

誘目性の高い講義コンテンツを作成する自動編集システム

中村 亮太[†] 井上 亮文^{††} 市村 哲^{††}
岡田 謙一^{†††} 松下 温^{††}

近年、大学などの教育機関では、学習用コンテンツとして講師映像と講義資料を組み合わせたものが制作されているが、単調な表示方法であるため学習者を飽きさせてしまうという問題がある。そこで本稿では、学習者にとって飽きにくい講義コンテンツを自動的に作成することができるシステム「MINO: Multimedia system an Instructor needs Not Operate」を提案する。著者らは映像の単調さを改善するために、誘目性（目が惹き付けられる）に着目し、画面に並列表示された講師映像と講義資料の表示サイズを交互に拡大縮小することで提示映像のスイッチングを行い、講義資料中の重要語句（講師の発話と一致した語句）を誘目性の高い表示へ変換することが自動的にできるシステムの開発を行った。MINO では、音声認識によりテキスト化した講師の発話情報と講義資料内の文字列とをマッチングさせることで重要語句の特定とともに、映像の切替えタイミングとフォントの変換タイミングを自動的に決定することができる。評価実験の結果、本提案手法は従来の提示方法に比べ、学習者を飽きさせないという評価を得ることができた。本稿では従来の提示方法について分析した結果を示し、開発したシステムの設計、実装、評価について述べる。

An Educational System that Automatically Generates Conspicuous Contents

RYOTA NAKAMURA,[†] AKIFUMI INOUE,^{††} SATOSHI ICHIMURA,^{††}
KENICHI OKADA^{†††} and YUTAKA MATSUSHITA^{††}

Recently, e-learning contents that combine the speaker video with supporting materials are produced in educational institutions such as universities. However, there is a problem that those systems make learners become tired because produced contents are monotonous. In this paper, we propose the system “MINO: Multimedia system an Instructor needs Not Operate” that can automatically edit the recorded speaker video and supporting materials. MINO allows users to automatically convert words in the supporting materials into conspicuous ones according to the utterance of the speaker. We used speech recognition to convert voice of lecturer into character string. Their data are matched up words into the supporting materials so that speaker video is synchronized with supporting materials. Through evaluations of the system, we verified the effectiveness of our system.

1. はじめに

近年、大学などの教育機関では、多くの教育資源がデジタル化され、インターネット上で共有されるようになった¹⁾。その一環として、講師映像と講義資料を組み合わせたコンテンツも制作されており²⁾、いつでもどこでも自主学習が可能な環境が整いつつある。しかし、そのようなコンテンツを制作するには映像と資

料を同期させるなどの手作業が必要であり、制作者にとって負担が大きいという問題がある。このためそれらのコンテンツは講師映像と講義資料を並べただけの単調なものであることが多く、学習者に飽きが生じてしまう傾向がある。

そこで本稿では、講師がどの部分を説明しているのか分かりやすく、また学習者を飽きさせないコンテンツへ自動的に編集できるシステム「MINO」(Multimedia system an Instructor needs Not Operate)を提案する。著者らは、自主学習時における学習者を飽きさせないための工夫として、目が惹き付けられる特性「誘目性」に着目し、この知識に基づいてMINOの構築を行った。MINOでは講師の発話と講義資料の同期を自動的に行うために、音声認識と形態素解析を

[†] 慶應義塾大学大学院理工学研究科
Graduate school of Science and Technology, Keio University

^{††} 東京工科大学
Tokyo University of Technology

^{†††} 慶應義塾大学理工学部
Faculty of Science and Technology, Keio University

利用している。音声認識によりテキスト化した講師の発話情報と講義資料内の文字列とを形態素解析したうえでマッチングを行う。MINO を利用することで、講師の発話に合わせて講義資料の表示サイズや資料内の文字の大きさ・色が誘目性の高い表示に変化するコンテンツを作成することができる。MINO によって作成されたコンテンツは蓄積型コンテンツとしてサーバに蓄えられ、学習者はネットワークを介していつでもどこでも利用することが可能である。

