

## 距離付き n-gram インデックスによる 認識誤りと未知語に頑健な高速検索法

岩見圭祐<sup>†1</sup> 藤井康寿<sup>†1</sup>  
山本一公<sup>†1</sup> 中川聖一<sup>†1</sup>

ニュースや新聞記事のようにテキスト情報を含むものであれば既存のテキスト検索エンジンを用いることで、欲しい情報を高速に検索することができる。しかし、現在のところ音声ドキュメントに対しての有効な検索手法は確立されていない。その理由として挙げられるのが、未知語や認識誤りといった音声ドキュメント特有の問題である。本研究ではこれらの問題を解決し、音声ドキュメントに対して有効な検索手法について検討する。すなわち、音声ドキュメントを音節単位の認識で音節ラティス化し、音節の置換誤り、挿入誤りを考慮した n-gram でインデックス化しておく。クエリは脱落誤りを考慮した n-gram で検索する。また、認識結果の尤度に基づいてインデックスを削減する方法を提案する。

### High-Speed Spoken Document Retrieval method for Out-of-Vocabulary Term by n-gram Index with Distance

KEISUKE IWAMI,<sup>†1</sup> YASUHISA FUJII,<sup>†1</sup>  
KAZUMASA YAMAMOTO<sup>†1</sup> and SEIICHI NAKAGAWA<sup>†1</sup>

We can find the information with an existing textual search engine if the target data consist of text information such as news and newspaper, but efficient spoken document retrieval (SDR) method is not currently the established, because spoken document has specific problems such as some recognition errors and out-of-vocabulary(OOV) terms. The aim of this study is to develop a robust SDR method. Spoken document is transformed to a syllable lattice by continuous syllable recognition and, then it is represented by the n-gram index with distance considering substitution errors and insertion errors. Deletion errors in recognition results are considered by the modification of query. In addition, we propose a method to reduce the index based on likelihood of recognition result.

#### 1. はじめに

本稿では、音声ドキュメント内の未知語を高速で検索する手法を提案する。未知語を検索可能にするために音節単位での認識結果を用いる。認識した結果から音節ラティスのトライグラムを作成し、そのトライグラムを用いてインデックスを構築する。インデックスは辞書順にソートしておくことで高速な検索が可能となる<sup>1)2)</sup>。認識誤りへの対策としては、今までも複数候補を用いて置換誤りに対処する手法が提案されているが<sup>3)19)</sup>、本稿ではこれに加え、挿入誤りと脱落誤りに対して対処する手法を実装した。また、これらの認識誤り対策の結果、インデックスのサイズと検出候補が増加するという問題に対して、音声認識の音響尤度を利用し、認識結果で尤度の低い部分だけ複数候補を用いることによってインデックスの圧縮おこなうことと、高速に検索を行うために認識誤り対策をどの程度おこなったのかという情報を距離として導入し、この距離をインデックスに付随させる手法によってインデックスと検出候補を削減を実現した。

#### 2. 音声ドキュメント検索手法

##### 2.1 大語彙連続音声認識の利用

従来の音声ドキュメント検索手法として、最も簡単な方法は、大語彙連続音声認識の書き起こしの結果に対して単語単位のテキスト検索を行う方法である。既知語に対しては高精度な検索が可能である。しかし、この方法では、未知語や、音声認識誤りの問題に対処することができない。未知語とは、大語彙連続音声認識の辞書にない単語のことである。辞書にない単語は、認識結果に現れることがないため、単語単位のテキスト検索では、未知語を検出することは不可能である。また、置換、挿入、脱落の認識誤りの問題もある。置換、脱落の認識誤りによって、辞書に登録されている単語であっても認識結果に現れない場合があり、その場合はテキスト検索を使用しても検索することができなくなってしまう。そのため、大語彙連続音声認識システムの性能によってテキスト検索の性能も決まってしまう。

##### 2.2 サブワード単位認識の利用

未知語に対しては、サブワード列として音節単位で認識した結果を使用する。音節単位の認識のおいても、単語単位の認識と同様に3種類の誤りが生じる。ドイツ語に対しては、

<sup>†1</sup> 豊橋技術科学大学  
Toyohashi University of Technology

5000 個の音節を用いた音節同士の重み付きレーベンシュタイン距離に基づく検索方法が提案されている (およそ、半分の単語が 1 音節語)<sup>4)</sup>。なお、中国語は音節数が 416 個で少ないため、検索の基本単位としてよく用いられる<sup>5)</sup>。また、置換、挿入、脱落誤りを考慮した音節列同士のマッチングに基づく検索方法も試みられている<sup>6)</sup>。音素の n-gram を用いた検索方法も種々提案されているが<sup>2)</sup>、基本的には bag of words の使い方で、音素認識誤りは考慮されていない。<sup>7)8)</sup>

