

推薦論文

板書の意識的な強調を利用した 復習用コンテンツ自動生成システム

井上 亮文^{1,a)} 品田 良太² 市村 哲¹ 星 徹¹

受付日 2011年4月18日, 採録日 2011年10月3日

概要: 本論文では, 講師がふだんから行う板書の強調を利用した講義動画のインデックスを自動生成するシステムを提案する. 講師は固定されたビデオカメラの前で通常どおり講義をする. その際, 板書の見出しや強調したい重要箇所にカラーマグネットを貼って強調する. システムでは, 強調の位置・時刻を利用して画像を切り出し, 講義内容を表す見出しや動画のシーンアクセスのためのインデックスとして利用する. 模擬講義を用いたシステムの使用感の評価実験では, マグネットの付与は講義の進行を妨げないという結果が示唆された. また, システムの性能評価では講義中に付与されたマグネットから, 89%のインデックスを自動生成することができた.

キーワード: 教育支援, 講義収録, E ラーニング

An Indexing System for Chalk Talk Videos Based on Conscious Highlighting

AKIFUMI INOUE^{1,a)} RYOTA SHINADA² SATOSHI ICHIMURA¹ TOHRU HOSHI¹

Received: April 18, 2011, Accepted: October 3, 2011

Abstract: In this paper, we propose an indexing system for chalk talk videos by detecting conscious highlighting marks on a blackboard. All a lecturer needs to do in his/her lecture is attach colored magnets near the important part of the notes, such as the title of a block. The system detects each pair of the magnets from the video, and then clips out the region inside the magnets as a still image. The image is used as an image tag so that a student can easily see what was talked on the lecture. The image is also used as an index to jump to the corresponding speech in the video. The experimental result by mock lectures showed that attaching the magnets didn't disturb the progress of a lecture. System performance evaluation showed that 89% of the magnet pairs could be detected correctly.

Keywords: education support, lecture archiving, E-learning

1. はじめに

E-learning の一環として, 講義動画の配信が行われている. すでにいくつかの大学では, 学内だけでなく YouTube

などでも講義動画を公開している [1], [2], [3]. 講義動画は, その講義を初めて受講する学生による新規視聴と, 受講済みの学生による復習視聴の 2 種類の形態が考えられる. 本論文では後者を扱うこととする.

学生がある講義を受講済みの場合, その概要は把握している. よって復習視聴では, 新規視聴のように動画を最初から再生することは少ない. それよりも, 重要な公式や講

¹ 東京工科大学コンピュータサイエンス学部
School of Computer Science, Tokyo University of Technology, Hachioji, Tokyo 192-0982, Japan

² 東京工科大学大学院バイオ・情報メディア研究科
Graduate School of Bionics, Computer and Media Sciences, Tokyo University of Technology, Hachioji, Tokyo 192-0982, Japan

^{a)} akifumi@cs.teu.ac.jp

本論文の内容は 2010 年 7 月のマルチメディア, 分散, 協調とモバイル (DICOMO2010) シンポジウムにて報告され, グループウェアとネットワークサービス研究会主査により情報処理学会論文誌ジャーナルへの掲載が推薦された論文である.

義中に分かりにくかった説明を選択して視聴する。必要な説明に即座に移動できたり、その説明を時間・場所を問わず繰り返し確認できたりすれば、より効果的な復習が可能になる。

スライド講義では、事前に作成済みのスライドに沿って説明をしていく。1枚のスライドは見出しと内容から構成され、スライドの切替わりは内容の切替わりを示す。講義の進行と教材の構造は定型化されており、電子的に処理がしやすい。そのため、講義中の見たいスライドへ移動するなどのランダムアクセスが可能な動画を自動生成する研究や製品は多い [4], [5], [6]。

これに対し板書講義では、講師のノートを基準にしつつも、学生の理解度に応じて内容や進行が変化する。見出しや内容をどのように表現するかは講師によって異なる。黒板の時間的・空間的な使い方も一定ではない。次の説明を直前の板書の近くに続けて書いたり、離れた余白部に書いたり、思い出した説明を以前の板書に追記したりする。このように、書式やレイアウトが定型化されておらず、内容の切れ目を電子的に処理することは難しい。板書講義は数学や語学など学力の基礎となる科目で多く行われているにもかかわらず、今なお、講義動画の再生位置を手動で探したり、インデックスを手動で作成したりしているのが現状である。

本稿では、講師に特別な負担をかけることなく復習視聴のためにアクセシビリティを高めた動画（以下、板書講義コンテンツ）を自動生成するシステムを提案する。本システムでは、講師は板書の“見出し”、“図表”、“説明”をマグネットで強調しながら講義を行い、その様子を撮影しておく。システムはこの動画から強調箇所を認識して画像ベースのインデックスを自動生成し、動画の任意の再生位置と関連付ける。学生はシステムが作成したインデックスから、動画の任意のシーンへの頭出しができる。また、提案システムは Full-HD 対応ビデオカメラ 1 台で利用でき、撮影設備がない一般的な教室でも利用できる。プロトタイプシステムを用いた評価の結果、システムは実用上問題のない精度でインデックス生成ができること、講師はふだんの講義スタイルを大きく変えることなくシステムを利用できることが示唆された。

本論文の構成を以下に示す。2 章では、関連研究に触れながら、既存システムの課題を整理する。3 章では、提案システムについて述べ、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。5 章では、プロトタイプシステムを用いた評価とその結果について述べる。最後に 6 章でまとめと今後について述べる。

2. 関連研究

これまでの板書講義配信システムでは、講義をいかにして見やすく配信するかに重きがおかれてきた。芦川らは、

板書中のチョーク音から講師が板書している付近を自動的にズームするシステムを提案している [7]。米川らは、ビデオカメラで撮影した黒板の文字や図形をベクトルデータ化して読みやすくするシステムを提案している [8]。これらの研究は動画における板書の視認性向上を目的としている。本研究が着目する復習視聴用のインデックス生成は考慮していない。

Dickson らは電子スライドを映したスクリーンとホワイトボードを撮影した講義動画からインデックスを自動生成・配信するシステムを提案している [9]。このシステムで生成されるインデックスはホワイトボードやスライドの全体画像である。全体画像は大きすぎると一覧性が悪く、小さすぎると内容の把握が困難であり、復習視聴用のインデックスとしては適していない。

