

英語ニュース音声を対象とした音声自動要約

堀 智織 古井 貞熙

東京工業大学 情報理工学研究科 計算工学専攻
〒 152-8552 目黒区大岡山 2-12-1
{chiori, furui}@furui.cs.titech.ac.jp

あらまし 本稿では、英語ニュース音声を対象とした音声自動要約法と要約結果を報告する。これまで、日本語を対象として、音声認識結果から、特定の割合(要約率)で部分単語列を抽出する自動要約手法を提案してきた。自動要約文に抽出される部分単語列は、単語重要度、言語尤度、信頼尺度、係り受け SCFG に基づく要約スコアが最大となるように動的計画法により決定する。本稿では、英語の自動要約に適用するため、係り受け SCFG(Stochastic Context Free Grammar) に基づく単語間遷移確率を推定するモデルを、英語の係り受け構造を推定できるモデルに拡張した。また、複数の発話から構成された音声を要約する手法として、要約する文数の増加に伴い増加する計算量を削減するため、2 段 DP(Dynamic Programming) による複数発話自動要約手法を提案した。

キーワード 音声自動要約, 英語ニュース音声, 単語重要度, 言語尤度, 信頼尺度, 係り受け SCFG, 2 段 DP

Automatic Speech Summarization for English Broadcast News Speech

Chiori HORI and Sadaoki FURUI

Graduate School of Information and Engineering, Tokyo Institute of Technology,
2-12-1 Ookayama, Meguro-ku, Tokyo, 152-8552 Japan
{chiori, furui}@furui.cs.titech.ac.jp

Abstract This paper reports an automatic speech summarization method and experimental results using English broadcast news speech. In our proposed method, a set of words maximizing a summarization score consisting of word significance measure, linguistic likelihood, confidence measure and SDCFG(stochastic dependency context free grammar) is extracted from automatically transcribed speech. This extraction is performed using a Dynamic Programming (DP) technique according to a target compression ratio(summarization ratio). In order to apply our method to English, the model of estimating word concatenation probabilities based on a dependency structure in the original speech given by a Stochastic Dependency Context Free Grammar (SDCFG) is modified. In order to reduce the amount of computation, a summarization method for multiple utterances using two-level DP technique is also proposed.

key words speech summarization, English broadcast news speech, word significance measure, linguistic likelihood, confidence measure, stochastic dependency context free grammar, two-level dynamic programming

1 はじめに

これまで、我々は単語抽出による音声自動要約手法を提案してきた [1][2]。日本語を対象として、音声認識結果から、文字数または単語数を基準とした特定の割合(要約率)で、要約文としての尤もらしさを示す要約スコアを最大とする部分単語列の抽出を行う自動要約手法を提案してきた。提案手法における要約スコアは、コーパスに基づく確率モデルを適用していることから、日本語に特化した手法ではなく、他言語への応用が可能であると考えられる。本稿では、英語のニュース音声を対象として、提案手法による自動要約を試みた。

英語の自動要約の研究として、テキストを対象とした文単位の自動要約手法を Knight らが提案している [3]。この手法は、確率付きの書き換え規則で表現された SCFG(Stochastic Context Free Grammar) に基づき、単語を削除する事により文圧縮(sentence compression)を行う自動要約手法である。つまり、原文と要約文が対になったコーパスを用いて、原文の構文木の書き換え規則が要約文で別の書き換え規則に変換、または削除される確率を学習し、自動要約に適用する。しかし、音声認識結果の自動要約を目的とした場合、必ずしも文法に当てはまらない音声認識結果を構文解析することは困難である。さらに、誤った構文解析結果に基づき音声認識結果を要約することは適切ではない。また、要約文と原文が対になっている大規模なコーパスを構築することも容易ではない。

一方、我々が提案している音声自動要約手法は、話題語を中心として言語的に尤もらしく単語を抽出するため、原文と要約文が対になったコーパスを必要とせず、既存のテキストコーパスを適用することが可能である。さらに、構文解析済みのコーパスから単語単位の係り受け SCFG(SDCFG:Stochastic Dependency Context Free Grammar) を学習することにより、確率的な構文情報を組み込むことが可能である [4]。但し、英語の係り受け構造は日本語と異なることから、原文の係り受け関係に基づく単語間遷移確率を推定するモデルを拡張した。一方、複数の発話から構成された音声を要約する手法として、発話単位の自動要約手法を拡張し、文間では単語間の接続を制約するスコアを適用せず、複数の発話を一文として要約する手法をこれまで提案してきた [5][6]。本稿では、要約する文数の増加に伴い増加する計算量を削減するため、各発話を可能性のある全ての要約率で要約後、全体として目的的要約率となるよう各文の要約文を組み合わせる 2 段 DP(Dynamic Programming) による複数発話自動要約手法を提案する。