日本語の音節数は 100 余種類と比較的少なく扱いやすい。音節列として認識することによって、認識の際に単語辞書を使用しないので、文法の制約を無視でき、未知語の発音をそのまま認識できる可能性がある<sup>9)</sup>。そこで、音節単位で認識した音節ラティスをサブワード列として用意しておき、音節ラティスの n-gram を用いる。手島らは音節認識結果をサフィックスアレイとしてテーブル化しておき、検索時に置換、挿入、脱落誤りを許しながらテーブルを探索する方法を提案しているが<sup>10)</sup>音節ラティスへの適用は困難である。

### 2.3 DP マッチングによる検索 (ベースライン)

誤りを含んだテキストに対しての検索として、一般的に用いられる連続 DP をベースラインとして使用する。本稿で扱う連続 DP では、サブワード単位で検索対象と検索クエリとの音節間距離を置換、挿入、脱落誤りを考慮して定義する。本稿では以下の 4 つの距離を用いた。

- ① 音節系列の編集距離
- ② 音節系列を音素系列に変換したもので、音素系列の編集距離
- ③ 音素間の編集距離を考慮した音節系列間の編集距離
- ④ 音節間の距離として対応する音節 HMM のバタチャリヤ距離<sup>11)</sup>

③に関しては、/ka/と/ga/の距離は 1、/ka/と/go/の距離は 2、/a/と/ka/の距離は 1 とするものである。

## 3. 認識誤り・未知語に頑健な高速検索手法

### 3.1 n-gram に基づく未知語検索法<sup>11)</sup>

本稿では、未知語に頑健な検索手法として、音節ラティスを使用して検索の際に認識誤りを考慮して検索を行う。本手法の概略を図 1 に示す。検索対象の音声ドキュメントに対して大語彙連続音声認識と連続音節認識を行い、インデックス化する。

未知語を検索可能にするため、サブワード列として音節ラティスの上位 m ベストを使用

する。そして、音節ラティスのデータを保持させておくデータ構造として、主にテキスト検索で用いられる n-gram インデックスを用いる。ここで用いる n-gram インデックスでは、音声ドキュメント内での出現位置情報と出現する n-gram の情報を保持させておく。n-gram インデックスの作成方法の概要を図 2 に示す。

この n-gram を辞書順に並べておけば、2 分探索で高速に検索できる。同じ n-gram が複数個連続してインデックス表に格納されることがあるが、これに対しては種々の改良法がある。また、本稿では n=3 としてトライグラムという固定長に限定しているため、トライグラムの種類とインデックスの位置を 1 対 1 に対応しているため、この関係を用いれば 2 分探索よりも高速に検索できる。4 音節長以上の検索後に対しては、3 音節の複数の組に分割し、それぞれで検索し、検索候補結果を連続性を考慮して、構成できるかどうかで、検索後の検索候補を出力する。

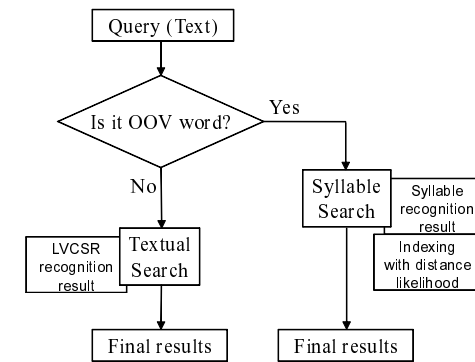


図 1 提案手法のフローチャート

### 3.2 認識誤りに関する対策<sup>11)</sup>

#### 3.2.1 置換誤り対策

置換誤り対策としては、文献<sup>11)</sup>に示されているように、m=3 として音節ラティスの上位 3 ベストを用いる。本稿では索引を構築する際に音節ラティスの上位 3 ベストを組みあわせ、トライグラムを作成する。(図 2 参照)つまり、1 つのインデックスに対して  $3 \times 3 \times 3 = 27$  個のトライグラム情報を持たせる。例えば、「フーリエ変換」の 1 ベストの認識結果が「フェキエヘンカン」になっていたとしても、3 ベスト中に正しい音節が含まれていれば検索す

