

VOD 講義のための字幕強調や短縮表示法

則本達哉^{†1} 小山登^{†1} 小林伸行^{†2}
椎名広光^{†3} 北川文夫^{†3}

教育システムとして、e-Learning の利用が増加している。その中でもスライドと動画を VOD で配信する講義や自主学習システムが利用されてきている。特に、現状のシステムではパワーポイントと映像のみを表示していることが多い。そのため、日本語を母国語としない留学生などの利用者には映像からの音だけでは話題の理解が難しいと考えられる。また、字幕を話者の話す速さで表示するシステムにおいても、同様に理解が難しいと考えられる。本研究では、VOD システムの映像部分に字幕を付すだけでなく、字幕内の重要単語の強調法と字幕の表示量を削減手法について提案する。

Emphasis and sentence contraction of Subtitles for a VOD learning system

TATSUYA NORIMOTO,^{†1} NOBORU KOYAMA,^{†1}
NOBUYUKI KOBAYASHI,^{†2} HIROMITSU SHIINA^{†3}
and FUMIO KITAGAWA^{†3}

Recently, a use of e-Learning system increases on Internet and Intranet. In particular, VOD learning systems with slides and videos have been used for self-learning and lectures. Therefore, for foreign students who are non-native Japanese language, we consider that understanding of topics is difficult only by slides and sounds. In this study, we attached subtitles to video parts of the VOD system. And, to make match subtitles with Japanese understanding of users, we emphasized important words in subtitles and contracted subtitles by Japanese dependencies and important words.

1. ま え が き

現在、Web 教材を用いた e-Learning、すなわち WBT (Web-Based Training) と対面授業を組み合わせたブレンディッドラーニングやスライドと講義の動画を配信する VOD (Video On Demand) による e-Learning 講義などさまざまな大学で e-Learning の利用が行われている。岡山理科大学では e-Learning 講義を 2004 年度から行っている⁸⁾。

これまで我々は、教材として提供されている VOD システムの映像に字幕を付加し、その字幕を利用して検索単語の頻度のヒストグラムを提示と単語の頻度分布から映像区間を推定するシステムを開発を行ってきた¹⁰⁾。その手法には、ヒストグラムに二次関数をあてはめこむ方法^{11),12)}、混合正規分布を EM アルゴリズム^{1),9)} であてはめる方法等¹³⁾ を提案してきている。

一方、現在提唱されている VOD システムとしては、映像、音声、スライドからなる構成^{3),5),6),8)} となっており、ユーザのより理解力を手助けするには映像と共に字幕を添付は、大きな手助けになると考えられる。しかしながら、単に話者の音声を字幕に付しただけでは、読んで理解するには表示速度があまりにも早いという問題が浮上してくる。

また、VOD の利用者は日本を母国語とする日本人学生だけとは限らず、多様な背景を持つ学生のニーズをた捉えて、VOD システムを作成する必要がある。日本語を母国語とする日本人学生でさえも講演者の音声をそのまま字幕しただけでは、読み切れない。ましてや、日本語を母国語としない留学生にとっては、理解の困難さが増すのは当然のことである。

そこで本研究では、学習者の理解を深めるための字幕表示を、字幕の添付だけでなく、字幕を表示するとき、重要単語の強調と字幕の表示量を抑えることによって、表示スピードを遅くすることを目的として、システムを開発を行っている。重要単語の強調については、 $tf \cdot idf$ を重要語を取り出す範囲を表示している字幕の前後を対象として、計算を行う変更を行っている。また、字幕量の削減には、字幕の係り受け解析を行った後に、係り受けの関係と重要語の $tf \cdot idf$ 値から文節の削減対象として判断を行う手法を用いている。

なお、文短縮の既存研究として三上²⁾、牧野⁷⁾ の方法があるが、リアルタイム性や汎用性

†1 岡山理科大学大学院 総合情報研究科
Graduate School of Informatics, Okayama University of Science