2 音声自動要約手法

音声認識結果から、文字数または単語数を基準とした特定の割合(要約率)で、要約文の尤もらしさを示す要約スコアを最大とする部分単語列を動的計画法により決定する。但し、本論文では形態素を単語と定義する。

2.1 要約スコアの定義

要約スコアは、単語重要度スコア I と言語スコア L 、信頼度スコア C 、および、単語遷移スコア T に基づき、次式のように定義する。 N 個の単語からなる認識単語列 $W = w_1, w_2, \dots, w_N$ から要約文として M ($M < N$) 個の単語を抽出し接合した単語列 $V = v_1, v_2, \dots, v_M$ の要約スコアは次式によって示される。

$$S(V) = \sum_{m=1}^M \{L(v_m | \dots v_{m-1}) + \lambda_I I(v_m) + \lambda_C C(v_m) + \lambda_T T(v_{m-1}, v_m)\} \quad (1)$$

但し、 λ_I 、 λ_C 、 λ_T は各スコアのバランスをとるための重み係数である。

認識された単語列より抽出された部分単語列を $V = v_1, v_2, \dots, v_M$ ($M < N$) とするとき、要約処理は (1) 式で表される要約スコアを最大にする \hat{V} を求める問題となる。

2.1.1 言語スコア

$L(v_m | \dots v_{m-1})$ は、要約文内の単語連鎖の適正度を示すスコアである。本研究では、単語 Ngram 確率の対数値を適用する。本稿では単語 3gram と単語 2gram を適用し、比較検討を行う。

2.1.2 単語重要度スコア

$I(v_m)$ は、原文における相対的な単語の重要度を示すスコアである。単語を発話の重要な情報を担う話題語とそれ以外の単語に分類し、話題語の重要度スコアとして、単語の出現頻度に基づく情報量を適用する。一方、話題語以外の単語の重要度スコアは、一定値とする。さらに、冗長性を削減するため、2 回目以降の話題語のスコアも一定値とした。

話題語にのみ単語の相対的な重要度を示すスコアを用いることにより、それ以外の単語は言語スコアにより制御され、『文』を構成する役割を果たす。

日本語では名詞(動作を示す名詞を含む)のみを話題語としたが、英語では、動詞も話題語とする。

2.1.3 信頼度スコア

$C(v_m)$ は、認識結果に含まれる認識誤りを要約文に抽出しないよう、音響的、言語的に信頼度の低い単語に対しペナルティを与えるものである。デコーダから出力された単語グラフに付与された音響尤度および言語尤度に基づく各単語に対する事後確率の対数値を、信頼度スコアとして用いる。

2.1.4 単語間遷移スコア

単語間遷移スコア $T(v_{m-1}, v_m)$ は、要約文内の単語連鎖を原文の係り受け関係に基づき制約するスコアである。日本語の要約では、文節内の単語の接続と文節を越えた単語間での接続を区分し、文節内単語間遷移では単語間遷移規則、異文節単語間では文節間の係り受け

確率に基づき、単語間遷移スコアを定義した。しかし、英語では、文節の単位の設定が困難であることから、単語間の係り受け関係を直接推定し、単語間遷移を規定する。単語間の係り受け関係は、SDCFG の確率により推定する。

図 1 に示すように、英語には、前方から後方への係り受けと後方から前方への係り受けが存在する。

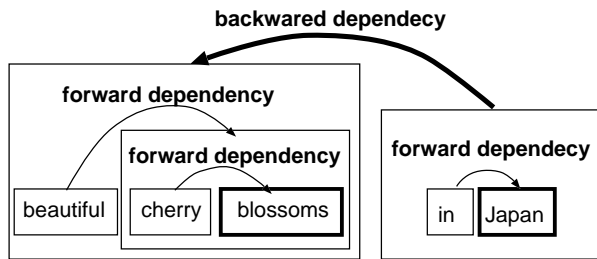


図 1: 英語の係り受け構造の例

英語の SDCFG は、以下の規則からなる。

- $\alpha \rightarrow \beta\alpha$ (forward dependency structure)
- $\alpha \rightarrow \alpha\beta$ (backward dependency structure)
- $\alpha \rightarrow w$

