Vol.2014-NL-216 No.13
Vol.2014-SLP-101 No.13
2014/5/23

類似音節バイグラムリストを用いた音声中の検出語検出

大内一揮[†1]　鳴海司郎[†1]　菅原翔太[†1]　李時旭[†2]　田中和世[†3]　伊藤慶明[†1]

音声中の検索語検出方式における未知語の検索語の高速な検索のために，我々は既に音節バイグラム事前検索結果を用いた高速検索方式を提案した．音節バイグラム事前方式ではインデックスの空間計算量が大きくこれの削減を図るため，本稿では類似音節バイグラムリストを用いた事前検索方式を提案する．類似音節バイグラムリストと各音節バイグラムに一致した音節バイグラムを有する区間から事前検索結果を再構成する．さらに音節バイグラム間照合による類似音節バイグラムリストを構築する方式を提案する．評価実験の結果，提案方式では事前検索方式と比べ，コア 177 講演を検索対象とした場合，検索精度の低下無しで検索時間を 0.95sec から 0.24sec(3.96 倍)の高速化を実現し，インデックスの空間計算量を 1.2GB から 29.7MB と約 1/41 に削減できた．また全 2702 講演検索を検索対象とした場合，検索精度の低下無しで 18.22sec から 2.75sec(6.63 倍)に，インデックスの空間計算量を 3.1GB から 120.8MB と約 1/26 に削減できた．音節バイグラム間照合で構築した類似音節バイグラムリストを用いることで，コア 177 講演，全 2702 講演ともに事前検索方式および事前検索結果から類似音節バイグラムリストを構成する方式と同等の検索性能が得られることを確認でき，音節バイグラム間照合により類似音節バイグラムリストが構築できることを示した．

Spoken Term Detection Using Similar Syllable Bigrams

KAZUKI OOUCHI[†1]　SHIRO NARUMI[†1]　SHOTA SUGAWARA[†1]
SHI-WOO LEE[†2]　KAZUYO TANAKA[†3]　YOSHIAKI ITOH[†1]

This paper proposes a new method for spoken term detection to improve an index size without deteriorating the detection accuracy and the detection speed. The proposed method utilizes the similarity between syllable bigrams. The evaluation experiments demonstrated the proposed method reduced the index size with the same detection accuracy and the same detection speed.

1. はじめに

近年，大容量記憶デバイスが普及し，撮りためた大量のビデオデータの中から特定のキーワードで検索する機能が求められている．この機能の実現に対し，音声中の検索語検出 (STD : Spoken Term Detection) が有効であると考えられ，STD に関する研究が盛んに行われるようになった．現在の STD では音声認識システムを用いて検索対象となる音声ドキュメントの音声認識を行いその書き起こしデータに対して検索を行う方式が一般的となっている[1][2][3][4]．単語単位の認識結果を用いる場合，認識時に用いる単語辞書内にある単語であれば高い精度で認識・検索が可能であるが，辞書内に無い単語，即ち未知語に対しては検索が困難である．そのためサブワード単位の認識結果を用いて，未知語の検索が行われる[5][6]．未知の検索精度は一般に既知語の検索精度に劣るため，未知語の検索語の場合検索精度の向上が重要である．

未知語検索のための我々のサブワード STD システムでは，予め検索対象となる音声ドキュメントに対してサブワード認識を行い，サブワード系列で保持しておく．検索語はサブワード系列に変換し，サブワードレベルで音声ドキュメントと連続 DP (Continuous Dynamic Programming) によって照合を行い，発話区間の選出を行う．この方式では，検索語は全ての音声ドキュメントに対して連続 DP を行わなければならず，検索時間は検索対象となる音声ドキュメント量に比例して増加する．そこで我々は音節バイグラム事前検索結果を用いた音声中の検索語検出における上位候補の高速検索[7]を提案した．検索語内の音節バイグラムがある発話中の音節列に出現していれば，その発話内で検索語が話されている可能性が高いと考える．全ての音節バイグラムによって音声ドキュメントを事前に検索(事前検索)しておき，その事前検索結果を保持しておく．検索語が与えられると，音節バイグラムを 1 音節ずつずらして得られる検索語内の音節バイグラムそれぞれの事前検索結果を参照し，検索語が発話されている可能性が高い発話のみを抽出する．抽出された発話区間に対してのみ連続 DP によるスコアリングを行うことで，検索精度の低下を抑えた上で検索の高速化を実現した．しかしこの方式では，検索精度を低下させないためには事前検索結果の大きなインデックスが必要となる．