†2 山陽学園大学 総合人間学部
Faculty of Human Sciences, Sanyo Gakuen University

†3 岡山理科大学 総合情報学部
Faculty of Informatics, Okayama University of Science

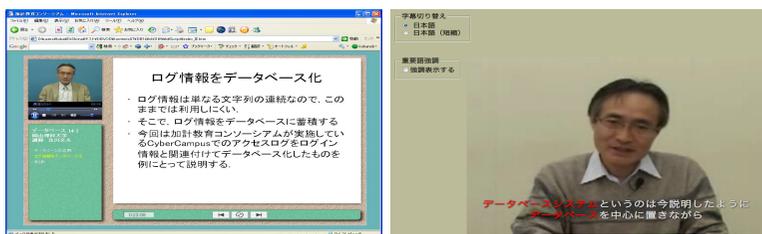


図 1 VOD 実行画面

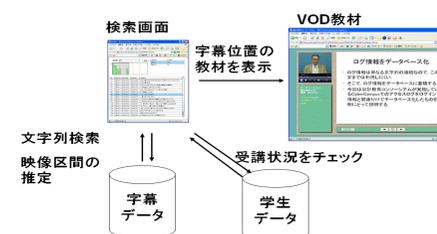


図 2 VOD システムの構成

が必要であるのに対して、本研究ではリアルタイム性が不要で VOD 講義という特殊なものを対象としているので、用いる手法も相違すると考えられる。

2. VOD システムによる e-Learning 講義システム

本研究で作成しているシステムは、岡山理科大学を含む関連 6 大学で構成している教育コンソーシアムにおける単位互換制度を利用した VOD による e-Learning 講義のシステム⁸⁾上(図 1(左))に別途追加する形で開発している(図 2)。2007 年度データベースの講義では、1 回の講義は 3 つのセクションに分かれており、1 つのセクションは 20~30 分程度となっている。ただし、14 回のみ 2 セクションとなっている。また、各セクションの最後に講義内容に関する課題があり、講義内容の理解を確認するために用いられている。

本研究で別途開発した VOD の実行画面は、図 1(右)のような構成で、次の 2 つの機能をラジオボタンで選べるようにしている。

- (1) 話者の話をそのまま字幕にし、重要語を強調して出力
- (2) 字幕の文を削減と重要語を出力

3. 字幕の重要単語

VOD 講義で講演者の字幕を表示する場合、より理解を増すということを考えると重要な単語を強調するのがよい。また、VOD 講義の全体からの重要語以外にも、今話している付近から得られる重要語も必要であると考えられる。そこで、 $tf \cdot idf^4)$ を表示している字幕の文の前後から計算するように変更する(図 3)。

まず、通常使われている文書 D_1, D_2, \dots, D_N に対して、単語 w_i の $tf \cdot idf(i, D_j)$ の定

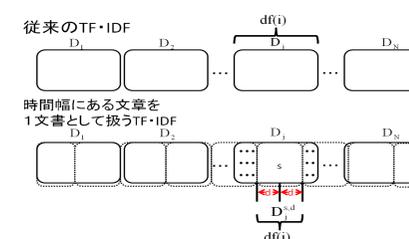


図 3 $tf \cdot idf$ の変更

義を以下に示す。

[$tf \cdot idf$ の定義⁴⁾]

単語 w_i が文書 D_j に含まれる数 $tf(i, D_j)$ 、単語 w_i を含む文書の数 $df(i)$ とするとき、

$$tf \cdot idf(i, D_j) = tf(i, D_j) \cdot \log \frac{N}{df(i)}.$$

これに対して、今話している付近から得られる重要語を求めるために、表示している字幕の前後を基準にして表示している字幕文内の重要語を強調することにする。そこで、字幕を表示している講義 D_j の文 s の前後 d 時間の文章を 1 つの文書 $D_j^{s,d}$ とみなして、 $tf \cdot idf(i, D_j^{s,d})$ を次のように定義する。

[文の前後を用いた $tf \cdot idf$ の定義]

