

係り受け SCFG に基づく音声自動要約法の改善

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、係り受け SCFG(Stochastic Context Free Grammar) を用いた発話単位の音声自動要約手法の改善、および複数の発話から構成された主題のある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た複数発話自動要約手法を提案する。これまで我々は、音声認識結果から、文字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音声自動要約手法を提案している。本研究では、原文の文意に整合しない要約文の生成を防ぐため、原文における係り受け構造に基づき、要約文の単語連鎖を制約する単語間遷移スコアを要約スコアに組み込む手法を新たに提案する。単語間遷移スコアとして係り受け SCFG に基づく単語間遷移確率を用いる。さらに、提案する発話文単位の要約手法を複数発話で構成された主題のある音声を要約に拡張し、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。

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

Improvements in Automatic Speech Summarization Based on Stochastic Dependency Context Free Grammar

Chiori HORI and Sadaoki FURUI

Tokyo Institute of Technology

2-12-1 Ookayama, Meguro-ku, 152-8550 Japan

Abstract This paper proposes an improved method of summarizing speech, sentence by sentence, in which a word transition probability determined by a dependency structure in original sentences given by Stochastic Dependency Context Free Grammar is incorporated in the summarization score. This paper also proposes a new method for summarizing a set of utterances with a single topic using an extended summarization technique. In our previous speech summarization method, a set of words were extracted from an automatically transcribed sentence according to a target compression ratio so that a summarization score consisting of a significance measure and a confidence score of each word and a linguistic likelihood of word concatenation is maximized using a dynamic programming (DP) technique. In order to alleviate the meaning alteration caused by word strings, we propose a new method in which word transition score determined by a probability of dependency in original sentences based on Stochastic Dependency Context Free Grammar is incorporated in the summarization score in our new method.

key words speech summarization, word significance measure, linguistic likelihood, confidence measure, stochastic dependency context free grammar, dynamic programming

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 [2][3][5]。

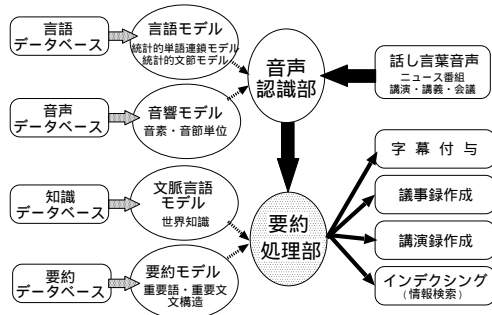


図 1: 音声自動要約システム

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][3]、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)[4][5]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。

しかし、この手法では原文における単語の係り受け構造が考慮されていないため、原文の文意に整合しない自動要約文が生成される可能性がある。この問題を改善すべく、文節重要度と文節間係り受け整合度の総和が最大になる部分文節列を原文から選択する手法が提案されている [6]。しかし、この方法は文節を単位としているため、助詞などの機能語が誤認識される傾向のある音声認識の結果を要約する手法として、必ずしも適切ではない。本研究では、原文の文意を保持するため、原文における係り受け構造に基づき、要約文内の単語連鎖を制約する単語間遷移規則と、係り受け SCFG を用いた単語間遷移スコアを要約スコアに組み込む手法を新たに提案する。

さらに、提案する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。この手法は、講演録等の抄録を作成する際、段落単位での音声自動要約に応用することが可能である。

2 発話単位の音声自動要約手法

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、新たに提案する単語遷移スコアに基づき、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。

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

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 $V = v_1, v_2, \dots, v_M$ ($M < N$) 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 [2][3]。

【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話題語らしさを示す話題語スコアを適用する。話題語スコアは名詞のみに定義し、名詞以外の単語には一定値を与えるものとする。

【言語スコア】

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、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。