本稿では、従来のコンテンツにおける問題を分析した結果を示し、構築した MINO システムの設計、実装、評価について述べる。

2. 背景

2.1 従来研究

従来、講義コンテンツを作成するには時間と手間がかかるため、コンテンツを自動編集するシステムの実現が求められており、講師 1 人でも作成可能な講義自動収録システムが提案されている^{3)~6)}。これら多くの講義自動収録システムは、ビデオカメラで撮影した講師映像と講義資料を用い、2 画面表示などで学習者に提示するようになっている。講義コンテンツにおいて、講師映像と講義資料を同期させる必要があるが、柳沼は SMIL⁷⁾ を用いた映像教材作成支援ツールの開発し、講義映像とスライドを同期した映像教材を作成した³⁾。石井は自動的に講義映像をシーンで区切り、復習を効率的に短時間で行うために映像のインデキシング化を行った⁵⁾。しかし、これらのシステムは講師映像と講義資料を同期させ、講義のアーカイブ化を行っただけとどまっており、本研究の目的である受講者を飽きさせないための工夫や講師が講義資料内のどの箇所を説明しているのかが直感的に分かる提示方法を行っていない。一方、講師が説明資料中のどの箇所を説明しているのか受講者に示すシステムとして、レーザーポインタとマウスを用いた研究がある^{8),9)}。菅原らは講義中に講師がレーザーポインタで説明資料を指し示した位置を記録し、コンテンツ再生時には記録したレーザーポインタの位置を矢印として表示し、説明資料の上のどの箇所を説明しているのか示すことを行っている⁸⁾。しかし、レーザーポインタでは説明箇所のみを安定した動きで指し示すことは難しく、受講者に混乱を招く恐れがあると思われる。また講師が毎回レーザーポインタを用いて説明するとは考えられず、本システムのように講師の発話情報を用いた手法の方が自然であり、説明箇所の案内を毎回漏らすことなく的確に行うことができると思われる。



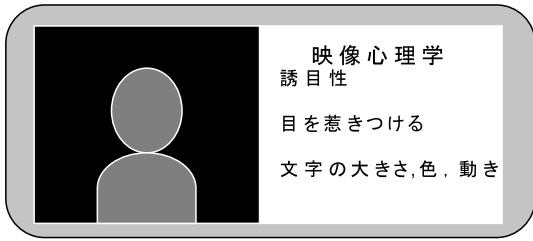
図 1 NHK 教育番組、放送大学の映像構成
Fig. 1 The video structure of the educational program of the NHK and the University of the Air.

2.2 誘目性

心理学では、映像が学習に及ぼす影響として、テキストと同時に映像を表示した場合、テキストだけの場合に比べて、学習者の事後テストの成績が向上することが報告されている¹⁵⁾。また、色彩が加わったときの効果として、白黒とカラー映像を見た場合、後者の方が見る人により多くの注意や興味を与えることができると報告されている¹¹⁾。映像心理学において、無意識的に目が惹き付けられる特性のことを誘目性という。たとえば、「信号や広告など色の变化（色相）が大きいもの」は誘目性が高く¹⁰⁾、また、「明暗の変化（明度）が大きなもの」、「サイズが大きなもの」、「動きがある」といったものも誘目性が高いという特性がある。実際にテレビを視聴しているときの眼球運動を測定した研究では、視聴者が映像内の動きのある対象物に注視しやすく、ズームングやショットの切替えといった映像独自の動きや変化のある操作に対しても注視しやすことが分かっている¹¹⁾。以上の知識を映像提示方法に加え、講師が話している講義資料内の語句を誘目性の高い表示に変えるようにした。

2.3 従来の講義コンテンツの映像構成

現在、遠隔講義コンテンツはテレビやインターネットを通して様々な現場で利用されている。ここで、従来のコンテンツの表示形式を次の 2 通りに大別する。テレビで放送されている放送大学や NHK 教育番組では図 1 のように 1 つ画面に講師映像や VTR、黒板やスライドなどの文字表示が切り替わって表示される構成になっている。一方、インターネットを介して配信されている SOI (WIDE university)¹²⁾ の映像構成は図 2 のように講師映像とスライドが横に並べられた構成になっており、講義の進行に合わせてスライ



講師映像 スライド

図 2 SOI (WIDE university) の映像構成

Fig. 2 The video structure of the SOI (WIDE university).

表 1 アンケート項目ならびにアンケート調査結果
Table 1 The result of questionnaire survey.

アンケート項目	放送大学	NHK 教育番組	Wilcoxon 符号付順位和 検定 p 値
映像が単調であると感じた	4.14	1.93	0.0000157
視聴していて飽きを感じた	4.07	2.21	0.0000548
また視聴したいと思った	2.07	3.51	0.0002885

(N=14)

ドが切り替わるようになっている。

2.4 放送大学と NHK 教育番組の比較

著者らは、どのような映像構成が受講者に単調さや飽きを感じさせてしまうのか、放送大学と NHK 教育番組と SOI を対象にして調査を行った。

まず、1 画面で映像を切り替える形式である放送大学と NHK 教育番組とを比較するために、2 つのコンテンツを視聴した経験がない被験者 14 名 (大学生) に映像を視聴させ、アンケート調査した。比較の対象としたコンテンツは大学・高校生を対象とした外国語科目やコンピュータ技術などであり、公平に評価を行うために難易度の高い放送大学の講義は対象から除外した (n = 40 分 × 20 講義)。そして映像視聴後、被験者にアンケートを行った。各項目に対して、強く賛同したときを 5 点満点とし、1~5 点の点数によって回答させ、平均値を比較した。アンケート項目ならびにアンケート結果を表 1 に示す。

