

要約筆記品質向上支援システム

Quality Improvement support function in Quality Evaluation System of Summary Transcript

高尾 哲康†
Tetsuyasu Takao

1. はじめに

聴覚障害者や高齢者への情報保障手段である要約筆記には「PC 要約筆記」と「手書き要約筆記」があり、いずれも要約筆記者が講演や番組などを聞き取り、リアルタイムで要約を行ない、キーボードや手書きで入力する。一般に日本語の発話速度は 200～400 文字/分であり、要約筆記者による入力量は PC の場合で 100～200 文字/分、手書きの場合で 40～80 文字/分となっている。要約筆記者は「速く」、「正確に」、「読みやすく」の 3 原則をもとに、技術の向上を目指してさまざまな研修プログラムで訓練を重ねる。個々の研修プログラムでは要約筆記の品質の尺度として、要約筆記利用者からのフィードバックや意見・要望を受けることが多い[1]。これらのフィードバックは個々の事例として受けることが多く、定量的な品質評価を受けることはほとんどなかった。そのため、長期間の研修を経ても要約筆記の品質向上の実感が得られにくくなっていた。これまで筆者らは講演者の発話内容のテキストと要約筆記者が入力したテキストをもとに定量的な評価ができるシステムを試作した[2][3]。要約評価計算には重み付き編集距離単位(主に形態素基本形と品詞)列の編集距離(Levenshtein Distance)計算に基づく方式を提案した。これにより個々の文や段落など局所的範囲での評価に有効性を確認できた。さらに、要約筆記されたテキストどうしの関連性や要約筆記者それぞれのタイプやくせを見つけ出すために多次元尺度法や潜在意味解析などの全体的評価も行なえるようにした。本論文では、要約評価計算の過程で得られた情報をもとに要約筆記者が書き下したテキストについて、よりよい要約テキストとなるような書き換え候補を提示・確認する機能とともに、システムの未熟性のために評価できなかった部分からよりよい要約表現を抽出する機能を紹介する。これにより、要約筆記者がよりよい要約ができるように支援することを目指している。

2. 品質評価に利用した要約筆記データ

要約筆記研修プログラムで使用した発話テキスト(T と表わす)と要約筆記者 8 名(手書き 4 名、PC 4 名、それぞれ、H1～H4、P1～P4 で表わす)でリアルタイム要約筆記したテキストを利用した。詳細を表 1 に示す。発話テキストには観光ガイド(約 4 分)を利用した。文字数には句読点や記号(矢印記号「→」、項目を表わす中黒「・」)、繰り返し記号(「〃」など)、削除記号(訂正線)などを含めている。手書き要約筆記の場合は PC による要約筆記とは異なり、二次元的な表現や複数行にわたる括弧記号や行を越えた矢印記号、横方向の波括弧付きの挿入文字などが含まれていることがある。その場合には、計算機可読テキストにする際にはその意図がわかるように XML タグ付きテキストにした。

3. 要約筆記品質評価システム

本システムはテキストアライメントモジュールと品質評価計算モジュールから構成される(図 1)。テキストアライメントモジュールは発話テキストと要約筆記テキストを入力とし、統計情報と言語情報をもとに、動的計画法を利用して対応する文や段落を関連づけるモジュールである(m 文対 n 文の対応付け)。アラインメント単位ごとに発話テキストと要約筆記テキストのペアが作成される。これにより品質評価計算対象範囲を狭くすることにより、後段の品質評価計算モジュールにおける評価計算精度を高めることができる。品質評価計算モジュールは、表記のゆれ(漢字の読みひらがな・カタカナ表記など)や要約筆記特有の省略表現などを吸収して正規化した形態素解析結果の形態素列に対し(形態素解析ツール MeCab を利用)、単語コスト、品詞コスト、単語間接続コスト、重複出現コスト(出現のたびに単調減少)を統計処理することにより、要約の品質評価(要約評価)の計算を行なう[2][3]。さらにテキスト全体およびアラインメント単位で形態素列の N-gram に基づいた多次元尺度法と潜在意味解析を行なう。多次元尺度法は、個体間の類似度をそれぞれの個体から得られる多変量の数値をもとに個体間距離を計算し、類似性の高いものが近くに配置するようにする方式である。潜在意味解析は単語の持つ多義性と多様性に対処した統計的技法であり、多変量ゆえに高次元となる単語文書行列を低次元に圧縮して近似的に表現する。これにより関連性の強い単語どうしが強調され、文書間の意味的類似度が明確になる。この処理には、統計解析用ソフト R とそのライブラリ群を利用した。

