

推薦論文

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

市村 哲[†] 福井 登志也[†]
井上 亮文[†] 松下 温[†]

講師が講義前にビデオカメラを 1 台設置しておくだけで、その収録映像から自動的に講師の動画映像と板書の静止画とを作成し、即座にインターネット配信可能とする講義自動収録システム ChalkTalk-OSS を構築した。板書静止画については、画像処理を施して講師の姿を消去する工夫を行っている。1 台の固定ビデオカメラのみで黒板全体を記録できるように、解像度の高い民生用ハイビジョンカメラを用いて実装した。本論文では、講義室環境に適した講師位置検出処理、人間が行うカメラワークに近い講師動画作成処理等について提案する。特に、1. 黒板領域自動検出、2. 講師領域自動検出、3. 板書静止画と講師動画の作成、および、4. 板書静止画と講師動画の配信の各方法について述べ、それぞれの実装と実験評価について議論する。

A System that Records Chalk Talks for Web-based Learning

SATOSHI ICHIMURA,[†] TOSHIYA FUKUI,[†] AKIFUMI INOUE[†]
and YUTAKA MATSUSHITA[†]

We developed ChalkTalk-OSS that automatically produces E-learning materials from videotaped chalk talks. The system separately extracts a lecturer's image and writing on blackboard from video images recorded with a single high-definition digital camcorder, and stores the former as a streaming video and the latter as a series of snapshot images. Methods for locating the lecturer in the class room environment and creating lecture-videos with smooth camera-work are described. Especially, methods for detecting the blackboard, detecting the lecturer's area, creating a lecturer's movie and animation of writing, and delivering the lecturer's movie and animation of writing are discussed.

1. はじめに

電子プレゼンテーションが普及した現在でも、大学や予備校等の教育機関において黒板を用いた講義は根強い支持を得ている。たとえば、電子プレゼンテーション中心の講義であっても、重要箇所の説明の際や受講者の理解ペースに合わせて説明する必要がある際に板書が頻りに用いられている。また、大学や予備校において、数学や英語では、ほとんどが黒板を用いて授業が行われているのが現状である。一方、近年、インターネットを利用した新しい情報サービスとして WBT (Web-Based Training) 等の E-ラーニングが注目されている。いつでもどこでも自分の進捗状況に合わせて学習を進められることや、通学費や施設費等の経費を削減できる等の利点がある。

以上のような背景から、著者らは、黒板の板書を用いた講義を E-ラーニング教材として提供できるようにすることが必要であると考え、これを目的とした講義自動撮影システムの試作および提案を行ってきた^{1),2)}。本試作システムを用いれば、講師が講義前に民生用ハイビジョンカメラ (HDV カメラ) を 1 台設置しておくだけで講義を自動撮影できるため、専属カメラマンが不要となり、講義を担当する講師自らがきわめて簡単に講義撮影できる利点がある。また、講義映像をインターネット経由で配信することを考慮し、講師近辺の映像は動画として配信し、板書全体は静止画として配信できるようにした。板書静止画は、板書内容に変化があったときのみ作成され静止画アニメーションとして Web ブラウザに配信されるため、ハイビジョン画質の板書映像であっても必要ネットワーク帯域が少な

[†] 東京工科大学
Tokyo University of Technology

本論文の内容は 2005 年 11 月のグループウェアとネットワークサービスワークショップ 2005 にて報告され、GN 研究会主催により情報処理学会論文誌への掲載が推薦された論文である。

いという利点がある。くわえて、画像処理を施すことで板書静止画から講師の姿を消去し、講師の陰のない全体板書画像をつねに学習者に配信できるようにした。

本論文では、過去において提案した上記試作システムの実用性と信頼性を向上させ、Linux 上で動作するオープンソースソフトウェアとして再構築した「ChalkTalk-OSS」システムの提案を行う。今回新たに、講師の姿や動きを検出する画像処理アルゴリズムを改良し、実際の講義映像で発生するノイズに対してロバストネスの高い講師検出処理を実現した。講師領域検出の結果は、ハイビジョン映像から講師近傍動画を切り出す処理、および、板書静止画から講師の姿を消去する処理に使用されるため、極力誤りを含まないことが要求されるが、上記試作システムにおいてはその誤りの多さが問題となっていた。今回、講義室において板書時に発生するノイズの原因を特定し、この原因に対して対策を講じることで講師領域検出の正確さを向上させた。