アンケート調査の結果、3 つの項目すべてにおいて有意差が認められ、放送大学の方が映像の単調さを感じ、飽きやすいという結果が得られた。そこで、映像の単調さについて分析するために、放送大学と NHK 教育番組の講師映像について表示持続時間ごとの頻度確率を求め、比較を行った。

講師映像の持続時間を比較した結果、図 3 のように放送大学では講師映像が 120 秒以上続く確率が最も高いが、NHK では 30 秒以内で別の映像に切り替わる

表示持続時間分布

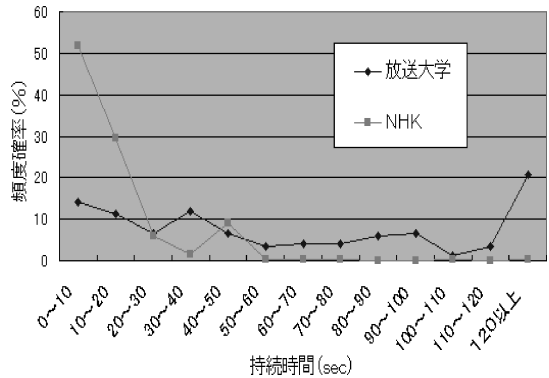


図 3 講師映像の持続時間確率分布

Fig. 3 The probability distributions on the videos of the lecture.

確率がきわめて高いことが分かる。この結果からも放送大学では講師映像が長く表示され続けることによって、単調になりやすいことを確認することができた。

2.5 映像の単調さと視聴者の心理状態

従来から映像をより効果的に視聴者に提示するためには、同一の映像を提示し続けるのではなく、何らかの映像表現を加える必要があると考えられてきた^{13),14)}。たとえば、テレビ番組のカメラワークの知識に基づいた多人数遠隔会議システムでは、会議参加者を様々なショットで撮影し、それらを適切なタイミングでスイッチングした映像を遠隔の参加者に提示することで、視聴者が退屈しない映像を作り出すことに成功している¹³⁾。また、映像の心理学において、映像操作と視覚行動について調査した結果によると、視聴者は映像情報の違いや変化に対応して、能動的に映像を見ようとする¹⁵⁾。さらに視覚的メディアである映画・映像においては、ショットとショットとを編集する際に、刺激の強弱・断続・連続などの視覚的リズムが考慮されている¹⁶⁾⁻¹⁸⁾。以上のように、映像を視聴者に提示するには映像表現に工夫が必要であり、映像が単調であると視聴者を飽きさせてしまう恐れがある。したがって、何らかのタイミングで映像を切り替える必要があることが分かった。

2.6 SOI と NHK 教育番組の比較

テレビのように 1 つの画面で映像が切り替わる形式と SOI のように 2 つの画面を同時に表示させる形式のどちらが良いか、これらのコンテンツを視聴した経験がない被験者 14 名 (大学生) に映像を視聴させ、簡単なアンケート調査を行った。比較の対象としたコンテンツは放送大学と NHK を比較したときと同様、大学・高校生を対象とした外国語科目やコンピュータ

表 2 アンケート項目ならびにアンケート調査結果 (SOI, NHK)
Table 2 The result of questionnaire (SOI, NHK).

アンケート項目	SOI	NHK 教育番組	Wilcoxon 符号付順位和 検定 p 値
映像が単調であると感じた	4.00	2.14	0.000055
自由度が大きいと感じた	3.36	2.36	0.011695
また視聴したいと思った	3.07	3.14	0.903572

(N=14)

などの情報通信の科目などであり、公平に評価を行うために難易度の高い SOI の講義は対象から除外した ($n = 40 \text{分} \times 20 \text{講義}$)。各項目に対して、強く賛同したときを 5 点満点とし、1~5 点の点数によって回答させ平均値を比較した。また、同時に被験者にインタビューし、感想を聞いた。

アンケート調査の結果、表 2 のように「映像が単調であると感じた」「自由度が大きいと感じた」の 2 つの項目で有意差が認められた。SOI では NHK のように映像のスイッチングは行われず、つねに講師映像と資料が並列表示されているために、映像の切替えはなく、それが被験者を見ていて単調に感じたようであった。さらに、SOI では画面に講師映像と並載するために資料を大きく表示することができず、スライド内の細かい文字を読み取ることができないという意見や講師がスライド中のどの箇所を話しているのかわからなくなるという意見が数多くあった。一方、SOI は操作性などで自由度が高く、たとえば、ユーザがコンテンツに対し自由に再生や停止などのコントロールが行えること・講師映像と資料を選択して視聴することができるという点で勝っているという意見もあった。

