

文クラスタリングによる複数言語モデルを用いた誤認識文の推定

藤永 勝久^{†,††} 小窪 浩明[†] 山本 博史[†] 菊井玄一郎[†] 下平 博^{††}

[†] ATR 音声言語コミュニケーション研究所 〒619-0288 京都府相楽郡精華町 2-2-2

^{††} 北陸先端科学技術大学院大学 情報科学研究科 〒923-1292 石川県能美郡辰口町旭台 1-1

E-mail: †{katsuhisa.fujinaga,hiroaki.kokubo,hirofumi.yamamoto,genichiro.kikui}@atr.co.jp,

††sim@jaist.ac.jp

あらまし 本稿では、音声認識結果の発話単位の正解判定法について提案する。近年、誤認識単語の推定の方法として複数の認識システムの出力結果の共通部分を用いる方法が提案されている。この方法は、複数のシステムが同じ単語を出力している部分は正解である可能性が高いという ROVER 法の考え方に基づくものであり、その単語を認識結果のうち信頼性の高い部分として出力するものである。本稿では、これを認識結果全体の信頼性を得るための方法に応用することとする。ROVER 法では複数の認識システムを用意しておき、それらを並列に実行させなければならないが、(1) 認識システムを複数用意することが困難、(2) 計算コストがシステム数に応じて増加、という問題点がある。本稿では最初の問題に対しては、コーパスの自動クラスタリングにより任意の数の言語モデルを生成し、2 番目の問題に対しては、リスクアリングを用いる。本手法に対し、大語彙連続認識実験の認識結果の正解判定による評価を行った。その結果、正解判定を行わない場合と比較して、認識結果に含まれる正解文を 10% 捨てることで 18 ポイント、20% 捨てることで 24 ポイント高い適合率が得られた。また、クラスタ数などの条件の違いに対する評価も行った。キーワード ROVER、文クラスタリング、信頼度尺度、誤認識文検出

Mis-recognized Utterance Detection Using Multiple Language Models Generated by Clustered Sentences

Katsuhisa FUJINAGA^{†,††}, Hiroaki KOKUBO[†], Hirofumi YAMAMOTO[†], Genichiro KIKUI[†], and
Hiroshi SHIMODAIRA^{††}

[†] ATR Spoken Language Translation Research Laboratories 2-2-2 Seika, Soraku, Kyoto, 619-0288
Japan

^{††} School of Information Science, Japan Advanced Institute of Science and Technology 1-1 Asahidai,
Tatsunokuchi, Ishikawa, 923-1292 JAPAN

E-mail: †{katsuhisa.fujinaga,hiroaki.kokubo,hirofumi.yamamoto,genichiro.kikui}@atr.co.jp,

††sim@jaist.ac.jp

Abstract In this paper, we propose a new method that detects mis-recognized utterances, based on voting scheme like ROVER. ROVER has two serious problems, 1) it is difficult to construct multiple speech recognition systems (SRSs), 2) calculation cost increases according to the number of SRSs. In contrast to the conventional ROVER, the proposed method uses multiple language models (LMs), general LM and sub LMs generated by clustered sentence, instead of different SRSs. Speech recognition with sub LMs is proceeded by rescoring, instead of parallel decoding. Through experiments, the proposed method resulted in 18-point higher precision with 10% loss of recall from baseline, and 22-point higher precision with 20% loss of recall.

Key words ROVER, sentence clustering, confidence measure, mis-recognized utterance detection

1. はじめに

近年、音声認識に対して統計的モデルの研究が進んできており、条件によっては認識結果に対してかなり高い精度が期待できるようになってきている。しかし、現状の音声認識において誤認識の問題は避けられない。そのため、音声認識を用いたシステムにおいては、あらかじめ誤認識が生ずることを前提とした対策を用意しておくことが必要となる。特に、音声翻訳の入力部分として音声認識を用いる場合は、たった一語の認識誤りが翻訳に対して大きな影響を与え、結果として全く異なる文に翻訳されてしまう可能性があり、誤認識への対策はきわめて重要な問題である。