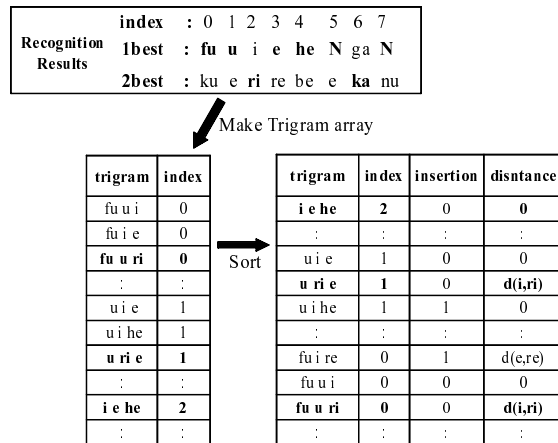


図 2 トライグラムアレイ作成手順

ることができる。3 ベスト中に含まれない場合でも、次節に示すの挿入誤り対策と脱落誤り対策の併用で対処できる。

### 3.2.2 挿入誤り対策

挿入誤りに関しては、索引構築時に 1 音節飛ばしたトライグラムを作成することで対処する(図 2 参照。3 連続音節に対して 1 箇所)。先頭の音節を飛ばしたトライグラムは作成しない。したがって、1 つのインデックスに対し、2 つのトライグラムが索引に追加される。

### 3.2.3 脱落誤り対策

脱落誤りに関しては上記 2 つの対策とは違い、検索時にクエリを数音節脱落させたものを含めて検索することで対処する。したがって、1 つのクエリに対して複数回検索をおこなう。本実験では 4 音節以上のクエリに対しては 1 つの脱落、7 音節以上のクエリに対しては 2 つの脱落を許す。但し、脱落は連続する 3 音節以内に 1 箇所に限る。

### 3.3 距離付き n-gram インデックスによる検索<sup>15)</sup>

前述した認識誤り対策をおこなうことで認識誤りに頑健な検索をおこなうことができる。しかし、同時に湧き出し誤りが増加するという問題がある。そこで、トライグラムインデックスの候補検出後に DP マッチングによる後処理による検索法を提案した<sup>11)</sup>。しかし、DP マッチングによる検出候補の削減では、検出候補すべてに対して DP マッチングをおこなわなければならないため、検索対象音声が増えたり、検出候補が増えたりと処理時間が長くなるという問題があった。そこで、認識誤り対策をおこなう際にどれだけの誤りを許容したかという情報(距離)を用いて、検索時に検出候補の削減をおこなう方法を提案した<sup>15)</sup>。距離を用いることで DP マッチングのように複雑な計算を行わずに、閾値との比較のみで検出候補の絞込みが高速におこなえる。

置換誤りの距離は 1 ベストからの音節間距離 (Bhattacharyya 距離<sup>11)</sup>) を使用した。1 ベストのみから生成されたトライグラムを基準とし、置換誤り対策によって生成されたトライグラムの各音節との距離の合計を置換誤りの距離とする(図 2 参照)。つまり、1 ベストの音節の距離は 0 となる。次に、挿入誤りの距離は挿入誤り対策をおこなったかの有無を 0, 1 の 2 値で表現する(図 2 参照)。最後に脱落誤りの距離に関しては、検索時に脱落させた音節数を脱落誤りの距離とする。本実験では最大 2 音節の脱落を許すため(3 連続音節につき 1 箇所)、0, 1, 2 の 3 値となる。

### 3.4 認識尤度に基づくインデックスの削減法

認識誤り対策をおこなうことで生じるもう一つの問題として、複数の誤りのパターンを考慮するため、インデックスのサイズが大きくなるのが挙げられる。前述したとおり、3 ベストまで考慮した置換誤り対策では一つのインデックスに対して 27 通りのトライグラムを生成し、挿入誤り対策では各トライグラムに対して、挿入の有りと無しとの 2 つのパターンを生成するため、インデックスのサイズは通常の約 54 倍ものサイズになってしまう。この問題に対処するため、音節認識の音響尤度に基づいたインデックスの削減方法を提案する。これは、音節認識結果から得られたフレーム数で正規化した音節の尤度を用いて、1 ベストとの相対的な比較をし、尤度の差が一定の閾値  $\alpha$  以内のトライグラムのみインデックスとして採用するというものである。