3. 講義映像の提示方法

2.6 節のアンケート結果と下記の従来研究¹⁹⁾ に従い、著者らは SOI のように 2 つの画面にそれぞれ講師映像と講義資料を表示させる形式で、画面の大きさを切り替えることで学習者に注目してほしい画面を提示する形式を採用した。視覚的フォーマットとその提示様式が情報の学習度に及ぼす影響について調査を行った結果によると学習内容に関係した映像と顔映像という 2 種類の視覚フォーマットが混在した映像提示様式の方が、いずれか 1 種類の等質型提示様式よりも学習効果を高めることが報告されている¹⁹⁾。

2 つの表示サイズは同じではなく、つねに一方は大きく、もう一方は小さく表示させることによって学習者に注目してほしい画面を提示する。これは NHK 教育番組とは見せ方は異なるが、映像を切り替えるとい

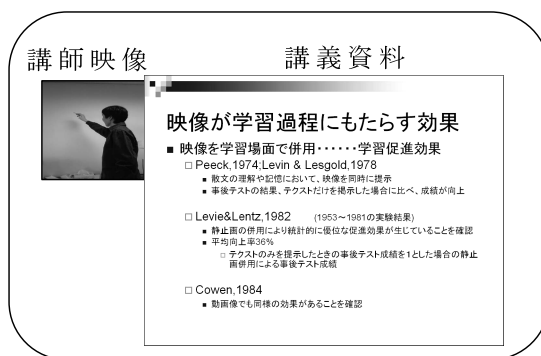


図 4 映像形式 1 (講義資料の拡大)

Fig. 4 The part1's video format (scale up the lecture materials).



図 5 映像形式 2 (講師映像の拡大)

Fig. 5 The part2's video format (scale up the lecturer).

う点で共通の形式であるといえる。この画面が切り替わる基準として、著者らは講師の発話内容に着目した。講師が講義資料の内容を説明しているときに講義資料を大きく、逆に講師映像は小さく表示する (図 4)。一方、講師が講義資料にない内容を長く話している場合、学習者は講義資料を注目する必要はないため、講師映像を大きく、講義資料は小さく表示する (図 5)。

放送大学と NHK 教育番組の映像比較でも示したとおり、同一の映像を長く提示し続けることは、視聴者に退屈感を与えてしまう恐れがあるため、あるタイミングで映像を切り替える必要がある。そこで、著者らは NHK 教育番組では、同一の映像が約 30 秒以上続くことがきわめて少なく、頻りに映像が切り替わるということに着目した。切替えのないひと続きの映像の持続時間と、映像の内容伝達との関連についての実験によれば、映像の持続時間が一定より長いと、それ以降の内容は見る者に伝わらないといわれている^{14), 20), 21)}。以上のことから、短い間隔で映像を切り替えることが視聴者に退屈さを感じさせにくくする暗黙のルールの一つではないかと考え、我々のシステムに適用できる

か検討した。

視聴中に 30 秒で講師映像から講義資料へ映像がスイッチされたとしても講師の声さえ聞こえていれば、特に問題はないと考えられるため、NHK 教育番組の知識を用いることが可能であると思われる。しかしながら逆のパターン（講義資料から講師映像へ）を考えた場合、やはり学習者が資料内の文章を読んでいるにもかかわらず、途中で講師映像に切り替わってしまうことは極力避けるべきである。そこで我々は学習者が講義資料 1 ページ分を読み終わるためにどのくらい時間を要するのかを調査し、30 秒のルールを適用することが可能であるか検討した。

SOI の講義資料 ($n = 578$) を 5 名の被験者に目を通させたところ、96.1% の確率で 30 秒以内に目を通すことが可能であることを確認した。ここで、映像の切替えタイミングを考えた場合、目を通すのに 30 秒以上かかるようなスライドが稀にあるということ considering、ある程度長く表示させ続けた方が望ましいように思われるが、著者らは学習者を飽きさせないことに重点を置いているため、強制的に 30 秒で映像を切り替えることにした。また、30 秒以上表示させたい場合も考えられるが、本システムはオンデマンド型であるため学習者はコンテンツの制御を行うことができる。したがって、再生ツールの機能として備わっている“巻き戻し”または“一時停止”ボタンをマウスでクリックすることで対応することができる。以下に映像の切替えルールを示す。

<形式 1：講義資料の拡大>

- 講師が講義資料内の語句を話した。
- スライドページが変わった。
- 講師拡大映像が 30 秒以上続いた。

<形式：講師映像の拡大>

- 講師が講義資料内の語句を話していない。
- 講義資料拡大映像が 30 秒以上続いた。

4. プロトタイプシステムの実装

4.1 MINO システム概要

MINO システム構成を図 6 に示す。本システムは、4 つのツールから成り立っており、講義中に音声認識ツール、PowerPoint 記録ツールが用いられ、講義終了後に重要語句検出ツール、コンテンツ生成ツールが用いられる。以下にコンテンツ生成までの手順を示す。