電子スライドを撮影した動画から見出しなどのインデックスを自動生成する研究も行われている。Mukhopadhyay らは、電子スライドのタイトル部分を講義動画の見出しとして配信するシステム “Lecture Browser” を提案している [10]。井上らは、講義スライドのフッタ部分に付与したラベルから、講義動画の見出しとなるインデックスを自動生成している [4]。これらの研究は、講義の進行が定型化された電子スライド講義を対象としている。本研究が対象とする進行がインタラクティブに変更される板書講義には対応できない。

3. 板書講義コンテンツ生成システム

3.1 システム概要

提案システムの概要を図 1 に示す。(1) 講義前、講師は教室の黒板全体が映る位置にビデオカメラを設置して撮影を開始する。(2) 講師は板書の重要箇所マグネットを付けながら講義をする。(3) 講義後、講師は撮影した動画をコンテンツサーバにアップロードする。(4) コンテンツサーバは動画中のマグネットを認識し画像として切り出す。(5) 画像を講義動画のインデックスとして配信。

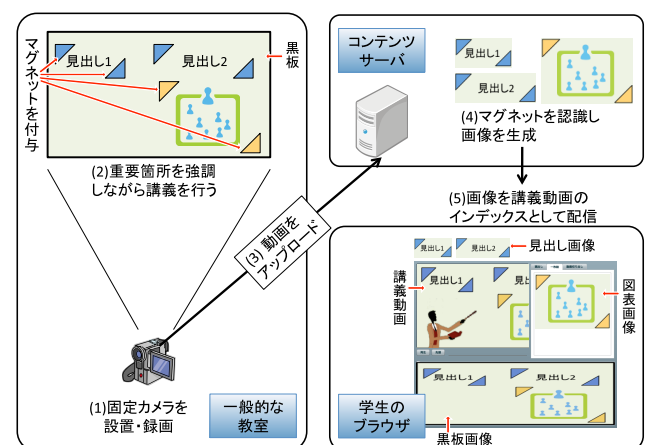


図 1 板書講義コンテンツ生成システムの概要

Fig. 1 Overview of the system for reviewing chalk talk lecture videos.

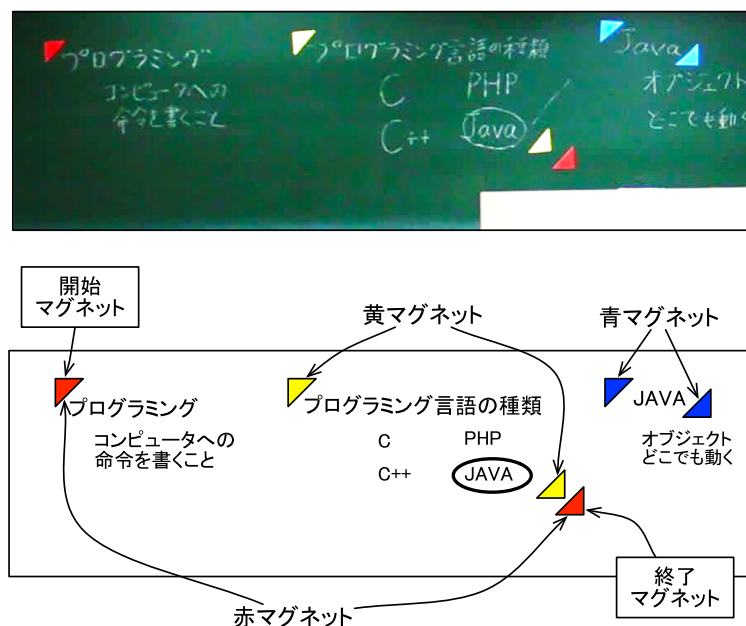


図 2 マグネット付与の例 (実写 (上部) とその模式図 (下部))

Fig. 2 An example of magnet highlights.

コンテンツサーバは生成された画像を動画の再生時刻と関連付け、インデックスとして使えるようにしたコンテンツを生成・Web 配信する。

本システムは座学形式の講義を対象としている。座学形式の講義では、講師は講義中につねに黒板を使いながら説明をし、学生はその説明をノートを取りながら学習する。語学や数学といった基礎教養科目の多くはこれにあてはまる。黒板を多用しない演習や、学生間のコミュニケーションが主体となるグループワークは想定していない。

本システムは、これまでの研究で考慮されていなかった板書講義動画への自動的なインデックス生成を実現している。これ以外に次に示す 3 つの特徴を持つ。

- 特別な機材を必要としないこと
 - 講師への負担が少ないこと
 - 復習視聴に必要な機能を提供できること
- 以降、これらの特徴について述べる。

3.2 撮影機材

撮影機材は、Full-HD 対応ビデオカメラ 1 台とカラーマグネットシート数枚だけである。一般に、講師は教室後方の座席からも視認可能な文字サイズで板書を行う。この文字サイズの板書は HDV1080i 対応ビデオカメラで黒板全体^{*1}が収まるように撮影した映像中においても十分に視認できることが知られている [11]。今回用いた Full-HD 対応ビデオカメラはその上位規格にあたり、同じ条件下で撮影された映像中の板書をより鮮明に視認することができる。それに加えて、Full-HD の 16 : 9 の映像は、横長の黒板を

無駄なく撮影するのに適している。

これまでの講義収録システムの多くは特殊な設備を必要とする。本システムでは一般家庭でも普及している機材で容易に撮影設備を整えることができるため、導入にかかるコストは非常に低い。

3.3 講義方法

講師は講義開始前、黒板全体が映る位置にビデオカメラを設置して録画ボタンを押す。講義終了後にはこの動画をコンテンツサーバにアップロードするだけでよい。講義の撮影に関して、講師には特別な知識や技術は必要ない。

講師は講義中、図 2 のようにカラーマグネットシートで板書を囲いながら講義を進める。講師はどの板書にどのマグネットを付与するかを意識する必要があるが、この動作はチョークによる下線や色の使い分けと類似している。マグネットを貼り付けるタイミングは、板書を書きながらでも、後になってからでもよい。講義方法による講師への負担はチョークのみの場合と比較して大きく変わらないと考えられる。

3.4 復習視聴用インデックス

本システムでは、講義中に講師がつけたマグネットの位置・時刻・色から復習視聴用インデックスを生成する。

マグネットは、図 2 に示すように、貼られた位置や向きにより開始と終了の意味を持つ。開始マグネットは囲う領域の左上に、終了マグネットは右下に付与する。システムはこの 2 つのマグネットのペア (以下、マグネット対) を動画から認識し、その内部の領域を画像として切り出す。この画像を動画アクセスのためのインデックスとして利用

*1 200 人が入れる規模の教室で横長の黒板 (約 7.5 m × 1.2 m)。

表 1 生成するインデックスの種類
Table 1 The classification of generated index.

