論文 / 著書情報
Article / Book Information

論題(和文)	講演音声自動要約の試み
Title	Challenge to Automatic Lecture Speech Summarization
著者(和文)	堀智織, 古井 貞熙
Author	SADAOKI FURUI
出典(和文)	話し言葉の科学と工学ワークショップ予稿集, Vol. , No. , pp. 165-172
Journal/Book name	, Vol. , No. , pp. 165-172
発行日 / Issue date	2001, 3

講演音声の自動要約の試み

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，話し言葉音声の認識を目指して平成 11 年度に開始したプロジェクトで構築された話し言葉コーパスを用いて、講演音声を音声認識し自動要約を試みた結果を報告する．これまで，我々は音声自動要約手法として，音声認識された各発話文から相対的に重要な単語を，特定の要約率 (認識結果の文字数に対する要約文の文字数の割合) で抜き出し，それらを接合することによって要約文を生成する音声自動要約の枠組を提案してきた．この発話単位の自動要約手法は，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最適な自動要約文として動的計画法により決定する．要約スコアは，自動要約文に抽出された各単語の単語重要度 (重要度スコア)，単語連鎖の言語尤度 (言語スコア)，音声認識時における各単語の音響的，言語的信頼度 (信頼度スコア)，および原文の係り受け構造に基づく単語間遷移確率 (単語間遷移スコア) の累積スコアに基づき定義される．さらに，この発話単位要約手法を複数発話から構成される主題のある音声の要約に拡張し，全体として特定の要約率となるよう，各発話文を可変的な要約率で要約する複数発話自動要約手法を提案した．本稿では，この複数発話自動要約手法を用いて講演音声を自動要約し，自動要約文を被験者により主観評価を行うとともに，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き評価を行った結果を報告する．

キーワード　音声自動要約，単語重要度，言語尤度，信頼尺度，係り受け SCFG，動的計画法，複数発話自動要約，正解要約文単語ネットワーク

Challenge to Automatic Lecture Speech Summarization

Chiori HORI　and　Sadaoki FURUI

Tokyo Institute of Technology
Department of Computer Science
2-12-1　Ookayama, Meguro-ku, 152-8550 Japan
{chiori,furui}@furui.cs.titech.ac.jp

Abstract

This paper reports a new method and experiments for automatic summarization of lecture speech using the spontaneous speech corpus which has been constructed by our national project started in 1999. We have proposed a method of automatic speech summarization, applied sentence by sentence, in which a set of words maximizing a summarization score is extracted using a dynamic programming technique from automatically transcribed speech and concatenated according to a target compression ratio. The summarization score consists of a word significance measure, an acoustic and linguistic confidence measure, linguistic likelihood, and a word concatenation probability determined by a dependency structure in the original speech given by Stochastic Dependency Context Free Grammar (SDCFG). In this paper, the automatic summarization technique for each utterance is extended to a set of utterances with consistent meanings.

One lecture in the spontaneous speech corpus is summarized by our proposed speech summarization method and the summarized lecture is evaluated by the comparison with manual summarization by human subjects. Various manual summarization results are merged into a word network and used for calculating the accuracy of automatic summarization.

Key words　speech summarization, word significance measure, linguistic likelihood, confidence measure, stochastic dependency context free grammar, dynamic programming, multiple utterances, word network of summarization results

1 はじめに

近年，大語彙連続音声認識 (LVCSR) システムの進展に伴い，音声付き画像データへの自動字幕付与，講演や講義および会議等の音声データに対する講演録や議事録等の抄録自動生成，および情報検索のための自動インデクシングなど，LVCSR システムの種々の応用が検討されている．しかしながら，LVCSR システムの出力結果には，自然発話による冗長な情報や認識誤りによる不要な単語が含まれている．そのため，ユーザの要求に応じて音声データから必要な情報のみを抽出することが求められている．これまで我々は，図 1 に示すような，LVCSR システムの出力結果からユーザの要求に基づき自動要約を行い，情報を出力する音声自動要約システムを提案している [1][2][3]．