<講義開始前の作業>

- ① DV カメラの録画を開始する。
- ② 音声認識ツールを起動させ録音を開始する。
- ③ PowerPoint 記録ツールを起動させ「Record」ボ

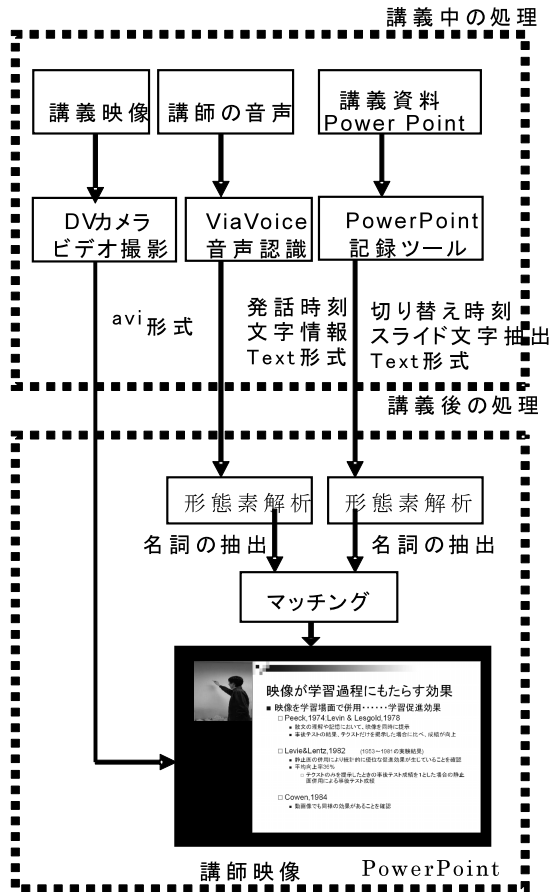


図 6 MINO システム構成図 Fig. 6 The structure of the MINO system.

タンを押す。

- ④ 講義用 PowerPoint を開き、スライドショーを開始する。

<講義終了後の作業>

- ⑤ 音声認識ツールと PowerPoint 記録ツールで記録した情報をテキスト保存する。
- ⑥ DV テープに記録した講師映像を PC にキャプチャし保存する。
- ⑦ 形態素解析ツールでテキストデータを読み取り、重要語句を決定し、テキスト出力する。
- ⑧ コンテンツ生成ツールを起動し、撮影した講師動画と講義に用いた PowerPoint ファイルと形態素解析ツールで出力したテキストファイルを読み取り、コンテンツを作成する。

4.2 講義中のイベントの取得

講義は 1 台の PC を用いて行われ、講師の映像はデジタルビデオカメラによって DV テープに記録される。なお講師映像、音声情報、スライド情報の同期をとる



図 7 音声認識ツール

Fig. 7 The automatic speech recognition tool.



図 9 形態素解析ツール

Fig. 9 The tool for morphological analysis.



図 8 PowerPoint 記録ツール

Fig. 8 The tool for record the events of PowerPoint.

ために、事前にビデオカメラの時計を PC の内部時計に合わせておく。また、PC に接続したマイクによって講師の音声録音され、発話は ViaVoiceSDK²²⁾ を用いて作成した音声認識ツール(図 7)によって文字列に変換される。PowerPoint スライド操作は、作成した PowerPoint 記録ツール(図 8)によってイベントとして記録される。記録される情報は、PowerPoint スライドショーの開始時刻、終了時刻、スライドページが変わった時刻、スライドの切り替わった順序、スライド内の文字列である。これら情報の取得には Microsoft Office の VBA オートメーションの仕組みを利用している。

音声情報は図 7 のように文字列に変換され、各語句にはそれぞれ時刻が付加される。PowerPoint 記録ツールではスライド単位で時間情報を管理している。

4.3 重要語句の決定

重要語句の決定は音声認識ツールと PowerPoint 記録ツールで出力したテキストデータを比較して行われる。スライド内の語句を講師が発話した場合、これを重要語句として扱うようになっている。そして、語句のマッチングを行うために形態素解析ライブラリ(ChaSen²³⁾)を用いた解析ツールを作成した(図 9)。形態素解析を用いると、文脈上意味を持つ最小の単位(形態素)を抽出することができる。抽出された語句

は形態素辞書と照合され、品詞に分けられる。ここで、マッチングする語句を名詞に絞り、形態素解析によって分けられた品詞の中から名詞だけを抽出することにした。これは講師の発話内容に注目すると、ポイントとなる名詞は必ず発話されるが、動詞や助詞に関しては、資料とは異なった動詞や助詞が用いられる傾向があったためである。

マッチングは講師が発話した名詞と PowerPoint スライド内の名詞の出現時刻の比較によって行われる。まず、発話された名詞の出現時刻とスライドが切り替えられた時刻とが比較される。その後で文字列のマッチングが行われ、同一の名詞が存在した場合、その名詞は重要語句として決定される。

