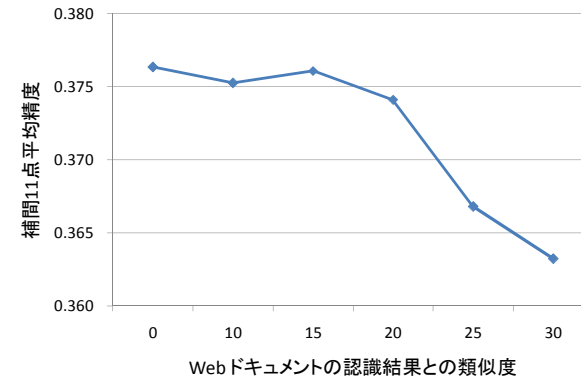


図5 Web ページの妥当性判定を取り入れた時の検索性能 (全 39 クエリ)
Fig.5 Spoken retrieval performance with Web page filtering.(39queries)

る。 ”RecogWeb” は、音声ドキュメントを Web ページで拡張し、統合時に重みを変化させたときの結果を表している。 ”OOV” は、式 (2) において、 α の値をクエリ毎の未知語率を用いて動的に決定したときの結果である。さらに、 ”Ideal” は、クエリ毎に最も検索精度が高くなるように α を決定したときの結果で、いわば最適な統合が行われた時の理想となる結果である。なお、4.2 節で述べた Web ページの妥当性判定を行っていないときの結果である。

図 3 が全てのクエリに対する結果だったのに対し、図 4 は、音声認識辞書に対する未知語を含んでいる 11 クエリに対する実験結果となっている。

図 3 の実験結果をみると、 $\alpha = 0.0$ 時 (Web 拡張を行わない) のベースラインに比べ、Web ページによりドキュメント拡張を行うことで検索精度が改善している。 $\alpha = 0.2$ の時に最も精度が良くなり、ベースラインに比べて 2.5% の改善が得られた。一方で、未知語率により動的に重みを変化させたときは、ベースラインに比べて 1.2% の改善に留まった。しかし、図 4 に示す、未知語のクエリだけに着目してみると、動的な重みの使用によりベースラインに比べて大幅な改善がみられた。

以上の結果から、音声ドキュメント検索において、Web ページによるドキュメント拡張の有効性を示すことができた。特に、認識辞書に対する未知語を含んでいるクエリに対しては、その改善幅は 4.2% にもなっている。ただ、最適な重みを選択したときの理想値の精度をみると、まだ改善の余地が残っているため、今後は動的な重み決定手法を工夫する必要がある。

ある。

続いて、図5は、Web インデックスとして使用する Web ページを類似度尺度により選択した場合の検索精度を示している。図5を見ると、類似度の閾値を0、すなわち類似度により Web ページを選択するよりも、収集した Web ページ全てをインデックスとして用いた場合の検索精度の方が良い結果となった。

これは今回、Web 検索用クエリを比較的簡単な手法により構成したことにより、音声ドキュメントの内容と無関係の Web ページが多く集まってしまったこと、また、類似度計算の方法も単語の出現頻度のみを使用していることが原因であると考えられる。従って、今後は、より適した Web ドキュメントを収集する方法を開発する必要がある。

7. おわりに

本稿では、音声ドキュメント検索のテストコレクションを用いて評価実験を行うことにより、類似性の高い Web ページによる音声ドキュメントの拡張の有効性について示した。

実験の結果、音声認識結果のみをインデックスとした場合よりも、Web ページを用いることで、39 個全て検索クエリに対しては、最大で 2.5%の改善、さらに未知語を含む 11 個の検索クエリに対しては、最大で 4.2%の検索精度の改善を確認できた。

今後は、Web 検索用クエリの”質”が Web インデックスの”質”を決める上で重要となってくるため、Web 検索用クエリの構成法の再検討を行い、Web インデックスの”質”を改善していくとともに、検索結果の統合時に使用するパラメータの動的な決定方法などを検討し、検索精度を改善していく予定である。

