

講師の意図を考慮した板書講義コンテンツの自動生成

品田 良太[†], 井上 亮文^{††}, 星 徹^{††}

[†] 東京工科大学大学院 バイオ・情報メディア研究科

^{††} 東京工科大学 コンピュータサイエンス学部

板書講義配信システムでは、講師の除去や追跡をするカメラワークなど、講義の見せ方に重きが置かれてきた。本稿では、板書講義の内容にそってアクセシビリティを高めたコンテンツを自動生成する手法を提案する。講師は固定されたビデオカメラの前で通常通り講義をする。この際、板書の重要箇所カラーマグネットを貼り付けていく。この動画からマグネットが貼られた位置、時間、付近の画像を切り出し、内容をあらわす見出しやシーンアクセスのためのインデックスとして利用する。本論文では、プロトタイプシステムの作成とその実装について述べる。

Automatic Contents Generator for a Chalk Talk Reflecting the Lecturer's Intention

Ryota SHINADA[†] Akifumi INOUE^{††} Tohru HOSHI^{††}

[†] Graduate School of Bionics, Computer and Media Sciences, Tokyo University of Technology

^{††} School of Computer Science, Tokyo University of Technology

E-learning systems for chalk talk lectures have focused on the appearance such as clearing and tracking the lecturer's figure. In this paper, we propose a method to generate contents focused on the accessibility. In this method, a lecturer has a chalk talk in front of a fixed video camera as always. The lecturer may stick some color magnets as necessary near to the important part of the notes. The system detects the positions and appearances of the magnets in the recorded movie. Then the system clips the area between a pair of the magnets as an image. Those images are used as headlines and index of the corresponding scene of the lecture. The prototype and its demonstrations are described.

1 はじめに

電子スライドを用いた講義では、スライドの切り替わりが内容の区切りを表す。一方で板書を用いた講義では、1枚の黒板を時間的・空間的に連続して説明をする。内容の区切りは主として口頭で伝達され、これをシステムが自動的に認識するのは困難である。

このような問題も影響してか、これまでの板書講義配信システムでは、板書の一部や全体をいかに見やすくするかに重きを置いてきた。電子スライド講義の配信では一般的な、特定シーンへ移動するイン

デックスの生成など、講義の内容に踏み込んだコンテンツの生成はほとんど考慮されていない。

結果として受講者には、シラバスにあるような一般的な情報や、シークバーや単位時間でのシーンアクセスしか提供されていない。特に復習を目的とする受講者にとって、これら貧弱な手段のみで見たい講義や見たいシーンを探し出すのは困難である。これでは学習を支援するどころか、動画を見ようという気持ちも削いでしまいかねない。

本稿では、講師が普段から行う自然な方法で、話の区切りや板書中の重要箇所をシステムが認識しやすくする手法を提案する。この際、特別な機材を用

いづとも容易にシステムを利用できるようにする。

本稿の構成を以下に示す。第2章では関連研究に触れながら、既存システムの問題点を挙げる。第3章ではその解決方法の提案と提案するシステムについて述べ、第4章ではプロトタイプシステムの実装方法について述べる。第5章では考察を行い、最後に第6章でまとめと今後について述べる。

2 関連研究

講義を撮影して配信する研究には、スライド講義を前提としたものと板書講義を前提としたものがある。

スライド講義を対象とした研究では、井上らはスライドのフッタに付与したラベルから、動画のインデックスを作成している¹⁾。中村らはスライド中の語句や講師の発話から、講義映像コンテンツを生成している²⁾。これらの研究では、シーンアクセスのためのインデックスの生成や、講師の言葉からスライドの文字を強調して表示など、スライドの内容まで踏み込んで構成・配信を行っている。

板書講義を対象とした研究では、市村らは講師の動画と講師の姿を除去した黒板画像を同時に配信している³⁾。芦川らは板書中のチョークの音を検出して、システムが自動的にズームなどの撮影動作をしている⁴⁾。大西らは板書講義の動画から板書の内容をブロック分割して、画像を切り出して Web ページとして配信している⁵⁾。これらの研究では、板書から講師を除去して板書を見やすくする、板書の内容を切り出して Web ページとして配信するなど、板書講義を見やすく配信することに重きが置かれている。

板書講義でも内容に踏み込んだ配信ができれば、より学習者の支援につながると考える。板書中のキーワードの見出しやインデックスを生成できれば、講義中の内容の把握や動画のシーンアクセスを実現できる。しかし、これまでは手動による動画編集によって実現されており、人的コストの問題や成果物にばらつきが出る問題があった。