4.4 重要語句の強調

講師の発話に同期してスライドの文字を強調表示するために以下の処理を行うようにした。重要語句が決定すると同時にその語句が発話された時刻が記録される。この時刻は重要語句が強調されるタイミングとなる。しかしながら、スライド内に複数の同一の語句が存在した場合、不都合が生じてしまう。そこで、発話語句とスライド内の語句を 1 対 1 で比較するのではなく、前後の語句を含めて比較することにした。これによって複数存在した語句ごとに何番目に出現した語句であるかが決定され、それぞれ発話時刻が記録されるようになっている。決定された重要語句とそのタイミングはテキストファイルとして出力される。

スライド内の文字の強調は、抽出された語句が含まれる文字列全体で行われる。誘目性を考慮した場合、背景色との明暗の差が大きく、動きがあるものに自然に目が惹き付けられる。この条件に基づいて、本実装では、重要な文字列は赤色に変わり、徐々に拡大され、しばらくして縮小して元に戻るといった動きをつけることにした。なお、一度強調した箇所は再度強調しないようにした。これは講師が話すたびに同じ箇所が何度も強調されるのは煩わしいと考えたためである。

最後に、システムはコンテンツ生成ツールにて講師

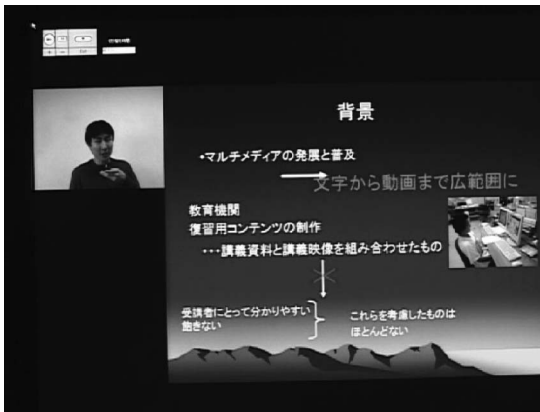


図 10 MINO が作成したコンテンツ映像
Fig. 10 The image of contents that MINO made.

映像の動画ファイルと講義に用いた PowerPoint ファイル、そして重要語句検出ツールで出力したテキストファイルを用いて、講義映像コンテンツの作成を行う。

図 10 に MINO システムによって作成した講義映像コンテンツの再生画面を示す。画面左上には再生・停止・一時停止ボタンがあり講義の進行をコントロールできるようになっている。その下には講師映像があり、縮小表示されている。画面右側には大きく講義資料が表示されている。この状況は講師が講義資料内の語句を話している状況である。講師が発話すると同時に講義資料内の当該語句の文字列フォントがアニメーションとともに徐々に拡大され、また赤く変化していく。

5. 評価と考察

講師映像と講義資料を単純に並列に表示させた SOI 形式と、提案した表示方法の 2 通りのコンテンツを作成し、評価実験を行った。被験者 10 名は大学 4 年生（そのうち遠隔講義受講の経験がある者は 2 名）であり、被験者に視聴させた映像は卒業研究を題材とし、その研究を担当した学生に本システムを利用してプレゼンテーション 10 分間を行わせた。その後、講師映像と講義資料を単純に並列に表示させた SOI 形式と、著者らが提案した表示方法の 2 通りのコンテンツを作成し、被験者に視聴させた。また、意図的に集中してしまうと、アンケート項目の 1 つである「飽きる」という要素を正確にとらえることができないため、被験者には特に講義に集中させず、ごく自然に視聴を行わせた。そして、各項目に対して 1~5 点の点数によって回答させ比較を行った。

アンケートの結果、表 3 に示すように「自然である」という項目を除いて、すべての項目で本システムの方が良いという回答が得られた。ウィルコクソン符

表 3 評価結果

Table 3 The result of evaluations of the system.

アンケート項目	A	B	Wilcoxon 符号付順位和 検定 p 値
理解しやすい	3.3	4.0	0.06833
見易い	1.8	3.8	0.00009
飽きにくい	2.5	4.5	0.00019
魅力がある	2.5	3.8	0.00088
自然である	4.0	3.0	0.00668
また視聴したいと思う	1.4	4.1	0.00006

A: 従来の 2 画面形式講義映像 (SOI 形式)