ここで、 α, β は任意の非終端記号、 w は終端記号 (単語) を表す。これらの規則が適用される確率を基に、Inside-Outside 確率を計算し、二つの単語が係り受け関係にある確率 (係り受け確率) を求める。

図 2 に示すように、非終端記号 S から、 L 単語から成る文 w_1, \dots, w_L が生成された場合、 w_m と w_l が係り受け関係にある確率は、 $\alpha \rightarrow \beta\alpha$ が適用され、 β から文節 $w_i \dots w_k$ が生成し、 w_m がその文節の主辞 (β に直接関連する単語) となり、かつ、 α から文節 $w_{k+1} \dots w_j$ が生成され w_l がその文節の主辞となる確率を、すべての α, β, i, j, k に関して総和をとることで計算できる ($\alpha \rightarrow \alpha\beta$ の場合も含む)。しかし、この係り受け確率をそのまま w_m から w_l への単語間遷移確率として用いると、要約時に w_m から $w_{k+1} \dots w_{l-1}$ のいずれの単語にも遷移することができない。このため、単語間遷移確率は、 w_m と w_l が係り受け関係にある確率と、 w_m と $w_{k+1} \dots w_{l-1}$ の各単語が係り受け関係にある確率の総和で定義する。 w_m と w_l が係り受け関係にある確率を $d(w_m, w_l, i, k, j)$ とした場合、 w_m と w_n の単語間遷移スコアは次式のように単語間遷移確率の対数値で定義される。

$$T(w_m, w_n) = \log \sum_{i=1}^m \sum_{k=m}^{n-1} \sum_{j=n}^L \sum_{l=n}^j d(w_m, w_l, i, k, j) \quad (2)$$

本研究で用いる SDCFG は、非終端記号数のみを決定し、全ての非終端記号の組み合わせによる規則が再帰的に適用される。非終端記号には、名詞節といった特殊な機能は設定されていない。全ての規則の適用確率は、学習データに基づき統計的に推定される。学習データで出現回数の多い規則の適用確率が高く、出現回数の少な

い規則は低くなる。音声認識結果のように文法的に誤った文であっても、この SDCFG により、係り受け関係を頑健に推定することが可能である。

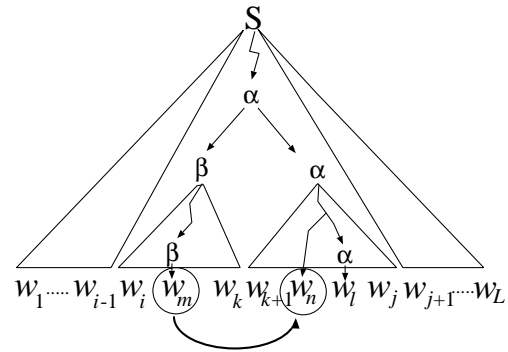


図 2: SDCFG に基づく英語の単語間遷移スコアの原理

2.2 動的計画法による各発話要約と 2 段 DP による複数発話要約

音声認識結果の各文に対し、設定された要約率に基づき、要約スコアが最大となる部分単語列を動的計画法により決定し、自動要約文とする [1][2]。さらに、複数発話要約手法として、文間でのスコアの適用を制約することにより、全文をまとめて自動要約する手法をこれまで提案してきた [5][6]。主題のある複数の発話を対象とするこの複数発話要約手法では、重要な情報を多く含む文は文長が長く、そうでない文は短くなるか完全に削除される。これにより、限られた文字数でより情報量の多い要約文を生成することが可能となる。この複数発話要約手法は、従来の重要文抽出に基づく要約手法と、発話単位の要約手法を統合した手法と解釈できる。

しかし、この手法は、要約対象となる発話数の増加に伴い、単語の組み合わせの数が増加し、計算量が増大してしまう。本研究では、複数発話の要約手法として、2 段 DP による要約手法を提案する。本手法は、第一段階として、各発話要約手法により可能な全ての要約率で各文を要約する。さらに、第二段階として、全体が目的の要約率となるよう各文の要約文を組み合わせ、その中から要約スコアが最大となる組み合わせを動的計画法により決定する。この手法は、単語単位で複数文を要約する処理に比べ、計算量の増加を抑制しつつ、全体に渡る重要度により各文を伸縮するという同様の効果を実現できる。