種類	対象	用途
講義インデックス	講義 1 回分	復習したい内容を含む講義動画の特定
説明インデックス	講義中の時刻	講義動画中の特定シーンへの移動
区間インデックス	講義中の区間	講義動画中の説明区間の切り出し

する。

本システムが生成するインデックスの種類を表 1 に示す。講義インデックスは、講義 1 回分と関連付けられる。たとえば全 15 回からなる「確率」の講義一覧の中で、「マルコフ連鎖」について解説している回を識別するために用いる。講義インデックスとして使える板書は、その回を特徴付ける語句や図などである。説明インデックスは、講義中の時刻と関連付けられる。たとえば「有限状態マルコフ連鎖」の説明開始時刻に移動するために用いる。説明インデックスとして使える板書は、タイトルや見出しのように内容の区切りを示すものである。区間インデックスは、講義中の区間と関連付けられる。たとえば「マルコフ連鎖の状態 j が再帰的であるための必要十分条件の証明開始から終了までの区間」を視聴するために用いる。区間インデックスとして使える板書は、数式の証明のように一部分だけでコンテンツとして成立するものである。

これらインデックスを用いることで、復習に効果的な機能を実現できる。講義インデックスにより、複数の動画を 1 つ 1 つ再生しなくても復習したい内容を含む動画の特定が容易になる。説明インデックスにより、再生カーソルを調整しなくても復習したい説明に即座にアクセスできる。区間インデックスにより、特定の説明だけを繰り返し視聴することができる。この区間動画を iPod などの携帯端末へ転送すれば、いつでもどこでも復習ができる。

4. 実装

4.1 実装環境

講義を撮影するためのビデオカメラには、 $1,920 \times 1,080$ ピクセル (約 230 万画素) で撮影できる SANYO 製の Full-HD 対応デジタルビデオカメラ Xacti DMX-HD1010 を用いた。板書の強調には、1 辺約 10 cm の直角二等辺三角形のカラーマグネットシートを用いた。アップロードされた動画の変換には FFmpeg [12] を用いた。講義動画中のマーク認識や画像の切り出しには C++ で利用できる画像処理ライブラリ OpenCV [13] を用いた。生成した板書講義コンテンツの配信には PHP と Flex を用いた。

4.2 コンテンツサーバ

アップロードされた講義動画からのコンテンツ生成と配信はすべてコンテンツサーバが行う。コンテンツサーバの処理の流れを図 3 に示す。以降、各手順の詳細について説

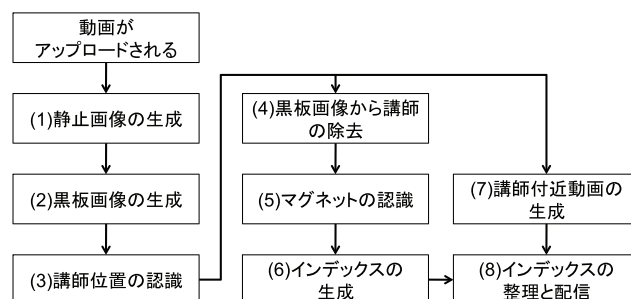


図 3 コンテンツサーバの処理手順

Fig. 3 Procedure of the contents server.

明する。

4.2.1 静止画像の生成と黒板画像の生成 (手順 1, 2)

コンテンツサーバは撮影した講義動画を受け取ると、FFmpeg を用いて数秒ごとに静止画像を生成する。この際、固定カメラを用いて黒板全体が映るように撮影をすると、黒板以外の部分も撮影されてしまう。そこで、HSV 表色系の HSV 比を用いて静止画像を黒板の色で 2 値化、ラベリングする。その結果、ラベリング領域が最大の部分を黒板領域と判断し、領域の上下左右の座標から黒板領域を切り出した黒板画像を生成する。色の閾値は黒板によって異なるため、実験的に決定した。

4.2.2 講師位置の認識と黒板画像から講師の除去 (手順 3, 4)

固定カメラで板書講義を撮影すると、講師による板書へのオクルージョンが発生する。システムは板書の一部をインデックスとして用いるため、講師が写っていない板書だけの画像が必要になる。本実装では市村ら [11] の手法を簡略化した方式を用いて講師位置の認識と講師の姿の除去を行った。この方式は、現在の画像と前の画像とを比較し、その差分から講師の位置を認識する。認識した講師の位置を前の画像の同じ位置で上書きすることで、講師の姿が映っていない黒板画像 (以下、講師除去画像) を生成する。

4.2.3 マグネットの認識 (手順 5)

本実装では、黄・青・赤の 3 色のマグネットを用いた。表 2 に各マグネットから生成されるインデックスの種類を示す。

黄マグネットは数学における関数のグラフ、物理における回路図、英語における「頻出語 take の使用例 (1) ... (2) ... (3) ...」といった図表・まとめ板書に付与する。これらはその回の講義を特徴付ける板書であることが多い。その内側の画像を講義インデックスとして利用することで、

表 2 マグネットの色と生成されるインデックスの関係
Table 2 Relationship between magnet color and index.