B: 本システムが作成した講義映像

号付き順位和検定²⁴⁾によると、「理解しやすい」という項目以外で有意差 ($p < 0.01$) が認められた。「理解しやすい」という項目で有意差は認められなかった原因としては表示方法に工夫を加えても内容が難しい場合、理解しにくい箇所は理解しにくいままであるため、評価結果に差が表れなかったのではないかと考えられる。しかし、その他の項目では、従来の提示方法よりも高い評価を得ることができたので、本稿の目的であった「講師がどの部分を説明しているのかが分かりやすい」と「学習者を飽きさせない」ということを達成することができたと考えている。特に「飽きにくい」という項目では評価が高く、著者らが目標としていた「学習者を飽きさせないコンテンツ」を作成することができたのではないと思われる。また、「どちらの講義映像をまた視聴したいか」という項目では被験者全員が、本提案手法のコンテンツを選択した。このほか、画面の切替えと講義資料内の文字列が変化することによって、注視点を誘導してくれるため、「講師の発話に対応した箇所を自分で探すという作業をしなくてもよい」という感想を多くの被験者から得た。そのほかの意見としては、「本当に必要な語句のみを強調させてほしい」、「どちらの映像に注目すべきか画面の大きさだけでなく、案内表示が欲しい」といった意見が得られた。

今回実装したプロトタイプシステムでは、スライドに載せられた語句は基本的にすべて重要な語句であるとし、発話語句すべての強調を行っていた。しかしアンケートにもあったように本当に重要な語句のみを強調させるほうが望ましい。その方法として、文字のフォントの違いによって検出する方法が考えられる。重要語句に関しては、スライドの作成者はその語句を枠で囲ったり、フォントをまわりの文字と区別して用いたりすることがある。したがって PowerPoint 記録ツールに新たに文字のフォントプロパティの取得機能を追加することで、より重要な語句を抽出できるので

はないかと考えられる。

6. おわりに

本稿では、従来の講義コンテンツの映像構成を分析した結果から講義映像の新しい提示方法を提案し、講義映像コンテンツ作成支援システム MINO を構築した。従来のコンテンツに比べ講師が資料中のどこを話しているかが分かりやすく、また画面の切替えが自動的に起こるため、見ていて飽きにくい映像を作成することができた。今後の課題として、講義資料の重要語句の決定、さらに文字列だけでなく図や表などの強調なども行いたいと考えている。

参考文献

- 1) 仲林 浩：e-Learning の要素技術と標準化，情報処理学会論文誌，Vol.43, No.4, pp.401-406 (2002).
- 2) 経済産業省商務情報政策局情報処理振興課：eラーニング白書，オーム社 (2005).
- 3) 柳沼良知：SMIL を用いた映像教材作成支援ツールの開発，情報処理学会研究報告，CE-73 (2004).
- 4) James, D. and Hunter, J.: A streamlined system for building online presentation archives using SMIL December 2000, *Proc. Australasian conference on Computing education ACSE* (2000).
- 5) 石井光雄：講義録画を利用した復習用 LIVE 教科書の開発，情報処理学会研究報告，CE-74 (2004).
- 6) Hauptmann, A.G. and Smith, M.A.: Text, Speech, and Vision for Video Segmentation: The Informedia Project, *AAAI Fall Symposium* (1995).
- 7) SMIL. <http://www.w3.org/TR/REC-smil/> (2005/4/20 アクセス).
- 8) 菅原，三木：レーザポインタを用いた遠隔講義システム，電気通信大学紀要，0000444962 JPN pp.117-123 (2002).
- 9) 岡部成玄：知識社会における情報通信技術を活用した高等教育—高等教育改革に資するマルチメディアの高度利用に関する研究，平成 14 年度研究成果報告 (2004).
- 10) 福田忠彦：感覚の整理と心理，Wide University School of Internet. http://gc.sfc.keio.ac.jp/class/2003_14454/slides/05/index_108.html (2005/5/1 アクセス).
- 11) Wogalter, M.S.: Effect of study to test maintenance and change of photographic mode and pose, *Applied Cognitive Psychology*, Vol.1, pp.241-253 (1987).
- 12) Wide University School of Internet. <http://www.soi.wide.ad.jp/> (2005/4/20 アクセス).
- 13) 井上，吉田，平石，重野，岡田，松下：映画の映像理論に基づく対面会議シーンの自動撮影手法，情報処理学会論文誌，Vol.45, No.1, pp.212-221 (2004).
- 14) 井上，岡田，松下：テレビ番組のカメラワークの知識に基づいた TV 会議システム，情報処理学会論文誌，Vol.37, No.11, pp.2095-2104 (1996).
- 15) 中島義明：映像の心理学，サイエンス社，pp.194-195 (1996).
- 16) 岡田 普：映像学・序説，p.260, 九州大学出版会 (1996).
- 17) ジェイムズ・モナコ (著)，岩本憲児ほか (訳)：映画の教科書，p.438, フィルムアート社 (1983).
- 18) 国際電気通信基礎技術研究所：視聴覚情報科学，p.271, オーム社 (1994).
- 19) Berry, C. and Brosius, H.-B.: Multiple effects of visual format on TV news learning, *Applied Cognitive Psychology*, pp.519-528 (1991).
- 20) 松井洋泰，福岡 崇：簡易番組制作の実証とデジタル映像制作技術に関する研究，京都府中小企業技術センター技報，No.32 (2004).
- 21) 山岸達児：映像の発想，p.240, 教育出版センター (1992).
- 22) IBM ViaVoiceSDK. http://www-306.ibm.com/software/voice/viavoice/dev/sdk_linux.html (2005/4/20 アクセス).
- 23) chasen. <http://chasen.aist-nara.ac.jp> (2005/4/20 アクセス).
- 24) 成田 滋，筱 更治：Excel によるデータ解析，エコノミスト社 (2001).