3 正解要約文単語ネットワークに基づく自動要約法の評価

自動要約文は、被験者の作成した要約文を正解とし、正解要約文に対する類似度により評価を行う。ただし、被験者の作成した正解要約文は、被験者により単語の組み合わせが異なる。このため、正解と考えられる要約文

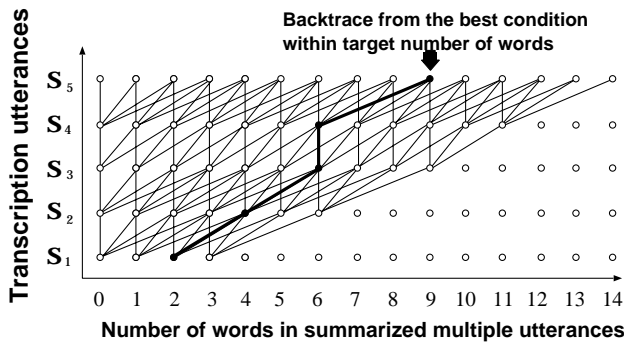


図 3: 複数発話の音声要約のための動的計画法の計算領域の例

を全て網羅することは困難である。さらに、自動要約文の正解要約文を一意に決定することができない。そこで、被験者の作成した正解要約文を単語ネットワークで表現し、ネットワーク上で自動要約文に最も類似している単語列を正解として、正解精度によって自動要約文を評価する手法を提案した [5][6]。正解要約文の単語間の連鎖をネットワークにまとめることにより、全ての可能性のある正解要約文の単語連鎖を近似的に網羅できる。

図 4 に、正解要約文単語ネットワークの例を示す。ただし、単語ネットワーク上の右矢印は単語間の接続を示している。このネットワーク上で、文頭記号<s>から文末記号</s>まで右矢印で結ばれた単語列は、要約文として重要な情報を保持し、言語的に正しく、原文の文意を保持した単語連鎖であることから、正解要約文と仮定できる。

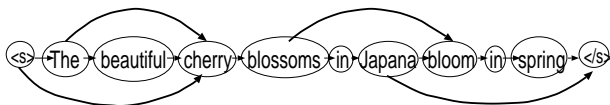


図 4: 正解要約文単語ネットワークの例

そこで、正解要約文ネットワーク上で自動要約文自身に最も近い単語列を正解として、式 (3) で表される要約正解精度により自動要約文を評価する。

$$Sum_acc = \frac{Len - Sub - Ins - Del}{Len} \times 100[\%] \quad (3)$$

ただし、 Sum_acc は要約正解精度、 Sub は置換誤り、 Ins は挿入誤り、 Del は削除誤り、 Len はネットワーク上で最も類似している単語列における単語数を示している。

要約正解精度は、正解要約文単語ネットワークを用いることにより、重要な情報の抽出、文意の保持、要約文としての適正について、まとめて一元的に評価できる特徴がある。

4 評価実験

4.1 実験条件

NIST主催の Topic Detection and Tracking(TDT) のタスクから、CNN ニュースの 5 ニュースの音声認識結果と音声認識誤りを含まない正解書き起こし文を、自動要約の対象とした。単語には品詞情報を付加して各モデルを学習し、評価時にも品詞情報を付加した。品詞付けには Brill tagger を用いた [7]。5 ニュース中の 50 発話文 (平均 30 形態素) について、70% と 40% の要約率で各文要約を行った。さらに、5 ニュースをニュース単位で、70% と 40% の要約率で複数文要約を行った。ただし、要約率は原文の単語数に対する自動要約文の単語数の割合によって定義される。

生成された自動要約文を、英語を母国語とする 17 人の被験者の作成した正解要約文の単語ネットワークに基づき評価を行った。

4.2 音声認識部の構成

英語のニュース音声を認識するため、JRtk(Janus Speech Recognition Toolkit) [8] を用いて音声認識を行った。認識システムの諸条件を以下に示す。