誤認識への対策のためには、まず誤認識個所の推定が必要となる。上記の音声翻訳等への応用を考えた場合、どの単語が誤っているかではなく、認識結果に誤りが含まれるかどうかが重要と考えられる。そこで本稿では誤認識単語の推定ではなく、得られた認識結果が正解文であるかどうかの推定を行うことを目的とする。

2. 従来法による認識結果の正解判定

2.1 従来法の文正解判定への応用

誤認識単語の推定の方法として複数の認識システムの出力結果の共通部分を用いる方法が提案されている[1]。この方法は、複数のシステムが同じ単語を出力している部分は正解である可能性が高いという ROVER 法[2]の考え方に基づくものであり、その単語を認識結果のうち信頼性の高い部分として出力するものである。本稿では、これを認識結果全体の信頼性を得るための方法に応用することとする。すなわち、複数のシステムが全く同じ認識結果を出力しているならば正しい認識結果である可能性が高いとするものである。

2.2 従来法における問題点

ROVER 法における最大の問題点は複数の認識システムを用意しておき、それらを並列に実行させなければならない点にある。用いる認識システムの数が多いほど効果が見込まれるが、性格の近いシステム同士の組み合わせや、他に比べて性能の低いシステムとの組み合わせでは効果が薄いことから、適切な認識システムの組み合わせをそろえる必要がある。しかし、システムの数が増えると、認識システムの構築にかかるコストが増大する。また、複数の認識システムを並列に実行させなければならないため、実行に必要な計算コストもまたシステム数に応じて増加する点も大きな問題となる。

3. 提案手法

我々は、2.1 節で挙げた ROVER 法の考えに基づき、言語モデルが互いに異なる複数の認識システムの認識結果の共通部分を用いて、発話単位で正解判定をする手法を提案する。提案法では、基準となる認識システム及び正解判定に用いる複数の認識システムにより音声認識を行い、正解判定用の認識システムの認識結果の多くが基準となる認識システムの認識結果と同じであれば、基準の認識結果を正解と判定する。複数のシステム

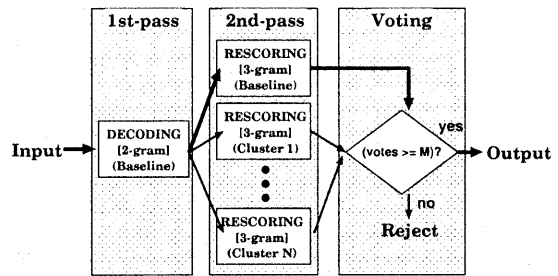


図1 認識システムの概要

を用いる場合には2.2 節で挙げたような問題が生じるが、認識システムの構築の問題に対しては、コーパスの自動クラスタリングにより任意の数の言語モデルを生成し、計算コストの問題に対しては、リスコアリングを用いることで解決する。

3.1 コーパスのクラスタリングによる言語モデルの生成

提案法では、与えられたコーパスの全てのデータを用いて学習したベースライン言語モデル及び、コーパスの自動クラスタリングによって得られたクラスタ毎に学習したクラスタ言語モデルを用いる。本稿では、クラスタリング手法としてエントロピーの総和を最小とするクラスタリングを用いた[3]。

ベースライン言語モデルはクラスタ言語モデルより学習データが多いため、これを用いた認識システムは、クラスタ言語モデルを用いたものより高い音声認識性能が期待される。このため、基準となる認識システムにはベースライン言語モデルを用い、正解判定に用いる認識システムにはクラスタ言語モデルを用いる。

クラスタ数の増加に従い個々のクラスタの学習データ量が減少し、クラスタ言語モデルの性能が低下する恐れがある。そのため、ベースライン言語モデルとクラスタ言語モデルを以下のように線形補完して認識に用いる。

$$\hat{P}(W_i | \cdot)_{C_n} = (1 - \lambda)P(W_i | \cdot)_{base} + \lambda P(W_i | \cdot)_{C_n} \quad (1)$$

ここで $P(W_i | \cdot)_{base}$ はベースライン言語モデルの N-gram 確率、 $P(W_i | \cdot)_{C_n}$ はクラスタ言語モデルの N-gram 確率、 λ は線形補完係数である。