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

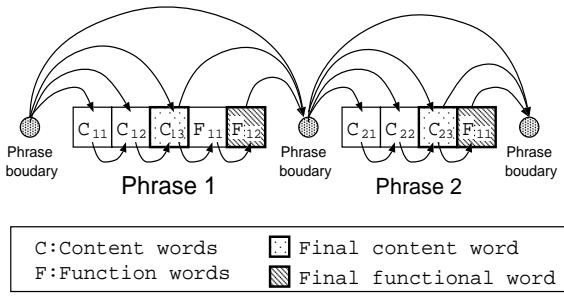


図 2: 単語間遷移規則

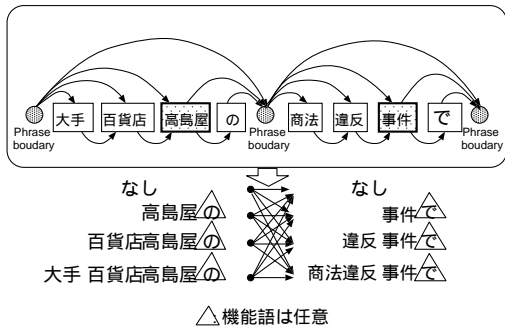


図 3: 単語間遷移規則に基づく要約例

0 個以上の機能語からなる単語列」とする。単語間遷移規則に基づく要約例を図 3 に示す。

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 [7] 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, \dots, 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。

$$Tr(w_k, w_l)$$

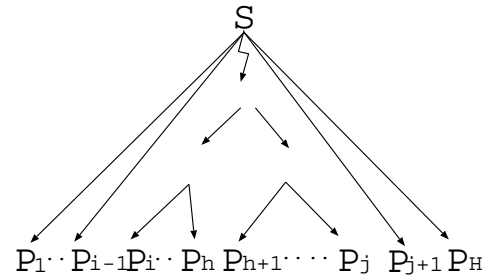


図 4: Inside-Outside 確率

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

ここで、 α と β は、係り受け SCFG の非終端記号を表す。 $g(\alpha \rightarrow \beta \alpha; i, h(w_k), j)$ は、図 4 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 $P_i \dots P_h$ が生成され、 α から $P_{h+1} \dots P_j$ 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。

3 自動要約法に対する評価尺度

人間によって作成された正解要約文に関して、日本語としての適正度、原文の文意の保持という点から、要約の良さを定量的に評価する [8]。

単語連鎖の適合率

日本語としての尤もらしさと原文の文意の保持を評価するため、自動要約文内に含まれる特定の長さの単語連鎖が正解要約文に含まれるか否かを評価する。自動要約文内に含まれる特定の長さの各単語連鎖が、被験者の作成した正解要約文のいずれかに含まれる割合を単語連鎖適合率とする。但し、原文の異なる位置に出現する単語は、たとえ表記が同一であっても別の単語として扱う。

4 発話単位自動要約手法に関する評価実験

4.1 実験条件

放送ニュース音声 419 文を大語彙連続音声認識システムを用いて音声認識し、得られた認識結果を用いて提案する発話単位自動要約手法により 20%, 40%, 60%, 70%, 80% の 5 段階の要約率で要約文を生成した。評価データとして、語彙サイズ 20000 で未知語率 1.7%、単語正解精度 90% 以上の放送ニュース文 (計 50 文) を用いる。

音声認識結果を自動要約する際、係り受け SCFG に基づく単語間遷移スコアを用いた自動要約文 (SCFG_REC) と用いない自動要約文 (NOSCFG_REC) の 2 種類を生成し、係り受け SCFG に基づく単語間遷移スコアによる効果を評価した。比較のため、書き起こし文に対しても同様に 2 種類の自動要約を行った (SCFG_TRS, NOSCFG_TRS)。

各自動要約文は、25人の被験者によってニュース音声の書き起こし文を用いて作成された正解要約文 (SUB_TRS) に基づき、単語連鎖適合率により評価した。各被験者の要約文 (SUB_TRS) をその他の24人の被験者の要約文を正解として評価し、自動要約文の目標値とした。また、被験者10人が60%~70%の範囲で音声認識結果を用いて要約した要約文 (SUB_REC) も、同様に SUB_TRS に対する評価を行った。要約率60%~70%の範囲で書き起こし文および音声認識結果から、単語をランダムに抽出した単語連鎖適合率 (RDM_TRS, RDM_REC) を比較のため示す。

さらに、被験者による音声認識結果に対する要約文 (SUB_REC) を正解として、認識結果に対する自動要約文の性能の評価を行った。各被験者の要約文 (SUB_REC) をその他の9人の被験者の要約文を正解として評価し、自動要約文の目標値とした。

4.2 音声認識システムの構成

【特徴抽出】

音声データを16kHz, 16bitでデジタル化し、フレーム長32ms, フレーム周期8msで対数パワーと12次元のメルケプストラムおよびメルケプストラム(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】