特徴抽出

音声データを 16kHz、16bit でデジタル化し、13 次元のメルケプストラム (MFCC) を抽出し、声道長正規化 (VTLN) および、クラスタ毎のケプストラム平均正規化を行う。次に、7 フレームを 1 セグメントとして特徴ベクトルの単位とし、各セグメントを LDA (Linear Discriminant Analysis) に基づいて 42 次元に圧縮する。

音響モデル

2000 個のガウス分布コードブックを 6000 状態の各混合重み分布で共有した、計 105k のガウス分布を持つ quinphone HMM を用いる。学習データには Wall Street Journal (WSJ)、English Spontaneous Scheduling Task (ESST)、Broadcast News (BN)、Crossfire and Newshour TV news show を用いている。

言語モデル

言語モデルとして、BN コーパスから学習した語彙サイズ 40k の単語 trigram を用いる。

デコーダ

JRtk[8] を用いて、単語グラフを中間表現とする以下の 3 パスからなる探索を行う。第 1 パスでは triphone と bigram を用いて木構造辞書に基づくフレーム同期のビームサーチを行い、単語グラフを生成する。第 2 パスでは単語グラフ上の単語からフラットな辞書を構成し、quinphone と trigram を用いてビームサーチを行い単語グラフを再構成する。第 3 パスでは、単語グラフを最小化した後で trigram を用いてリスコアを行い最終的な認識結果を得る。

4.3 要約処理部の構成

単語重要度, 要約用言語モデル, および単語間遷移確率に適用する SDCFG は, Penn Treebank コーパスに含まれる Wall Street Journal と BROWN コーパス 10681 文 (約 3.5M 単語) を用いて算出した [9]. SDCFG は, Penn Treebank に付与された非終端記号を適用せず, 括弧付きのコーパスとして非終端記号数のみを固定して学習した. ただし, 非終端記号数は 100 である.

要約用言語モデルの性能を評価するため, 被験者が作成した正解要約文に対するパープレキシティ, および未知語率を図 5 に示す. 被験者による各文要約 (EACH) と複数文要約 (MULTIPLE), および要約前の書き起こし (TRS) を比較する. %値は要約率を示す.

要約用言語モデルは要約文より学習した言語モデルを適用することが望ましいが, 大量の要約文で構成されたコーパスは現存しない. そこで, ニュース音声の要約文の代替として新聞記事を用いた. 正解要約文に対するパープレキシティの値が高いことから, 新聞記事に基づく要約用言語モデルが, 要約文の特徴を表現するには不十分であることが分かる. さらに, 正解要約文に対するパープレキシティが, 3gram に比べ 2gram が低いことから, 要約用言語モデルが新聞記事から学習された場合, 3gram では単語連鎖の制約が強すぎ, 要約文の精度の劣化を招いてしまう恐れがある. そこで, 本研究では 2gram を適用し, 自動要約文を生成する.

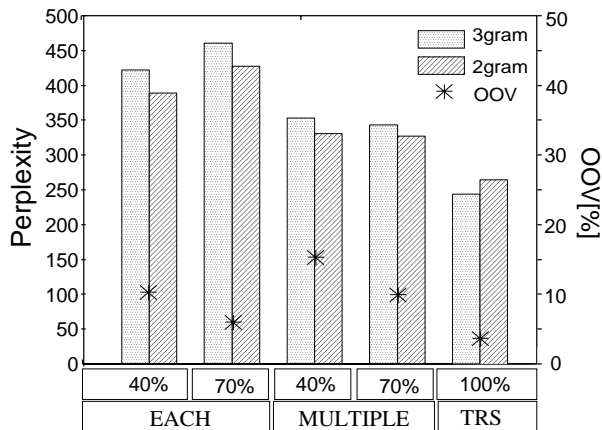


図 5: 要約用言語モデルの正解要約文に対する Perplexity および未知語率

4.4 評価結果

音声認識結果 (REC) に対する自動要約文を, 正解単語ネットワークにより評価した例を, 表 1 に示す. 図 6 に各文要約, 図 7 に複数文要約の正解要約文単語ネットワークに基づく単語正解精度を示す. 要約用言語モデルとして 3gram を適用した結果と要約正解精度を比較したところ, 2gram を用いた方が最大約 2% 精度が向上することが確認された.

