

1 はじめに

我々は、これまで単語抽出に基づく音声自動要約の手法 [1][2] を提案してきた。しかし、すべての文を一樣に扱い、単語抽出により要約を行う従来の手法を用いて要約された文は、要約率が小さいときに文として不完全なものを生じやすく、要約精度が低下してしまう課題があった。

そこで本稿では、あらかじめ原文から重要な文を抽出しておき、残った文に対して従来の単語抽出による要約を行う手法を用いることで、要約文をできるだけ自然な文章にし、要約文の高精度化をはかる。この手法を用いて講演音声を自動要約し、複数の被験者により作成された正解要約文単語ネットワークに基づく評価 [3] を行った結果を報告する。

2 音声自動要約手法

まず、ユーザーによって与えられた要約率から、重要文抽出手法と単語抽出手法でそれぞれどのくらい要約するかの割合を決定する。定められたそれぞれの要約率に従い、認識文を重要文抽出、単語抽出の順で要約する。

2.1 重要文抽出部

認識された原文から認識率の低い文、理解困難な文をあらかじめ要約文から除き、重要文を抽出する。

1 文が N 個の単語からなる認識単語列 $W = w_1, w_2, \dots, w_N$ のときの、重要文抽出に用いる要約スコア $S_s(W)$ を以下のように定義する。

$$S_s(W) = \frac{1}{N} \sum_{n=1}^N \{L_s(w_n) + \lambda_{I_s} I_s(w_n) + \lambda_{C_s} C_s(w_n)\} \quad (1)$$

個別のスコアとして、これまで要約文の生成に有効であることが確認されているものから、以下に説明する言語スコア L_s 、重要度スコア I_s 、信頼度スコア C_s を選んで用いた。 λ_{I_s} 、 λ_{C_s} は各スコアのバランスをとるための重み係数である。

言語スコア

言語スコア $L_s(w_i)$ は各文の単語連鎖の適正度を表すスコアであり、以下の式で表される。

$$L_s(w_i) = \log P(w_i | \dots w_{i-1}) \quad (2)$$

$P(w_i | \dots w_{i-1})$ の計算には 1.5M 形態素からなる講演の書き起こしに前処理を施して論説調の表現に変換したものと、60 講演の予稿集から作成した単語 trigram を用いた。

このスコアは認識誤りによって生じる文中の言語的に不自然な単語連鎖に対して、より小さい値が付けられる。

重要度スコア

重要度スコア $I_s(w_i)$ は原文の中での相対的な文の重要度を表すスコアであり、単語の出現頻度に基づく情報量から求める。

名詞、動詞、形容詞、未定義語の内容語に、重要度スコアとして式 (3) で表される、単語の出現頻度に基づく情報量を適用する [4]。

$$I_s(w_i) = f_i \log \frac{F_A}{F_i} \quad (3)$$

w_i : 音声認識結果に含まれる内容語

f_i : 音声認識結果中の内容語 w_i の出現頻度

F_i : 大規模コーパス中での内容語 w_i の出現頻度

F_A : 大規模コーパス中での総内容語数 ($= \sum_i F_i$)

講演の書き起こし (1.5M 形態素)、60 講演の予稿集、WWW 上の講演録 (2.1M 形態素)、形態素の NHK のニュース原稿 (22M 形態素)、毎日新聞 (87M 形態素) および「音声情報処理」(51k 形態素) のテキストを用い、出現した全約 120k 種類の単語の各々の出現頻度を求めたコーパスを用いた。

このスコアはキーワードとなる重要な単語に対して大きい値が付けられ、逆に、認識誤りの単語のように原文の内容と関係のない単語には小さい値が付けられる。

信頼度スコア

信頼度スコア $C_s(w_i)$ は、音響的、言語的な信頼度を表すスコアである。デコーダから出力された単語グラフに付与された、単語仮説 w_i が出現する事後確率の対数値を各単語に与える。このスコアは音響尤度および言語尤度から計算され、音響的、言語的に信頼度の低い単語には小さい値が付けられる。

2.2 単語抽出部

認識精度の低い文、重要度の低い文をある程度削除した後の文に、従来の単語抽出 [1] の手法を用いて最終的に出力される要約文を作成する。

この手法は言語スコア、単語重要度スコア、信頼度スコア、および単語間遷移スコアを各単語に付け、これらを用いて 2 段 DP による要約を行う。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。

* Use of the important sentence extraction in automatic speech summarization

By Tomonori Kikuchi, Sadaaki Furui (Tokyo Institute of Technology), and Chiori Hori (NTT Communication Science Laboratories)

言語スコア, 単語重要度スコア, 信頼度スコアはそれぞれ, 2.1 重要文抽出部で使用したものと同一コーパスを用いて求めた. また, 単語間遷移スコアは毎日新聞約 4 万文の構文解析済みの京大テキストコーパスを用いて, 構文木制約付きの Inside-Outside アルゴリズムにより, 係り受けパラメータの推定を行ったものから求めた.

3 評価実験

3.1 要約実験条件

話し言葉コーパス中の男性話者 (AS99SEP097) 1 名による単語正解精度約 70% の音声認識結果を, 提案手法を用いて 70% と 50% の要約率で自動要約を行った. 音声認識システムは文献 [3] と同じものを用いた.