提案法では用いるコーパスは1つであるため、コーパスを収集するコストが低い。また、自動クラスタリングにより、容易に任意の数の言語モデルを生成することが可能である。

3.2 複数の言語モデルを用いた発話単位の正解判定

提案法を用いた正解判定システムの概要を図1に示す。正解判定システムは2パスデコードと投票による正解判定器によって構成される。はじめに、1パスの処理において、ベースラインの2-gram 言語モデルを用いて入力音声の認識を行い、単語ラティスを得る。次に、2パスの処理として、ベースラインと各クラスタそれぞれにおいて、3-gram 言語モデルを用いて単語ラティスのリスコアリングを行い認識結果の1-bestを得る。最後に、ベースラインの認識結果に対し、各クラスタの認識結果を用いて信任投票を行う。本報告では、各クラスタの認識結果がベースラインと同じ場合に1票投じた。得票数が閾値(M)より多い場合は正解と判定して受理し、少ない場合は誤りとし

文数	171,894
平均文長	6.9 単語
総単語数	1,183,175
語彙数	20,561

表 1 コーパスの統計値

標準化周波数	16KHz
分析フレーム長	25ms
分析フレーム周期	10ms
特徴量	12 次 MFCC+12 次 Δ MFCC+ Δ power
音響モデル	性別依存, 1,400 状態, 5 混合 HMnet
言語モデル 1 パス	単語 2-gram
言語モデル 2 パス	単語 3-gram
デコーダ	ATRSPEC [4]
単語辞書サイズ	36,810

表 2 実験条件

て棄却する。

ROVER 法では各システムが独立して認識を行うため、システム数に比例して計算コストが増加する。これに対し、提案法では計算コストの最もかかる 1 パスの計算を 1 度しか行わないため、計算コストを大幅に削減できる。

4. 提案法の評価: 実験条件

大語彙連続音声認識結果に対する正解判定により提案法の評価を行った。

4.1 実験条件

我々は、旅行会話の自動翻訳のための音声認識部の開発を進めている。そのため、音声認識タスクとして旅行会話タスクを選択した。評価に用いるコーパスとして、旅行者向けのフレーズブックに現れるような旅行会話の基本表現を大量に集めたものを用いた。使用したコーパスの統計値を表 1 に示す。各発話には、それがどのような状況で用いられるかにより、「ビジネス」「留学」など 10 個のトピックの 1 つが振られている。評価実験においては、161,744 文を用いてモデルの学習を行い、学習に用いない 510 文を用いて評価を行った。その他の実験条件は表 2 に示す。

ベースライン言語モデルを用いた場合に得られた認識率は、発話正解精度 (utterance accuracy) が 68.63%, 単語正解精度 (word accuracy) が 88.89% であった。

4.2 評価尺度

一般に、正解判定では、入力 of 正解, 不正解をどれだけの精度で認識できるかを評価する。しかし、音声翻訳を目的とする正解判定の場合、正解発話をどれだけの精度で受理したかが重要になるため、評価尺度として次式で計算した適合率と再現率を用いて評価を行う。

$$\text{適合率 (Precision)} = \frac{\text{受理された正解発話数}}{\text{受理された発話数}}$$

$$\text{再現率 (Recall)} = \frac{\text{受理された正解発話数}}{\text{ベースラインの正解発話数}}$$

正解判定を行わず全ての認識結果を受理した場合、適合率はベースラインの発話正解精度と等しく、再現率は 100% である。

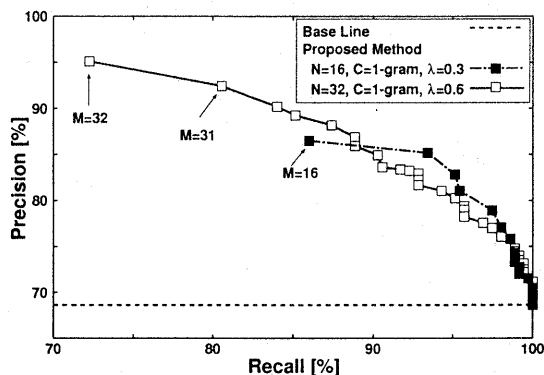


図 2 提案法の正解判定性能

音声翻訳を目的とした場合、適合率が高いことが最も要求されるが、再現率が低いと受理されない発話が増加する。そのため、本稿では、一定以上の再現率においてどれだけの適合率を得るかを重視して評価を行う。