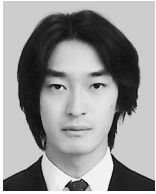
(平成 17 年 6 月 1 日受付)

(平成 17 年 11 月 1 日採録)



中村 亮太 (学生会員)

2002 年東京工科大学工学部情報通信工学科卒業。2004 年同大学院前期博士課程終了。現在、慶應義塾大学大学院理工学研究科後期博士課程在学中および工科大 Linux オープンリサーチセンター研究員。ヒューマンインタフェース、マルチメディアオーサリングの研究に従事。2004 年 DICOMO 優秀プレゼンテーション賞。2005 年 DICOMO 最優秀プレゼンテーション賞および優秀論文賞受賞。



井上 亮文 (正会員)

1999年慶應義塾大学理工学部計測工学科卒業。2001年同大学院前期博士課程修了。2005年同後期博士課程修了。博士(工学)。現在、東京工科大学コンピュータサイエンス学部助手。マルチメディアオーサリング、実世界志向インタフェース、ネットワークセキュリティの研究に従事。



市村 哲 (正会員)

1989年慶應義塾大学理工学部計測工学科卒業。1994年同大学院理工学研究科博士後期課程修了。博士(工学)。同年富士ゼロックス(株)入社。1997~1999年富士ゼロックスパロアルト研究所(FXPAL)駐在。2002年より東京工科大学助教授。グループウェア、ネットワークサービス、生体情報活用等の研究に従事。『IT TEXT 基礎 Web 技術』、『IT TEXT 応用 Web 技術』(オーム社)。DICOMO 2003 & DICOMO 2005 優秀論文賞受賞。ACM, 電子情報通信学会各会員。



岡田 謙一 (フェロー)

慶應義塾大学理工学部情報工学科教授, 工学博士。専門は, CSCW, グループウェア, ヒューマン・コンピュータ・インタラクション『ヒューマンコンピュータインタラクション』(オーム社)、『コラボレーションとコミュニケーション』(共立出版)をはじめ著書多数。情報処理学会誌編集主査, 論文誌編集主査, GW研究会主査等を歴任。現在, 情報処理学会 GN 研究会運営委員, BCC 研究グループ幹事, 日本 VR 学会 CS 研究会副委員長。1996 および 2001 年情報処理学会論文賞, 2000 年情報処理学会 40 周年記念論文賞, 日本 VR 学会サイバースペース研究賞, IEEE SAINT 04 最優秀論文賞受賞。電子情報通信学会, 情報処理学会フェロー, IEEE, ACM, 人工知能学会各会員。



松下 温 (フェロー)

1963年慶應義塾大学工学部電気工学科卒業, 1968年イリノイ大大学院コンピュータサイエンス専攻修了, 1989~2002年慶應義塾大学理工学部教授, 2002年より東京工科大学教授, 2003~2005年東京工科大学コンピュータサイエンス学部長。マルチメディア通信, コンピュータネットワーク, グループウェア等の研究に従事, 情報処理学会理事, 同学会副会長, マルチメディア通信と分散処理研究会委員長, グループウェア研究会委員長, 電子情報通信学会情報ネットワーク研究会委員長, MIS 研究会委員長, パーチャルリアリティ学会サイバースペースと仮想都市研究会委員長, 情報処理学会 ITS 研究会委員長などを歴任。郵政省, 通産省, 農水省, 建設省, 都市基盤整備公団, 行政情報システム研究所等の委員長, 座長, 委員を多数歴任。特に国土交通省, 住宅情報化標準策定委員会委員長, 経済産業省総合エネルギー調査会電子計算機と磁気ディスク委員会委員長, 経済産業省総合エネルギー調査会ルータ装置基準委員会委員長, 最高裁判所専門委員を務める。『やさしい LAN の知識』(オーム社)、『201x 年の世界』(共立出版)等著書多数。1993年情報処理学会ベストオーサ賞, 1995 および 2000 年情報処理学会論文賞, 2000 年情報処理学会 40 周年記念論文賞, 2000 年日本 VR 学会サイバースペース研究賞, 2001 年情報処理学会功績賞受賞, 情報処理学会フェロー, 電子情報通信学会フェロー, 人工知能学会, ファジイ学会, IEEE, ACM 各会員。