そこで本論文では，音節バイグラム事前方式におけるインデックス化方式を応用し，類似音節バイグラムリストを用いたインデックス削減方式を提案する．事前検索方式では，各音節バイグラムについて音声ドキュメントを検索した結果上

†1 岩手県立大学
Iwate Prefectural University.
†2 独立行政法人 産業技術総合研究所
National Institute of Advanced Industrial Science and Technology
†3 筑波大学
Tsukuba University

ⓒ 2014 Information Processing Society of Japan

Vol.2014-NL-216 No.13
Vol.2014-SLP-101 No.13
2014/5/23

位 K 件の発話区間である発話 ID を保持している．ある音節バイグラム A の上位 K 件の発話区間を見ると，その音節バイグラム A に一致している音節バイグラム A がまず現れ，次に類似した音節バイグラム B が続けて複数現れる．音節バイグラム B の上位にも同様に音節バイグラム B，音節バイグラム A が現れ，重複して保持することになる．そこで，音節バイグラム A の上位 K 件中に出現する音節バイグラムを抽出し，それを音節バイグラム A の類似音節バイグラムリストとする．音節バイグラム A には，一致する音節バイグラム A のみを保持することとし，音節バイグラム A の事前検索結果は類似音節バイグラムリストと各音節バイグラムの少数の一致する候補から再構成する．各音節バイグラムが保持する件数を最小限にすることで，上位 K 件の事前検索結果を保持する場合と同等の検索精度と検索時間を維持しつつ，インデックスの空間計算量を大幅に削減できると考える．

一方，この方式では事前検索結果を作成せざるをえず，その作成に時間を要する．さらに検索対象ごとにリストを用意する必要がある．そこで事前検索結果を作成せずに，検索対象に依らない類似音節バイグラムリストを構築するための音節バイグラム間照合方式を提案する．

2. 類似音節バイグラムリストを用いた事前検索方式の提案

本章では，音節バイグラム事前検索方式の空間計算量を削減するため，類似音節バイグラムリストを用いた事前検索方式を提案し，その方式を詳述する．

2.1 音節バイグラム事前検索方式の事前検索結果の空間計算量

音節バイグラム事前検索方式では，事前検索結果としてあらゆる音節バイグラムに対して検索対象と照合を行い上位 K 件までの候補を予め保持しておく必要がある．そのため，事前検索結果の空間計算量が大きくなる問題がある．発話 ID のみを候補の情報として保持した場合の空間計算量は CSJ コア 177 講演で 1.2GByte，2702 講演では 3.1GByte となる．事前検索結果の空間計算量は音声ドキュメントの量の増加に伴い増加すると考えられるため，事前検索結果の空間計算量の削減は重要な課題と考える．

2.2 類似音節バイグラムリストの導入による空間計算量削減

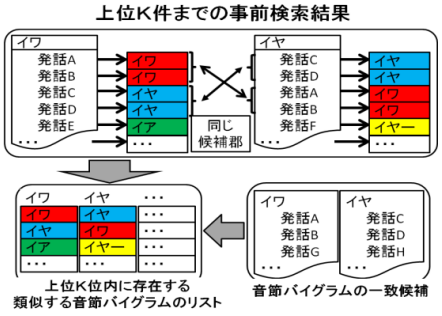


図 2.1 類似音節バイグラムリスト

提案方式のイメージ図を図 2. 1 示す．はじめにで述べたように事前検索方式において，「イワ」の事前検索結果の上位 K 件の発話区間の音節バイグラムを見ると，図のように最上位には「イワ」と一致する候補が複数現れ，次に「イワ」に類似する音節バイグラム「イヤ」「イア」などが現れ，それぞれ複数候補を有している．同様に「イヤ」の事前検索結果をみると「イヤ」と一致した音節バイグラムに続き「イヤ」に類似する音節バイグラム「イワ」「イヤー」などの音節バイグラムが現れ，それぞれに複数の候補が含まれる．この「イワ」「イヤ」を比較すると，両方の事前検索結果に同じ候補群が重複して保持されている．この重複保持の解消を狙う．そこで，各音節バイグラムに対し，類似音節バイグラムリストとそのバイグラムに一致する候補のみ保持することにより，検索精度と検索時間の低下を抑えつつ空間計算量を削減する．