不特定話者音素文脈依存HMM(1012状態)、混合ガウス分布混合のIPAの女性モデルを用い、評価話者の985発話(約2時間)を用いて最尤推定話者適応を行った。

【言語モデル】

単語 bigram, trigram を用いる。放送ニュース原稿テキスト5年分(1992年7月から1996年5月)の約50万文を、形態素解析システムJUMANを用いて形態素に分解し、「単語+読み+品詞」を一つの形態素として bigram と trigram を学習した。

【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。

4.3 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放送ニュース原稿テキスト5年分(1992年7月から1996年5月)の約50万文を用い、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求める。

【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。そこで、要約対象であるニュース音声と話題が重なっており、要約文に求められる簡潔な表現が多く含まれている新聞テキストを用いて要約用言語モデルの学習を行った。要約用言語モデルとして3年分の毎日新聞(1996-1998)のテキストを用い、音声認識に用いた言語モデルの2000語彙を用いて単語 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-Outside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120とした。

4.4 評価結果

実験結果を表1に示す。要約率70%における、SUB_TRSとSUB_RECのそれぞれを正解とした場合の単語連鎖

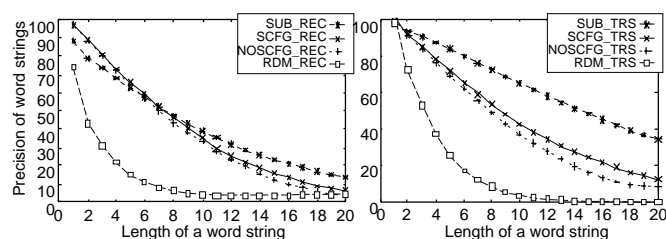


図5: 要約率70%時のSUB_TRSを正解とした場合の単語連鎖適合率と単語連鎖長の関係

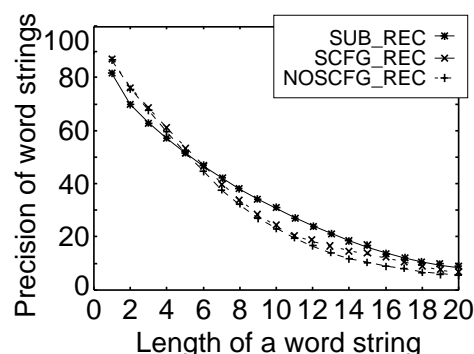


図6: 要約率70%時のSUB_RECを正解とした場合の単語連鎖適合率と単語連鎖長の関係

適合率と単語連鎖数の関係を図5と図6に示す。

図5より、音声認識結果、書き起こし文のどちらにおいても、SCFGに基づく単語遷移スコアを用いた手法が、単語連鎖適合率が高く、原文の文意をより保持することができたと考えられる。さらに、被験者が作成したにも関わらず、音声認識結果に対する要約文 (SUB_REC) は、SCFGを用いた書き起こし文に対する要約文 (SCFG_TRS) よりも単語連鎖適合率が低く、認識誤りを含む認識結果を要約する難しさが示されている。これは、被験者による要約文作成の際、被験者が認識誤りを検出することができないという問題と、重要単語が誤認識されている場合、重要箇所を抽出できないという問題によるものである。