3 提案

講師が普段から行う行動（囲みや下線などのマーキング）から、コンテンツを生成する手法を提案する。

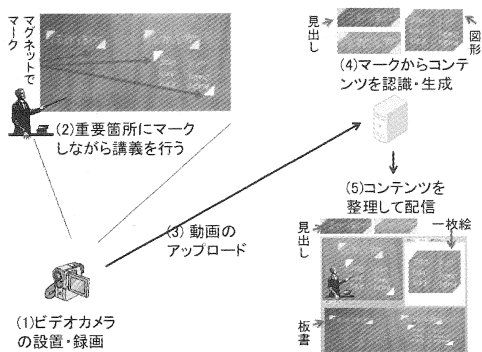


図1 システムの流れ

3.1 提案システム

提案手法を用いたシステムの流れを図1に示す。(1) 講師は講義前に黒板全体が映る位置にビデオカメラを設置し、撮影を開始する。(2) 講義中に任意の箇所にマグネットを付けながら講義を行う。(3) 講義終了後に撮影した動画をシステムにアップロードする。(4) システムでは、講義中に講師が付けたマグネットを認識してマグネットに応じたコンテンツの生成を行う。(5) 生成されたコンテンツと板書講義動画を整理して Web で配信する。

提案システムにおいて講師が行う作業は(1)(2)(3)だけである。これらの作業はどれも非常に簡単なものであり、特別な機械操作の知識を必要としない。そのため、電子機器の慣れ不慣れに関わらず、誰でもシステムを利用できる。

システムではアップロードされた板書講義の動画から画像処理などを用いて、内容に踏み込んだコンテンツを生成する。それらを整理して動画とともに配信する。動画の編集はシステムがすべて自動で行ってくれるため、講師1人でも容易に講義を配信することができる。

3.2 撮影機材

撮影に必要な機材は、Full-HD 対応のビデオカメラ1台とカラーマグネットシート数枚だけである。したがって、撮影設備を備えていない普通の教室でも、手軽にシステムを利用することができる。Full-HD 対応の理由としては、人が板書の文字を認識するのに必要な解像度を確保するためである。また、黒板全体が映るようにカメラを設置して撮影し

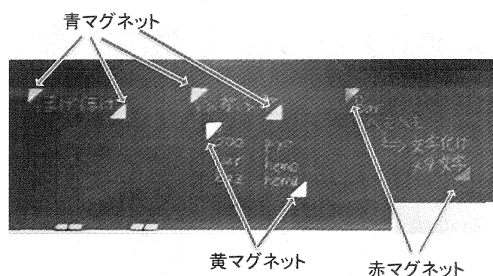


図2 マグネットが付与された黒板

アップロードするだけでよい。ため、動画の撮影や編集を行う人間は不要である。

3.3 板書講義中のマグネットの付与

講師は図2のように、重要だと思う箇所、覚えてほしい図表・数学の公式、説明の区切りなどを、カラーマグネットシートで囲むだけである。一般に講師は板書の内容を強調する際、チョークで囲んだり下線を引いたりする。マグネットの付与はそのような一般的慣習と類似している上、負担も同程度である。マグネットで囲むタイミングも任意でよい。

3.4 コンテンツの生成と配信

システムは動画中に講師が付けたマグネットの色と形状・位置・時間からコンテンツの生成を行う。マグネットには囲いの開始を意味するものと終了を意味するものがあり、この2つを認識してコンテンツは作られる。コンテンツの種類はマグネットの色によって定義される。

講師がマグネットを付けた位置や時間から、見出しやインデックスとなる画像、まとまった板書の一枚絵画像、任意の区間を切り出した動画が生成される。講師が明示的にマグネットを付けることで、講師の考えや意図を考慮したコンテンツを生成する。

配信ページでは、生成されたコンテンツを整理して表示する。インデックスにより、動画の特定のシーンへの移動を実現する。見出しや一枚絵画像により、講義の大まかな内容の取得を支援する。区間の動画により、任意の説明だけの学習を支援する。また、講義中の講師付近のみを切り出した動画も配信する。