事前検索方式において各音節バイグラムの上位K件を保持すれば精度低下がないとする．音節バイグラム A の上位 K 件に含まれる音節バイグラムのラベル(A_1～A_m)を抽出し，それを音節バイグラム A の類似音節バイグラムリストとする．音節バイグラム A には，完全に一致する音節バイグラム A を含む発話 ID 群のみを保持する．音節バイグラム A の K 件の事前検索結果は類似音節バイグラムリスト中の音節バイグラム A_1 から A_m までの一致候補を積み上げることで再構成・近似できると考える．各音節バイグラムが保持する件数を最小限にし，事前検索結果の上位 K 件を再構成することで空間計算量の大幅な削減を実現する．

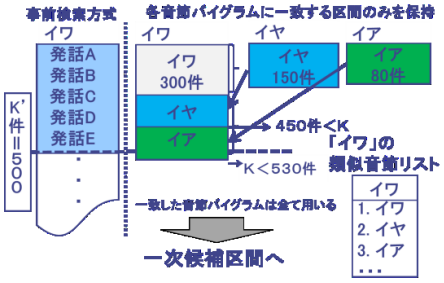


図 2.2 音節バイグラムリストを用いた一次候補区間の再構成方式

検索語が与えられた後，事前検索結果を再構成し一次候補の抽出方法を例を用いて説明する．検索語が「イワテ」であった場合，検索語に含まれる全て(2 つ)の音節バイグラム「イワ」と「ワテ」を抽出する．

図 2.2 に音節バイグラム「イワ」を対象にした場合の一時候補区間再構成のイメージを示す．図 2.2 の左側は事前検索方式における，「イワ」の事前検索結果上位 K 件を示している．「イワ」の事前検索結果を再構成する場合，まず音節バイグラム「イワ」と一致する発話区間を抽出する(図では 300 件)．次に「イワ」の類似音節バイグラムリストを参照し，最も類

ⓒ 2014 Information Processing Society of Japan

似する「イヤ」から，「イヤ」と一致する発話区間(150 件)を抽出し，上位 K'件未満の候補数であれば，次の類似音節バイグラム「イア」の発話区間(80 件)を抽出する．このようにして候補数が K'件以上になるまで候補を抽出する．これらの発話区間を事前検索結果と見なし一次候補区間として，後段の連続 DP により詳細な照合，ランク付けを行う．

提案方式は，音節バイグラム事前検索方式における手順を以下のように変更する．

(1) 全ての音節バイグラムを検索語として，音声ドキュメントのサブワード系列と照合し，事前検索結果として保持する．

(2) 各音節バイグラムに対し上位 K 件までの事前検索結果を参照し，その音節バイグラムに類似する類似音節バイグラムリスト(類似度が高い高整合度順)を作成する．

(3) 各音節バイグラムに対して以下の 2 つを検索時のデータとして保持しておく
- 類似音節列バイグラムリスト：各音節バイグラムに対し，事前検索結果上位 K 件までに含まれる類似する音節バイグラムをリスト化したもの．
- 各音節バイグラムと一致した発話区間群：音声ドキュメントの認識結果である音節列から各音節バイグラムと一致する音節バイグラム発話区間を有する発話区間のみ，その ID を保持する．

(4) 検索語が与えられると，検索語内に含まれる各音節バイグラムに対して上記方式で作成した類似音節バイグラムリストを参照し，リストの上位から順に音節バイグラムを選択し，その音節バイグラムと一致した音節バイグラムを有する発話区間を抽出する．抽出した件数の合計が K'件以上になるまで繰り返し，一次候補区間とする．この K'は実験により検索精度の低下が無いように決定する．