マグネットの色	適用する板書の例	生成されるインデックス
黄	図表, まとめ板書	講義インデックス
青	見出し	講義インデックス, 説明インデックス
赤	証明, 例題の解法, 和文英訳	区間インデックス

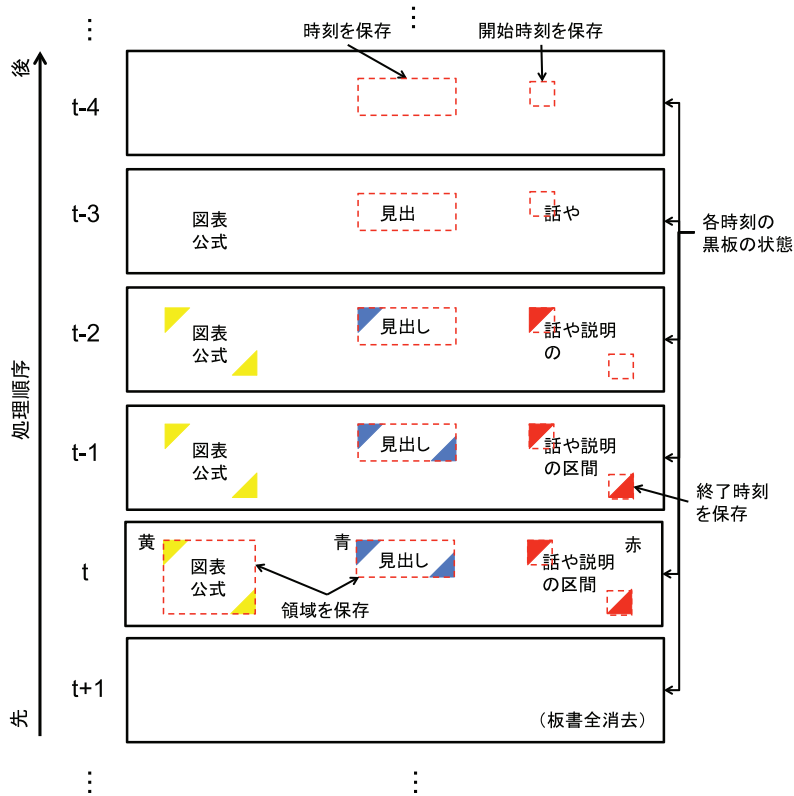


図 4 インデックス生成の処理
Fig. 4 The procedure of generating each index.

復習をしたい内容を含む講義動画の特定を容易にする．青マグネットは「関数の微分」「キルヒホッフの法則」「比較級・最上級」のように説明の区切りとなる見出しや単語に付与する．その内側の画像を講義インデックスと説明インデックスとして利用し，講義動画中への特定シーンへの移動を容易にする．赤マグネットは公式の証明，例題の解法，和文英訳問題のように，その導出過程が重要となる板書に付与する．その内側の画像を区間インデックスとして利用し，講義動画の一部分だけを別コンテンツとして生成する．

各マグネットの認識には RGB 表色系の RGB 比を用いて 2 値化とラベリングを施し，一定の大きさの領域をマグネットと判断した．色の閾値および大きさは，マグネットを実際に撮影して実験的に決定した．

開始マグネットと終了マグネットの認識には，マグネットの形状を用いた．マグネットの形状は直角二等辺三角形をしており，図 2 に示すように囲う板書の外側に直角がくるように付与される．システムは認識したマグネットの直角が左上にある場合は開始マグネット，右下にある場合は

終了マグネットと認識する．

4.2.4 インデックスの生成（手順 6）

講師除去画像からそれぞれのインデックスを生成する処理を図 4 に示す．実線の矩形 6 つは，時刻 ($t-4$) から時刻 ($t+1$) までの板書の状態を示しており，講師は時刻 ($t-3$) から板書を書き始め，時刻 (t) ですべての板書を書き終えている．時刻 ($t+1$) では黒板全体を消去した状態になっている．システムによるインデックス生成処理は現在の時刻 ($t+1$) から過去の時刻 ($t-4$) にさかのぼりながら行う．

講義インデックス（黄マグネット：図 4 左側）の場合，講師は時刻 ($t-3$) から板書を書き，時刻 ($t-2$) でマグネットの付与を行っている．システムがマグネット対を最初に認識するのは，時刻 ($t+1$) からさかのぼって時刻 (t) である．このときのマグネット対内側の領域を講義インデックスの画像として保存し，動画全体と関連付ける．システムが時刻 ($t-1$)，時刻 ($t-2$)， $\dots$ ，とさかのぼったときに同じ座標で認識するマグネット対内部の画像は時刻 (t) で保存された画像と同じであるため，時刻 ($t-1$) 以降では無

視する。

説明インデックス（青マグネット：図 4 中央）の場合，講師は時刻 $(t-3)$ で板書を書き始め，時刻 $(t-2)$ で板書が書き終わった後で開始マグネットを付与している．その後，時刻 $(t-1)$ で終了マグネットを付与している．システムがマグネット対を最初に認識するのは，時刻 $(t+1)$ からさかのぼって時刻 (t) である．このときのマグネット対の内側の領域を講義インデックス，説明インデックスの画像として保存する．説明インデックスと関連付ける時刻は，システムが処理の過程で開始マグネットを最初に認識した時刻 (t) から時刻をさかのぼり，その内側の板書がなくなった時刻 $(t-4)$ とする．

区間インデックス（赤マグネット：図 4 右側）の場合，講師は時刻 $(t-3)$ から板書を書き始め，途中の時刻 $(t-2)$ で開始マグネットを付与し，時刻 $(t-1)$ では板書を終えて終了マグネットを付与している．システムがマグネット対を最初に認識するのは，時刻 $(t+1)$ からさかのぼって時刻 (t) である．このときのマグネットの内側の領域を区間インデックスの画像として保存する．区間インデックスと関連付ける説明の開始時刻は，システムが処理の過程で開始マグネットを最初に認識した時刻 (t) から時間をさかのぼり，その付近の板書がなくなった時刻 $(t-4)$ とする．区間インデックスと関連付ける説明の終了時刻は，システムが処理の過程で終了マグネットを認識した時刻 (t) から時間をさかのぼり，終了マグネットを最後に認識した時刻 $(t-1)$ とする．この時刻 $(t-4)$ から時刻 $(t-1)$ までを切り出した動画と区間インデックスとを関連付ける．

インデックスをマグネット付与時刻と関連付けてしまうと，講師は板書とマグネットとの対応を強く意識しながら講義をすすめる必要があり，大きな負担となる．本システムはインデックスを板書の開始時刻と関連付けるため，マグネットを付与するタイミングは（板書を消す前ならば）いつでもよい．図 4 の区間インデックスのように貼り忘れたものを後から付与することができるため，講師の負担は少ない．また，後から重要性が増した用語に対しても適用できるなど，システムの柔軟性も高い．

4.2.5 講師付近動画の生成（手順 7）

講師は自身の近くの板書を解説しながら講義を進めるため，黑板全体の動画を配信するよりも，講師付近の動画を配信する方がより学習の支援になると考えられる．そこで，撮影した板書講義から講師の付近だけを切り出した動画（以下，講師付近動画）を生成する．まず，各板書画像から 4.2.2 項で認識した講師位置を用いて，図 5 に示すように講師位置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，トリミングした講師付近の画像を時系列に並べて画像のみの動画を生成する．最後に，生成した動画に元の音声合成し講師付近動画を生成する．