5. 提案法の評価: 実験結果

5.1 提案法の正解判定性能

本節では提案法の正解判定性能の評価を行う。クラスタ数 (N)=32, エントロピー法でクラスタリングを行う際に用いるエントロピーの計算に用いる N-gram (C)=1-gram, 言語モデルの線形補完係数 (λ)=0.6 の場合と, N=16, C=1-gram, λ =0.3 の場合の実験結果を図 2 に示す。前者は再現率が 80%, 後者は再現率が 90% において最も高い適合率が得られた実験条件である。図中の各点は信任投票の閾値 (M) 毎の実験結果を表す。

提案法は閾値の増加に従い、再現率が低下し適合率が増加した。ベースラインに対し、90% 程度の再現率で 18 ポイント程度、80% 程度の再現率で適合率が 24 ポイント程度向上した。

提案法は、 λ などの条件の違いにより、適合率の最大値や同じ適合率に対する再現率などが大きく異なる傾向が確認できた。このため、必要とする適合率、再現率に応じて、条件を決定する必要があることがわかる。

5.2 従来法との比較

本節では提案法を従来法と比較する。従来法として、(a) 正規化尤度を用いた方法、(b) 正規化尤度, 事後確率, 認識結果の第一候補と第二候補の正規化尤度の比を SVM で統合した方法、(c) ROVER 法, の 3 つを用いた。正規化尤度を用いた方法では、各発話のフレーム単位の正規化尤度が閾値以上の場合に正解と判定して受理した。SVM で統合した方法では、学習データや評価データに含まれない 1010 発話を用いて学習した SVM を用いて正解判定を行った。ROVER 法では 3 つの認識システムを用いて正解判定を行った。本来、ROVER 法は音声認識精度を向上する手法であり、正解判定法ではないが、本稿では発話単位で多数決を行い、選ばれた認識候補の得票数が閾値以上の時に正解と判定して受理した。ROVER 法に用いた認識システムの概要と個々の認識結果を表 3 に示す。SYSTEM 1 は提案法におけるベースライン言語モデルを用いた認識システ

	音響モデル	言語モデル		認識率 [%]	
		1 パス	2 パス	発話正解精度	単語正解精度
SYSTEM 1	性別依存, 1400 状態, 5 混合 HMnet	単語 2-gram	単語 3-gram	68.63	88.89
SYSTEM 2	性別依存, 1400 状態, 5 混合 HMnet	多重クラス複合 2-gram [5]	—	65.69	88.55
SYSTEM 3	性別非依存, 2000 状態, 16 混合 HMM	単語 2-gram	単語 3-gram	67.25	88.89

