

論文 / 著書情報
Article / Book Information

論題(和文)	重要文抽出と文圧縮による音声自動要約手法の英語音声への適用
Title(English)	
著者(和文)	Heie Matthias, 岩野 公司, 古井 貞熙, 堀 智織
Authors(English)	Matthias Hietland Heie, Koji Iwano, SADAOKI FURUI
出典(和文)	日本音響学会 2004年春季講演論文集, Vol. , No. 2-8-10, pp. 79-80
Citation(English)	, Vol. , No. 2-8-10, pp. 79-80
発行日 / Pub. date	2004, 3

重要文抽出と文圧縮による音声自動要約手法の英語音声への適用*

△ハイエ マティアス ©岩野 公司 古井 貞熙 (東工大) 堀 智織 (NTT CS 研究所)

1 はじめに

我々はこれまでに、重要文抽出と文圧縮の2段階処理に基づく、音声自動要約手法を提案し、日本語話し言葉コーパス(CSJ)の音声を用いた実験により、提案手法の有効性を確認している[1]。この手法では、第一段階で予め音声認識結果から認識率の低い文、理解が困難と判断される文を除き、第二段階で、単語抽出による要約[2]により文圧縮を行う。単語抽出による要約に重要文抽出を組み込むことで、単語を単位とした自由度の高い要約文生成を実現しつつ、出現位置の離れた単語の連結による不自然な要約文の生成を抑制することが可能となる。本稿では、この音声自動要約手法を英語ニュース音声に適用した結果について報告する。我々は既に、文圧縮のみに基づく要約手法を、英語音声に適用した結果について報告を行っている[3]。そこで、ここでは、重要文抽出手法の組み込みによる効果について論ずる。

2 音声自動要約手法

重要文抽出と文圧縮による音声自動要約システムを図1に示す。

まず、ユーザーは要約率を設定する。要約率は、原文の単語数に対する要約文の単語数の割合として定義される。また、重要文抽出と単語抽出、それぞれでどれだけの要約を行うか、割合を設定する。システムは、音声認識を行ったのち、認識文からフィラー単語を削除する。残された文について重要文抽出、単語抽出の順で、それぞれの要約スコアをもとに要約処理を行う。

2.1 重要文抽出

重要文抽出では、入力文ごとに以下で定義される要約スコアを求め、スコアの上位となる文を設定した要約率になるまで選択する。

1文が N 個の単語からなる認識単語列 $W = w_1, w_2, \dots, w_N$ に対する、要約スコア $S_s(W)$ は以下のように定義される。

$$S_s(W) = \frac{1}{N} \sum_{n=1}^N \{L(w_n) + \lambda_I I(w_n) + \lambda_C C(w_n)\} \quad (1)$$

L, I, C はそれぞれ言語スコア、重要度スコア、信頼度スコアであり、 λ_I, λ_C は各スコアのバランスをとるための重み係数である。以下、個々のスコアについて詳しく説明する。

言語スコア

言語スコア $L(w_i)$ は各文の単語連鎖の適正度を表すスコアであり、以下のように単語trigramを用いて算出する。

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

重要度スコア

重要度スコア $I(w_i)$ 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算出される。

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

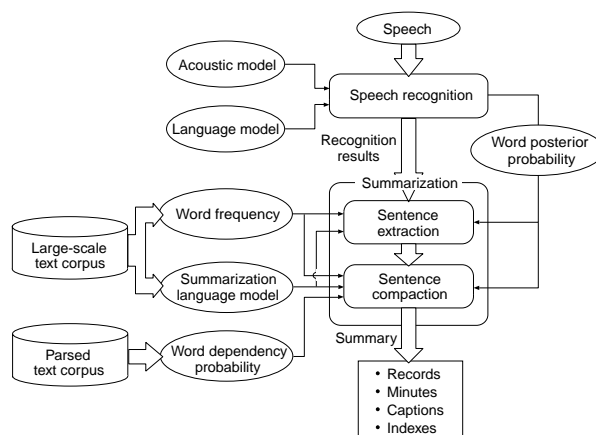


図1. 重要文抽出と文圧縮による音声自動要約システム

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 ($= \sum_i F_i$)