### 3.5 検索性能

#### 3.5.1 認識誤りに対する頑健性<sup>11)</sup>

ここでは、提案する誤り対策の手法による検索語の検出率を推定する。図 3 に 3~4 つの音節からなる音節列を認識した際に考えられる結果を示す。ここで、"Cor" は正しく認識された音節を示し、"InCor" は誤って認識された(置換, 挿入, 脱落)音節をあらわしている。"p" は音節認識の正解精度とする、たとえば、本研究で用いたデータでは  $p=0.86$ (表 1) であり、この時、3~4 音節のクエリであれば約 85% (recall) 検出できると推定される。しかし、適合率 (precision) の推定値の算出は困難である。

#### 3.5.2 メモリ使用量と検索時間

n-gram インデックスを構築する際に必要とされるメモリ使用量は式 (1) であらわされる。

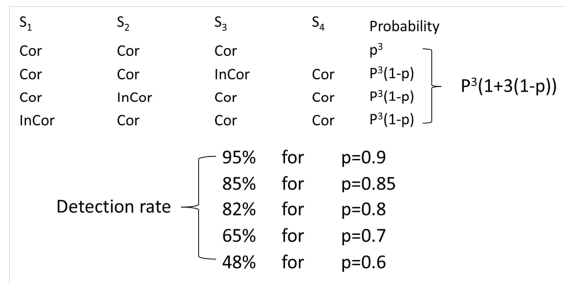


図 3 検出率の推定

ここで、 $n$  は  $n$ -gram を示し、 $m$  は用いる認識結果の候補数である。 $m^n$  は  $m$ -ベストまでを用いて  $n$ -gram を構築する際に可能な音節の組み合わせの数となる。 $N$  は認識した音声ドキュメントの総音節数となる。 $Tr$  は 1 つの  $n$ -gram に必要とされるサイズであり、式 (2) のようになる。

$$M = m^n \times n \times (N - 2) \times Tr \quad (1)$$

$$Tr = N\text{-gram}(\text{int}) + \text{index}(\text{int}) + \text{挿入距離}(\text{int}) + \text{置換距離}(\text{float}) = 16[\text{byte}] \quad (2)$$

## 4. 評価実験

### 4.1 実験データ

実験データには CSJ(日本語話し言葉コーパス) のコアデータ 44 時間分を用い、本研究室で開発された SPOJUS++<sup>14)</sup> による音節認識結果を対象とし、検索、評価をおこなった。大語彙連続音声認識 (LVCSR) の辞書 (20000 語) には、コア講演以外の CSJ2702 講演を学習データとし、カットオフを 4 とした。したがって、出現回数が 4 回以上あるものは未知語にはならない。今回検索語セットは伊藤らの報告<sup>12)</sup> にあるものを使用した。連続音節認識には音節の 4 グラムの言語モデルを用いた。連続音節認識による音節認識率と LVCSR の結果による単語認識率と音節列に変換後の音節認識率を表 1 に示す (単語正解率は 72%)。連続音節認識の第 1 候補の結果と LVCSR の結果から変換した音節認識結果は、ほぼ同程度である。連続音節認識結果の第 3 候補までを考慮すると音節認識結果の正解率は 89% とかなり高い。

### 4.2 既知語検索結果

既知語クエリ 39 種類 478 個を用いて性能評価をおこなった。LVCSR の認識結果を音節

表 1 音節認識率 (%)

output	Del	Ins	Subs	Corr	Acc
音節 (1best)	3.9	3.6	12.5	83.6	80.0
音節 (3best)	3.9	2.2	6.9	89.1	86.9
大語彙 音節	4.1	2.3	12.5	83.3	81.1
単語認識結果	5.4	4.6	22.7	71.9	67.3

表 2 既知語検索結果 (大語彙認識)

既知語	大語彙 → 音節列						テキスト検索
	対策無し	①置換	②挿入	③脱落	②+③	DP	
検出数	249	-	249	253	255	305	250
正解数	235	-	235	239	239	257	239
再現率	0.48	-	0.48	0.48	0.48	0.52	0.48
適合率	0.94	-	0.94	0.94	0.94	0.84	0.96
F 値	0.63	-	0.63	0.64	0.64	0.64	0.64