この音声自動要約システムに用いる要約手法として，LVCSR システムを用いて認識された各発話文から相対的に重要な単語を，特定の要約率 (認識結果の文字数に対する要約文の文字数の割合) で抜き出し，それらを接合することによって要約文を生成する音声自動要約の枠組を提案した．この発話単位の要約手法は，要約文に抽出された各単語の単語重要度 (重要度スコア) と単語連鎖の言語尤度 (言語スコア)[1][2]，および音声認識時における各単語の音響的，言語的信頼度 (信頼度スコア)[3]，および原文の係り受け構造に基づく単語間遷移確率 (単語間遷移スコア)[4] の累積スコアを要約スコアと定義し，これを最大とする部分単語列を動的計画法により決定する．

この発話単位の自動要約手法を用いてニュース音声を対象として自動要約を行った結果，相対的に重要な情報を包含し，日本語として尤もらしい要約文を生成することができた．この自動要約手法は，話し言葉の特徴である言い誤り，言い直しや言い淀みといった冗長な情報を除去しつつ，重要な情報を担う単語列を抽出する点で，重要箇所の抽出のような indicative な要約においても，抄録のような informative な要約においても必要不可欠な基礎技術になると考えられる．さらに，全体として特定の要約率となるよう，発話単位要約手法を複数発話から構成される主題のある音声の要約に拡張し，各発話文を可変的な要約率で要約した．これにより，各発話の重要度が考慮され，相対的に重要な情報を多く含む文の文長は相対的に長く，冗長な情報を多く含む文は短く要約されるないし削除される．この複数発話要約手法は，重要文抽出と発話単位の自動要約を組み合わせた要約手法ととらえることができる．本稿では，複数発話自動要約手法を用い，講演録作成を目的として講演音声の自動要約を試みる．

2 音声自動要約手法

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 Tr 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 $W = w_1, w_2, \ldots, w_N$ 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 $V = v_1, v_2, \ldots, v_M$ の要約スコアは次式によって示される．

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

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 $V = v_1, v_2, \ldots, v_M (M < N)$ 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 $\hat{V}$ をを求める問題となり，動的計画法を用いて解くことができる．さらに，同一の認識結果から生成された要約率の異なる要約文から最適な要約文を決定するため，単語数に基づく正規化要約スコアを定義した [1][2]．

【単語重要度スコア】

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．本研究では，単語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．ただし，重要な単語は主として名詞であるので，名詞以外の単語には一定値を与える．

【言語スコア】

言語スコア $L(v_m|\ldots v_{m-1})$ 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，統計的言語モデルである単語 trigram を用いる．

【信頼度スコア】

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を含む要約文候補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．

【単語間遷移スコア】

単語間遷移スコア $Tr(v_{m-1}, v_m)$ は，要約文内の単語連鎖が原文において係り受け関係にあるか否かを示す単語間遷移確率の対数値で定義され，係り受け関係にない単語連鎖にペナルティを与えるものである．以下この算出法について説明する．

単語間遷移規則

要約文内の単語連鎖は，図 2 に示す原文の係り受け構造に基づく単語間遷移規則に従うものと規定する．この単語間遷移規則では，文節内の最終実質語または最終機能語のみが文節境界を越え，かつ後続の文節内の任意の実質語にのみ遷移する．さらに，文節境界から次の文節境界への遷移を許すため，離れた文節間での単語遷移が可能である．但し，個々の文節は「1 個以上の実質語と 0 個以上の機能語からなる単語列」とする．

単語間遷移確率

文節内の単語間での遷移確率は，前方から後方の単語に係るという正規文法を定義し，係り受け関係のある単語間の遷移確率を 1，係り受け関係のない単語間の遷移確率を 0 とする．また，文節境界を越える単語間の遷移確率は，各単語が属する文節の係り受け構造に基づき単語間遷移を規定する．構文木が一意に定まる場合，文節間の係り受け関係の有無も一意に定まり，文節境界を越え