2.3 性能評価

本節では､インデックスの空間計算量について評価を行う．まず 2.3.1 節ではインデックスの空間計算量の評価を行う．2.3.2 節で CSJ177 講演を用いて，2.3.3 節で CSJ2702 講演を用いて検索精度と検索時間について評価を行う．

2.3.1 インデックスの空間計算量についての評価

音節バイグラム事前方式と 2.2 節で述べた提案方式とをインデックスの空間計算量で評価する．検索対象は CSJ コア 177 講演と全 2702 講演を用いた．音節バイグラム事前方式の K は検索精度の低下がない場合とし，コア 177 講演では K=5,000，全 2702 講演では K=12,500 とする．提案方式で用いられる各音節バイグラムと一致した音節バイグラムを有する発話区間の総保持件数はコア 177 講演では 763,572 件，全 2702 講演では 19,079,772 件となった．

インデックスの空間計算量の計算を行う際，発話 ID の 4 Byte のみを保持すれば良い．音節バイグラム事前方式と提案方式のインデックスの空間計算量の計算式を(3)と(4)に，各々の空間計算量を表 2.1 に示す．

$$4B(\text{発話 ID})\times K\text{ 件}\times\text{音節バイグラムの種類数}(261^2) \tag{3}$$

$$\begin{aligned}&\text{各音節バイグラムと一致した発話区間のインデックス}\\&\quad+\ \text{バイグラムリスト}\\&=4B(\text{発話 ID})\times(\text{音声ドキュメント内の音節数}-\text{発話数})\\&\quad+\ \text{バイグラムリスト}\end{aligned} \tag{4}$$

表 2.1 インデックスサイズの比較

	音節バイグラム事前方式		提案方式
	Top-K 件 (検索精度低下なし)	SIZE	SIZE
177 講演	K=5,000	1.2GB	29.7MB
2,702 講演	K=12,500	3.1GB	120.8MB

提案方式では，インデックスの空間計算量は音声ドキュメント中の音節バイグラムの数と類似音節バイグラムリストの和となる．音節バイグラム数は，各発話で，(その発話中の音節数－1)であり，合計すると(音声ドキュメント中の音節数－発話数)となる．これは音声ドキュメント中の音節数に近く，一定となる．一方，類似音節バイグラムリストは音節バイグラムの種類数×バイグラムリスト長となる．音節バイグラム事前方式と比べ，提案方式のインデックスサイズはコア 177 講演の場合 1.2GB から 29.7MB と約 1/41，全 2702 講演の場合 3.1GB から 120.8MB と約 1/26 とインデックスの大幅な削減に成功した．

2.3.2 CSJ コア 177 講演を対象とした場合の検索性能

検索対象に CSJ コア 177 講演を用いた場合の検索精度と検索時間を表 2.2 と図 2.3 に示す．実験では類似音節バイグラムリストは全事前検索結果を参照し(k=ALL(53892))､一旦全類似音節バイグラムリストを用意し，検索語が与えられて再構成時に K'までを抽出した．

類似音節バイグラムリストを用いる方式では，検索精度の低下がない保持件数は K'=2,000 だった．これは，音節バイグラム事前検索方式での上位 5,000 件の下位に含まれる不適切な音節バイグラムが排除され，不要な候補抽出が抑制されたためと考える．

ⓒ 2014 Information Processing Society of Japan

Vol.2014-NL-216 No.13
Vol.2014-SLP-101 No.13
2014/5/23

表 2.2 類似音節バイグラムリストを用いた検索性能
(CSJ コア 177 講演)

保持件数 K’	ALL	6000	2000	1000	500	300	0
MAP (%)	68.4	68.4	68.6	68.2	67.6	67.3	63.2
Time(sec.)	0.95	0.43	0.24	0.16	0.12	0.09	0.05