単語重要度スコア (I) のみを用いた場合と言語スコ

表 2: 要約文中の認識誤りと認識誤りを含む要約文数

要約前	各文要約		複数文要約	
	180 単語 (45 文)	326 単語 (94 文)	40%	70%
要約率	40%	70%	40%	70%
I	42 (27)	111 (40)	99 (56)	199 (71)
IL	44 (28)	87 (37)	86 (53)	166 (69)
ILC	23 (15)	49 (22)	34 (28)	82 (47)
ILT	46 (27)	84 (37)	90 (56)	173 (69)
$ILCT$	22 (13)	51 (24)	25 (17)	80 (47)
RDM	82 (30)	87 (21)	89 (45)	169 (65)

() 内は認識誤りが含まれる要約文の数を表す.

ア (L) と組み合わせた場合 (IL) を比較する. さらに単語間遷移スコア (T) を考慮した場合 ($ILT, ILCCT$) と考慮しない場合 (IL, ILC) を比較する. また, 信頼度スコア (C) を用いた場合 ($ILC, ILCCT$) と用いない場合 (IL, ILT) を比較する. 音声認識結果 (REC) については, 全 5 種類の自動要約文を生成した. 言語スコアおよび信頼度スコアの重み係数 $\lambda_L, \lambda_C, \lambda_T$ および単語重要度スコアにおける一定値 I_{const} は, 様々に変化させ求められた最適値を用いた.

さらに, 音声認識の精度が 100% であった場合の音声自動要約の性能を調べるため, 人手により音声を書き起こした正解書き起こし文 (TRS) について自動要約を行った. 正解書き起こし文 (TRS) の自動要約では信頼度スコアが存在しないので, 上記 5 種類の自動要約文のうち信頼度スコアを考慮しない 3 種類を生成した.

提案手法の有効性を検証するため, 自動要約文と等しい要約率で単語をランダムに抽出した要約文 (RDM) に対して評価を行った. さらに, 上限として, 被験者各 17 人の被験者の正解要約文を, 他の 16 人の正解要約文で作成した要約文単語ネットワークに基づき評価した平均正解精度 (SUB) を示す.

全ての自動要約条件において, ランダムに単語を抽出した場合 (RDM) と比較して, 要約正解精度が有意に高くなることが示された. 各文要約, 複数文要約のどちらにおいても, 単語重要度 I に比べ言語尤度を組み合わせた IL の精度が高い. さらに, 単語間遷移スコアを組み合わせた ILT の要約精度が最も高いことが示されている. 音声認識結果に対する自動要約では, 信頼度スコア C を組み合わせることにより, さらに改善されている. 以上の点から, 原文の係り受け関係を考慮し, 認識誤りの自動要約文への抽出を抑制することにより, 誤要約が削減されたものと考えられる. ただし, 被験者の作成した正解要約文の精度にはまだ至っていない.

音声認識誤りのうち置換誤りと挿入誤りが, 複数文要約のテストセット 125 文 (5 ニュース) 中 94 文に計 362 単語, 各文要約のテストセット 50 文中 45 文に計 180 単語の誤りが含まれている. これらの認識誤りが要約文に抽出された場合, 発話意図とは異なる要約文が生成されてしまう. 提案手法による自動要約文に抽出された誤認識単語の数と, その誤認識単語が含まれる自動要約文の数を表 2 に示す. 表 2 より, 提案手法が, 誤認識された単語が含まれていることによる誤要約を削減できていることが分かる. また, 信頼度スコアにより, 顕著に認

表 1: 正解要約文単語ネットワークによる音声自動要約結果の評価例

音声認識結果	VICE PRESIDENT AL GORE SAYS THE GOVERNMENT HAS A PLAN TO AVOID THE INEVITABLE PROSPECT OF INCREASED AIRPLANE CRASHES AND FATALITY IS
70% 要約文	VICE PRESIDENT AL GORE SAYS THE GOVERNMENT HAS A PLAN TO AVOID THE INCREASED AIRPLANE CRASHES
	VICE PRESIDENT AL GORE SAYS THE GOVERNMENT HAS A PLAN TO AVOID INCREASED AIRPLANE CRASHES
40% 要約文	<INS> THE GOVERNMENT HAS A PLAN TO AVOID THE INCREASED AIRPLANE CRASHES
	GORE THE GOVERNMENT HAS A PLAN TO AVOID THE INCREASED AIRPLANE CRASHES
	THE INCREASED AIRPLANE CRASHES