スコアは内容語のみに付与され、それ以外の単語についてはスコアを0と定義した。なお、本研究では、名詞と動詞を内容語とした。

信頼度スコア

信頼度スコア $C(w_i)$ 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で定義される。

2.2 単語抽出

単語抽出は、文献[3]と同様に行った。この要約手法は、2段DP法に基づき、要約スコアが最大となる部分単語列を決定するものである。要約スコアとしては、重要文抽出と同様の「言語スコア」「重要度スコア」「信頼度スコア」の他に、要約文の単語間の係り受け構造を考慮した「単語間遷移スコア」を用いる。そのため、構文解析済みのテキストコーパスから単語単位の係り受けSCFGを学習しておく。

3 評価実験

3.1 実験条件

NIST主催のTopic Detection and Tracking (TDT) タスクの、CNNニュースの5つのニュース(128発話文)を要約対象とした。これらを、要約率40%、70%で要約する。

評価には、要約正解精度[2]を用いる。本実験では、英語を母国語とする17名の被験者が作成した要約正解文から、単語ネットワークを生成し、要約正解精度を求めた。

文献[3]では、同じ評価データ・実験条件における、単語抽出のみによる要約正解精度を報告している。単語抽出における4つの要約スコアを、最適な重みで融合して用いたときに最も高い性能を示し、要

* Application of automatic speech summarization based on sentence extraction and compaction to English speech
By Matthias Heie, Koji Iwano, Sadaoki Furui (Tokyo Institute of Technology), and Chiori Hori (NTT Communication Science Laboratories)

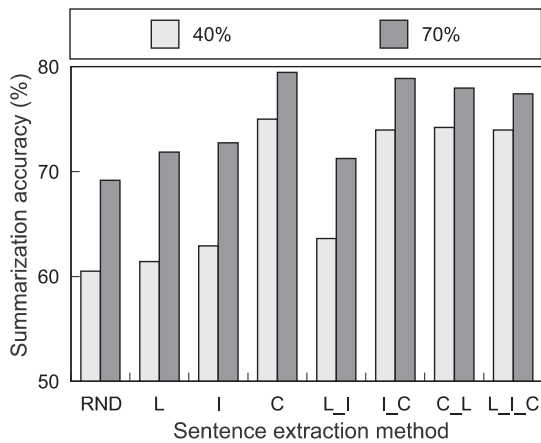


図 2. 要約スコアごとの重要文抽出による要約正解精度

約正解精度は 40% 要約で 54.1%, 70% 要約で 71.1% であった。本実験では重要文抽出を組み込んだ効果についての評価を行うため、単語抽出部はこの最適化された条件で固定しておく。

3.2 音声認識部

音声認識には JRTk (JANUS Speech Recognition Toolkit)[4] を用いた。使用した音響・言語モデルは、文献 [3] と同様である。なお、評価データに対する単語正解精度は 78.4% であった。

3.3 要約部

言語スコア算出のための trigram、重要度スコア算出のための単語頻度情報は、Penn Treebank コーパス [5] に含まれる Wall Street Journal と Brown コーパス (約 175k 文, 4.7M 単語) から作成した。単語抽出に用いる SDFG は、Brown コーパス中の約 11k 文を使って学習した。

3.4 実験結果

まず、重要文抽出における最適な要約スコアの組み合わせを検証した。要約スコアとして、言語スコア (L)、重要度スコア (I)、信頼度スコア (C) のみを使用した場合、それぞれのスコアを組み合わせで使用した場合 (L_I , I_C , C_L)、全てを組み合わせで使用した場合 (L_I_C) について実験を行った。重み係数 λ_L, λ_C は、ニュースごとに、他の 4 つのニュースを用いた実験結果が最良となるように最適化を行った。また比較用として、ランダムに重要文抽出を行った場合の結果を RND として示した。40%, 70% 要約における重要文抽出のみでの要約正解精度を図 2 に示す。ランダム選択と比較して、全ての要約スコア条件で要約正解精度の向上が確認できる。特に、信頼度スコア C が有効であることがわかる。