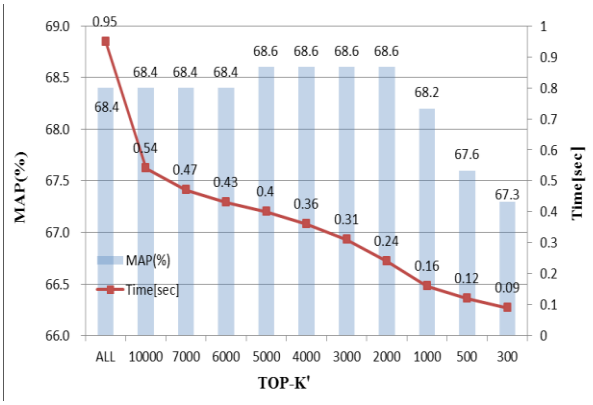


図 2.3 類似音節バイグラムリストを用いた検索性能
(CSJ コア 177 講演)

音節バイグラム事前方式では，検索精度の低下無しの K=5,000 の時に検索時間を 0.95sec から 0.39sec へと 58.95%削減できていたが，類似音節バイグラムリストを用いる方式では検索精度の低下無しの K'=2,000 の時，検索時間の削減は 0.95sec から 0.24sec へと 74.74%削減し，検索速度の向上を確認した．このとき，検索精度は 0.2 ポイント向上した．これは上位の発話区間に含まれる不正解発話区間数が減少し，正解の候補数が減少しなかったためと考えられる．また向上した 0.2%は誤差の範囲と考える．

2.4 CSJ 全 2702 講演を対象とした用いた場合の実験結果

検索対象に CSJ 全 2702 講演を用いた場合の検索精度と検索時間について表 2.3 と図 2.4 に示す．

全 2702 講演を検索対象とした場合，類似音節バイグラムリストを用いる方式では，検索精度の低下がない保持件数は K’=15,000 だった．

表 2.3 類似音節バイグラムリストを用いた検索性能
(CSJ コア全 2702 講演)

保持件数 K’	all	15000	12500	5000	3000	300	200	0
MAP (%)	63.9	64.0	63.8	63.4	63.6	63.2	62.6	62.4
Time(sec.)	18.22	2.75	2.60	1.72	1.44	1.00	0.96	0.90

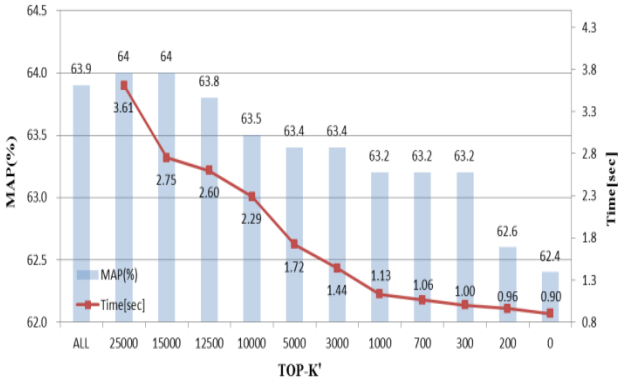


図 2.4 類似音節バイグラムリストを用いた検索性能
(CSJ 全 2702 講演)

検索対象に全 2,702 講演を用いた場合，検索時間は all の 18.22sec と比べ，検索精度低下無し(K’=15,000)の時 2.75sec，事前検索方式では検索精度低下無し(K=12,500)の時 18.22sec から 1.56sec であり，この場合提案方式は事前検索方式より若干検索時間を要した．K=K’=12,500 とした場合，提案方式は同一の音節バイグラムを全て抽出するため 12,500 より多くの一次候補区間を抽出したのが検索時間を要した要因と考える．

3. 音節バイグラム間照合による類似音節バイグラムリストの構築方法の提案

2 章で提案した方式では，類似音節バイグラムリストの作成には時間を要する問題があった．このリストの作成に事前検索結果を必要とし，例えば 177 講演の場合，事前検索結果の作成に約 1 日，2702 講演の場合は約 16 日を要する．また，検索対象ごとにリストを別々に用意する必要がある．例えば 177 講演の類似音節バイグラムリストを 2702 講演で用いる場合，リストに抜けが生じており，検索精度の低下が予想される．そのため検索対象毎に類似音節バイグラムリストが必要となり，それぞれの事前検索結果を作成する必要がある．類似音節バイグラムリストの作成は 2702 講演の場合約 3 日掛かり，全体で約 19 日を要する．