一方、図6では、認識結果を用いた自動要約文が被験者による要約文と同等の値を示しており、提案手法が有効であることがわかる。

5 複数発話を対象とする音声自動要約手法

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

表 1: 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	ジュネーブで開かれている地球温暖化対策の国際会議で日本政府は西暦二千年以降先進各国が GDP 国内総生産に合わせた二酸化炭素の排出削減に努めるという新たな国際目標を日本としては今回初めて提案することを決めました
要約率 80%	ジュネーブで開かれている地球温暖化対策の国際会議で日本政府は 先進各国が 国内総生産に 二酸化炭素の排出削減 目標を日本としては今回初めて提案することを決めました
要約率 70%	ジュネーブで開かれている地球温暖化対策の国際会議で日本政府は 先進各国が 二酸化炭素の排出削減 目標を日本として 提案することを決めました
要約率 60%	ジュネーブで開かれている地球温暖化対策の国際会議で日本政府は 目標を日本として 提案することを決めました
要約率 40%	地球温暖化対策 会議で日本政府は 目標を 提案することを決め
要約率 20%	二酸化炭素の排出削減 目標を 提案すること
音声認識結果	年 で開かれている月いう温暖化対策の国際会議で日本政府は西暦二千年以降先進各国が GDP 国内総生産に合わせた二酸化炭素の排出削減に努めるという新たな国際目標を日本としては今回初めて提案することを決めました
要約率 80%	開かれている 温暖化対策の国際会議で日本政府は 先進各国が 二酸化炭素の排出削減に努めるとい新な国際目標を日本としては今回初めて提案することを決めました
要約率 70%	温暖化対策の国際会議で日本政府は 二千年以降先進各国が 国内総生産に 二酸化炭素の排出削減 目標を日本としては今回初めて提案することを決めました
要約率 60%	温暖化対策の国際会議で日本政府は 先進各国が 二酸化炭素の排出削減 目標を日本としては今回初めて提案することを決めました
要約率 40%	温暖化対策の国際会議で 各国が 二酸化炭素の排出削減 目標を 提案することを決めました
要約率 20%	二酸化炭素の排出削減 目標

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。

5.1 動的計画法による音声要約

J 個の発話文 S_1, \dots, S_J ($S_j = w_{j1}, w_{j2}, \dots, w_{jN_j}$) より、 M ($M < \sum_j N_j$) 単語からなる部分単語列 $V = v_1, v_2, \dots, v_M$ 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
 $\langle s \rangle$: 文頭記号
 $\langle /s \rangle$: 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 $\langle s \rangle, w_{11}, \dots, w_{jl}, w_{jn}$ の要約スコア。但し、 $(0 \leq l < n \leq N_j)$)

$G_j(m)$: 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

$$G_0(m) = \begin{cases} \log P(w_{jn} | \langle s \rangle) + \lambda_I I(w_{jn}) + \lambda_C C(w_{jn}) & \text{if } 1 \leq n \leq N_j \text{ \& } m = 0 \\ -\infty & \text{otherwise} \end{cases}$$

$$B_0(m) = \phi$$

(3) 漸化式計算

for $j = 1$ to J

[文頭の計算]

$$g_j(m, 0, n) = \begin{cases} G_{j-1}(m-1) + \log P(w_{jn} | \langle s \rangle) + \lambda_I I(w_{jn}) + \lambda_C C(w_{jn}) & \text{if } 1 \leq n \leq N_j \\ -\infty & \text{otherwise} \end{cases}$$

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

[文内部の計算]

for $m = j \times 2$ to N_j

for $n = 2$ to N_j

for $l = 1$ to $n-1$

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

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

[文末の計算]

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

$$(\hat{n}, \hat{l}) = \operatorname{argmax}_{\substack{0 < n \leq N \\ 0 \leq l \leq N - 1}} g_j(m, l, n) + \log P(\langle /s \rangle | w_{jl} w_{jn})$$

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

(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

$\hat{l} = l'$

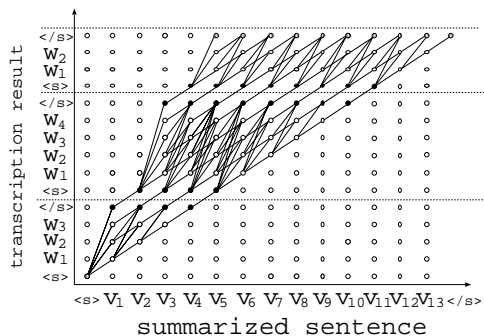


図 7: 複数発話の音声要約のための動的計画法の計算領域.

```

m = m - 1
else
  vm-1 = </s>
  vm-2 = <s>
  (n̂, l̂) = Bj-1(m - 2)
  m = m - 3
  j = j - 1

```

動的計画法の処理過程を図 7 に示す.

