

Title	音声データベースに対する情報検索
Author(s)	前田勇希
Citation	
Issue Date	2001-09
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1550
Rights	
Description	島津明, 情報科学研究科, 修士

音声データベースに対する情報検索

前田勇希

北陸先端科学技術大学院大学 情報科学研究科

2000 年 8 月 15 日

キーワード: 音声検索, 音声, 音声認識, ニュース音声.

記憶デバイスの大容量化やネットワークの広帯域化に伴い音声の電子化が行なわれはじめている．しかし，この音声を検索する手段についてはまだよいものがないのが現状である．このような中で音声認識を用いた音声検索手法に関する研究が進められているが，音声認識可能な語彙数の制約の問題などがある．そこで，語彙数の制約を取り除くため単語の音素に着目し，その音素間の類似度から単語間の類似度を求める研究が行われている．

本研究では，ニュース記事読み上げ音声を対象とした，音声認識の語彙の制約のない検索手法を提案し，システムを実装してその評価を行った．このシステムでは，confusion matrix (CM) を用いた照合手法によりクエリー音素に合致した記事を検索する．

音声をかな漢字交じり文に書き起こすためには，音声中で用いられている語彙が音声認識器の辞書に含まれていなければならない，また同音異義語や形態素区切りを適切に処理する必要がある．しかし，実際には常に新しい語が出現するという問題がある．本研究では，通常のかな漢字交じり文ではなく，発音を示す記号である音素を対象とした検索を試みる．音素認識には辞書が必要でないため，語彙の制限のない検索が可能である．ここでは，RWCP ニュース音声コーパスと大語彙日本語音声認識器 *julius* を用いて音素列を書き起こすこととする．なお，この組み合わせでの音素認識の誤り率は約 34%であった．

検索のためのクエリーは音素列として与えられる．記事の朗読音声から音声認識で生成した音素列とクエリーの音素列の距離を計算し，関連があるか否かを判定する．しかし，認識誤りが混入するため，誤りなく認識された音素列だけでなく，認識誤りによって長さが変動した音素列も検索できる必要がある．

認識誤りのモデル化のため，訓練用の音声と手で書き起こした音素列を用いて，ある音素が音声認識によって観測されたときにそれが実際はどの音素であるかという確率を CM という行列へとまとめておく．これを用いてクエリー中の音素と記事中の音素との距離を計算する．また，比較のため CM を用いず，音素が同一のとき 1，異なるとき 0 とする binominal method (BM) についても計算する．

音声認識に用いられている連続 DP マッチング法を用いて，クエリーの音素列と記事の音素列の間の距離計算を行う．また，連続 DP マッチング法を確率が扱うことが出来るように拡張した確率連続 DP マッチング法をについても，クエリーと記事との距離を計算する．これらの距離に基づいて，クエリーと記事との関連の有無を判定する．

検索手法の有効性を検証するため，音声全文検索エンジン JAISearch を，Web サーバ上で動作する CGI として実装した．

モデル化が有効であるか否かを実際に検索を行うことによって評価する．6 話者によって発声された全 246 記事から各話者ごとに記事数を平均するようにして 91 記事を任意に選び誤りの傾向を知るための訓練用データとし，89 記事を任意に選び評価用のデータとした．評価用データセット対して，音素列クエリー 50 個を用意した．1 クエリーあたりの参照記事数は約 11 記事で平均音素数は 6.46 である．

評価に先立ち，音声認識で得られた音素列に対して認識誤りの存在を考慮せず，クエリーの音素列を完全に一致する音素列を抽出するという方法で検索を行った．認識誤りのため，*recall* 32%，*precision* 77%，*F-measure* 24% となった．これに対し，連続 DP と BM では *recall* 51%，*precision* 59%，*F* 29%，連続 DP と CM では *recall* 65%，*precision* 52%，*F* 31% であった．確率連続 DP と CM では *recall* 62%，*precision* 68%，*F* 36% であった．

音素認識に基づく音声検索として，BM と CM，連続 DP マッチング及び確率連続 DP マッチング法について実験結果を示した．CM と確率連続 DP マッチングの組み合わせがもっともよい *F-measure* を示した．