単語 w_i が文書 $D_j^{s,d}$ に含まれる数 $tf(i, D_j^{s,d})$ 、時間区間 $2d$ で区切った文書数 N 、単語 w_i を中心とした時間区間 $2d$ で区切った文書の数 $df(i)$ とするとき、

$$tf \cdot idf(i, D_j^{s,d}) = tf(i, D_j^{s,d}) \cdot \log \frac{N}{df(i)}.$$

3.1 字幕文における重要語の例

前節の $tf \cdot idf(i, D_j^{s,d})$ を 2007 年度データベースの講義 1 回目セクション 1 に適用し、対象となる単語に下線、横に $tf \cdot idf(i, D_j^{s,d})$ 値を示したものを表 1 に示す。特に No.112 の前後の時間区間 $d = 2.5$ 分を中心にして示している。また、 $tf \cdot idf(i, D_j^{s,d})$ 値は、文書として扱う前後の時間区間内の単語の数の違いにより、No.106 から No.119 の間でも、同じ単語でも $tf \cdot idf$ 値に違いのある単語「説明」とない単語「データベースシステム」が存在している。

表 1 強調する単語 (講義名「データベース」1 回目 1 セクション)

No.	開始	終了	字幕
1	00:01	00:02	みなさん _(5.68) 、こんにちは
			~
106	13:04	13:13	上から説明 _(6.39) しますと ユーザ _(3.38) と書いてあるのは、これは皆さん使う人ですね 電話 _(2.22) を使う人 なら 皆さん _(5.27) のひとりひとり
107	13:14	13:19	それから会社 _(5.37) の中なら会社 _(5.37) の事務 _(5.37) をする人ひとりひとりということになります
108	13:20	13:30	そういう人たちが実際に先ほど言ったいろいろな データベース _(8.04) にアクセス _(2.67) する場合の構造 _(6.57) を、
109	13:30	13:33	おおまかな構造 _(6.57) を表しています。
110	13:36	13:47	右のほうに データベースシステム _(11.40) と書いてあり、全体を示しているということが書かれていますが、
111	13:47	14:01	データベースシステム _(11.40) と呼んだ時は 応用プログラム _(5.47) 、それから データベース管理システム _(5.01) これは DBMS _(3.20) と省略 _(2.62) されますが
112	14:01	14:16	データベース管理システム _(5.01) 、それから データベース _(9.65) この三つを全体を指して データベースシステム _(11.40) と呼びます。
113	14:18	14:26	では、次からはそれぞれについて説明 _(8.52) していきます。
114	14:28	14:37	データベースシステム _(11.40) というのは今説明 _(8.52) したように データベース _(9.65) を中心 _(2.57) に置きながら
115	14:38	14:48	応用プログラム _(5.47) そして、データベース管理システム _(5.01) を加えた全体のことを指します。
116	14:48	14:55	特徴 _(6.40) としてはそこに四つ黒丸 _(4.25) で書いてありますが、
117	14:55	15:03	この四つが特徴 _(6.40) になります。
118	15:03	15:16	一つ目は、大量のデータ _(12.7) を効率 _(9.44) 良く蓄積 _(3.20) できることです。
119	15:16	15:25	先ほどもいいましたように例えば 電話 _(2.22) だったならば何百万人 _(3.22) というようなデータ _(12.7) が蓄えられているわけですから
			~
182	24:47	24:53	以上で第一回目 _(7.63) の第一セクション _(4.04) を終わります。

表 2 字幕の強調語の評価 上位 1 位 (講義名「データベース」1 回目 1 セクション)

	強調すべき単語を含む文	強調すべきでない単語を含む文
強調した単語を含む文	116	28
強調しなかった単語を含む文	163	206

表 3 字幕の強調語の評価 上位 2 位まで (講義名「データベース」1 回目 1 セクション)

	強調すべき単語を含む文	強調すべきでない単語を含む文
強調した単語を含む文	168	83
強調しなかった単語を含む文	51	151

表 4 字幕の強調語の評価 上位 3 位まで (講義名「データベース」1 回目 1 セクション)

	強調すべき単語を含む文	強調すべきでない単語を含む文
強調した単語を含む文	192	130
強調しなかった単語を含む文	27	104

表 5 字幕の強調語の評価 上位 5 位まで (講義名「データベース」1 回目 1 セクション)