そこで、信頼度スコア C を含む種々の要約スコアを利用した場合について、重要文抽出・単語抽出を組み合わせた手法の要約性能の評価を行う。重要文抽出による要約の割合を 0 ~ 1 まで 0.1 刻みで変化させた時の 40%, 70% 要約時の要約正解精度の変化をそれぞれ図 3, 4 に示す。横軸が重要文抽出による要約の割合であり、0 が単語抽出のみを行った場合、1 が重要文抽出のみを行った場合に相当する。重み係数には、重要文抽出のみによる実験において最適化された値をそのまま用いている。

40% 要約では、重要文抽出のみを行った場合に最も性能がよく、単語抽出のみによる結果と比較すると絶対値で約 21%、要約正解精度が向上している。しかし、両手法を組み合わせた効果は得られていな

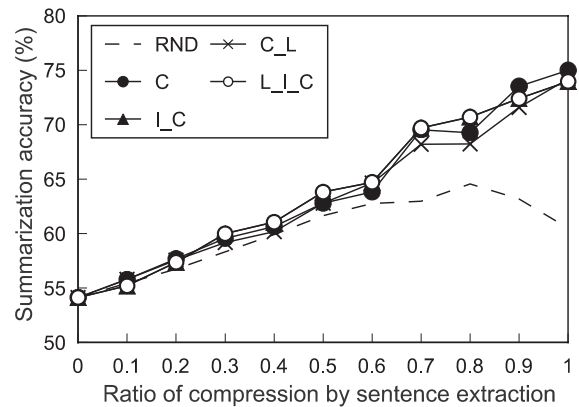


図 3. 40% 要約における重要文抽出による要約の割合と要約正解精度

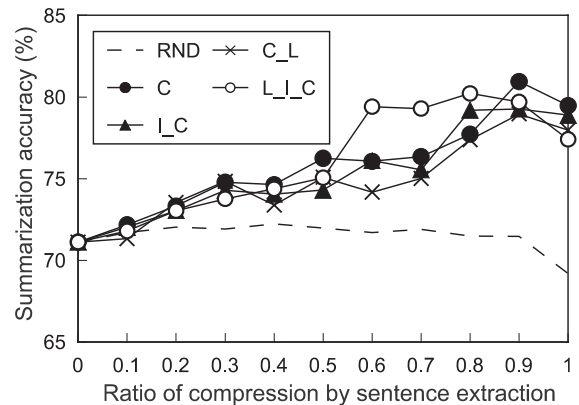


図 4. 70% 要約における重要文抽出による要約の割合と要約正解精度

い。70% 要約では、重要文抽出と単語抽出の組み合わせ効果が得られ、単語抽出のみの結果から最高で約 10% の要約正解精度の向上が見られる。最適な重要文抽出の割合が 0.8 ~ 0.9 となっていることから、70% 要約の場合においても重要文抽出の効果が大きいことがわかる。

4 おわりに

重要文抽出と単語抽出の 2 段階処理に基づく音声要約手法を英語ニュース音声に適用した結果について報告した。実験結果からは、英語ニュース音声の要約に対し、特に重要文抽出の効果が大きいことが示されたが、日本語話し言葉の要約結果 [1] には、これほど顕著な傾向は見られない。これは、言語の違いだけでなく、放送ニュースと話し言葉というスタイルの違いにも起因している可能性がある。今後は、日本語ニュース音声、あるいは英語話し言葉を対象とした要約実験の結果との比較や、被験者の作成した正解要約文間の特徴比較を行い、その背景を明らかにする必要がある。

参考文献

- [1] 菊池 他, 音講論, vol.1, pp.97-98 (2002-9).
- [2] 堀 他, 信学論 D-II, vol.J85-D-II, no.2, pp.200-209 (2002-2).
- [3] 堀 他, 音講論, vol.1, pp.69-70 (2001-10).
- [4] A. Waibel, et al., Proc. HLT 2001, pp.11-13, San Diego (2001-3).
- [5] <http://www.cis.upenn.edu/~treebank/>