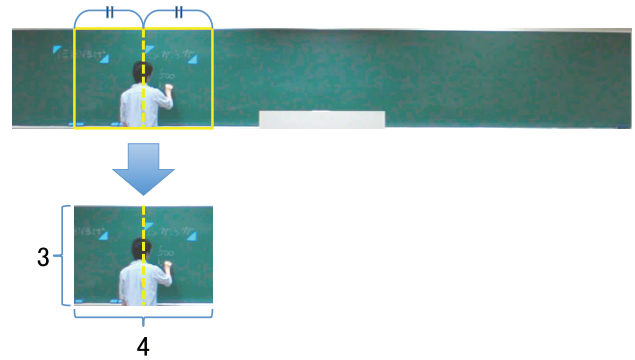


図 5 講師付近の画像の切り出し

Fig. 5 Clipping the image near the lecturer.

4.2.6 インデックスの整理と配信（手順 8）

システムは，手順 1~7 で得たインデックスと動画から復習視聴用のコンテンツを生成・配信する．図 6 にその外観を示す．トップページである講義一覧ページ (a) には，アップロードされた講義の一覧が表示されている．各講義には，講義のサムネイルや講義インデックスが表示されている．学生は講義インデックスを見ながら目的の内容を含んだ講義を探すことができる．

講義を選択した後に表示される講義配信ページ (b) には，各講義の配信ページが表示されている．ページ上部 (1) には，その講義の見出しを表す講義インデックスが表示されている．これにより学生は講義内容の概略を把握することができる．ページ中央左 (2) には，講師だけをズームアップした講師付近動画が配信されている．これにより講師が説明している箇所を把握できる．ページ中央右 (3) には，システムが生成したインデックスを表示している．(3) はタブ構造になっており，説明インデックス，講義インデックス，区間インデックスの一覧を切り替えて表示することができる．ページ下部 (4) には，講師の姿を除去した講師除去画像が表示されている．黑板に何が書かれているかを，講師の影を気にすることなく見ることができる．この画像は (2) の動画と同期しており再生時刻の講師除去画像が表示される．

説明インデックスの動作を図 7 に示す．図 7 上部では動画の再生時間が 4:15 を示している．この状態でインデックス表示部にある説明インデックスの見出し画像を選択すると，その見出しを書き始めた再生時間である 0:18 にジャンプしていることが確認できる．

区間インデックスの動作を図 8 に示す．図 8 右上にある区間インデックスの画像を選択すると，説明インデックスと同様，その説明の開始時間である 14:12 へ再生カーソルが移動する．また，区間インデックスの画像下部には 00:14:12 - 00:18:48 のように，説明の開始時刻と終了時刻が表示されている．“PodCast” のリンクを選択するとこの区間だけを切り出した動画の RSS が配信される．これを iTunes などに登録すればこの区間動画を携帯端末で持ち出

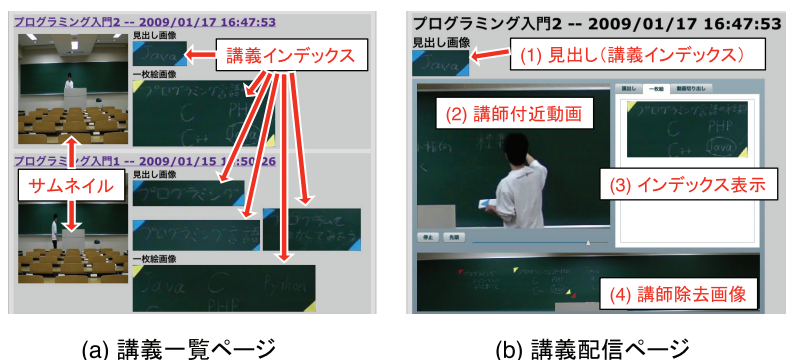


図 6 講義動画配信ページと講義インデックスの役割

Fig. 6 Example of the lecture contents.



図 7 説明インデックスの動作

Fig. 7 The behavior of an explanation index.



図 8 区間インデックスの動作

Fig. 8 The behavior of an interval index.

し、いつでもどこでも繰り返し復習をすることができる。

5. 評価

5.1 システムの利用評価実験

5.1.1 方法

提案システムがどの程度講師の負担になるかを知るために、講義歴 10 年以上の男性講師に、実際にシステムを利用した講義を行ってもらった。講義後、システムについてのアンケートとインタビューを実施した。

アンケートの項目を表 3 に示す。アンケートでは、マグネットを貼る・剥がすなどの行為に対して「かなり〇〇した・〇〇した・あまり〇〇しない・〇〇しない」の 4 段階で答えてもらった。また、項目 (1), (2), (3) ではその選択肢を選んだ理由も記入してもらった。このアンケートにより、講師がマグネットを付与しながら講義を行うことをどのように感じたかを把握する。インタビューでは、講師とともに撮影した動画を視聴しながら講義を振り返った。これにより、ある板書に対してマグネットを貼ったことの是非や、講義を振り返って気になったことを聴取した。

講義の科目は数学とし、1 回 20 分の講義を異なる内容で計 3 回行った。1 回目は、講師自身が講義を円滑に行うために板書内容を事前に整理したノート（以下、講義ノート）を用いて講義を行った。2 回目は、講義ノートを用いたうえで問題を解くための演習時間を設けて講義を行った。3 回目は、講義ノートを用いずに講義を行った。これにより、講義方法の違いによるマグネットの付与方法の違いを比較する。アンケートは 1 回目の講義の後に行った。インタビューは各講義の終了後に行った。

講義は 250 人程度が入れる教室で行った。黒板のサイズは幅 7.2m × 高さ 1.2m である。Full-HD 対応ビデオカメラは黒板全体が撮影できるように黒板から 11m 後方に設置した。

5.1.2 結果および考察

表 3 のアンケート結果から「マグネットによる編集」をあまり意識しないという回答が得られた。この理由としては「数学の講義ではある程度講義資料がまとめてあり、そ

表 3 アンケートの内容と結果
Table 3 The questionnaire and its answers.