本節では事前検索の結果を用いない類似音節バイグラムリストの作成方式を提案する．

3.1 音節バイグラム間照合方式

提案する音節バイグラム間照合方式では，各音節バイグラムに対してどの音節バイグラムが類似しているかを全バイグラムと DP 照合することにより求め，各音節バイグラムの類似している順に音節バイグラムをリスト化する．これにより類似音節バイグラムリストを作成する．

ⓒ 2014 Information Processing Society of Japan

Vol.2014-NL-216 No.13
Vol.2014-SLP-101 No.13
2014/5/23

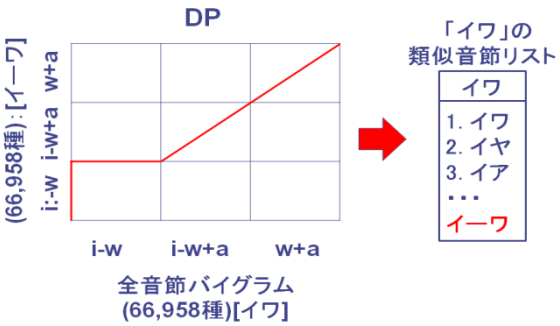


図 3.1 「イワ」の音節バイグラム間照合

提案方式は，図 3.1 のように 2 つの音節バイグラムをそれぞれ triphone 列に直した上で DP 照合を行う．局所距離には triphone 間の音響間距離を用いる．

音節バイグラム間照合方式の手順を以下に示す．

1. 各音節バイグラムに対して全音節バイグラムと，ともに triphone に展開した上で DP 照合を行う．その際，局所距離には triphone 間の音響間距離を用いる．
2. 各音節バイグラムについて DP 距離の小さい順に類似音節バイグラムをソートする．
3. 各音節バイグラムについて，類似する順に並んだ音節バイグラム ID からなる類似音節バイグラムリストを作成する．

3.2 音節類似バイグラムリストの空間計算量

音節バイグラム間照合によって構築した類似音節バイグラムリストの全インデックスの空間計算式を以下の式(5)に示す．

$$\text{音声ドキュメントの音節バイグラム数} + 4\,\text{B(音節バイグラム ID)} \times \text{全音節バイグラム}(261^2) \times \text{全音節バイグラム}(261^2) \quad (5)$$

音節バイグラム 1 種類につき約 0.26MB を必要とし，全体で約 16.7GB となる．この全体のインデックスはシステム構築時のみに必要であるためハードディスクや外部ディレクトリに保存しておけば良い．検索時にはメモリ上に各音節バイグラムの類似する音節バイグラムを必要最低限保持すれば良い．例えば K’=15,000 のとき，各音節バイグラムの候補が 15,000 件以上になるまでの類似する音節バイグラムを保持すると，インデックスの空間計算量は約 184.8MB となり約 1/93 に削減できた．このとき各音節バイグラムの類似音節バイグラム数は平均 440 件であった．音節バイグラム事前検索方式での空間計算量は約 120.8MB で，音節間バイグラム間照合による類似音節バイグラムリストの方が大きいが，現在のコンピュータで動作させる上では大きな差ではないと考える．

3.3 性能評価

提案した新しい類似音節バイグラムリスト作成方式を用いた場合の検索性能を評価実験により検証する．

3.3.1 CSJ コア 177 講演を対象とした場合の検索性能

検索対象に CSJ コア 177 講演を用いた場合の検索精度と検索時間について表 3.1 と図 3.1 に示す．

表 3.1 音節バイグラム間照合による類似音節リストを用いた検索性能(CSJ コア 177 講演)

保持件数 K’	ALL	10000	2000	1000	500	300	0
MAP (%)	68.4	68.4	68.4	68.7	67.9	67.5	63.2
Time(sec.)	1.02	0.68	0.28	0.19	0.13	0.10	0.05