	強調すべき単語を含む文	強調すべきでない単語を含む文
強調した単語を含む文	213	190
強調しなかった単語を含む文	6	44

3.2 字幕文における重要語の評価

前節の字幕文の前後を用いる $tf \cdot idf(i, D_j^{s,d})$ を字幕文における重要語に用いることの評価として、各文の $tf \cdot idf(i, D_j^{s,d})$ 値の高い上位 1 位、2 位まで、3 位まで、5 位までに対して、どれだけ重要語として網羅できているのかを、人手で評価したものを表 2,3,4,5 に示す。ここで表 2,3,4,5 を評価するもとして、次式で計算される再現率、精度から求められる F 値を示す。

$$\text{再現率} = \frac{\text{本手法で強調した単語が正しい文数}}{\text{人間が選んだ強調単語を含む文数}}$$

$$\text{精度} = \frac{\text{本手法で強調した単語が正しい文数}}{\text{本手法が強調した単語を含む文}}$$

$$F \text{ 値} = \frac{2 \times \text{再現率} \times \text{精度}}{\text{再現率} + \text{精度}}$$

上位 1 だけを重要語として強調した場合の再現率 = 0.530, 精度 = 0.806, F 値 = 0.639, 上位 2 位までを重要語として強調した場合の再現率 = 0.767, 精度 = 0.669, F 値 = 0.715, 上位 3 位までを重要語として強調した場合の再現率 = 0.877, 精度 = 0.596, F 値 = 0.710,

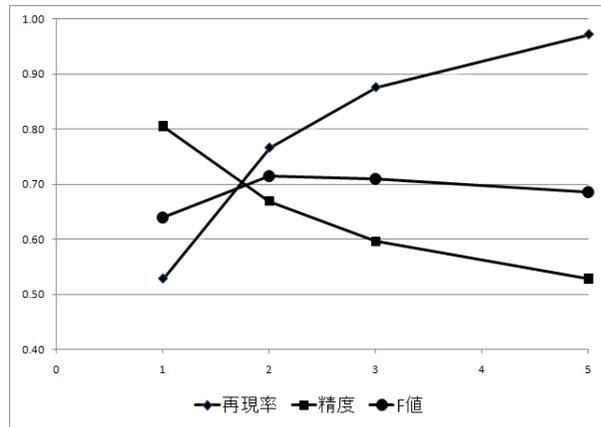


図 4 字幕の強調語の評価

上位 5 位までを重要語として強調した場合の 再現率 = 0.973, 精度 = 0.528, F 値 = 0.684 となる。

上位 1 だけでは, 精度がよいが再現率が低いため使い勝手が悪く, 上位 5 位まででは, 反対に再現率が良いが精度が低い。F 値からも分かる通り, 中間的な 2,3 位を表示するのが良いと考えられる。図 4 は再現率, 精度, F 値を上位何位までごとに表したものである。

4. VOD 講義における字幕の係り受けと強調単語による短縮

字幕をそのまま表示するのでは, 字幕の流れる速さが早い。そのため字幕だけで内容を理解することは難しいと考えられる。前節の重要単語を表示することも重要ではあるが, それ以上に字幕の量を減らして必要な部分だけを表示させることが有効であると考えられる。

本研究では, 字幕の文の係り受け解析し, 係っている側の文節の重要度を前節で計算した $tf \cdot idf(i, D_j^{s,d})$ 値から求め, 文節を削除する順序を計算する。そして表示する文字量から削除の優先順位が高い文節を削除した後の文を新しい字幕として表示する。以下ではそれらについて述べる。

4.1 字幕の文節の係り受けと強調単語による文節評価

字幕の文の係り受け解析し, 係っている側の文節に含まれる単語の前節で計算した $tf \cdot idf(i, D_j^{s,d})$ 値を係っている先の文節に順に足し合わせて評価値 $T(i)$ を計算する。次に, その評価値の大きい重要語を含む文節を残し, 反対に評価値の小さい値の文節を削除するよ