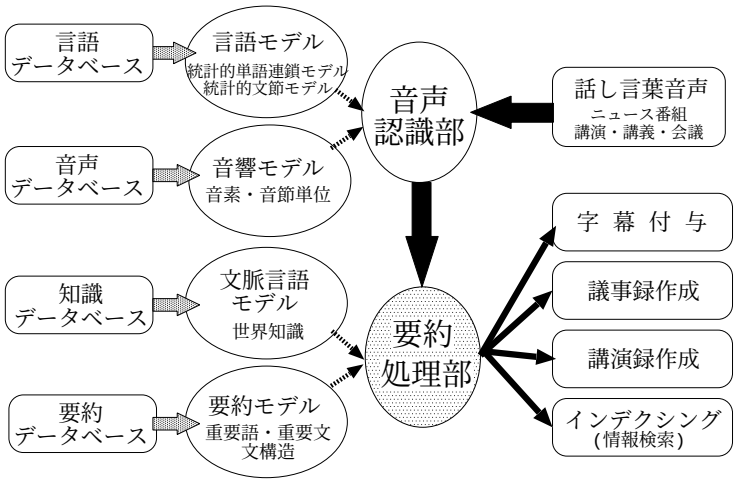

図 1: 音声自動要約システム

]

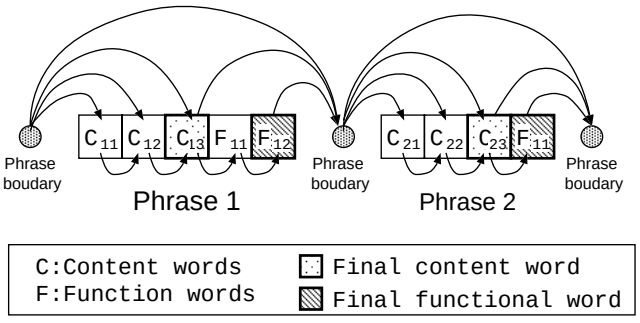

図 2: 単語間遷移規則

る単語間の遷移も文節内と同様，係り受け関係にある文節に属する単語間遷移確率は 1，係り受け関係に無い文節に属する単語間の遷移確率は 0 となる．しかし，構文木には曖昧性があるため，文節間の係り受け関係を確率的に推定し，文節境界を越える単語連鎖の単語間遷移確率として適用する．本研究では，係り受け SCFG の確率を単語間遷移確率として適用する．

係り受け SCFG に基づく単語間遷移スコアの定義

文節単位の係り受け SCFG[5] に基づき，文節境界を越える遷移を考慮した単語間遷移スコアを定義する．1 文は H 個の文節 $P_1, \ldots, P_H$ により構成される．

k 番目の単語 w_k と l 番目の単語 w_l の単語間遷移スコアは，w_k が文節 $P_{h(w_k)}$ に属し，w_l が $P_{h(w_l)}$ に属している際に，$h(w_k) = h(w_l)$ ならば文節内の遷移に基づく規則 $(R(w_k, w_l) = 0, 1)$ を用い，$h(w_k) < h(w_l)$ ならば 2 文節 $P_{h(w_k)}$,$P_{h(w_l)}$ が係り受けの関係にある確率 (の対数) を用いて，次式のように定義する．

$$Tr(w_k, w_l) = \begin{cases} \log \sum_{i=1}^{h(w_k)} \sum_{j=h(w_l)}^{H} \sum_{\alpha,\beta} g(\alpha \rightarrow \beta\alpha; i, h(w_k), j) & if\ h(w_k) < h(w_l) \\ \log R(w_k, w_l) & if\ h(w_k) = h(w_l) \end{cases} \quad (2)$$