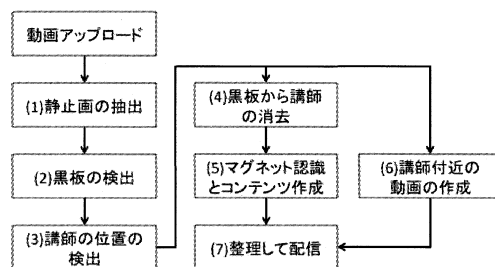


図3 実装システムの動作概要

4 実装

4.1 システムの動作概要

プロトタイプシステムの動作概要図を図3に示す。(1) システムでは、板書講義の動画がアップロードされると動画の中から数秒毎に静止画を抽出する。(2) 抽出した静止画から黒板領域のみを切り出した画像を作成する。(3) 黒板画像から講師が存在する位置を特定する。(4) 特定した講師を黒板画像から除去する。(5) 講師を除去した黒板画像からマグネットを認識し、コンテンツを生成する。(6) 黒板画像中の講師付近のみを切り出した動画を作成する。(7) (5) で生成されたコンテンツと (6) で作成された動画を整理して配信する。

4.2 黒板の抽出と講師の除去

全体画像から黒板のみの画像の抽出には RGB 表色系を用いた。今回の撮影環境で黒板の成分を分析した結果は式(1)となった。この値を用いて2値化、ラベリングをする。その結果、もっとも領域が大きい部分を板書と判断し、領域の上下左右の座標から黒板画像を切り出す。

$$\text{黒板領域の RGB 値} = \begin{cases} R < 40 \\ G < 70 \\ B < 70 \end{cases} \quad (1)$$

切り出された板書画像から講師の位置の特定と除去には、フレーム間差分法⁶⁾を用いた。講師の位置抽出と除去の流れを図4に示す。講義撮影を開始した直後は黒板領域に誰も映っていないと考えられる。その時の画像を最初の画像 (img[i]) とし、次の講師画像 (img[i+1]) との差分をとることで、講師の位置を特定する。追加された板書・黒板消しの

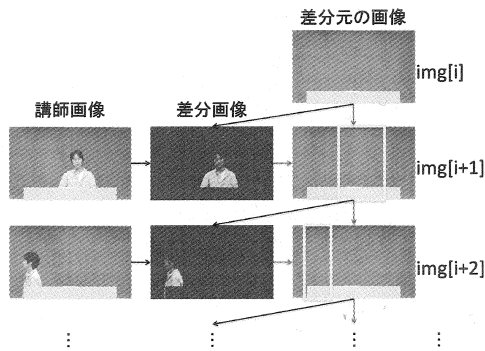


図4 講師の認識と除去

移動，ノイズなどと講師を区別するために，ラベリング処理を行う．もっとも変化量が多い領域を講師と判断する．認識された講師の領域を，元の画像 ($img[i]$) の画像で上書きして，新たな画像を作成する．以降，新たに作成された画像を次の差分元の画像 ($img[i+1]$) として，次の講師画像 ($img[i+2]$) と差分をとり，講師の位置の特定と除去を繰り返す．

4.3 講師付近を切り出した動画の生成

板書講義において講師がいる場所が，現在説明中の箇所であると考えられる．黒板全体の動画を配信するよりも，講師付近を切り出した動画を配信するほうが，より学習の支援になるといえる．そこで，撮影した板書講義から，擬似的に講師付近のみを切り出した動画を作成する．

まず図5の様に，フレーム間差分法で認識した講師の座標を中心に，縦横比が4:3になるように画像をトリミングする．次に，トリミングを行った画像を時系列に並べてつなげることで，画像だけの動画を作成する．最後に，作成した動画に元の音声を合わせることで，講師付近を切り出した動画を作成する．

4.4 マグネットの認識

講師が付けるマークには，直角二等辺三角形のカラーマグネットシートのペアを用いた．マグネットの色は青，黄，赤を用いて表1のように定義を行った．青のマグネットは，講義中に登場する単語の中で，見出しやインデックスとなる部分を囲うのに用いる．黄のマグネットは，講義中に登場する図表や数学の公式など，まとまった箇所を囲うのに用い

