

講義再現システムにおけるスライドへの重要度付与法の検討*

2S-3

松田和彦[†] 田口亮[‡] 桂田浩一[†] 山田博文[‡] 新田恒雄[†][†]豊橋技術科学大学大学院 工学研究科 [‡]豊橋技術科学大学 工学部

1 はじめに

近年、計算機を用いた講義支援システムの研究が盛んである。これらの研究の中で講義を再現する方法が検討されている[1][2]。講義再現の主要目的は、復習を通して学生に講義への理解を深めさせることである。そこで、再現の際に個々のスライドに対する重要性を提示できれば、学生は重要箇所に沿って短期間に復習を行なうことができると考える。我々は、教師が重点を置いて講義を行なっている箇所、即ち重要なスライドをシステムが自動的に見つけ、学生に提示する方法を検討している[3]。

本報告では、講義再現システムの概要を述べ、続いて、講義中に収集した非言語情報とスライド中の言語情報に基づく、スライドへの重要度自動付与法を説明した後、評価実験結果を述べる。

2 講義再現システム

講義再現システムの画面例を図 1 に示す。図の中央は講義スライドになっている。講義再現システムは講義スライドを表示すると共に、音声と同期して再生する。再生時にはマウス位置及びスライドへのデジタルインク書き込みも、スライドと音声に同期して表示される。

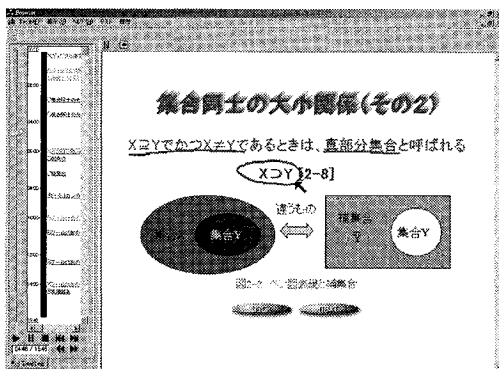


図 1：講義再現システムの画面例

図の左端はコントロール部である。コントロール部には縦方向を時間軸に、講義スライドのタイトルと重要度を表示している。タイトルは重要度に応じて色分けされており、学生は重要度を容易に判別することができる。コントロール部にはこの他、タイムスライダ、コントロールボタン、経過時間などが表示されている。

3 講義スライドへの重要度自動付与

3.1 重要度

重要度とは教師が各スライドに付ける重要性に対する主観的評価である。今回は、教師が講義スライドを「非常に重要」、「重要」、「やや重要」、「重要でない」の 4 段階で評価している。

3.2 非言語情報を用いた重要度付与法

我々は[3]で非言語情報を用いた重要度付与法を提案した。この手法は、講義スライドの重要度に関係が深いと思われる、以下の非言語情報 8 項目を講義中に収集し、多層ニューラルネット (MLN) を学習した後、その出力に基づいて重要度を判定する。

- ① 表示時間 ② 発話時間
- ③ マウス移動量 ④ デジタルインク描画量
- ⑤⑥ 音声のパワー (平均と分散)
- ⑦⑧ 音声のピッチ (平均と分散)

実験の結果、重要なスライドには比較的時間を掛け、デジタルインクを多用する傾向が見られた。そのため、非言語情報の①から④までの値が高い場合、MLN はそのスライドが重要であると判定する「クセ」を持つ。

重要度付与の精度が良くない講義を具体的に調べると、教師が「非常に重要」とした箇条書きのスライドに対し、(デジタルインクを使用せず) 短時間で説明を終わらせた後、次の具体例を示すスライドで時間を掛けて説明していた。このため、MLN は「箇条書きのスライドを重要でない。具体例のスライドは重要である。」と判定した。これらの結果から、MLN の持つクセを補正するには、スライド中の言語情報を加味する必要があると考えられる。

*An examination of an importance attachment to the slides in lecture review system
K.Matsuda[†], R.Taguchi[‡], K.Katsurada[†], H.Yamada[‡], T.Nitta[†]
[†]Graduate School of Eng., Toyohashi University of Technology
[‡]Faculty of Eng., Toyohashi University of Technology