ここで，α と β は，係り受け SCFG の非終端記号を表す．$g(\alpha \rightarrow \beta\alpha; i, h, j)$ は，図 3 に示すように，開始記号 S から対象となる文が生成された際，$\alpha \rightarrow \beta\alpha$ の規則が適用され，さらに β から $P_i \ldots P_h$ が生成され，α から $P_{h+1} \ldots P_j$ が生成される事後確率を表す．この事後確率は inside 確率と outside 確率を用いて推定する．

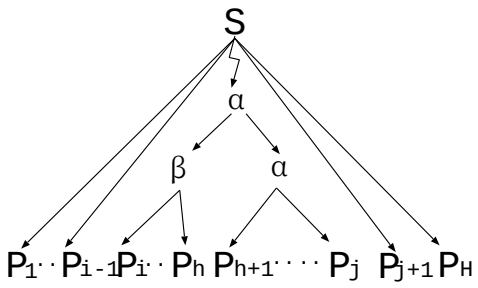

図 3: Inside-Outside 確率

3 複数発話を対象とした音声要約手法

発話単位の要約手法を用い，複数の文を一文として扱うことにより複数発話の要約手法に拡張する．但し，文境界の遷移では単語間遷移スコアを用いず，さらに前文の文末記号と後続する文の文頭記号を必ず遷移する

と規定する．要約スコアには単語重要度スコアが含まれているため，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．この手法は従来の重要文抽出に基づく要約手法と，これまで提案してきた発話単位の要約手法を統合した新しい要約手法であり，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．

J 個の発話文 $S_1, \ldots, S_J$ $(S_j = w_{j1}, w_{j2}, \ldots, w_{jN_j})$ より，$M(M < \sum_j N_j)$ 単語からなる部分単語列 $V = v_1, v_2, \ldots, v_M$ を，式 (1) で与えられる要約スコアが最大となるように決定するアルゴリズムを以下に示す．

(1) 記号と変数の定義

$s_j(k,l,n)$ ：１単語あたりの要約スコア

$$s_j(k,l,n) = \log P(w_{jn}|w_{jk}w_{jl}) + \lambda_I I(w_{jn}) + \lambda_C C(w_{jn}) + \lambda_T Tr(w_{jl}, w_{jn})$$

$P(w_{jn}\|w_{jk}w_{jl})$	：言語スコア
$I(w_{jn})$	：重要度スコア
$C(w_{jn})$	：信頼度スコア
$Tr(l,n)$	：単語遷移スコア
<s>	：文頭記号
</s>	：文末記号
$g_j(m,l,n)$	：局所最適スコア

(文１の先頭から文 j 内の単語列 w_{jl},w_{jn} で終る m 単語から成る部分単語列 <s>$, w_{11}, \ldots,$ w_{jl}, w_{jn} の要約スコア．但し，$(0 \le l < n \le N_j)$)

$G_j(m)$ ：文末の局所最適スコア
(文１の先頭から文 j の終端までの m 単語から成る部分単語列の要約スコア)

$b_j(m,l,n)$	：バックポインタ
$B_j(m)$	：文末のバックポインタ

(2) 初期設定

$$G_0(m) = \begin{cases} \log P(w_{jn}|\texttt{<s>}) + \lambda_I I(w_{jn}) + \lambda_C C(w_{jn}) \\ \qquad if\ 1 \le n \le N_j \ \ \& \ \ m = 0 \\ -\infty \qquad otherwise \end{cases}$$

$$B_0(m) = \phi$$

(3) 漸化式計算
for $j = 1$ to J
［文頭の計算］

$$g_j(m,0,n) = \begin{cases} G_{j-1}(m-1) + \log P(w_{jn}|\texttt{<s>}) \\ \quad + \lambda_I I(w_{jn}) + \lambda_C C(w_{jn}) \\ \qquad if\ 1 \le n \le N_j \\ -\infty \qquad otherwise \end{cases}$$