参 考 文 献

- 1) Cheng, Y.-J. and Chen, H.-H.: Aligning Words from Speech Recognition and Shots for Video Information Retrieval, *Proc. of the TRECVID2004* (2004).
- 2) Akiba, T., Aikawa, K., Itoh, Y., Kawahara, T., Nanjo, H., Nishizaki, H., Yasuda, N., Yamanashita, Y. and Itou, K.: Construction of a Test Collection for Spoken Document Retrieval from Lecture Audio Data, *IPJS Journal*, Vol.50, No.2, (2009).
- 3) 西崎博光, 中川聖一: 音声認識誤りと未知語に頑健な音声文書検索手法, 電子情報通信学会論文誌, Vol.J86-D-II, No.10, pp.1369-1381 (2003).
- 4) 岩田耕平, 伊藤慶明, 小嶋和徳, 石亀昌明, 田中和世, 李時旭: 語彙フリー音声文書検索手法における新しいサブワードモデルとサブワード音響距離の有効性の検証, 情

報処理学会論文誌, Vol.48, No.5, pp.1990-2000 (2007).

- 5) Wechsler, M., Munteaun, E. and Schauble, P.: New Techniques for Open-Vocabulary Spoken Document Retrieval, *Proc. of ACM SIGIR'98*, pp.20-27 (1998).
- 6) Ng, K. and Zue, V.W.: Subword-based approaches for spoken document retrieval, *Speech Communication*, Vol.32, No.3, pp.157-186 (2000).
- 7) cheng Pan, Y., lin Chang, H., Chen, B. and shan Lee, L.: Subword-based Position Specific Posterior Lattices (S-PSPL) for Indexing Speech Information, *Proc. of the INTERSPEECH 2007*, pp.318-321 (2007).
- 8) 胡新輝, 吳友政, 柏岡秀紀: Confusion Network を用いた音声ドキュメントの検索及び評価に関する研究, 第2回音声ドキュメント処理ワークショップ講演論文集, 豊橋技術科学大学メディア科学リサーチセンター, pp.85-90 (2008).
- 9) Kim, W. and Hansen, J. H.L.: Advances in SpeechFind: Transcript Reliability Estimation Employing Confidence Measure based on Discriminative Sub-word Model for SDR, *Proc. of the INTERSPEECH 2007*, pp.2409-2412 (2007).
- 10) Akiba, T. and Yokota, Y.: Spoken Document Retrieval by Translating Recognition Candidates into Correct Transcriptions, *Proc. of the INTERSPEECH 2008*, pp.2166-2169 (2008).
- 11) 前沢慎吾, 杉本樹世貴, 西崎博光, 関口芳廣: 音声ドキュメント検索用テストコレクションにおける音節インデックスの効果, 第71回全国大会講演論文集第2分冊, 情報処理学会 E3-10 (2009).
- 12) Terao, M., Koshinaka, T., Ando, S., Isotani, R. and Okumura, A.: Open-Vocabulary Spoken-Document Retrieval Based on Query Expansion Using Related Web Documents, *Proc. of the INTERSPEECH 2008*, pp.2171-2174 (2008).
- 13) 宇野有, 伊藤彰則, 牧野正三: 音声ドキュメントの索引付けに向けたウェブ検索を用いたデータ収集における未知語率の検討, 春季講演論文集, 3-Q-30, 日本音響学会, pp.275-276 (2009).
- 14) Mamou, J. and Ramabhadran, B.: Phonetic Query Expansion for Spoken Document Retrieval, *Proc. of the INTERSPEECH 2008*, pp.2106-2109 (2008).
- 15) Singhal, A. and Pereira, F.: Document Expansion for Speech Retrieval, *Proc. of ACM SIGIR'99*, pp.34-41 (1999).
- 16) 高野明彦, 西岡真吾, 丹羽芳樹: 連想に基づく情報アクセス技術: 汎用連想計算エンジン GETA を用いて, 情報の科学と技術, Vol.54, No.12, pp.634-639 (2004).
- 17) 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, 伊藤克亘: 音声ドキュメント検索テストコレクションの試作と基本検索性能評価, 第1回音声ドキュメント処理ワークショップ講演論文集, 豊橋技術科学大学メディア科学リサーチセンター, pp.73-80 (2007).
- 18) Singhal, A., Buckley, C. and Mitra, M.: Pivoted document length normalization, *Proc. of ACM SIGIR'96*, pp.21-29 (1996).