項目	回答
(1) 講義をする際に「マグネットによる編集」を意識したか？	あまり意識しない
(2) 講義中にマグネットを貼る行為は負担になるか？	あまり負担にならない
(3) 講義中にマグネットを剥がす行為は負担になるか？	あまり負担にならない
(4) その他, 自由意見を記入	後述

の時点で構造化が終わっているため、それに沿ってマグネットを貼り付ければよいから」ということであった。「マグネットを貼る・剥がす行為」は負担になるかという問いに対しては、「マグネットを貼らないことにしたことはないが、あまり負担にならない」という回答が得られた。

次に、講義ノート（事前準備）の有無で結果を比較した。講義ノートありでは、ノート上で内容の構造化がされているため講師は迷うことなくマグネットを貼り付けていた。また、講義後のインタビューでも「講義ノートがあるとマグネットの貼り忘れが発生しない」との意見が得られた。一方、講義ノートなしでは、「マグネットの色や位置を迷う」や「(マグネットでの編集を) 多少意識する」という意見が得られた。実際にマグネットで囲った領域を、マグネットを貼り直すことで拡張する事例が見られた。

講義内容がインタラクティブに変化する場合、講義ノートを用いてもマグネットの付与で迷う可能性がある。これは講義回数を重ねることによる「慣れ」で発生頻度が減少していくと考えられる。インタビューでも「回数を重ねるごとに慣れてくる」という意見を聞くことができた。

以上の結果から、マグネットを付与しながら講義を行うことはそれほど負担にならないと考えられる。講義ノートを用いた場合と用いない場合とでは、講義ノートを用いた方がよりうまくシステムを利用できることが分かった。講義ノートを用いない場合やインタラクティブに講義が進行する場合でも、システムの利用回数が増えることで「慣れ」が発生し、より自然にマグネットを付与しながら講義が行えるようになっていくと考えられる。

5.2 システムの性能評価実験

5.2.1 方法

講義動画をシステムに入力して生成されたインデックスの定量評価を行った。評価に用いた動画は、5.1 節で用いた動画が3本と、大学生7人に行ってもらった模擬講義動画が15本の計18本である。動画から静止画像を切り出す周期は6秒に設定した。この6秒という時間は、システムでの処理後の容量と動画再生時のバランスを考え実験的に決定した。

まず、システムが動画からコンテンツを生成するまでにかかる時間を測定した。実行環境はCPU 2.4GHz, 4GB

RAM, OSはMac OSX 10.6である。システム全体の処理時間はインデックスの数に依存するため、測定対象は18本のうち最もインデックス数の多いもの(20分1.72GBで10個のインデックス)とした。

次に、講義中に付与されたマグネット対の認識精度を評価するために、各インデックスの適合率 P_m と再現率 R_m を求めた。全動画中で、マグネット対で強調した箇所(正解数)を C_m 、システムが抽出した強調箇所を N_m 、 N_m のうち正しく抽出された数を M_m とすると、適合率 P_m と再現率 R_m は次の式(1)および(2)で求められる。この P_m および R_m を、各マグネットの色ごとに求めた。

$$P_m = \frac{M_m}{N_m} \times 100(\%) \quad (1)$$

$$R_m = \frac{M_m}{C_m} \times 100(\%) \quad (2)$$

説明インデックスでは説明の開始時刻の差について評価を行った。人間が判断した説明開始時刻を T_a 、システムが判断した説明開始時刻を T_s とすると、再生開始時刻の差 T_{diff} は $T_{diff} = T_a - T_s$ で求められる。システムが認識したすべての説明インデックスに対してこの T_{diff} を求め、その平均 T_{ave} と標準偏差 T_{stdev} を求めた。これらの値が小さいほど、時間的に正確な説明インデックスが生成できたといえる。

区間インデックスに関しては、時間幅を持つ要素に特化した適合率と再現率を求めた[14]。手動で区間インデックスと判断した時間帯の合計を T_i 、システムが区間インデックスと判断した時間帯の合計を T_m 、 T_m が T_i と一致する時間帯の合計を T_c とすると(図9)、区間適合率 P_s と区間再現率 R_s は次の式(3)および(4)で求められる。区間適合率が高いほど、講師がつけた区間インデックスを再現しているといえる。区間再現率が高いほど、システムが生成した区間インデックスに無駄がないといえる。

$$P_s = \frac{T_c}{T_i} \times 100(\%) \quad (3)$$

$$R_s = \frac{T_c}{T_m} \times 100(\%) \quad (4)$$

5.2.2 結果および考察

システムがコンテンツの生成に要した時間は、動画のアップロード時間を除くと約5分であった。処理時間は動画中のインデックス数や出力動画の画質に左右されるが、それを考慮しても90分の講義動画1本あたり30分程度で

表 4 インデックス生成の適合率と再現率の結果

Table 4 The precision and recall of the generated indexes.

	講義インデックス (黄)	説明インデックス (青)	区間インデックス (赤)	全体
Cm	30	25	19	74
Nm	40	24	19	83
Mm	28	22	16	66
Pm (適合率)	70%	92%	84%	80%
Rm (再現率)	93%	88%	84%	89%

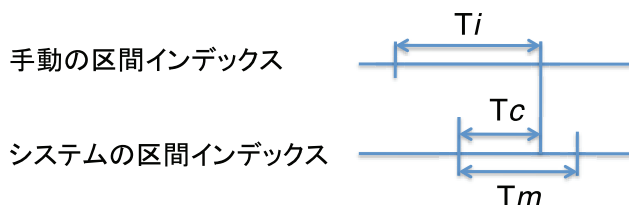


図 9 区間インデックスの検出例

Fig. 9 An example of the segment index.