$$b_j(m,0,n) = \phi$$

［文内部の計算］
for $m = j \times 2$ to N_j
for $n = 2$ to N_j
for $l = 1$ to $n-1$

$$g_j(m,l,n) = \max_{0 \le k < l} \{g_j(m-1,k,l) + s_j(k,l,n)\}$$

$$b_j(m,l,n) = \mathrm{argmax}_{0 \le k < l} \{g_j(m-1,k,l) + s_j(k,l,n)\}$$

［文末の計算］

$$G_j(m) = \max_{\substack{0<n\le N_j \\ 0\le l\le N_j-1}} g_j(m,l,n) + \log P(\texttt{</s>}|w_{jl}w_{jn})$$

$$(\hat{n}, \hat{l}) = \mathrm{argmax}_{\substack{0<n\le N \\ 0\le l\le N-1}} g_j(m,l,n) + \log P(\texttt{</s>}|w_{jl}w_{jn})$$

$$B_j(m) = (\hat{n}, \hat{l})$$

(4) トレースバック
$j = J$
$m = M$
while $m > 0$
$v_m = w_{\hat{n}}$
$l' = b_j(m, \hat{l}, \hat{n})$
$\hat{n} = \hat{l}$
if $l' \neq \phi$ then
$\hat{l} = l'$
$m = m - 1$
else
$v_{m-1} =$ </s>
$v_{m-2} =$ <s>
$(\hat{n}, \hat{l}) = B_{j-1}(m-2)$
$m = m - 3$
$j = j - 1$

動的計画法の処理過程を図 4 に示す．

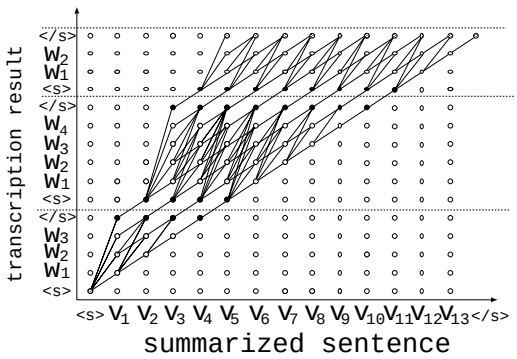

図 4: 複数発話の音声要約のための動的計画法の計算領域．

4　要約文の評価法

被験者よって作成された正解要約文を基準として，日本語としての適正度，原文の文意の保持という点から，自動要約の良さを定量的に評価する手法として，これまで単語連鎖適合率 [3] を用いてきた．この単語連鎖適合率は，被験者間で正解要約文が異なり，さらに被験者による正解要約文が全ての正解要約文を網羅していないということから，長さの異なる部分単語連鎖による適合率を用いることにより，全ての可能性のある正解要約文に対する評価を行おうとした尺度である．

一方，本稿で提案する正解要約文単語ネットワークは，被験者の作成した正解要約文を単語間の連鎖を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している．自動要約文は，このネットワーク上で最も近い要約文に対して単語正解精度により一元的に評価できる．図 5 に正解要約文単語ネットワークの例を示す．

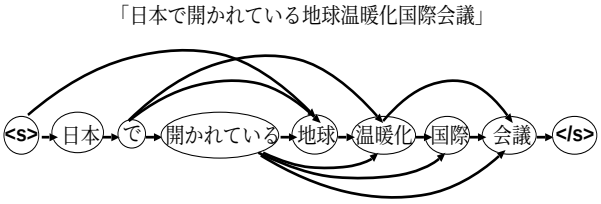

図 5: 正解要約文単語ネットワークの例

5　評価実験

5.1　実験条件

話し言葉コーパス中の男性話者 (AS99SEP097)1 名による講演音声の書き起こし文 (TRS) および大語彙連続音声認識システムによる音声認識結果 (REC) を提案手法を用いて 70–80%と 40–50%の要約率で自動要約を行った．