に変換したデータを用いた実験結果を表 2 に示す。①が置換誤り対策、②が挿入誤り対策、③が脱落誤り対策を示している。それぞれ 3-gram による検索である。大語彙認識は 1 ベストのみでしか実験をおこなっていないため、置換誤り対策の部分の結果は載せていない。DP は音節単位で連続 DP によるワードスポッティングをおこなった結果となっている (パタチャリヤ距離を利用)。テキスト検索は LVCSR の結果をそのまま用いて、単語検索した結果である。最もよい再現率を示したのは連続 DP であった。しかし、同時に湧き出し誤りも多く、適合率は提案手法やテキスト検索に比べ低い結果を示している。F 値でみると、最も高い値を示したのは、提案手法、連続 DP、テキスト検索とも同等であった。提案手法はテキスト検索と比較すると、多少適合率で劣っているが、明確な性能差は見られなかった。

表 3 に示すように、連続音節認識の 3 ベストの結果を用いて、全ての誤り対策をおこなった結果は、大語彙認識を音節列に変換した結果よりもやや低い性能を示している。このことから、既知語の検索には大語彙認識の結果を用いるほうが良いことがわかる。最終的な結果としては、既知語に関しては大語彙認識のテキストによる検索でも、音節列に変換してからの提案手法や連続 DP でもよいと考えられる。

### 4.3 未知語検索結果

未知語クエリ 43 種 197 個を用いて性能評価をおこなった。連続音節認識の結果の 3 ベストまでを用い、それぞれの対策をおこなった際の検索結果を表 4 に示す。ベースラインとなる連続 DP には音節認識結果を 3 ベストまで用いて、距離にパタチャリヤ距離を採用したものを載せている。このほかにも、1 ベストに対して、パタチャリヤ距離の結果 (F 値の最大値 = 0.47)、音節単位での編集距離の結果 (F 値の最大値 = 0.45)、音素単位での編集距離の結果

表 3 既知語の誤り対策別の検索結果 (連続音節認識)

未知語	音節認識							
	対策無し	①置換	②挿入	③脱落	①+②	①+③	②+③	①+②+③
検出数	139	167	142	224	190	266	267	275
正解数	136	156	138	186	161	212	197	214
再現率	0.28	0.32	0.28	0.38	0.33	0.43	0.40	0.43
適合率	0.98	0.93	0.97	0.83	0.85	0.80	0.74	0.78
F 値	0.43	0.47	0.43	0.52	0.47	0.56	0.52	0.56

表 4 未知語の誤り対策別の検索結果 (連続音節認識)

未知語	音節認識								
	対策無し	①置換	②挿入	③脱落	①+②	①+③	②+③	①+②+③	DP
検出数	34	61	34	67	66	107	78	119	101
正解数	34	58	34	58	59	80	66	88	80
再現率	0.17	0.29	0.17	0.29	0.30	0.41	0.34	0.45	0.41
適合率	1.00	0.95	1.00	0.87	0.89	0.75	0.85	0.74	0.79
F 値	0.29	0.45	0.29	0.44	0.45	0.53	0.48	0.56	0.54

(F 値の最大値=0.57), 音素間距離を考慮した音節単位での結果 (F 値の最大値=0.47), を図 4 に示している. キャプションの①~④はそれぞれ 2.3 の "DP マッチングによる検索" に挙げた距離の定義と対応している. これらの結果は距離を用いて絞り込みをおこない, F 値が最も高くなる値を採用している. ベースラインの連続 DP と比較しても, 提案手法はほぼ同等の検索性能を示している. この結果から, 全ての認識誤り対策をおこなうことで F 値で 0.56 と, 高精度に未知語を検出できることがわかる. (既知語に対する結果と同等)

また, 同様の条件で, 音節認識結果のラティス候補数を変化させた際の結果を表 5 に示す. 4 ベスト, 5 ベストと候補数を増やすことで, 再現率を増加させることができている. また, 距離を用いることでうまく湧き出し誤りが抑制し, F 値も増加している. しかし, 候補数を増やすことで, インデックスのサイズも増加してしまう問題がある. そこで, この表の 5 ベストのインデックスを認識結果の尤度に基づいて削減し検索をおこなった. その結果を図 5 に示す. 5 ベスト①と 5 ベスト②はインデックスの削減をおこなったものである. 図のキャプションに括弧に書かれている数字はインデックスのメモリサイズを表している. もともとの 5 ベストのインデックスサイズは 3GB 程度である. 本手法でインデックスを削減することで, 元々の大きさの 1/5 にしても性能はほぼ変わらないことがわかる. ほぼインデックスの同じ大きさの 3 ベストや 4 ベストのインデックスと比べても, 多少性能が良いことがわかる.