図 5 字幕の文節の評価

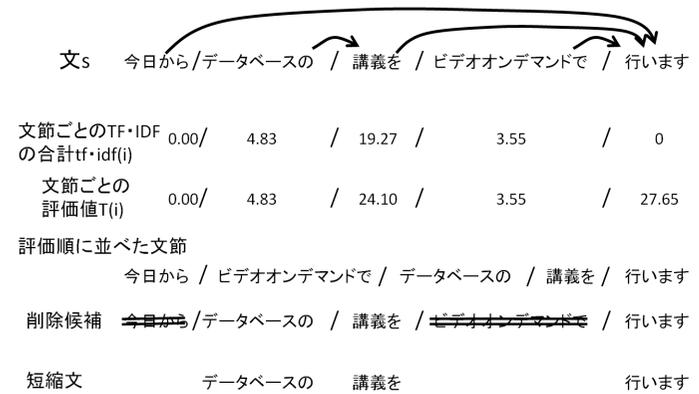


図 6 字幕の文節の評価と削除例

うに順位づけをおこなう。以下に文節評価アルゴリズムを示す(図 5 に図解)。なお, 字幕の文の係り受けには, Cabocha¹⁴⁾ を利用している。

[定義] 文 s , 文 s の文節 $C(s) = C_1, C_2, \dots, C_n$, 各文節 $C(i)$ がかかる先 $A(C_i)$, 各文節 C_i の $tf \cdot idf$ 値 $tf \cdot idf(i)$, 各文節 C_i の評価値 $T(i)$ とする。

[文節評価アルゴリズム]

(1) 各文節 C_i 内の単語 w_j の $tf \cdot idf$ の合計を計算する。

$$\text{for } i = 1..n \text{ do}$$

$$tf \cdot idf(i) = \sum_{w_j \in C_i} tf \cdot idf(j, D_j^{s,d})$$

表 6 字幕文短縮例 $m = 24$ (講義名「データベース」1回目 1 セクション)

No.	開始	終了	字幕 ($m = 24$)
1	00:01	00:02	みなさん、こんにちは
			~
106	13:04	13:13	説明しますと書いてあるのは、これは人ですね。 電話を使う人なら皆さんのひとりひとり、
107	13:14	13:19	事務をする人ひとりひとりということになります。
108	13:20	13:30	言ったデータベースにアクセスする場合の構造を
109	13:30	13:33	おおまかな構造を表しています。
110	13:36	13:47	
111	13:47	14:01	データベースシステムと呼んだ時は応用プログラム。 データベース管理システムこれは省略されますが、
112	14:01	14:16	それからデータベースこの三つ呼びます。
113	14:18	14:26	では、次からはそれぞれについて説明していきます。
114	14:28	14:37	説明したようにデータベースを中心に置きながら
115	14:38	14:48	指します。
116	14:48	14:55	書いてありますが、
117	14:55	15:03	なります。
118	15:03	15:16	一つ目は、データを効率良く蓄積できることです。
119	15:16	15:25	というようなデータが蓄えられているわけですから。
			~
182	24:47	24:53	以上で第一回目の第一セクションを終わります。

表 7 字幕文短縮例 $m = 30$ (講義名「データベース」1回目 1 セクション)

No.	開始	終了	字幕 ($m = 30$)
1	00:01	00:02	みなさん、こんにちは
			~
106	13:04	13:13	上から説明しますと書いてあるのは、これは皆さん使う人ですね。 電話を使う人なら皆さんのひとりひとり、
107	13:14	13:19	事務をする人ひとりひとりということになります。
108	13:20	13:30	人たちが言ったデータベースにアクセスする場合の構造を、
109	13:30	13:33	おおまかな構造を表しています。
110	13:36	13:47	書いてあり、示しているということが書かれていますが、
111	13:47	14:01	データベースシステムと呼んだ時は応用プログラム データベース管理システム DBMS と省略されますが、
112	14:01	14:16	それからデータベースこの三つ呼びます。
113	14:18	14:26	では、次からはそれぞれについて説明していきます。
114	14:28	14:37	説明したようにデータベースを中心に置きながら
115	14:38	14:48	データベース管理システムを加えた全体のことを指します。
116	14:48	14:55	特徴としては四つ黒丸で書いてありますが、
117	14:55	15:03	特徴になります。
118	15:03	15:16	一つ目は、大量のデータを効率良く蓄積できることです。
119	15:16	15:25	何百万人というようなデータが蓄えられているわけですから
			~
182	24:47	24:53	以上で第一回目の第一セクションを終わります。