但し，自動要約を行う前に，話し言葉と書き言葉の言い回しの違いを吸収するため,「えーと」や「あのー」等のフィラー単語を削除し，丁寧語や話し言葉特有の砕けた表現は論説調の書き言葉へ変換する前処理を行った．具体的な例として,「ございます」という丁寧語は「ある」に,「んです」という砕けた表現は「のである」に変換した．

生成された自動要約文を，6 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．

5.2　音声認識システムの構成

【特徴抽出】
音声データを 16kHz，16bit でディジタル化し，フレーム長 24ms，フレーム周期 10ms で対数パワーと 12 次元のメルケプストラムおよび Δ メルケプストラム (計 25 次元) を抽出する．さらに発話毎にケプストラム平均正規化を行う．
【音響モデル】
話し言葉コーパス中の自動要約対象となる講演者以外の男性話者による 59 時間 (338 講演) の音声データを用いて，16 混合ガウスの不特定話者音素文脈依存 HMM(3000 状態) の音響モデルを作成した．
【言語モデル】
単語 bigram，trigram を用いる．音響モデルを作成した際に用いた音声データの書き起こし文を，形態素解析システム JTAG により形態素に分解し，約 1.5M 形態素を用いて語彙 20k の言語モデルの学習を行った．但し,「単語＋読み＋品詞」を形態素の単位とした．
【デコーダ】
単語グラフを中間表現とする 2 パスデコーダを用いる．第一パスでは HMM と bigram を用いて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．このとき，単語間の音素文脈依存も考慮する．

5.3　要約処理部の構成

【単語重要度スコア】
音声認識用言語モデルを学習した 1.5M 形態素の講演の書き起こしを用い，全文における各単語の出現頻度に基づき重要度スコアを計算した．
【要約言語モデル】
要約用言語モデルは要約文における単語連鎖をモデル化したものであるが，言語モデルを学習できる要約文の大規模なコーパスは存在していない．そのため，話し言葉コーパスの書き起こし文を，自動要約における前処理を施して論説調の表現に変換し，単語 trigram を学習した．
【文節単位の係り受け SCFG】
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，構文木制約付きの Inside-Outeside アルゴリズムを用いて，係り受け SCFG のパラメータの推定を行った．但し，非終端記号数は 100 とした．

5.4　評価結果

実験結果を表 1 に示す．表は，書き起こし文 (TRS) と音声認識結果 (REC) を要約率 70–80% および 40–50%で要約した際の，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している．但し，括弧内は要約率を示している．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，被験者各 6 人の被験者の正解要約文を他の 5 人の正解要約文で作成した要約文単語グラフに基づき評価した平均単語正解精度 (SUB) を示す．参考までに，ニュース音声を 20–30%に自動要約した結果も合わせて示す．

また，要約方法として，各発話文を目標の要約率にする方法と，複数発話をまとめて要約する方法を検討した．複数発話を扱う場合は，計算量の点から全講演をまとめて要約することが困難であったため，講演の先頭から決まった個数の連続した文を抜き出して要約し，続いて抜き出す文セットを一文ずつシフトさせながら要約を繰り返す手法を試みた．本実験では 10 文をセットした．このとき，一つの発話文に対して複数の要約結果が得られることがあるが，その場合は最も要約スコアの高い結果を採用した．但し，ニュース音声の場合は，ニュース

表 1: 自動要約結果の正解要約文単語 グラフに基づく単語正解精度

	要約率	要約単位	書き起こし文			認識結果	
			RDM	TRS	SUB	RDM	REC
講演音声	40–50%	各発話	33.2(48)	49.5(48)	82.7(49)	22.9(46)	30.2(46)
		複数発話	27.8(41)	46.3(41)		19.2(42)	17.6(42)
	70–80%	各発話	68.5(76)	80(76)	97.2(80)	42.8(75)	42.7(75)
		複数発話	64.2(73)	73.9(73)		38.9(70)	39.4(70)
ニュース音声	20–30%	複数発話	-	74.6	82.3	-	70.7