#### 4.4 検索時間

表 3 の DP マッチングの平均検索時間が 500[ms] なのに対し, 距離付きトライグラムを

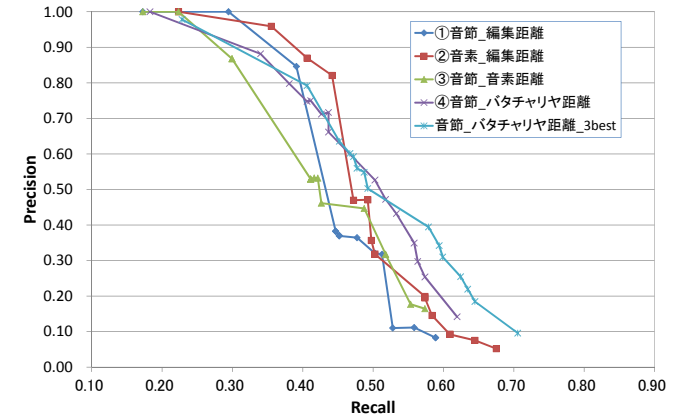


図 4 連続 DP マッチングによる未知語検索性能

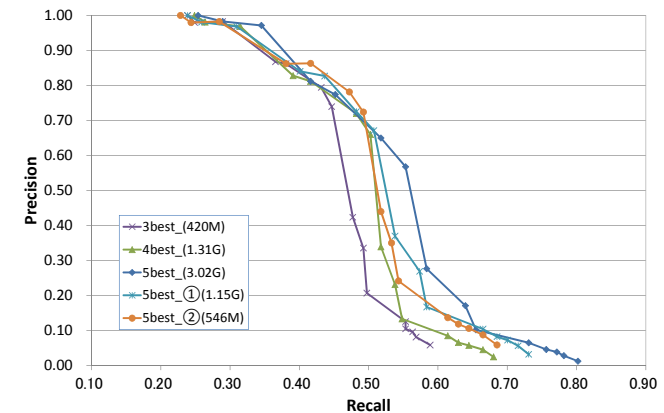


図 5 インデックス削減をおこなった未知語検索結果

表 5 音節ラティスの候補数別の未知語検索結果

未知語	音節認識ラティス				
	1best	2best	3best	4best	5best
検出数	78	89	119	132	157
正解数	66	72	88	95	102
再現率	0.34	0.37	0.45	0.48	0.52
適合率	0.85	0.81	0.74	0.72	0.65
F 値	0.48	0.50	0.56	0.58	0.58

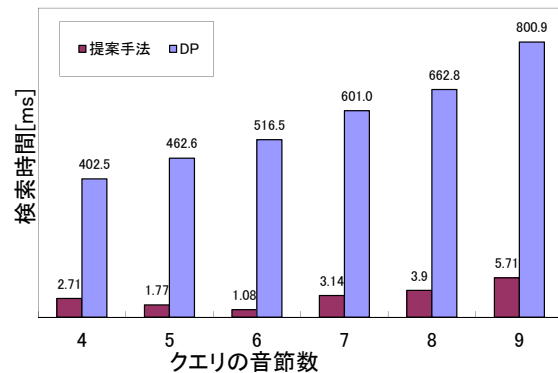


図 6 1 クエリあたりの検索時間の比較 (音声ドキュメント量:44 時間)

用いた場合の平均検索時間は 2.5[ms] となった。この結果から、検索対象が大規模の場合、連続 DP による検索よりも提案手法のほうがすぐれていることわかる。この検索時間の差は、検索対象音声の時間が大きくなると顕著になる(連続 DP は線形に増加、提案手法は対数スケールで増加)。