表 3 ROVER 法で用いた認識システム

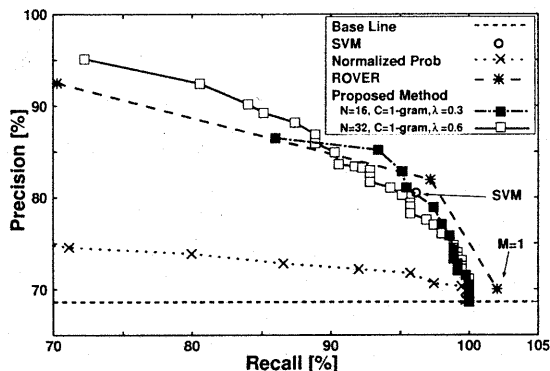


図 3 提案法と従来法の比較

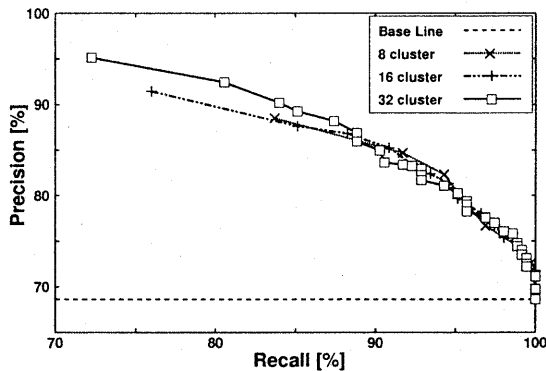


図 4 クラスタ数による比較: $C=1\text{-gram}$, $\lambda=0.6$

ムと等しい。また、SYSTEM 3 の音響モデルは情報処理振興事業協会 (IPA) の「日本語ディクテーション基本ソフトウェア」[6] 収録されているものを用いた。他の条件に関しては 4.1 節と共通である。

提案法と従来法の実験結果を図 3 に示す。提案法における実験条件は 5.1 節と同じである。ROVER 法において閾値が 1 の時に再現率が 100% を越えた。これは、多数決の結果、ベースラインの認識結果より多くの正解発話が選択されたためである。

はじめに提案法及び ROVER 法を他の手法と比較する。提案法及び ROVER 法は、正規化尤度を用いた方法より高い性能を得た。また、複数の信頼度尺度を SVM で統合したものと比較すると、同じ再現率において同等以上の適合率を得た。加えて、提案法及び ROVER 法では閾値を変更することにより再現率を落として適合率を向上させるなどの調整が可能だが、SVM を用いた方法では困難である。これらより、複数の認識システムの出力結果の共通部分を用いる方法は他の手法と比べて発話単位の正解判定に有効であると考えられる。

次に提案法を ROVER 法と比較する。実験の結果、95% 以下の再現率においては提案法が高い適合率を得た。また、提案法は計算コストが低いため、ROVER 法より多くの認識システムを用いることが可能である。そのため、閾値による再現率、適合率の調整が容易である。これらより、正解文を 5% 程捨てることを許容するならば、ROVER 法と比較して提案法が有効であると考えられる。

5.3 クラスタ数の違いによる比較

本節ではクラスタ数の違いによる比較を行う。 $C=1\text{-gram}$, $\lambda=0.6$ の場合の結果を図 4 に示す。

提案法では、クラスタ数の増加に応じて適合率の最大値が向上した。しかし、90% 以上の再現率における適合率は、殆んど変わらない傾向がみられた。また、クラスタ数の増加に応じて

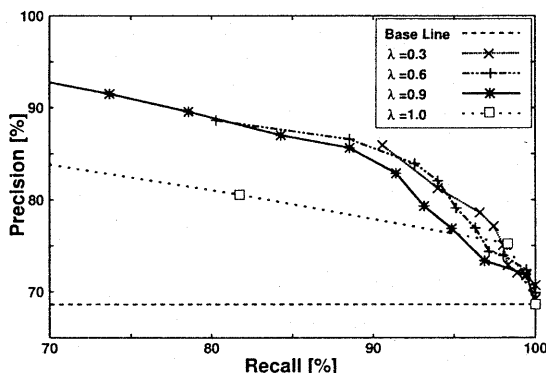


図 5 線形補完係数による比較: $N=10$, $C=1\text{-gram}$

閾値による適合率、再現率の変化が小さくなり、閾値による調整が容易になる傾向がみられた。

5.4 線形補完係数の違いによる比較

本節では提案法の線形補完係数の違いによる比較を行う。 $N=10$, $C=1\text{-gram}$ の場合の結果を図 5 に示す。

提案法では線形補完係数の増加に応じて適合率の最大値が向上し、同じ適合率を得る再現率は低下する傾向がみられた。その結果、再現率ごとに最大の適合率を得る線形補完係数が異なる傾向がみられた。また、線形補完を行わない場合 ($\lambda=1.0$) は大きく性能が低下した。これらより、各クラスタの言語モデルをベースラインのモデルと線形補完することは有効であり、線形補完係数は必要とする再現率に応じて決定することで高い適合率が得られることがわかる。

5.5 クラスタリング手法の違いによる比較

本節では提案法のクラスタリング手法の違いによる比較を行う。比較したクラスタリング手法は、エントロピー法 (1,2,3-gram) とランダム、コーパスに記述されたトピックによるクラ