(2) 各文節が係っている先に合計した $tf \cdot idf(i)$ を足し合わせる。

(2-1) for $i = 1..n$ do // \$T(i)\$ の初期化

$$T(i) = tf \cdot idf(i)$$

(2-2) for $i = 1..n$ do

if $A(C_i) = C_j$ then

$$T(j) := T(j) + T(i)$$

(3) $T(j)$ を小さい順に C_j を並び替え、並び替えた文節を $\kappa(s)$ とする。

文節評価アルゴリズムの処理の様子を図 5 に示す。図 5 では文 s が 5 文節 C_1, C_2, C_3, C_4, C_5 からなり、係り受けは、 C_1 が C_4 ($A(C_1) = C_4$)、 C_2 が C_3 ($A(C_2) = C_3$)、 C_3 が C_4 ($A(C_3) = C_4$)、 C_4 が C_5 ($A(C_4) = C_5$)、(1) での単語 $tf \cdot idf$ の合計を計算が、 $tf \cdot idf(1) = 10$ 、 $tf \cdot idf(2) = 3$ 、 $tf \cdot idf(3) = 5$ 、 $tf \cdot idf(4) = 1$ 、 $tf \cdot idf(5) = 0$ としている。次に、(2) 係っているもとの文節の $tf \cdot idf$ の合計に係る先に足し合わせによって、 $T(1) = 10$ 、 $T(2) = 3$ 、 $T(3) = 8$ 、 $T(4) = 19$ 、 $T(5) = 19$ と計算される。最後に、 $\kappa(s) = \{C_2, C_3, C_1, C_4, C_5\}$ と

削減リストが作成される。

また、VOD 講義「データベース」1回目第 1 セクションの No.2 「今日からデータベースの講義をビデオオンデマンドで行います。」に対する評価を図 6 に示す。図 6 では、評価値 $T(i)$ 値の低い順の“今日から” ($T(1) = 0.000$)、“ビデオオンデマンドで” ($T(4) = 3.55$)、“データベースの” ($T(2) = 4.83$)、“講義を” ($T(3) = 24.10$)、“行います。” ($T(5) = 27.65$) の順で削除候補となる。

4.2 字幕の文節の評価と文字数による短縮

最終的に短縮した字幕の表示は、表示する文字数との関係から決められる。そこで本研究では、前節の文節ごとの評価 $T(i)$ による削除候補に従って、表示する文字数に収まるようにしている。以下では、その手法について示す。なお、アルゴリズム中の $Top()$ は番号が小さい要素のを取り出す関数とする。

表 8 字幕文短縮例 $m = 36$ (講義名「データベース」1回目 1 セクション)

No.	開始	終了	字幕 ($m = 36$)
1	00:01	00:02	みなさん、こんにちは
			~
106	13:04	13:13	上から説明しますとユーザと書いてあるのは、これは皆さん使う人ですね。電話を使う人なら皆さんのひとりひとり、
107	13:14	13:19	それから会社の中なら会社の事務をする人ひとりひとりということになります。
108	13:20	13:30	人たちが先ほど言ったデータベースにアクセスする場合の構造を、
109	13:30	13:33	おおまかな構造を表しています。
110	13:36	13:47	書いてあり、示しているということが書かれていますが、
111	13:47	14:01	データベースシステムと呼んだ時は応用プログラムそれからデータベース管理システムこれはDBMSと省略されますが、
112	14:01	14:16	それからデータベースこの三つをデータベースシステムと呼びます。
113	14:18	14:26	では、次からはそれぞれについて説明していきます。
114	14:28	14:37	説明したようにデータベースを中心に置きながら
115	14:38	14:48	データベース管理システムを加えた全体のことを指します。
116	14:48	14:55	特徴としてはそこに四つ黒丸で書いてありますが、
117	14:55	15:03	この四つが特徴になります。
118	15:03	15:16	一つ目は、大量のデータを効率良く蓄積できることです。
119	15:16	15:25	何百万人というようなデータが蓄えられているわけですから
			~
182	24:47	24:53	以上で第一回目の第一セッションを終わります。