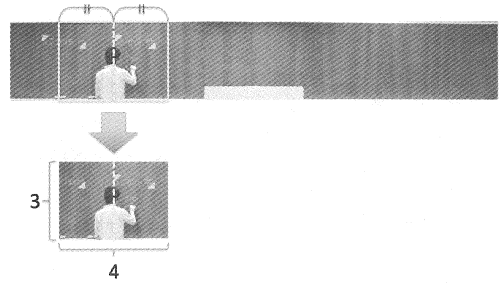


図5 講師付近の画像のトリミング

表1 マグネットの色と生成コンテンツの定義

マグネットの色	生成コンテンツの定義
青	説明や話の見出し・インデックス
黄	図表や数式の公式
赤	説明や話の開始と終了

る．赤のマグネットは，講義中の説明や話の区切りを認識するのに用いる．

マグネットの認識はRGB表色系でRGBの値を用いて行った．実際に成分分析をした結果，今回の実験環境では各色のRGB値は式(2)～式(4)となった．この値を用いて，画像を2値化しラベリング処理を行い，各マグネットを認識する．

$$\text{青マグネットのRGB値} = \begin{cases} B > 150 \\ B > R \times 2.0 \\ G > 100 \end{cases} \quad (2)$$

$$\text{黄マグネットのRGB値} = \begin{cases} R > 150 \\ R > B \times 1.4 \\ G > 150 \end{cases} \quad (3)$$

$$\text{赤マグネットのRGB値} = \begin{cases} R > 100 \\ R > G \times 1.3 \\ R > B \times 1.3 \end{cases} \quad (4)$$

4.5 コンテンツの生成

コンテンツ生成の概要図を図6に示す．コンテンツの生成処理は，講師の除去された画像を時系列的に現在から過去にさかのぼりながら行う．

見出しやインデックス (青マグネット) の場合は，最初にマグネットのペアを認識したとき (t) を，そ

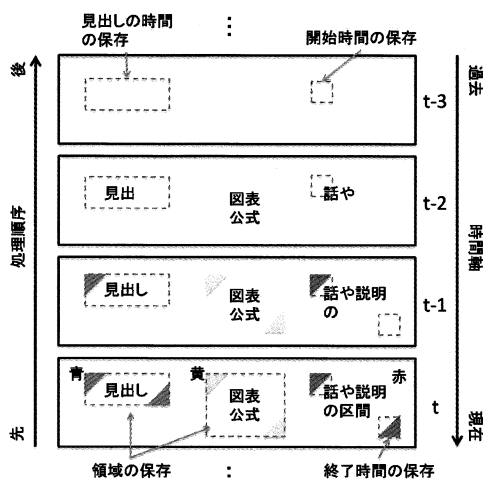


図 6 コンテンツ生成の概要

の見出しがすでに書かれた状態と判断し、(t)の時点の画像を保存する。時間をさかのぼり、マグネットで囲われた領域に書かれていた板書がなくなった(t-3])と判断したら、その時点(t-3])の時間を保存する。この時間が動画中のインデックスとなる。

一枚絵画像の切り出し(黄マグネット)の場合は、最初にマグネットのペアを認識したとき(t)を、その図表が完成された状態と判断し、(t)の時点の画像を保存する。以降の処理では同じ領域を認識しても無視をする。

動画区間の切り抜き(赤マグネット)の場合は、終了マグネットを認識(t)してそれが消えた(t-1)から消える直前(t)を説明の終わりと判断してその時間(t)を保存する。時間をさかのぼり、開始マグネットを認識して、開始マグネット付近の板書がなくなったとき(t-3)を説明の始まりと判断して、その時間(t-3)を保存する。このようにして得た開始時間と終了時間から、動画の区間の切り抜き処理を行う。

4.6 コンテンツの整理と配信

講義一覧ページの例を図 7 に示す。一覧ページではアップロードした講義動画が時系列に並んでいる。各講義には、アップロード時に設定するタイトルと講義の見出し画像、一枚絵画像が表示される。見出しや一枚絵画像が表示されることで、タイトル

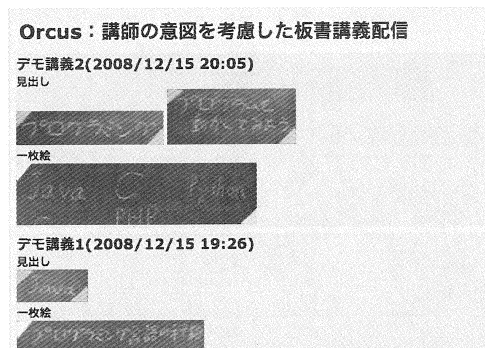


図 7 講義一覧ページの例

だけではわからない講義の内容を知ることができる。また、復習者には講義の内容を思い出す手がかりにもなる。

映像配信ページの例を図 8 に示す。ページ上部(1)では、青のマグネットから生成した見出し画像を、タグとして表示している。タグを見ることで、講義でどのような話がされたのかを、大まかに知ることができる。ページ中央左(2)では、講師付近を切り出した動画が配信されている。ページ下部(4)では、講師を除去した画像が表示されている。この画像は動画の時間によって自動的に切り替わる。ページ中央右(3)では、マグネットの認識から生成したコンテンツの配信を行っている。これにより、特定シーンへのアクセスや図表の一枚絵画像の表示、話の一区切りの動画配信を実現している。

