論文 / 著書情報
Article / Book Information

論題(和文)	音声自動要約におけるSVDを用いた重要文抽出手法の検討
Title(English)	
著者(和文)	広畑 誠, 新中 庸介, 古井 貞熙
Authors(English)	SADAOKI FURUI
出典(和文)	日本音響学会 2003年秋季講演論文集, Vol. , No. 2-6-17, pp. 93-94
Citation(English)	, Vol. , No. 2-6-17, pp. 93-94
発行日 / Pub. date	2003, 9

音声自動要約における SVD を用いた重要文抽出手法の検討 *

◎広畑 誠　新中 庸介　古井 貞熙 (東工大)

1　はじめに

我々はこれまで講演音声の自動要約手法を提案してきた [1]．本稿では，特異値分解 (Singular Value Decomposition; SVD) を用いて重要文抽出を行う要約手法として，潜在的意味分析に基づく手法と次元圧縮に基づく手法について検討する．得られた要約文に対し，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を行った結果を報告する．

2　重要文抽出による音声自動要約

2.1　文境界判定

重要文抽出を行うにあたり，認識結果を自動的に文単位に区切る必要がある．しかし，講演音声などの話し言葉には，テキストの場合とは異なり，明確な文の単位が存在しない．そこで，文境界として，一般に句点が入るとされる位置 (文境界 1) と，さらに体言止，並列節，条件節などの節境界を考慮した位置 (文境界 2) の 2 つを定義し，それぞれの文単位について検討を行った．文境界の判定は，全てのポーズ区間の前 2 単語 (w_1, w_2) と後ろ 2 単語 (w_3, w_4) を利用し，閾値 θ を用いた式 (1) により行う [2]．

$$-\frac{1}{4}\log P(w_1, w_2, w_3, w_4) + \frac{1}{5}\log P(w_1, w_2, \text{境界}, w_3, w_4) \geq \theta \quad (1)$$

2.2　潜在的意味分析に基づく手法 (手法 1)

テキスト重要文抽出において，その文書内で意味のある文を抜きだし，不要なものは選択しないようにするための手法として，潜在的意味分析に基づく手法が提案されている [3]．この重要文抽出手法では，潜在的意味分析を行うため，特異値分解を利用している．まず，行成分が単語 (名詞)，列成分が文の行列 $\mathbf{A} = [A_1 A_2 \dots A_N]$ を用意する．語彙数は M，文数は N であり，$\mathbf{A}$ は $M \times N$ の行列となる．ここで，文 n に対する列成分 $A_n = [a_{1n}, a_{2n}, \dots, a_{Mn}]^\top$ の要素 a_{mn} を文 n における単語出現頻度ベクトル $T_n = [t_{1n}, t_{2n}, \dots, t_{Mn}]^\top$ の要素 t_{mn} を用いて以下の式で定義する．

$$a_{mn} = L(t_{mn}) \cdot G(t_{mn}) \quad (2)$$

ただし，$L(t_{mn})$ は文 n の単語 m に対しての局所重み関数で，$G(t_{mn})$ は大域重み関数である．

局所重み関数としては，以下の 2 種類を用意する．

- No weight (N) : $L(t_{mn}) = t_{mn}$.
- Binary weight (B) : $t_{mn} > 0$ なら $L(t_{mn}) = 1$, そうでなければ $L(t_{mn}) = 0$.

また，大域重み関数として以下の 2 種類を用意する．

- No weight (N) : $G(t_{mn}) = 1$.
- Corpus weight (C) : $G(t_{mn}) = \log \frac{F_A}{F_m}$．ただし，$F_m$ は大規模コーパス中での単語 m の出現頻度で，F_A は大規模コーパス中での総単語数 $(= \sum_m F_m)$ である．

大規模コーパスには話し言葉コーパス (CSJ) の講演書き起こし (約 8M 形態素) のテキストを用い，出現した約 50k 種類の単語の出現頻度を求めた．

a_{mn} を要素に持つ行列 $\mathbf{A}(M \geq N)$ が与えられたとき，$\mathbf{A}$ は特異値分解によって，次のように 3 つの行列で表現できる．