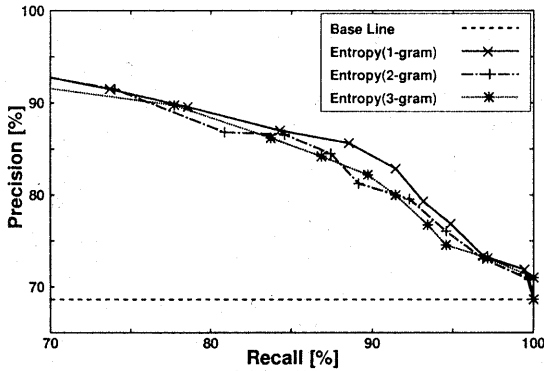


図 6 クラスタリング手法の比較 (1) : $N = 10$, $\lambda = 0.9$

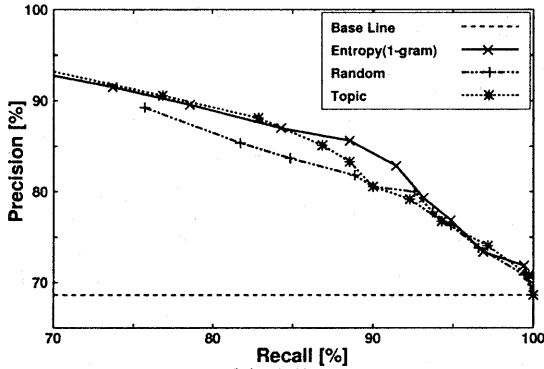


図 7 クラスタリング手法の比較 (2) : $N = 10$, $\lambda = 0.9$

スタリングである。 $N=10$, $\lambda=0.9$ の場合での実験結果を図 6, 7 に示す。

エントロピー法の計算に使う N -gram を比較すると、1-gram が最も性能が高く、3-gram が最も性能が低い傾向がみられた。エントロピー法 (1-gram) と他のクラスタリングを比較すると、エントロピー法が最も性能が高く、ランダムが最も低い傾向がみられた。しかし、ランダムの場合でも 90% 程度の再現率で適合率が 12 ポイント程度向上する結果が得られた。これより、クラスタリング手法に関わらず、複数の言語モデルを用いることにより高い精度で正解判定を行うことができることがわかる。

5.6 正解判定を行う認識結果の選択基準の比較

提案法では、ベースラインの認識結果に対し各クラスタの認識結果による信任投票を行い正解判定を行う。これに対し、ROVER 法と同様にベースラインと各クラスタの認識結果の多数決を行い、選択された認識結果の得票数が閾値以上の時を正解とする方法も考えられる。すなわち、提案法では正解判定を行う対象はベースラインの認識結果であるのに対し、多数決により正解判定を行う対象を決定する方法である。この方法は、ベースラインで誤認識をした発話を各クラスタで正確に認識することにより、提案法より多くの正解発話を得ることができる可能性がある。ベースラインと各クラスタの認識結果が全て同じもののみを受理する場合は、提案法と多数決法の受理する認識結果は同じとなる。