## 5. おわりに

本稿では、距離付きトライグラムインデックスによる高速検索法を提案し、音節正解率 84%(第 3 候補までで 89%) の音声ドキュメントから既知語に対して 0.56, 未知語に対して 0.58 の F 値を得た。湧き出し誤りについても距離付きトライグラム認識尤度を用いて検出候補を絞り込むことで高速に検索することができた。インデックスが大規模になる問題に関しても、音響尤度を用いることで解決することができることを示した。今後の課題として、検索精度と絞込みの精度の改善が挙げられる。検索精度に関しては大語彙連続音声認識の単語認識結果の信頼度の低い区間だけを未知語候補区間とする方法が考えられる<sup>18)</sup>。

## 参考文献

- 1) C.Allauzen, M.Mohri and M. Saracla, "General Indexation of Weighted Automata - Application to Spoken Utterance Retrieval", Proc. workshop on interdisciplinary approaches to speech indexing and retrieval, HLT/NAACL, pp.33-40(2004).
- 2) B.Chen, H.Wang, L-s.Lee, "Retrieval of broadcast news speech in Mandarin Chinese collected in Taiwan using syllable-level statistical characteristics", Proc. ICASSP, pp2985-2988(2000).
- 3) M.Saraclar, R.Sproat, "Lattice-based search for spoken utterance retrieval", Proc. HLT/NAACL, pp129-136(2004).
- 4) Martha Larson, Stefan Eickeler, "Using syllable-based indexing features and language models to improve German spoken Document Retrieval" proc. EuroSpeech, pp. 1217 - 1220(2003)
- 5) Hsin-min Wang, "Experiments in syllable-based retrieval of broadcast news speech in Mandarin Chinese " Speech Communication, Vol 32, pp. 49 - 60(2000)
- 6) Martin Wechsler, Eugen Munteanu, Peter Schauble, "New techniques for open-vocabulary spoken document retrieval" Proc.SIGIR ,pp. 20 - 27(2008)
- 7) Corinna Ng, Ross Wilkinson, Justin Zobel, "Experiments in spoken document retrieval using phoneme n-grams " Speech Communication, Vol 32, pp. 61 - 77(2000)
- 8) S Dharanipragada, S Roukos, "A multistage algorithm for spotting new words in speech" IEEE Transactions on Speech and Audio Processing, vol.10, No.8, pp. 542 - 550,(2002)
- 9) K.Ng, "Towards robust methods for speech document retrieval", Pro.ICSLP, pp.1088-1091(1998).
- 10) 手島茂樹, 桂田浩一, 新田恒雄, "Suffix Array を用いた音声文書の高速検索 " 第 3 回音声ドキュメント処理ワークショップ, pp. 29 - 32(2009)
- 11) 中川聖一, 高橋将史, 他, "未知語に頑健な音声ドキュメント検索手法の検討 " 第 3 回音声ドキュメント処理ワークショップ論文集, pp. 7 - 14(2009.2)
- 12) 伊藤慶明, 西崎博光, 他, "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, SLP-78 No.4(2009)
- 13) 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.86-DII, No.10, pp. 1369 - 1381(2003)
- 14) 藤井康寿, 山本一公, 中川聖一, "大語彙連続音声認識システムの改善:SPOJUS++" 第 4 回音声ドキュメント処理ワークショップ論文集, pp. 1 - 10(2010.2)
- 15) 中川聖一, 岩見圭祐, 藤井康寿, 山本一公, "連続音節認識結果の距離つきトライグラムアレイ化による未知語音声の超高速検索 " 第 4 回音声ドキュメント処理ワークショップ論文集, pp. 58 - 64 (2010.2)
- 16) H.M.Meng, W.K.Lo, Y.C.Li, and P.C.Ching, "Multi-scale audio indexing for chinese spoken document retrieval", Proc. ICSLP, vol.IV ,pp.101-104(2000).
- 17) J. Zhang, L. Wang, S. Nakagawa "LVCSR based on context-dependent syllable acoustic models", Asian Workshop on Speech Science and Technology, SP2007-200(2008).
- 18) H.Nishizaki and S.Nakagawa "Japanese Spoken Document Retrieval Considering OOV Keywords Using OOV Detection Processing and Word Spotting.", Proc. Human Language Technology , pp.144-151 (2002).
- 19) Yi-cheng Pan and Lin-shan Lee "Performance analysis for lattice-based speech indexing approaches using words and subword units" IEEE Transactions on Audio, Speech and Language Processing, vol.18, No.6, pp. 1562 - 1574,(2010)