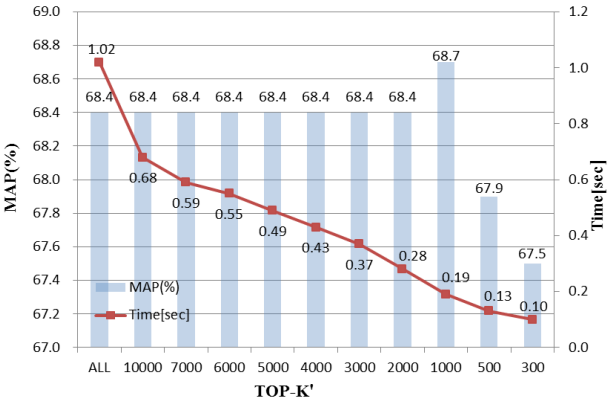


図 3.2 音節バイグラム間照合による類似音節リストを用いた検索性能(CSJ コア 177 講演)

表 3.1，図 3.2 より音節バイグラム間照合による音節類似リストを用いた方式では，検索精度の低下がない保持件数は K’=1,000 だった．この時検索時間が 0.37 秒となり，全照合と比べ 2.6 倍高速となった．

事前検索結果を用いて作成した類似音節バイグラムリストと比較する図 2.3 では K’=2,000 のとき検索時間は 0.24sec で検索精度は 68.6%，図 3.2 では K’=1,000 のとき検索時間は 0.19sec で検索精度は 68.7%となり，同等以上の検索性能が得られることを確認できた．

3.3.2 CSJ 全 2702 講演を対象とした場合の検索性能

検索対象に CSJ 全 2702 講演を用いた場合の検索性能と検索時間について表 3.2 と図 3.3 に示す．

表 3.2 音節バイグラム間照合による類似音節リストを用いた検索性能(CSJ 全 2702 講演)

保持件数 K’	ALL	15000	12500	5000	3000	1000	200	0
MAP (%)	63.9	64.0	63.9	63.7	63.9	63.5	62.9	62.4
Time(sec.)	13.75	2.54	2.33	1.45	1.27	1.04	0.91	0.88

ⓒ 2014 Information Processing Society of Japan

Vol.2014-NL-216 No.13
Vol.2014-SLP-101 No.13
2014/5/23

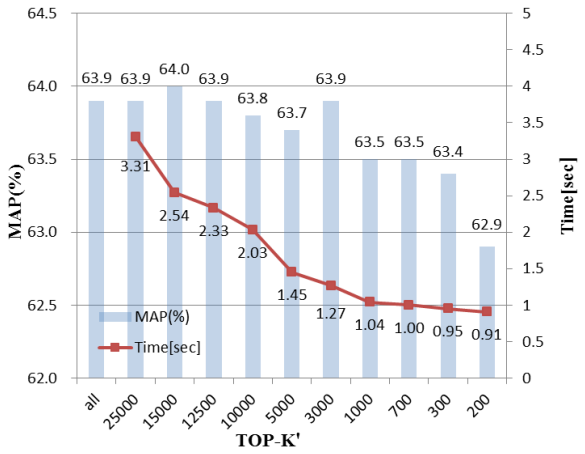


図 3.3　音節バイグラム間照合による類似音節リストを用いた検索性能(CSJ 全 2702 講演)

表 3.2，図 3.3 より，類似音節バイグラム間照合による類似音節リストを用いた方式では，検索精度の低下がない(検索精度 64.0%)保持件数は K’=15,000 だった．この時，検索時間は 2.54 秒となり 7.2 倍高速となった．K’=3,000 ではほぼ同じ検索精度 63.9%で検索時間は 1.27 秒となり 14 倍高速となった．CSJ177 講演と同様に事前検索結果を用いて作成した類似音節バイグラムリストと比較する．図 2.4 の事前検索結果を用いたリストでは K’=15,000 のとき検索精度が 64.0%で検索時間は 2.75sec であったのに対し，図 3.2 の音節バイグラム間照合による類似音節リストを用いた場合 K’=15,000 のとき検索精度が 64.0%で検索時間は 2.54sec と，177 講演と同様にほぼ同等の検索性能が得られることを確認した．