[定義] 表示する文字量 m , 対象とする字幕文 s , 字幕文 s の文節 C_1, C_2, \dots, C_n , 字幕文 s の文節 C_i の文字数 $L(C_i)$ 評価 $T(i)$ の小さい順の字幕文 s の文節集合 $\kappa(s)$ とする。

[文節削除アルゴリズム]

(1) 短縮する字幕の文字数に関して, $\sum_{C_i \in \kappa(s)} L(C_i) \leq m$ を満たすまで以下を繰り返す。

(1-1) $\kappa(s) := \kappa(s) - Top(\kappa(s))$ // $\kappa(s)$ の先頭の評価の低い文節を 1 つ削除

上記の文節削除アルゴリズムで、図 6 の例文の 2 文節を削除すると図 6 の下の様に字幕文が短縮できる。

表示する字幕量については、現実的には決めなければならない。そこで、動画表示画面に収まる文字数 1 行から 2 行分 (表示文字数 $m = 24$ から $m = 36$) の講義名「データベース」1 回目 1 セクションに対する短縮された字幕を表示文字数 $m = 24, 30, 36$ に分けて表 6, 7, 8 に示す。

表 9 短縮字幕の評価 (表示文字 講義名「データベース」1 回目)

対象文の文字数	21~24	25~30	31~36	37~48
字幕として適切な文	15	27	11	17
字幕として許容できる文	14	15	21	32
重要語を失っている文	9	5	2	15
非文になった文	2	7	8	18
(適切+許容)/対象文数	0.725	0.778	0.762	0.598

表 10 短い字幕文の評価

	文節数 2~3	1文 20 文字以下	1文 24 文字以下
必要 (短い単語)	34	51	86
不要 (意味のない字幕)	15	17	27

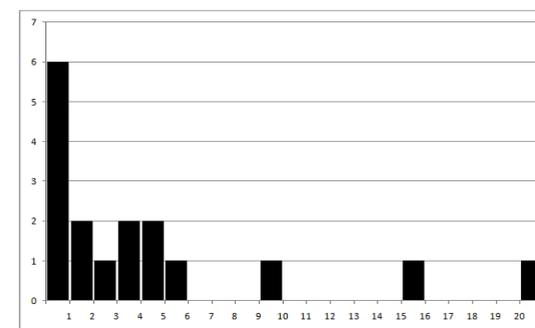


図 7 講義「データベース」1 回目第 1~3 セクションの短い文の要不用

4.3 字幕の短縮の評価

講義「データベース」1 回目 (第 1~3 セッション) に対する文節削除アルゴリズムの適用を、削除前の字幕文の文字数ごとに分け、短縮した字幕文が、適切な文の数、許容範囲の文の数、重要語を失った文の数、非文となった数を表 9 に示す。このうち重要語を失うものと非文になる場合を失敗と考えると、削除前の字幕文の文字数が、短いものと長いものに対して、失敗が多くなる傾向がある。

そこで、短い文については、短縮前の字幕文自身が、意味のないことを言っている場合と、短い単語を言っている場合の 2 種類で分類した結果を表 10 に示す。意味のないことを言っている場合については、字幕文中の単語の $tf \cdot idf$ の最大値の分布を図 7 に示す。図 7 から分かるように、 $tf \cdot idf$ 値の低いものがほとんどで重要語が含まれていないため、字幕