5 議論

今回、RGB 表色系を用いて黒板やマグネットの認識を行った。しかし、環境光の状況や用いるマグネットの条件によっては、提示した RGB 値でうまくいくとは限らない。そこで、RGB 表色系の代わりに HSV 表色系を用いる方法が考えられる。HSV 表色系では輝度の変化に影響されずに色を識別できる。また、利用環境に合わせて色を調節するキャリブレーション機能の実現も考えられる。撮影毎に色を調節することで、より高い精度で認識できると考える。

コンテンツの生成にはマグネットのペアを用いた。講師の負担をさらに減らすには、用いるマグ

デモ講義2(2008/12/15 20:05)

(1) 見出し

タグ:

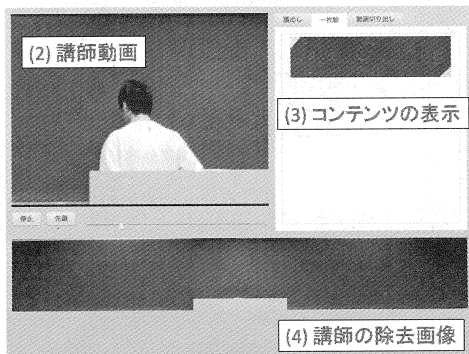


図 8 映像配信ページの例

ネットの数を減らすことが考えられる。これは、文献 5) のブロック分割手法を用いることで、実現できると考える。マグネットが付けられた位置からブロック分割を行い、領域を切り出すことで、1 枚のマグネットで画像を切り出すことができるようになる。と考える。

見出しの表示は、画像をそのまま用いているため検索には向かない。しかし文字認識のために、講師にきれいな字を書くことを意識させては、手軽にシステムを利用できない。この問題は、視聴者が画像の見出しを文字に変換できる仕組みを実現することで解決できる。また、任意の見出しを付けることができるようになれば、より講義内容を表したものになると考える。

6 おわりに

本稿では、講義中に講師が付与したマグネットから、板書の内容に踏み込んだコンテンツを生成する手法を提案し、プロトタイプシステムの試作を行った。システムは 1 台のビデオカメラと数枚のマグネットだけで利用できる。普段と同じように講義を行うだけで、自動的にコンテンツの生成から配信までを行ってくれる。今回は黒板の講義を対象としたが、マグネットの色が認識できれば、ホワイトボードなどでも利用できると考える。

今後はマグネットの色やマークの種類を増やすこ

とで、生成できるコンテンツの種類を増やしていく。また、アップロードされる講義が増えた場合に、分類と検索が必要になる。その際の表示方法についても検討をしていきたい。

参考文献

- 1) 井上宗徳, 下川俊彦. 講義スライドのフッターを用いたラベル付けによる講義映像のインデックス作成に関する研究. 電子情報通信学会技術研究報告. E.T, 教育工学, Vol. 107, No. 391, pp. 1-6, 20071208.
- 2) 中村亮太, 井上亮文, 市村哲, 岡田謙一, 松下温. 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. 情報処理学会論文誌, Vol. 47, No. 1, pp. 172-180, 20060115.
- 3) 市村哲, 福井登志也, 井上亮文, 松下温. Web 学習用講義コンテンツを自動作成する板書講義収録システム. 情報処理学会論文誌, Vol. 47, No. 10, pp. 2938-2946, 20061015.
- 4) 芦川平, 菅沼明, 谷口倫一郎. 黒板講義におけるチョーク音検出を利用した講義自動撮影システムの構築と評価. 情報処理学会九州支部研究会報告, pp. 60-67, 2003.
- 5) 大西正輝, 泉正夫, 福永邦雄. 講義映像における板書領域のブロック分割とその応用. 電子情報通信学会論文誌. D-I, 情報・システム, I-情報処理, Vol. 83, No. 11, pp. 1187-1195, 20001125.
- 6) 田中靖哲, 長谷部齊志, 菊池春秀, 下中智, 川合亘, 橋本周司, 富永将史, 渡辺博己, 清水早苗, 飛谷謙介. 『人物』検出手法に関する研究開発. ソフトピアジャパン共同研究報告書, 2007.