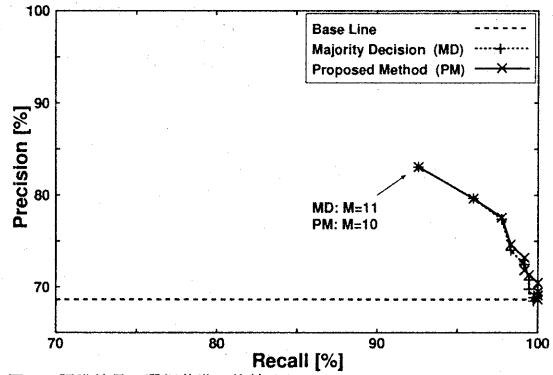


図 8 認識結果の選択基準の比較 (1) : $N = 10$, $C = 1$ -gram, $\lambda = 0.2$

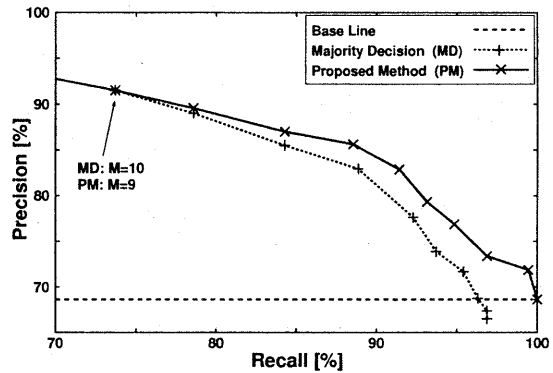


図 9 認識結果の選択基準の比較 (1) : $N = 10$, $C = 1$ -gram, $\lambda = 0.9$

本節では、提案法と多数決法を比較した。 $N=10$, $C=1$ -gram の場合の $\lambda=0.2$, 0.9 の結果を図 8, 9 に示す。 λ が小さい場合においては提案法と多数決法に有意な差は見られないが、 λ の増加に従い、多数決法の性能が低下する傾向がみられた。これは、各クラスタの線形補完前の言語モデルの性能がベースラインのモデルより低いいためクラスタ言語モデルの重みが増加するにつれて誤認識が増え、多数決により誤りを選択してしまったためだと考えられる。この結果より、ベースラインの認識結果に対する信任投票が有効であることがわかる。

5.7 提案法により得られる単語正解精度

以上の実験では受理された発話の発話正解精度を表す適合率で評価を行った。これに対し、本節では正解判定を行うことにより受理された発話の単語正解精度を調査した。5.1 節と同じ条件における提案法の実験結果を図 10 に示す。再現率は他の実験と同じく発話単位で求めたものである。

提案法では、90% 前後の再現率で 7 ポイント程度、80% 前後の再現率で 9 ポイント程度単語正解精度が向上した。これより、提案法で正解判定をすることにより、受理した発話の単語正解精度が向上することがわかる。

6. 実装方法の改良による高速化

本章では、提案法の更なる高速化について説明する。提案法は ROVER 法と比べ計算コストが低いですが、実装方法を改良することにより更に高速な処理が可能となる。改良が可能な点は以下

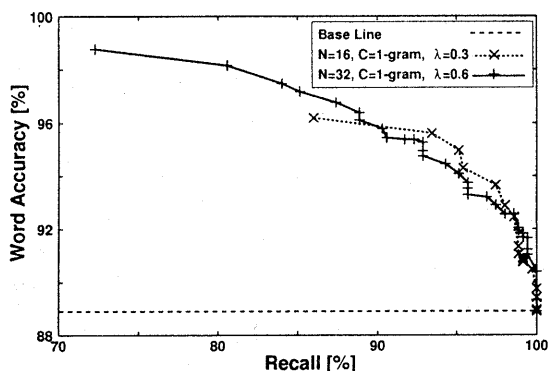


図 10 提案法で得られる単語正解精度

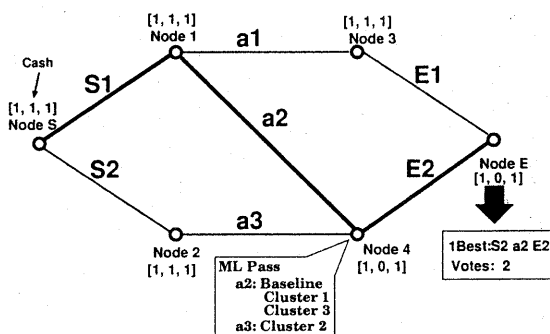


図 11 1-best の同時探索

下の 2 つである。

(1) 言語モデル読み込みの高速化

ベースライン言語モデルの学習データをクラスタリングして各クラスタ言語モデルを生成するため、ベースライン言語モデルはクラスタ言語モデルのエントリを全て含む。そのため、ベースライン言語モデルと各クラスタ言語モデルの言語スコアを並べて記述することにより、言語モデルの読み込みは一度に行うことができる。

(2) 1-best の同時探索

単語ラティスの各ノードに各クラスタがベースラインと同じパスを選択したか否かを記録するキャッシュを用意することにより、ベースラインの 1-best を求めると同時に同じ 1-best を出力するクラスタの数も求めることができる。1-best 探索の例を図 11 に示す。1-best 探索は以下の手順で行われる。