記事単位で複数発話要約を行っている．ニュース音声の音声認識結果に対する自動要約では，平均単語正解精度が 90%以上であるニュース記事対象としたが，本稿で用いた講演音声の平均単語正解精度約は 70%である．

講演音声の要約結果は，特に要約率が小さい場合にニュース音声との差が顕著である．これは，講演音声に現れる話し言葉特有の現象 (言い誤り，言い直し，曖昧な表現，文法を逸脱した表現) が，影響しているものと考えられる．音声認識結果を対象とした要約文では，さらに認識誤りにより性能の劣化が著しい．このため，書き起こし文を対象とした場合には，RDM に比べて有意に高い要約精度が得られているが，認識結果を対象とした場合には，要約率が 40%–50%の発話単位要約を除くと，今のところ RDM と同程度の精度しか得られていない．

表 2(次頁) に実験結果の例を示す.

6 まとめ

本稿では，講演音声の講演録作成や検索を目的として単語重要度 (重要度スコア)，単語連鎖の言語尤度 (言語スコア)，音声認識時における各単語の音響的，言語的信頼度 (信頼度スコア)，および原文の係り受け構造に基づく単語間遷移確率 (単語間遷移スコア) に基づき単語を抽出し連接することにより自動要約文を生成する手法を用いて，音声自動要約を試みた．今後，自然発話音声を対象として，相対的に重要な情報を抽出し，日本語として尤もらしく自動要約文を作成するためには，自然発話を対象とした音声認識性能の向上が必須であり，さらに，単語重要度，係受け関係の推定精度を向上させるためには，要約データベースの整備が不可欠である．

謝 辞

放送ニュースのデータベースを提供して下さった NHK 放送技術研究所に感謝致します．京大コーパスを提供して下さった京大言語メディア研究室に感謝致します．

参考文献

[1] 堀, 古井，情処学研報, 99-SLP-29, pp.103–108(1999). 信学技報, SP99-110, pp.103–108(1999).

[2] Hori and Furui, Proc. ICASSP2000, Istanbul, Vol.3, pp.1579–1582(2000).

[3] Hori and Furui, Proc. ICSLP2000, Baijing, Vol.4, pp.326–329(2000).

[4] 堀, 古井，信学技報, SP2000-95-116, pp.127–132(2000).

[5] 堀，加藤，伊藤，好田，信学論，Vol.J83-D-II，No.11, pp.2407–2417(2000).