表 1. 要約筆記テキストと要約評価の向上

	筆記者	文字数	入力速度 文字数/分	要約率 文字数比 (%)	要約評価の変化					
					要約筆記 原データ	各単位最大 1箇所修正	各単位最大 2箇所修正	各単位最大 3箇所修正	向上分	
発話	T1	972	249.2							
	H1	286	73.3	29.4%	0.4404	0.5753	25	0.6981	47	0.2577
	H2	238	61.0	24.5%	0.3318	0.5618	27	0.6792	50	0.3474
	H3	238	61.0	24.5%	0.3529	0.5481	26	0.6625	48	0.3096
手書き	H4	297	76.2	30.6%	0.3966	0.5737	27	0.7160	51	0.3194
	P1	533	136.7	54.8%	0.6432	0.7743	22	0.8281	36	0.1849
	P2	492	126.2	50.6%	0.5923	0.7384	26	0.8139	44	0.2216
	P3	370	94.9	38.1%	0.4849	0.6847	27	0.7968	49	0.3119
PC	P4	497	127.4	51.1%	0.6101	0.7216	23	0.8020	42	0.1919
	手書き平均	264.8	67.9	27.2%	0.3804	0.5647	26.3	0.6890	49.0	0.3085
PC平均		473	121.3	48.7%	0.5826	0.7298	24.5	0.8102	42.8	0.2276

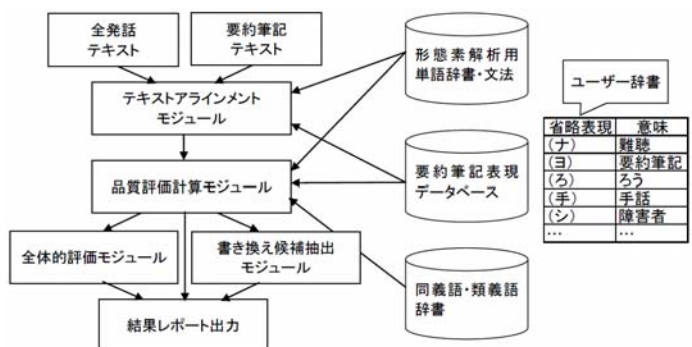


図 1. システム構成図

† 富山国際大学現代社会学部

要約評価計算には、発話と要約筆記それぞれの形態素列について、各形態素を 0~1 の重み付き編集単位要素とみなして編集距離を求めることで求められるものとする。編集距離とは列 A と列 B について、A を編集操作(削除、挿入、置換)して B にするときの必要最低限の操作数のことである。要約評価は編集操作コストを発話と要約筆記の形態素コスト値の総数で割り、数値の範囲を 0~1 に正規化した数値にした。0 に近ければ発話との相違が多く、1 に近ければ相違が少なくなる。表 2 に要約評価計算の例を示す。表において、行方向は発話テキストの各形態素の並び、列方向は要約筆記の各形態素の並びである。コスト値は、形態素解析用単語辞書に格納されている形態素コストを初期値とし、発話全体で繰り返し出現するごとに形態素コストが順次単調減少するようにしている。各セル値 E_{ij} の計算は表 2 の式にて全セルについて計算を行ない、表の最右下のセル値を 1 から引いた値がテキスト全体の要約評価となり、この値が 1 に近いほどよい要約になる。

要約候補の抽出は次のように行なう。表 2 の要約評価を算出するマトリクスにおいて、最右下のセルから最左上のセルまで評価値が最も小さくなる方向(上方、左方、左上方のいずれか)に順次たどることで発話と要約筆記との対応セルを求める。次に、発話と要約筆記との対応関係のうち相互にマッチしないもの(前後のセル間で評価値の差が大きい場合)を抽出する。表 2 の例では、

- ・「急性中耳炎は」(発話)と文頭(要約筆記)
- ・「で起こる。」(発話)と「で」(要約筆記)

が該当する。この際、直前のセルの評価値との差が大きいもの、発話内の形態素コスト値の大きいものから優先的に要約候補を提示する。なお、要約候補の提示における修正箇所は自立語を含む文節単位(文節境界判定は構文解析ツール CaboCha を利用)とし、文章としての形態素間のつながりが保たれるようにした。

4. 実験結果

筆者者ごとの要約筆記テキストを本システムに適用した結果を表 1 右側に示す。各筆記者が書き下した要約筆記テキスト(要約筆記原データ)の要約評価とともに、アラインメント単位ごとの要約候補の提示にしたがい、1 箇所および 2 箇所の修正をほどこした場合の要約評価の変化を図 2 に示す。要約筆記原データの要約評価が低い場合(要約筆記に十分習熟していない場合)ほど向上分が大きく、効果があることがわかる。向上分が大きい H2、P3 について、改良による効果を多次元尺度法および潜在意味解析にて図 3 に示した(H21 は各単位 1 箇所、H22 は各単位 2 箇所の修正後)。システムによる要約評価計算がうまくいかなかった部分については、人間固有の要約がされていると考え、アラインメント単位ごとに一定以上の要約筆記文字数にもかかわらず要約評価が低くなってしまった部分を抽出することにした。以下に例(発話→要約筆記)を示す。