各要約率で、上段：正解要約文，下段：自動要約文，_は認識誤り，INS は挿入，DEL は脱落を表す。

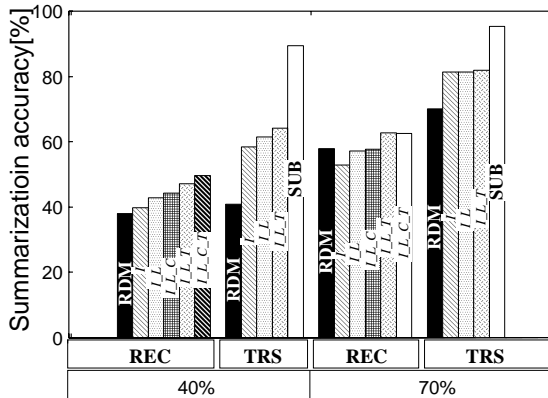


図 6: 40%,70%要約率における各文要約の要約正解精度

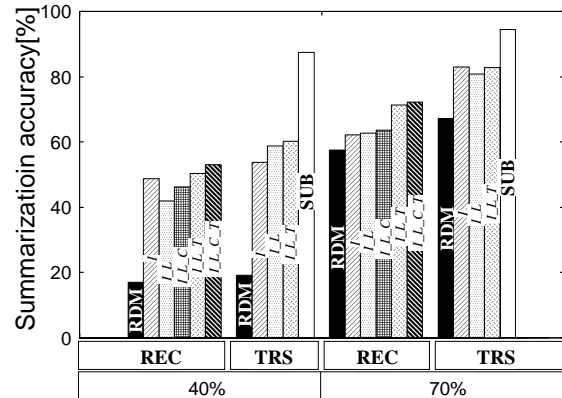


図 7: 40%,70%要約率における複数文要約の要約正解精度

識誤りが排除されていることが分かる。

5 まとめ

本稿では、英語のニュース音声の音声認識結果と書き起こしを対象として、発話文単位要約手法と複数発話要約手法により自動要約を試みた。英語のニュース音声の自動要約の結果、日本語のニュース音声の自動要約の結果と同様、提案手法が相対的に冗長な情報を削除しつつ重要な情報を抽出していることが示された。言語スコアの本来の目的は、要約文として尤もらしい単語連鎖を生成することであるが、同時に文脈に整合しない単語の自動要約文への抽出を削減できる。その結果、文脈に整合しない音声認識誤りを排除している。日本語のニュース音声の要約と比較して、英語のニュース音声の要約では信頼度スコアによる改善が大きい。これは、日本語のニュース音声の自動要約では、信頼度スコアは自動要約文中の認識誤りを削減する作用に留まっているのに対し、英語のニュース音声の自動要約では、さらに明確に発音された重要な単語を抽出する作用があるものと考えられる。

謝辞

音声認識装置を提供し、英語の音声認識に貢献してくださった Carnegie Mellon University の Alex Waibel 教授、Rob Malkin 氏、Hua Yu 氏に感謝致します。正解

要約文作成に協力してくださった Sheffield 大学の Yoshi Gotoh 氏に感謝致します。

参考文献

- [1] 堀 智織, 岩崎 淳, 古井 貞真, “話題語に着目したニュース音声の要約法の検討”, 音学秋季講論, Vol.1, 3-1-11, pp.117-118(1999).
- [2] C. Hori and S. Furui, “Automatic Speech Summarization based on Word Significance and Linguistic Likelihood,” Proc.ICASSP2000, Vol.3, pp.1579-1582 (2000).
- [3] K. Knight and D. Marcu, “Statistics-Based Summarization — Step One: Sentence Compression,” Proc. National Conference on Artificial Intelligence (2000).
- [4] 堀, 古井, “係り受け SCFG に基づく音声自動要約法の改善”, 信学技報, SP2000-95-116, pp.127-132 (2000).
- [5] 堀, 古井, “主題のある複数発話を対象とした音声自動要約法”, 音講論, Vol.1, pp.93-94 (2001).
- [6] C. Hori and S. Furui, “Advances in Automatic Speech Summarization,” Proc. EUROSPEECH2001, vol.III, pp.1771-1774, Aalborg (2001).
- [7] <http://www.cs.jhu.edu/~brill>
- [8] A.Waibel et al., “Advances in Meeting Recognition,” Proc. HLT2001, pp.11-13, San Diego (2001).
- [9] <http://www.cis.upenn.edu/~treebank>