表 2: 書き起こしと音声認識結果に対する自動要約結果

書き起こし文
えー パラ言語情報ということなんですが あ 簡単に最初にえー 復習をしておきたいと思います ま あのー こうやって あっ 話しておりますとそれは勿論 あの 言語的情報を伝えるということが一つの重要な目的 ん なんでありますが 同時にパラ言語情報そして非言語情報が伝わっておりますま この三分法は藤崎先生によるものでして えー パラ言語情報というのは 要は あの 意図的に制御できる話者がちゃんとコントロールして出してるんだけども言語情報と違って連続的に変化するから カテゴライズすることがやや難しいそういった情報 で ありますで そういったものが音声生成過程の中で畳み込まれて まー 一次元の音声波として実際は出ている訳ですで 従来我々の研究といたしましてはこの部分を まー 音響的な手法によって 分析していた訳ですがきょうの発表の眼目は少し音声生成過程を遡りましてえー 調音運動の次元 で えー 従来の結果と一致するようなか違いが 観測されるかどうかということを見ようそういうのがきょうの発表の目的眼目でございます で 従来どのような成果が得られまし得られてきているかということを簡単にまたこれも復習になりますが えー これは今回用いました後程説明いたします笹田がという文をえー 中立感心落胆そして疑いプラス反問というようなパラ言語情報を指定して えー ある話者が発音し分けたもの で ございます えー 具体的にどういうのかって言いますとえー 私が今ここで話者私ですんでやってみますとえー わ中立が笹田が感心が笹田が落胆が笹田が えー 疑いが笹田がというような形でありますえ 御覧のように えー F〇にも非常に大きな違いが観察されますしから デュレーションにも えー 非常に大きな違いが観察されます
書き起こし文自動要約結果
パラ言語情報ということなのだが最初にしたい やっているとそれは勿論言語的情報を伝えるということが一つ重要な目的なのだがパラ言 語情報非言語情報がいる この法は藤崎ものだがパラ言語情報というのは制御できる話者がコントロールしているの だが言語情報と違って連続的に変化するためカテゴライズすることがやや難しい情報である 音声生成過程の中で畳み込まれて一次元の音声実際は出ている 従来研究としてはこの部分を音響的な分析していたがきょうの発表の眼目は音声生成過程を 調音運動の結果と一致するような違いが観測されるかどうかということそれがきょうの発表の目的眼目である 従どのような成果が得られているかということを簡単にこれも復習になる これは後程説明する笹田がという文を中立感心落胆疑い反問というパラ言語情報を指定して話者が分けたものである 具体的にどういうものかと言うと私が今ここで話者私やって中立笹田感心笹田が落胆が笹田疑いが笹田いう形である F〇も非常に違いが観察されるデュレーションにも非常に違いが観察される
音声認識結果
えーパラ言語情報ということなんですが簡単に最初にえー復しゅうをしておきたいと思います まあのー声で話しておりますとそれは勿論あの言語的情報を伝えるということが一つの重要な目的がなりますが同時に パラ言語情報そして非言語情報が伝わったまず まこの三文ご藤崎先生によるものでしてえーは言語情報というのは今日はあのー意図的にそういうできる 話者がちゃんとコントロールした出してるんだけども言語情報と違って連続的に変化する肩ぐらいいーすることが や難しいそういったと答えます そういったものが温泉先生家庭の中で畳込まれてまー一次元の音素えーとして実際をやってる訳です で従来我々の研究で関しましてはこの部分をま音響的な情報によって分析していた訳ですが 表の発表の科目は少し音素えーせ下がってを遡るいましてえー調音運動の次元でえー従来のった結果と一致するような違いが 観測されるかどうかということを見ようそういうま表の発表の目的はもがございます で従来同様なそれがが得られました得られてきているかということを簡単にまたこれも復習になりますがえーこれは今回用いました あのちょうど説明いたします 刺さだかという文をえー注意する感心落胆そして疑いの三文とようなパラ言語情報してしてえある話者が発音してはけたものでございます えー具体的にどういうのかと言いますとえー私が今この話者が進んでいますとえー文字列がさ三番が関心がすその泊まったのは 落胆がそのそうだからえー疑いが三歳台がを有用かつあります えー御覧のようにえーFゼロにも非常に大きな違いが観察されますしがでしょうんでもえー非常に大きな違いが観察されます
音声認識結果の自動要約結果
パラ言語情報というなのが簡単に最初に復習 声でいるとそれは勿論言語的情報を伝えるということが一つの重要な目的がなるが同時にパラ言語情報非言語情報 この三文藤崎先生によるものは言語情報というのは今日は意図的に 話者がコントロールした言語情報と連続的に変化する肩ぐらいすることが難しいそういったと答える それらが温泉先生家庭の中で畳一次元の音素実際をして 従来我々の研究でしましてはこの部分を音響的な情報によって分析していた表の発表の科目は音素調音運動の次元で 従来の結果と一致するような違いが観測されるかどうかということを表の発表の目的はある 従来それがが得られているかということを簡単にまたこれも復習になる がこれは今回説明するだかという文を注意する感心落胆疑いの三文とようなパラ言語情報ある話者が発音してはけたものである 具体的にどういうのかと言うと私が今話者が進のでいると文字列が三番が関心がその泊まったのは落胆疑いが有用かつある 〇に非常に大きな違いが観察されるしが非常に大きな違いが観察される これは組織的な違い話者を越えて観察