完了するといえる。

収録した講義動画は、学生の復習を支援するためにも、遅くとも翌日までに編集を終えて公開されるのが望ましい。しかし、板書を使った講義は毎日複数科目実施されている。このすべてを1人で手動編集する場合、科目の実施数によっては翌日までに間に合わない可能性が高い。本システムを用いれば、その日の収録動画をまとめてアップロードしておくだけで、翌日までにすべての編集を確実に終わることができる。

各インデックス生成の適合率と再現率の結果を表 4 に示す。インデックス全体の適合率 P_m は 80%，再現率 R_m は 89% という結果が得られた。

適合率 80% は、システムが提示したインデックスに 2 割のノイズが混入することを意味する。1 回の講義を 90 分とすると、その板書で生成されるインデックスの総数は 20 程度^{*2}である。ノイズとして混入するインデックスの数は 1 講義あたり 4 程度である。この数は手動でも十分に除去可能なレベルであることから、実用上の問題は少ないと考える。

再現率 89% は、全インデックスの約 9 割を正しく検出できることを意味する。この精度は講義インデックスのように「復習したい講義動画を特定」するには実用上十分な精度であるといえる。一方、説明インデックスのように「講義動画中の特定時刻に移動」するにはすべてのインデックスを正しく検出可能な精度が望ましい。これに対しては人手で容易にインデックスを追加できるような管理インターフェースを提供する必要があるだろう。

講義インデックスの適合率が他のインデックスに比べて低い理由として、5.1.2 項で述べたマグネットの貼り直しが

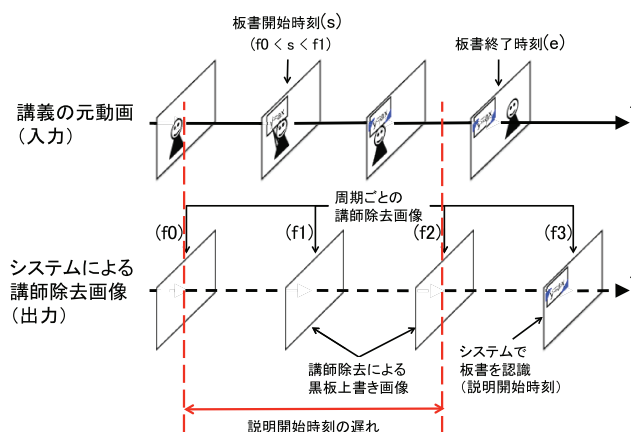


図 10 説明インデックスの遅れの原因

Fig. 10 Comparison of frame-sequences between input and output videos. Erasing the figure of the lecturer contributed to the delay of a comment index.

ある。本システムでは、マグネットの貼り直しについては考慮していなかった。4.2.4 項のインデックス生成手法では、終了マグネットの付与を契機としてマグネット対の認識を行う。そのため、講師が終了マグネットを貼り直すと、そのマグネット対を新たなマグネット対として再度認識する。結果として、ほぼ同じ内容の板書領域を複数回認識していた。この問題には 1 度認識した領域と新たに認識した領域とをテンプレートマッチングで比較し、閾値以上ならば新たにインデックスを生成しないといった処理が必要になる。また、区間インデックスでは照明ムラの影響で赤マグネットを RGB 比で認識できないことがあった。この問題には、マグネットを認識する閾値のキャリブレーションや HSV 表色系で色認識を行うことで対応できると考えられる。

説明インデックスの再生開始時刻の差の平均 T_{ave} は -6.9 秒、標準偏差 T_{stdev} は 4.2 であった。ほとんどの説明インデックスは、手動で判断した再生開始時刻よりもシステムで判断した時刻のほうが遅かった。

この理由を図 10 で説明する。図上部には講義の元動画の時系列の様子が描かれている。図下部には元動画をシステムに入力し、講師の姿の除去を行った後のフレーム時刻 ($f_0 \sim f_3$) の画像が描かれている。元動画では時刻 (s) で板書を書き始めている。システムは本来、フレーム時刻 (f_1) で板書を認識する。そして、それより前のフレーム時

^{*2} 表 4 より、20 分の模擬講義 18 本のインデックス総数は 74 個、1 本平均で 4~5 個、90 分換算で 18~22 個。

刻 (f_0) の時刻が説明インデックスの時刻になるはずである。しかし、板書の書き始めは板書と講師の位置が近いいため、フレーム時刻 (f_1, f_2) では講師の姿の除去とともに板書の除去もしてしまった。そのためシステムが板書を認識するのはフレーム時刻 (f_3) となってしまった。結果、システムが認識した説明インデックスの時刻は、板書を認識したフレーム時刻 (f_3) の前のフレーム時刻 (f_2) の時刻となってしまった。このように、講師が見出しとなる板書付近にいて、板書の認識が遅くなり説明インデックスの時刻が遅くなっていた。この問題は、講師の姿の除去で上書きする黒板領域を数ピクセル四方のような細かい領域ごとに行うことで改善できると考えられる。また、画像処理だけでなく講師の声やチョークの音など、音声情報と組み合わせて板書の開始時刻を認識することで解決できると考えられる。

区間適合率 P_s は 96%, 区間再現率 R_s は 90% という結果が得られた。

この精度を 5 分の区間動画で考えると、冒頭 14 秒が欠落し、末尾に冗長部分が 30 秒含まれる。我々はこの精度でも実用上問題ないと考えている。区間動画はその中間部分に主要な説明がある。冒頭部分は証明する定理の名称など板書を見れば分かる説明であり、これらが一部欠落したとしても本題には影響が少ない。末尾に追加される冗長部分も、視聴している学生が不要と思えば再生を停止すればよい。欠落部分の影響が出るような長時間の説明は、区間動画よりも講義動画全体として視聴すべきである。

区間が一致しなかった例として、システムが開始マグネット付近の板書を認識できず、マグネットを付与した時刻を区間動画の開始時刻と誤認していた。これにより、動画の開始時刻を手動で判断した場合よりも遅く判断したことがあった。この問題は、システム側で板書認識の閾値を変更することで対応できると考えられる。また、区間インデックスでも、手動で判断した時刻よりもシステムが判断した時刻の方が全体的に遅かった。理由も説明インデックスと同様のため、講師領域の上書きの単位を細かくしたり、音声情報と組み合わせることで改善できると考えられる。

5.2.3 その他の知見

システムが生成したインデックスのいくつかで、画像内の板書が滲んでしまい読み取りにくくなるケースが見られた。この滲みは動画から静止画への変換にともなう劣化ではなく、元動画の段階で発生していた。今回の実験では環境光の関係でフォーカスが不安定になることがあり、その影響が出てしまった。これは機種特有の現象であり、他の機材の多くで同様の問題がないことを確認している。

また、現段階ではシステムに全文検索インタフェースが用意されていない。履修中の講義の復習だけならば講義インデックスの目視でも問題はない。しかし何年分ものアーカイブを扱ったり、関連する未履修の講義を探したりする

場合には限界がある。一部のインデックスは手書き文字認識可能かもしれないが、板書は乱雑な書き方がされることも多い。文字認識のために丁寧な板書を強いるのは講師の負担となり、本システムの「講師の負担なく使える」という目標とは相容れない。それよりも、講師や学生がコンテンツに対して検索用のタグを追加する機能のほうが適していると考えている。