としては表示しなくてもよいとも考えられる．反対に必要とされる 2~3 の分節はほぼキーワードと考えてよく，成形して字幕として出力したほうが良い．言い換えると，字幕文の文字数が少ないものに対しては，本研究で提案した文節削除のアルゴリズムを利用するのはなく， $tf \cdot idf$ 値を閾値とした別な手法を利用したほうが良いと言える．

一方，長い文については，複文が多く文節が多いため有効に文節を削除できていない．これに対しては，文を単文に分けてから，文節削除アルゴリズムを適用したほうが良いと考えられる．

以上のことから，短縮前の字幕文の文字数が 20~48 文字に，本研究で提案した手法を適用するのが良く．短い字幕文や長い字幕文は，個別の手法を適用するのが良いと考えられる．

5. ま と め

本研究では，字幕を出力に，重要語の強調と，その重要語と係り受けを用いて字幕文を短縮する手法を提案した．係り受けの多い部分には有効であると考えられるが，文短縮全体と見ると文の短縮では，助詞などの文末の語を削除する手法を入れていないので削除量が少ないところもある．今後は，先行研究の手法を取り入れて行く予定である．また削除アルゴリズムだけでは，日本語として不自然な部分があるので，文の生成についても追加していきたい．

利用からの視点に関しては，現在の所，考慮が少ない．日本語を母国語としている学生のほかに，母国語としない留学生の利用される手法の考えていきたい．単語の難易度の構築や単語の類似度を用いることで，より深い意味での字幕表示システムを作成できるのではないかと考えている．

参 考 文 献

- 1) A.P.Dempster, N.M.Laird, and D.B.Rubin.: Maximum likelihood form incomplete data via the EM algorithm. Journal of the Royal Statistical Society series B, Vol. 39, No.1, pp.1-38, 1977.
- 2) 三上, 増山, 中川: ニュース番組における字幕生成のための文内短縮による要約, 言語処理学会論文誌「自然言語処理」, Vol.6, No.6, pp.65-81, 1999.
- 3) 伊藤, 藤井, 石川: 音声文書検索を用いたオンデマンド講義システム, 電子情報通信学会技術研究報告 SP 音声, Vol.101, No.523, pp.55-60, 2001.
- 4) 北, 津田, 獅々子: 情報検索アルゴリズム, 2002.

- 5) Haruo Yokota, Takashi Kobayashi, Taichi Muraki, and Satoshi Naoi.: UPRISE: Unified Presentation Slide Retrieval by Impression Search Engine. IEICE Trans. on Info. and Syst., Vol. E87-D, No. 2, pp. 397-406, 2 2004.
- 6) 森本, 室田, 清水: 教育用動画像検索システムと時間情報同期方法の開発. 電子情報通信学会論文誌 D-I, Vol. J88-D-I, No. 10, pp. 1515-1524, 2005.
- 7) 牧野, 池田, 山本: 類似用例文の部分的置換による文短縮 (解析, 分析), 情報処理学会研究報告 NL-173(4), pp. 21-28, 2006.
- 8) 北川, 大西: 対面講義と e-learning(LMS + VOD) とを併用した講義形式の実践と分析, 日本教育情報学会学会誌 Vol.22 No.3 pp.57-66, 2007.
- 9) 金森, 竹之内, 村田: R で学ぶデータサイエンス 5 パターン認識, 共立出版, 2009.
- 10) 小林, 椎名, 北川: 字幕データを用いた VOD 教材検索システムの提案, pp416-417, 教育情報システム学会第 31 回全国大会, 2009.
- 11) 小林, 椎名, 北川: 字幕付き VOD 講義に対する単語頻度分布による検索, 情報処理学会研究報告, Vol. 2009-DBS-149, No.7 pp.1-7, 2009.
- 12) 椎名, 小林, 北川: 言語処理学会第 16 回年次大会, VOD 講義の字幕を利用した話題検索, pp.547-550 2010.
- 13) 小林, 椎名, 北川: 字幕データ付き VOD 講義の単語頻度に対する混合正規分布モデルによる映像区間の推定, pp.306-307, 日本教育情報学会第 26 回年会, 2010.
- 14) Cabocha, <http://chasen.org/taku/software/cabocha/>