「この像が高村光雲によって建てられたのは明治 31 年」→「明治 31 年高村こうん立像」

「この門から多くの大名が城にのぼりました」→「大名の多くはこの門から登城」

「大手門は関東大震災の折にくずれました」→「関東大しんさいで大手門倒壊」

「テレビ放送が始まりました」→「TV 放映スタート」

5. まとめ

本実験から要約筆記支援のための機能(要約候補提示、候補を利用した場合の評価)の効果が確認できた。でシステムの要約候補提示機能の効果が確認できた。今後は、さまざまな要約筆記データを収集し、要約評価精度の向上や失敗箇所についての分析を進めるとともに、コスト計算手法やパラメータの最適化などを行なっていく。今回は発話テキストという教師データがある場合について考察したが、実際に要約筆記する場合においては複数人での要約筆記テキストをマージしてよりよい要約筆記テキストを構築するなどの応用も考えられる。

参考文献

- [1] 話しことばの要約、三宅初穂、全国要約筆記問題研究会 (2012)
- [2] 高尾哲康、要約筆記品質評価システムの改良、FIT2011、3Q-5、(2011)
- [3] 高尾哲康、要約筆記品質評価システムにおける要約候補文提示機能、FIT2012、2M-6、(2012)

表 2. 評価値計算と要約候補抽出

j \ i	コスト値	急性	中耳炎	は	風邪	が	原因	で	起こる	。
		3135	4000	10	2914	10	2326	10	2206	188
急性	0.000	0.156	0.356	0.356	0.502	0.502	0.618	0.618	0.728	0.738
中耳炎	0.145	0.302	0.501	0.502	0.356	0.357	0.473	0.473	0.583	0.593
は	0.146	0.302	0.502	0.502	0.357	0.356	0.472	0.473	0.583	0.592
風邪	0.262	0.418	0.617	0.618	0.473	0.472	0.356	0.357	0.467	0.476
が	0.262	0.419	0.618	0.618	0.473	0.473	0.357	0.356	0.466	0.476
原因	0.262	0.418	0.617	0.618	0.473	0.472	0.356	0.357	0.467	0.476
で	0.262	0.419	0.618	0.618	0.473	0.473	0.357	0.356	0.466	0.476
起こる	0.262	0.418	0.617	0.618	0.473	0.472	0.356	0.357	0.467	0.476
。	0.262	0.418	0.617	0.618	0.473	0.472	0.356	0.357	0.467	0.476

$$E_{ij} = \min(E_{i-1,j} + C_{i-1}/C, E_{i,j-1} + C_{j-1}/C, E_{i-1,j-1} + A)$$

$$A = \begin{cases} 0 & : i-1 \text{ と } j-1 \text{ の位置の形態素がマッチ} \\ & \text{(表記基本形、品詞、同義語)した場合} \\ (C_{i-1} + C_{j-1})/C & : \text{上記以外 (C:コスト値の総和)} \end{cases}$$

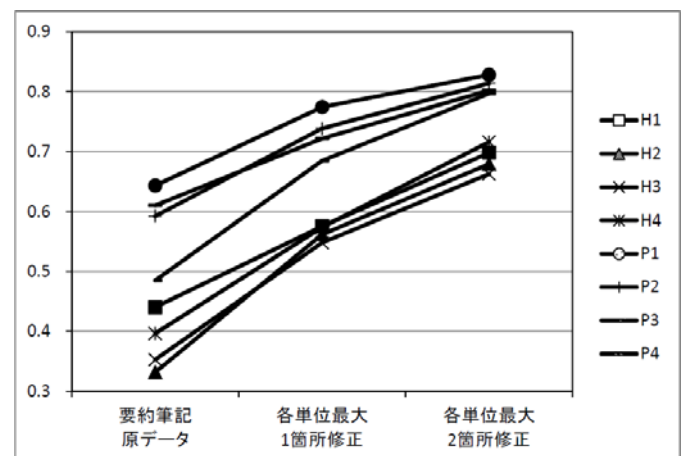


図 2. 要約候補選択時の要約評価の変化

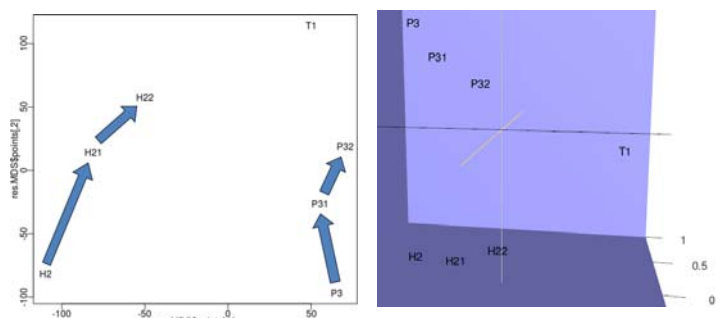


図3. 多次元尺度法(左)、潜在意味解析(右)による