また今回、講師近辺の映像をなめらかな動画として保存するための機構を新たに設け、人間が行うカメラワークに近い講師動画作成を可能とした。ユーザ評価実験の結果から、システムが自動生成した映像はきわめて自然であり、人が撮影した講義収録映像と比較しても劣らない見やすさを有することを確認した。くわえて、講師が何を指し示したか、または、どのようなジェスチャをしたかを記録するために、講師の胴体と手の位置関係を考慮し、講師の手部分と予測される領域と胴体部分と予測される領域とをつなげた範囲を講師動画として保存する機構を導入した。

2. 従来研究と研究の背景

ビデオカメラによって板書映像を記録しようとする、講師撮影映像の一部として板書を記録する必要があるが、黑板全体をカバーする領域を動画として保存またはインターネット配信しようとするネットワーク帯域を多く必要とするという問題点がある。このような理由から、講義における板書の自動撮影に関する研究の一環として、複数のカメラで講義を撮影して最も講義状況に適した映像を選択して記録する研究や、固定カメラとズーム制御可能なカメラとを併用し重要箇所に限定して板書を自動ズーム撮影する研究がなされてきた^{3)~7)}。しかしながら、装置が大がかりとなるため、どの教室にでも手軽に持ち運んで使えるシステムとすることが難しいという問題や、セットアップに手間がかかるという問題があった。くわえて、講義の場合は、少し前に説明された内容を参照しながら現在の内容を理解するということがしばしばあり、黑板全

体を同時に見渡したいという要求が少なからず存在する。このため、複数カメラを切り替えて撮影したり、黑板をパンして撮影したりした映像には不満がでることが予測される。

また、講義映像収録以外の分野において、たとえば、テレビ会議参加者を自動撮影するシステムが提案されている。市販されているテレビ会議製品の多くは1台の光学ズーム式カメラを備え、手動で方向や拡大率を制御できるようになっているが、先進的な研究としては、これらをすべて自動化しようとする動きがある^{9)~11)}。しかしながら、光学ズーム式カメラについては、首振り速度やズーム速度が間に合わなくなる恐れが常に存在する。たとえば、いったんズームする箇所を間違ってしまうと、それから復帰するためには多大な時間がかかるという問題がある。

そこで著者らは、過去において、講師近辺の映像を低画質動画として保存すると同時に板書全体を高画質静止画として保存する講義自動収録システムを提案した¹⁾。1台の固定ビデオカメラのみで黑板全体を明瞭に記録するためにハイビジョンカメラを用い、また、板書静止画については、画像処理を施して講師の姿を消去する工夫を施した。

従来のビデオカメラ(DVカメラ等)で黑板全体(著者らの大学において最も一般的な黑板のサイズは750 cm × 120 cm)を撮影した場合には解像度が絶対的に足りないために、解像度の高い民生用ハイビジョンカメラ(HDVカメラ)を利用した。たとえば、DVカメラ画像では判別困難な文字や講師の顔が、ハイビジョンカメラ画像では明確に判別できる。ハイビジョンカメラとしては、記録画素を1,440 × 1,080(約155万画素)確保できる民生用のSONY「HDR-HC1」⁸⁾を使用した。HDR-HC1は、HDV規格(ミニDVカセットテープにハイビジョン映像の記録再生を可能にするための規格)に準拠したデジタルビデオカメラであり、映像記録フォーマットはMPEG2-TSである。しかしながら、ハイビジョン映像をそのままインターネットに送出するときわめて多くのネットワーク帯域が必要になるという問題があるため、著者らは、講師近辺の映像のみ動画として保存し、板書全体は静止画として保存することで必要ネットワーク帯域を減らす工夫を行った。板書に変化があったときのみ静止画を作成し、板書を静止画アニメーションとしてインターネット配信できるようになっている。さらに、この試作システムでは、画像処理を施して板書静止画から講師の姿を消去し、講師の陰のない静止画を作成して保存するようになっている。動画であれば、板書の一部

が講師の陰となって見えない場合にも、少し待てば講師がその部分の前から立ち去るために問題とならないが、板書を静止画アニメーションとして記録してしまうと、その見えない部分が長時間にわたって読めない状態になってしまうためである。実際の講義を対象としたユーザ主観評価において、板書静止画は高画質であり、十分見やすい画像であるとの結果を得ることができた¹⁾。また、ハイビジョンカメラで黒板領域全体を動画として保存したファイル容量に比較して、試作システムが作成したコンテンツの総容量はきわめて小さく、有効に帯域を圧縮できることが確認できた¹⁾。