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top \quad (3)$$

特異値分解の特徴は単語同士の相互関係を捕らえ，潜在的な単語と文のクラスタリングができることにある．各特異値に対する最適な文を選択することを目的として，以下のアルゴリズムで重要文抽出を行う．

(1) $k = 1$ とし，$\mathbf{A}$ に対して SVD を行う．

(2) 右特異行列 $\mathbf{V}^\top = [V_1 V_2 \dots V_N]^\top$ に対し，右特異ベクトル $V_k = [v_{1k}, v_{2k}, \dots, v_{Nk}]^\top$ の成分で最も大きい値に対応する文を抽出する．

(3) 目標となる要約率に達したら終了し，そうでなければ $k = k + 1$ とし，ステップ (2) に戻る．

このアルゴリズムにより，類似した文が抽出されず，重要な話題を網羅するように文を抽出できる．

2.3　次元圧縮に基づく手法 (手法 2)

特異値分解により，各文は特異値によって重み付けられたベクトル空間に射影することができる．このとき，利用する特異値の数を減らすことで，射影するベクトル空間の次元を圧縮することが可能となる．このベクトル空間において，各文にスコアを与えて重要文抽出を行う手法を提案する．

特異値行列 $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_N)$ に対し，K 個の特異値のみを用いることで，文 i を K 次元空間にベクトル表現することができ，

$$\psi_i = [\sigma_1 v_{i1}, \sigma_2 v_{i2}, \dots, \sigma_K v_{iK}]^\top \quad (4)$$

となる．この K 次元空間に表現された文 i に対するベクトル ψ_i のノルムを，

$$\|\psi_i\| = \sqrt{\sum_{k=1}^{K} (\sigma_k v_{ik})^2} \quad (5)$$

で定義する．このノルムをその文のスコアと考え，要約率に達するまで，スコアの高い文から順に抽出する．

∗ A study on important sentence extraction methods using SVD for automatic speech summarization
By Makoto Hirohata, Yousuke Shinnaka, and Sadaoki Furui (Tokyo Institute of Technology)

表 1. 各手法による要約正解精度

	RDM	手法 1				手法 2 (K=1/K=2/K=4)			
		N-N	N-C	B-N	B-C	N-N	N-C	B-N	B-C
文境界 1	37.0	39.1	38.8	38.0	38.5	44.6/42.2/41.9	42.9/42.6/42.6	43.4/43.6/42.0	42.1/42.5/43.2
文境界 2	37.0	38.5	39.4	38.6	37.7	43.8/42.6/42.0	44.2/44.1/45.3	42.5/42.9/42.1	44.1/43.7/41.7

3 評価実験

3.1 要約実験条件

CSJ の男性話者 14 名，女性話者 6 名の計 20 名による講演音声に対し，文献 [4] の音声認識システムを用いて得られた音声認識結果について，50%の要約率で重要文抽出による自動要約を行った．音声認識結果の単語正解精度は約 76%であった．式 (1) での尤度計算は単語 trigram に基づいており，その学習には境界情報が入った 53 講演の書き起こしを用いた．ただし，式 (1) のパラメータ θ の値は実験的に求めた最適値を用いた．そのときの適合率/再現率は文境界 1 で約 78%/71%，文境界 2 では約 74%/60%であった．また，平均の文数は文境界 1 で 57.5 文，文境界 2 で 74.9 文であった．

局所重み関数と大域重み関数の N-N，N-C，B-N，B-C の各組合せに対し，手法 1，手法 2 を適用して重要文抽出を行った．また，重要文抽出手法の有効性を検証するため，文をランダムに抽出する手法 (RDM) の評価も行った．

要約文の評価は，3 名の被験者により重要文抽出され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 [5]．この手法では，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 (要約正解精度) を求める．ただし，この被験者の重要文抽出は，文境界 2 の単位に基づき区切られたものを対象としている．