5.2 複数発話の音声自動要約実験および実験結果

発話単位自動要約に用いたニュース音声の認識結果から、平均単語誤り率 90% 以上の平均 5 発話で構成された 5 つのニュースの書き起こし文を、提案手法により 20-30% の要約率で要約を行い、12 人の被験者の作成した要約文に基づき単語連鎖適合率で評価を行った。比較のため、SCFG に基づく単語遷移スコアを適用しない要約文 (NOSCFG) も生成了。被験者の要約文をその他の 11 人の要約文を正解として評価した SUB_ART を目標値とした。以下に実験結果の例を示す。

書き起こし文

次は大雨です。きょうの日本列島は大気の状態が不安定になったのに加えて九州では梅雨前線の活動が活発になった為九州や東北地方を中心に各地で激しい雨が降りました。また雷を伴ってひょうが降ったところもありまして各地で被害が出ました。記録的な豪雨に見舞われた九州の様子を福岡からまたひょうの被害も出た東北地方の様子を仙台からお伝えします。今回の豪雨は上空に強い寒気が入ってきたため大気の状態が非常に不安定になって短い時間に激しい雨がゲリラ的に降ったのが特徴です。九州では日豊線が今も宮崎県内の一部の区間で不通になっています。また九州自動車道と宮崎自動車道も一部区間で通行止めになっています。一方東北地方の J R でも秋田県などの一部の区間で不通になっています。気象庁によりますとこの強い雨は今夜半すぎまで続く見込みで九州と四国地方それに近畿から。東北地方にかけては一時間に四十ミリから六十ミリの雷をともなった激しい雨が降る恐れがあり九州や関東地方ではところによって八十ミリを超える恐れもあります。これから今夜半すぎまでの雨量は多いところでさらに百ミリから百五十ミリに達する見込みです。

自動要約文

大気の状態が不安定に九州では東北地方を中心に各地で雨が降り。雷をひょうところもあり。豪雨に見舞われたひょう。不安定に。宮崎県内の一部区間で不通になっています。九州自動車道通行止めになっています。東北地方の一部区間で不通になっています。夜半ともなった激しい雨が降る恐れがあります。夜半。

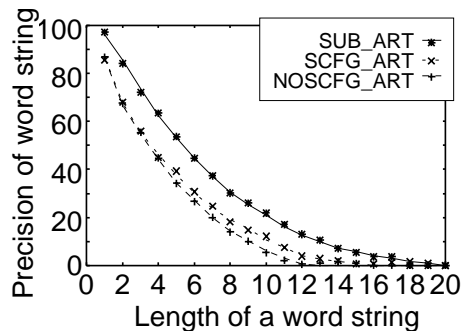


図 8: 主題のある複数発話要約文の単語連鎖適合率と単語連鎖長の関係

5.3 単語連鎖適合率による結果

単語連鎖適合率による評価結果を図 8 に示す。自動要約文 (SCFG_ART) は被験者の作成した要約文 (SUB_ART) には達していないが、要約結果の例に示すように日本語として尤もらしく、適宜重要な情報を抽出することができている。今後、音声認識結果を要約し評価を行う予定である。さらに、複数文の自動要約結果を評価するため、情報検索等のタスクによる評価を行いたい。

6 まとめ

本稿では、係り受け SCFG (Stochastic Context Free Grammar) を用いた発話単位の音声自動要約手法の改善、および複数の発話から構成された主題のある音声に要約する手法として、発話単位の自動要約手法を拡張した複数発話自動要約手法を提案した。本手法は、原文における係り受け構造に基づき、要約文の単語連鎖を制約する単語間遷移スコアを要約スコアに組み込むことにより、原文の文意をより正確に保持した要約文が生成可能である。さらに、提案手法が音声認識結果を用いた被験者による要約文と同等の性能を示した。複数発話自動要約手法は予備的な段階ではあるが、発話単位の自動要約手法への拡張の可能性を示した。

謝辞

放送ニュースのデータベースを提供して下さいました 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いまし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, 音講論, Vol.1, 3-1-11(1999).
- [2] 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).
- [3] Hori and Furui, Proc. ICASSP2000, Istanbul, Vol.3, pp.1579-1582(2000).
- [4] 堀, 古井, 音学論, Vol.1, 2-8-18, pp.63-64(2000).
- [5] Hori and Furui, Proc. ICSLP2000, Baijing, Vol.4, pp.326-329(2000).
- [6] 小黒, 尾関, 張, 高木, 言語処理学会第 6 回年次大会発表論文集, pp.133-136(1999).
- [7] 堀, 加藤, 伊藤, 好田, 信学論, Vol.J83-D-II, No.11, P2407-2417(2000).
- [8] 堀, 古井, 音講論, Vol.1, 2-8-18, pp.63-64(2000).