3.3.3 事前検索結果による類似バイグラムリストとの比較

2 章で述べた類似音節バイグラムリストで行った検索と比較する．CSJ コア 177 講演に対して 2 章の事前検索による類似バイグラムリストでは，K’=2,000 で検索時間 0.24sec，検索精度 68.6%．音節バイグラム間照合による方法では K’=1000 で検索時間 0.17sec，検索精度は 68.7 であり，図 2.3 と図 3.2 を比べてもほぼ同等の性能と考える．CSJ 全 2702 講演においても事前検索結果による類似音節バイグラムリストでは，K’=15,000 で検索精度 64.0%．検索時間 2.75sec であったのに対し，音節バイグラム間照合による類似音節バイグラムリストでは K’=15,000 のとき検索精度が 64.0%で検索時間は 2.54sec であり,図 2.4 と図 3.3 を比べると若干音節バイグラム間照合による方法が早いが，ほぼ同等の検索性能が得られることが確認できた．以上より音節バイグラム間照合を用いた類似音節バイグラムリストの構築方式の有効性が検証できた。また，検索時間の短縮の問題において本研究では音節を用いているが，文献[8]では音素 N-gram の照合を行うことで検索時間削減が可能であることが確認されている．そのため，適切なサブワードを用いる，あるいは各種サブワードの事前検索結果を適切に組み合わせて用いることで，更なる検索時間削減が図れると考えるがその検証は今後の課題としたい．

4.　おわりに

本論文では音節バイグラム事前方式におけるインデックスの空間計算量の削減を図るため，類似音節バイグラムリストを用いた事前検索方式を提案した．類似音節バイグラムリストと各音節バイグラムに一致した音節バイグラムを有する区間から事前検索結果を再構成することで，各音節バイグラムの保持件数を大幅に減らし，事前検索方式の検索精度と検索時を維持した上でインデックスの空間計算量の削減を図った．

実験の結果，提案方式によりコア 177 講演を検索対象とした場合，検索精度の低下無しで検索時間を 0.95sec から 0.24sec(3.96 倍)の高速化を実現し，インデックスの空間計算量を 1.2GB から 29.7MB と約 1/41 に削減でき全 2702 講演検索を検索対象とした場合，検索精度の低下無しで 18.22sec から 2.75sec(6.63 倍)に，インデックスの空間計算量を 3.1GB から 120.8MB と約 1/26 に削減できた．

さらに音節バイグラム間照合による類似音節リストを作成する方式を提案し，コア 177 講演，全 2702 講演ともに事前検索方式および事前検索結果から類似音節バイグラムリストを構成する方式で，ほぼ同じ検索性能が得られることを確認できた．これにより，本方式を用いることで事前検索結果を作成せずに，類似音節バイグラムリストが構築できることを示した．

謝辞　本研究の一部は文部科学省科学研究費補助金基盤(C)No. 24500124 を 受けて実施された．

参考文献

1)　Yoshiaki Itoh, et al. "Spoken Term Detection Results Using Plural Subword Models by Estimating Detection Performance for Each Query," INTERSPEECH, pp. 2117 -2120, 2011.
2)　名取 賢他, "任意語彙発話音声検索のための複数の認識モデルを利用した音節遷移ネットワークの構築" ,日本音響学会, 2009 年秋季研究発表会講演論文集, 1-R-27, pp.205-206,2009.9
3)　神田 直之他, 多段リスコアリングに基づく大規模音声中の任意検索語検出, 電子情報通信学会論文誌 D　Vol.J95-D　No.4 pp.969-981, 2012
4)　岡 隆一他, フレーム特徴の音素記号化に基づく語彙に依存しない音声検索, 電子情報通信学会論文誌 D-Ⅱ Vol.J86-D-Ⅱ, No.6, pp.764-775, 2003.
5)　Yuji Onodera, et al. “Spoken term detection by result integration of plural subwords using confidence measure”, WESPAC, 2009
6)　勝浦広大他,"放送大学の講義音声を対象とした高速キーワード検索の性能評価", 第 6 回音声ドキュメント処理ワークショップ, SDPWS2012-05, 2012.
7)　斉藤裕之他, N-音節事前検索結果を用いた音声中の検索語検出における上位候補の高速検索, 音講論 3-1-2, 2012.
8) 岩見圭祐, 山本一公, 中川聖一, "複数音声認識システムを併用した音節 n-gram 索引による検索性能の改善", 第 6 回音声ドキュメント処理ワークショップ, SDPWS2012-10, 2012.

ⓒ 2014 Information Processing Society of Japan