3.2 評価結果

評価結果を表 1 に示す．いずれも要約率 50%で，20 名の講演音声に対して自動要約を行ったときの要約正解精度の平均値である．

手法 1 では，RDM に対して約 2%程度の改善しか得られなかった．実験に用いた講演音声は，本質的に潜在的な意味クラスの数が少ないと考えられ，k の数が大きくなるにつれて，特異値 σ_k に対応する文には，講演の話題とは関係の小さいものが多く含まれるようになる．そのため，手法 1 の要約精度の低下に繋がったと考えられる．

図 1，2 に手法 2 における次元数 K と要約精度の関係を示す．これを見ると，特に 1〜5 次元の空間に文を射影することで，良い結果が得られることが分かる．表 1 の手法 2 の結果は，次元数 K が 1，2，4 のときの結果である．手法 2 では，RDM に対しては約 6%の要約精度の改善が見られ，手法 1 と比較しても約 4%の改善が見られた．

また，手法 1，2 において，文境界の単位の違いや重み付け関数の組合せの違いによる要約正解精度の大きな違いは見られなかった．

4 まとめ

本稿では，音声自動要約において，特異値分解を用いた重要文抽出を行う手法について検討した．重要な話題に関する文を抽出するため，特異値分解によって，各文を低次元での空間上にベクトル表現し，そのノルムを用いる手法を提案した．また，単語ネットワークによる評価により，手法の有効性を示した．

今後の課題としては，認識誤りによる文抽出への影響を低減するため，認識結果に対する信頼度や言語スコアの利用 [1] があげられる．

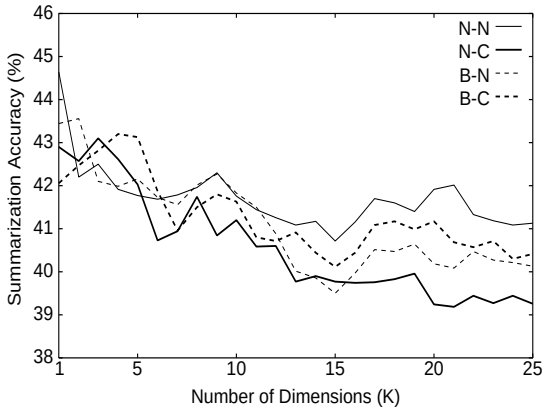

図 1. 手法 2 による要約正解精度 (文境界 1)

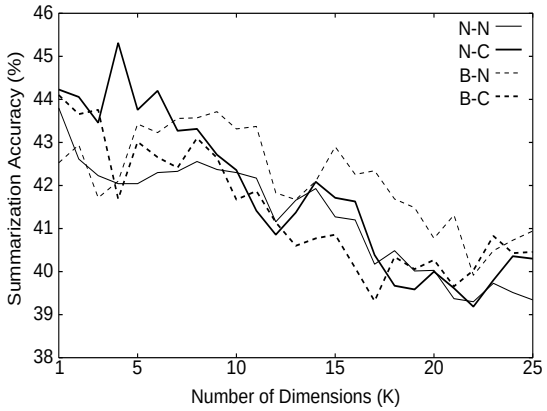

図 2. 手法 2 による要約正解精度 (文境界 2)

参 考 文 献

[1] 菊池智紀, 古井貞熙, 堀智織, “重要文抽出と文圧縮による音声自動要約,” 電子情報通信学会技術研究報告, NLC2002-81 / SP2002-158, pp.61-66, 2002.

[2] 北出祐, 南條浩輝, 河原達也, 奥乃博, “談話標識と話題語に基づく統計的尺度による講演からの重要文抽出,” 情報処理学会研究報告, NL-155-13 / SLP-46-2, pp.7-12, 2003.

[3] Y.Gong and X.Liu, “Generic text summarization using relevance measure and latent semantic analysis,” Proc. ACM SIGIR, pp.19-25, New Orleans, La, USA, 2001.

[4] T.Kawahara, H.Nanjo, T.Shinozaki, and S.Furui, “Benchmark test for speech recognition using the corpus of spontaneous Japanese,” Proc. SSPR, pp.135-138, 2003.

[5] 堀智織, 古井貞熙, “単語抽出による音声自動要約文生成法とその評価,” 電子情報通信学会論文誌, Vol.J85-D-II, No.2, pp.200-209, 2002.