- 発声始端 (Node S) のキャッシュを全て 1 とする。
- 発声始端以外では、ベースライン言語モデルを用いて、発声始端からあるノードに至るパスのうち、最大の尤度を持つパスを求める。また、ベースラインの最尤パスが各クラスタ言語モデルでも最尤であるかを求め、最尤なら 1、異なるなら 0 でベースラインの最尤パスの 1 つ前のノードのキャッシュの値と AND をとり、そのノードのキャッシュに入れる (例えば、Node 4 におけるベースラインの最尤パスが a2 である場合には、Node 1 のキャッシュと AND をとり Node 4 のキャッシュの値とする)。
- 全てのノードの計算が終了するまで (b) に戻る。
- 終端ノード (Node E) からバックトレースを行い、ベースラインの 1-best を求める。
- 終端ノードのキャッシュが 1 になるものは、ベースラインと同じ 1-best をもつ。

7. ま と め

本稿では、音声認識結果の発話単位の正解判定法として、言語モデルが異なる複数のシステムの共通部分を用いた発話単位の正解判定法を提案した。本手法は ROVER 法に対し次のような利点がある。

- 与えられたコーパスに対し自動クラスタリングを行い、任意の数の言語モデルを自動生成することにより、容易に複数

の音声認識システムを構築できる。

- 音声認識で最も計算コストの高い 1 パスの計算はベースライン言語モデルを用いた 1 度だけであるため、計算コストが非常に低い。

提案法を評価するために、大語彙連続音声認識結果に対する正解判定を行った。その結果、提案法はベースラインに対し 90% 程度の再現率で 18 ポイント程度、80% 程度の再現率で 24 ポイント程度適合率が向上し、他の手法と比較して同等以上の性能を得た。また、信任投票の閾値により、容易に再現率、適合率の調整が行えることが確認された。次にクラスタ数や線形補完係数など条件の違いによる評価を行い、次のような結果を得た。

- クラスタ数の増加に応じて適合率の最大値が向上した。
 - 線形補完係数の増加に応じて適合率の最大値が向上し、同じ適合率を得る再現率の値は低下した。
 - エントロピー計算に 1-gram を用いたエントロピー法によりクラスタリングを行った場合において、最も高い性能を得た。
 - ベースラインの認識結果への信任投票が高い性能を得た。
- 最後に、実装方法の改良による高速化について論じた。

文 献

- Y. Kodama, T. Utsuro, H. Nishizaki, S. Nakagawa. "Experimental Evaluation on Confidence of Agreement among Multiple Japanese LVCSR Models". Proc. EUROSPEECH 2001, pp. 2549-2552, 2001.
- J. G. Fiscus. "A Post-processing System to Yield Reduced Error Word Rates: Recognizer Output Voting Error Reduction (ROVER)". Proc. IEEE Workshop on Automatic Speech Recognition and Understanding, pp. 347-354, 1997.
- 清水 徹, 大野 見生, 樋口 宜男. "文セットのクラスタリングに基づく統計的言語モデル". 音響学会講演論文集, 1-6-14, pp. 31-32, 1998-3.
- T. Shimizu, H. Yamamoto, H. Masataki, S. Matsunaga, and Y. Sagisaka. "Spontaneous Dialogue Speech Recognition Using Cross-word context Constrained Word Graphs". Proc. ICASSP 1996, pp. 145-148, 1996.
- 山本 博史, 匂坂 芳典. "接続の方向性を考慮した多重クラス複合 N-gram 言語モデル". 電子情報通信学会論文誌, Vol. J83-D-II, No. 11, pp. 2146-2151.
- 河原 達也, 李 晃伸, 小林 哲則, 武田 一哉, 峯松 信明, 嵯峨山 茂樹, 伊藤 克巨, 伊藤 彰則, 山本 幹雄, 山田 篤, 宇津呂 武仁, 鹿野 清宏. "日本語ディクテーション基本ソフトウェア (99 年度版)". 日本音響学会誌, Vol. 57, No. 3, pp. 210-214, 2001.