3.3 言語情報を考慮した重要度付与法

最初に、スライド間の関連を言語情報から調べる方法を検討する。具体的には、スライド中から単語を抽出し、スライド間の類似度を求める。以下に手順を示す。

- 1) 全スライド中の文を形態素解析して名詞を抽出する。複数のスライドで使われた名詞を講義内容を表す単語（キーワード）とする。今回の実験では2枚以上で使用された単語とした。
- 2) スライドごとにキーワードの有無（有1，無0）を要素とするベクトル V_i を考え、このベクトルからそれぞれのスライド間の類似度を計算する。ここではベクトルの内積を用いる。

$$\text{sim}(V_i, V_j) = \frac{V_i \cdot V_j}{\|V_i\| \|V_j\|}$$

ここで、 $i, j=1, 2, \dots, P$ はスライドの番号に対応する。

- 3) 類似度が閾値以上であるスライド間には関連があるとすると、今回は閾値を0.65とした。

次に i 番目のスライドの非言語情報 $U_m (m=1, 2, 3, 4)$ で①から④の情報に対応)を補正する方法を述べる。

- A) 関連するスライドの非言語情報 U_{mj} と類似度を用いて、関連性を考慮した非言語情報 U'_m を次式から計算する。

$$U'_m = \frac{\sum_{j=1}^P \text{sim}(V_i, V_j) \cdot U_{mj}}{\sum_{j=1}^P \text{sim}(V_i, V_j)} \quad \begin{array}{l} \text{sim}(V_i, V_j) < 0.65 \\ \text{のとき} \\ \text{sim}(V_i, V_j) = 0 \end{array}$$

- B) U_m を U'_m に置き換える。ただし、MLNは重要なスライド (U_m が大) を検出することが目的であるので、 $U_m > U'_m$ ならばそのままとする。

重要度自動付与にはMLN（4層；各層は8，16，8，1のユニットを持つ）を適用する。学習段階では、入力信号に補正後の8つの非言語情報を、また教師信号として教師が付けた重要度を与える。

4 評価実験

本報告で提案した重要度付与方法を評価するため、非言語情報のみを用いた重要度付与方法との比較実験を行なった。教師7名が5つの講義コンテンツのうち2つ以上を用いて講義を行い、324枚のスライドからなる24の講義の非言語情報を取得した。重要度は講義を行なった教師が付与した。講義コンテンツごとに学習・評価を行ない、その結果を集計した。2

表1：非言語情報をを用いた重要度付与

教師	MLN				相関係数	再現率
	×	△	○	◎		
×	52	14	0	7	◎→◎ 63.1%	◎→◎○ 82.1%
△	23	19	17	24		
○	10	11	21	42		
◎	5	10	16	53		

表2：言語情報を考慮した重要度付与

教師	MLN				相関係数	再現率
	×	△	○	◎		
×	51	7	5	10	◎→◎ 63.1%	◎→◎○ 86.9%
△	27	18	17	23		
○	5	15	22	39		
◎	2	9	20	53		

（◎は「非常に重要」、○は「重要」、△は「やや重要」、×は「重要でない」を表す。）

つの手法の実験結果を表1，2に示す。

比較実験から、言語情報を考慮した重要度付与では、教師が「非常に重要」としたスライドのうち、MLNが87%を「非常に重要」または「重要」と判定し、再現率が向上した。

また、講義コンテンツごとに見ると、非言語情報のみの重要度付与で最も結果の悪かった講義では、再現率（◎→◎○）を69%から78%へと大幅に向上させることができた。

5 まとめ

非言語情報のみの重要度自動付与に対して、新たに言語情報を加えた方法を提案した。評価実験の結果、言語情報を用いてスライド間の関連を調べ、補正する手法が有効であることが示された。

今後はキーワードに対する検討、教師音声でのキーワード抽出などによって、重要度付与の精度を更に向上させたい。

参考文献

- [1] G. D. Abord et al, "Classroom 2000: An Experiment with the Instrumentation of a Living Educational Environment", IBM Systems Journal, Vol.38, No.4, pp.508-530 (1999).
- [2] K. Katayama et al, "Use of Action History Views for Indexing Continuous Media Objects", Proc. of AMCP'98, pp.344-355 (1999).
- [3] 山田他, "講義再現システムにおけるスライドへの重要度自動付与方法とその評価", 信学技報 ET2001-5, pp.25-32(2001).