認識結果について, 重要文抽出手法を用いないもの (NOS), 言語スコア (L_s), 重要度スコア (I_s), 信頼度スコア (C_s) をそれぞれ単独で用いた重要文抽出手法を取り入れたもの, それぞれのスコアを組み合わせた重要文抽出手法 (L_s-I_s , I_s-C_s , C_s-L_s), さらにこれら全てのスコアを組み合わせた重要文抽出手法を取り入れたもの ($L_s-I_s-C_s$) について, 全 8 種類の自動要約文を生成した. ただし, 言語スコア, 重要度スコア, 信頼度スコアの重み係数 λ_{L_s} , λ_{I_s} , λ_{C_s} , および, 重要文抽出手法と単語抽出手法での要約の割合は実験的に求めた最適値を用いた.

要約文の評価は 9 人の被験者の作成した正解要約文に基づき作成した単語ネットワークを用いて行った. 生成された自動要約文自身に最も近い単語列をネットワーク中から抽出し, その単語列を正解文として, 要約文の単語正解精度 (要約精度) を求める.

さらに, 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.

3.2 評価結果

実験結果を図 1, 2 に示す. 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と比較して, 従来の手法, 本手法とも要約精度が高くなり, 手法の有効性が示された. また, 要約率 50%, 70% のどちらの要約文においても重要文抽出手法を用いない場合より, 用いた場合の要約正解精度が高いことが示された. 重要文抽出手法と単語抽出手法で要約する割合がおよそ 2:1 のとき, 最も要約精度が良かった.

重要文抽出に用いる各スコアについては, 重要度スコア (I_s), もしくは信頼度スコア (C_s) を単独で用いたものについて, どちらも約 2% の要約精度の改善がみられた. また, これら 2 つのスコアの組合せ (I_s-C_s) による重要文抽出手法では, それぞれのスコアを単独に用いたときよりもさらに要約精度は高くなり, 重要文抽出手法を用いない場合に比べて約 3% の改善がみられた.

一方, 言語スコア (L_s) を単独で用いた重要文抽出手法についてはほとんど効果が得られなかった. そのため, 他の 2 つのスコアと L_s とを組み合わせた (L_s-I_s , C_s-L_s) 重要文抽出手法, 全てのスコアを組み合わせた ($L_s-I_s-C_s$) 重要文抽出手法においても,

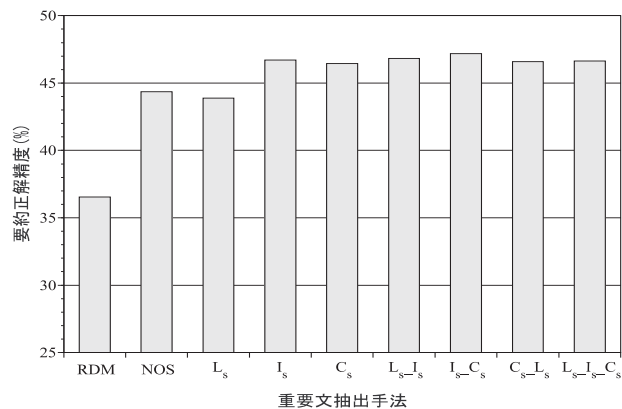


図 1. 要約率 50% の要約精度

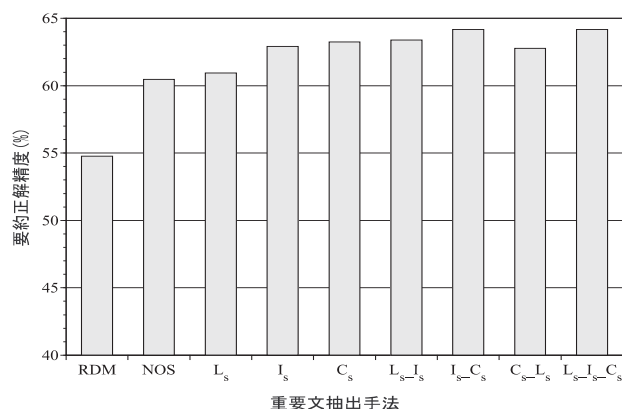


図 2. 要約率 70% の要約精度

I_s-C_s に比べて要約精度の改善はみられず, 言語スコアは要約文の生成には有効であるが, 文の重要性とは関係の少ないことが分かる.

4 まとめ

本稿では, 従来の単語抽出による要約文作成手法の前段階に重要文抽出手法を組み込んだ音声自動要約の手法を提案した. 講演音声の自動要約を行い要約精度を評価したところ, この手法の有効性が示された. また複数のスコアを組み合わせることにより, 要約精度をより改善することができた.

今後の課題として, より高精度な要約スコアの検討, 被験者による主観評価実験による種々の手法の比較検討を行っていきたい.

参考文献

- [1] 堀智織, 古井貞熙: “講演録作成を目的とした講演音声自動要約,” 日本音響学会 2001 年秋季講演論文集, 2-1-10, pp.67-68(2001-10).
- [2] 堀智織, 古井貞熙: “単語抽出による音声自動要約文生成方とその評価,” 電子情報学会論文誌 D-II, Vol. J85-D-II, No.2, pp.200-209 (2002-2).
- [3] 堀智織, 古井貞熙: “講演音声の自動要約の試み,” 話し言葉の科学と工学ワークショップ講演予稿集, pp.165-171(2001-3).
- [4] 岩崎淳, 古井貞熙: “ニュース音声からの話題抽出法の検討,” 日本音響学会 2001 年秋季講演論文集, 1-1-14, pp.27-28(1998-10).