6. おわりに

本論文では、マグネットを用いて意識的に強調した板書からインデックスを自動生成・配信するシステムの提案と実装を行った。システムでは、板書講義中に講師が付与するマグネットから講義インデックスや説明インデックス、区間インデックスを自動生成する。これら画像や動画のインデックスにより学生の復習学習を支援する。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のみで、撮影設備を持たない一般教室でも利用が可能である。

模擬講義によるシステムの評価実験では、講師がマグネットを付与しながら講義を行うことは大きな負担にならないということが示唆された。また、システムの性能評価では講義中に付与されたマグネット対から実用上問題のない速度と精度でコンテンツ自動生成を実現できた。

今後は、講師の声などの音の情報と組み合わせることでインデックス生成精度の向上を図っていく予定である。また、強調方法の簡略化のために、板書のブロック分割技術 [15] を用いて開始マグネットだけのインデックス生成や、マグネットの色や形状の多様化により生成できるインデックスの種類の増加、講師ごとのマグネット色のキャリブレーション、全文検索用タグ機能を考えている。

本システムは横方向に長い 1 枚の黒板を想定していたが、上下分割式や引き戸式の黒板も存在する。これらは書き終わった板書を上方や側方へ退避させるといった独特の使い方が存在する。今回は想定していなかったこのようなイベントを検知できれば、インデックスに別の意味を付加することも可能と考えられる。

本システムは復習用コンテンツの生成に特化しているが、教育効果をさらに高めるにはいくつかの機能を追加する必要があると考えている。1 つは講師と学生とのインタラクション支援機能である。動画中の説明が分かりにくかった箇所や関連する別講義の情報を講師・学生間のコミュニティで共有していけば、講師にとっては講義方法・内容双方の改善につながる。学生にとっては疑問の早期解決につながる。もう 1 つは学生主体のコンテンツ編集機能である。学生の視点から重要と思える板書に非公式なインデックスを付与できるようになれば、同じ講義を他者がどのように理解したかを視覚的にとらえることができる。最終的にはこれら機能を本システムに統合し、個別の復習支援の場か

ら複数人での協調学習の場へとシフトさせていきたい。

参考文献

- [1] 京都大学: OCW, 入手先 (<http://ocw.kyoto-u.ac.jp/>) (参照 2011-07-23).
- [2] 慶應義塾大学: SFC GLOBAL CAMPUS, 入手先 (<http://gc.sfc.keio.ac.jp/>) (参照 2011-07-23).
- [3] UC Berkeley: available from (<http://www.youtube.com/ucberkeley>) (accessed 2011-07-23).
- [4] 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, Vol.107, No.391, pp.1-6 (2007).
- [5] 小澤憲秋, 武部浩明, 勝山 裕, 直井 聡, 横田治夫: 文字認識を利用した講義動画中のスライド同定, 情報科学技術フォーラム 2002, pp.133-134 (2002).
- [6] echo360: The EchoSystem, available from (<http://echo360.com/echosystem/>) (accessed 2011-07-23).
- [7] 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).
- [8] 米川 輝, 橋本政宏, 若原裕範, 豊崎大基, 萩原靖久, 岩月正見: 通常教室における講義のデジタルコンテンツ自動作成システム (e-Learning と FD 支援/一般), 電子情報通信学会技術研究報告, ET, 教育工学, Vol.105, No.488, pp.19-24 (2005).
- [9] Dickson, P.E., Adrion, W.R. and Hanson, A.R.: Automatic creation of indexed presentations from classroom lectures, *ITiCSE '08: Proc. 13th Annual Conference on Innovation and Technology in Computer Science Education*, pp.12-16 (2008).
- [10] Mukhopadhyay, S. and Smith, B.: Passive capture and structuring of lectures, *MULTIMEDIA '99: Proc. 7th ACM International Conference on Multimedia (Part 1)*, pp.477-487 (1999).
- [11] 市村 哲, 井上亮文, 宇田隆哉, 伊藤雅仁, 田胡和哉, 松下温: Chalk Talk: 講師動画と板書静止画の同時記録が可能な講義自動収録システム, 情報処理学会論文誌, Vol.47, No.3, pp.924-931 (2006).
- [12] FFmpeg, available from (<http://ffmpeg.org/>) (accessed 2011-07-23).
- [13] OpenCV: Open Computer Vision Library, available from (<http://opencv.willowgarage.com/wiki/>) (accessed 2011-07-23).
- [14] 井上亮文, 吉田竜二, 平石絢子, 重野 寛, 岡田謙一, 松下温: 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).
- [15] 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, Vol.83, No.11, pp.1187-1195 (2000).

推薦文

本論文は板書講義の復習用教材の自動作成システムを提案している。講師が板書の見出しや強調したい箇所に「講義, 説明, 区間」の3種のカラーマグネットを貼ることで, システムはマーキングの位置と時刻情報を利用して画像を切り出し講義内容を表すインデックスとして利用する。模擬講義による評価実験においてマグネットの付与は講義の

進行を妨げないという結果が示唆され, マグネットから約89%のインデックスを自動生成することに成功している。論文のアイデアは興味深く, また既存研究に対する本研究の優位性・有用性を合理的な方法で検証している優れた論文である。

(グループウェアとネットワークサービス研究会主査
小林 稔)



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。2005 年同大学院後期博士課程修了。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講師。インタラクティブシステムの研究に従事。本会論文誌編集委員。ヒューマンインタフェース学会, ACM 各会員。



品田 良太

2008 年東京工科大学コンピュータサイエンス学部卒業。2010 年同大学院前期博士課程修了。現在, 三菱電機情報ネットワーク株式会社に勤務。在学中, 講義収録システムの研究に従事。



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。2011 同大学教授。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 2011 最優秀論文賞受賞。ACM, 電子情報通信学会各会員。



星 徹 (フェロー)

1969 年東京工業大学電気工学科卒業。
同年日立製作所入社。1975 年 UCLA
大学院修士課程修了。2003 年東京工
科大学コンピュータサイエンス学部教
授。2007 年同学部長。現在、東京工
科大学名誉教授。博士（工学）。本会

論文誌編集委員，GN 研究会主査，理事等を歴任，本会フェ
ロー。IEEE，ACM，電子情報通信学会，電気学会各会員。