しかしながら、上記試作システムは、実現可能性を検証するためのプロトタイプとして構築したものであり、次のような重要な課題を未解決のままとしていた。

- (1) 黒板領域を撮影のつど手作業で指定する必要があった。
- (2) 講師領域を誤って認識する状況が少なからず発生した。
- (3) 自動作成された講師動画が不自然であり見にくかった。

(1) に関し、試作システムでは、黒板の四隅座標（認識対象領域）をユーザが手作業で入力する必要があった。また(2)に関し、比較的単純な背景差分法とフレーム間差分法の組合せによって講師領域を検出していたために、認識システムが講師を見失ってしまう問題や、講師がほとんど動かなかった場合に講師位置の特定が曖昧になってしまうという問題があった。さらに(3)に関し、自動生成された講師動画は、一定間隔（4秒ごと）で、講師位置を中心とした映像に強制的に切り替える方式で作成されていたために、映像に不連続感があり見にくいという問題があった。

3. 実装と実験評価

ChalkTalk-OSS の典型的な使用手順について述べる。講師は、講義開始前に教室の後方にハイビジョンカメラ1台を設置し、黒板全体を撮影できるよう画角を調整して撮影を開始する。講義終了後、撮影に使ったビデオテープを巻き戻し、ハイビジョンカメラとLinux PCとをIEEE1394ケーブルで接続してMPEG2-TS形式でデジタルビデオキャプチャを行う（このビデオキャプチャツールについても開発を行った）。そして、Linux PC上に作成した動画ファイルを入力ファイルとしてChalkTalk-OSSを実行する。これだけの操作で、講師の姿を消去してつねに黒板の内容全体が見える状態の板書を静止画として保存し、同時に、講師を自動追尾した映像を動画として保存できる。くわえて、

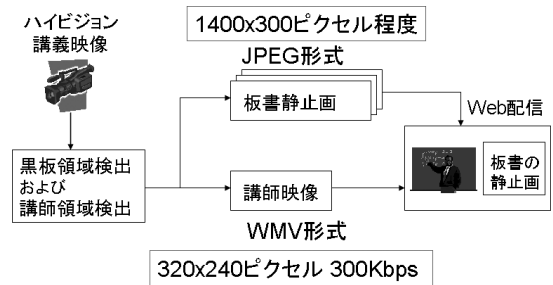


図1 変換処理の概要

Fig. 1 Outline of the conversion process.

これらのコンテンツをストリーミング配信するためのWebページも自動作成されるため、Webサーバにコピーしてただちにコンテンツ配信ができるようになっている。図1に変換処理の概要を示す。

なお、動画ファイルを入力してから板書静止画と講師動画を生成してWebサーバに配置するまでにかかる処理時間は、3GHz Pentium4を用いた場合に動画ファイル再生時間の約1.5倍である。

以下、1. 黒板領域自動検出、2. 講師領域自動検出、3. 板書静止画と講師動画の作成、および、4. 板書静止画と講師動画の配信に節を分け、実装と実験評価について述べる。

3.1 黒板領域自動検出

以前の試作システムでは、黒板の四隅座標（認識対象領域）をユーザが手作業で指定する必要があり手間がかかるという問題があった。そこで今回、黒板領域を自動検出する機能をChalkTalk-OSSに新たに設けることで、撮影準備作業を簡略化することを目指した。

黒板を検出する方法として、黒板の形状と色に着目して検出するように方針を定めた。まず色に関し、RGB表色系から黒板領域の特徴を抽出することを試みた。しかしながら、RGBの各要素の画素値を調べた結果、同じ黒板を撮影したシーンでも、部屋の明るさの変化等によってRGBの値が大きく異なり、その際の法則性が分かりにくいという結果となった。

そこで次に、HSV表色系¹²⁾から黒板の色を特定することを試みた。HSV表色系では、色合いを表す色相：H、色の鮮やかさを表す彩度：S、色の明るさを表す明度：Vで色が表現される。Hの値（色相）は、輝度の変化に依存せずに色を識別できることが特徴であり、今回の実験でも、部屋の明るさがかなり大きく違う環境でもHの値に大きな差は見られなかった。実験の結果、黒板としては、 $90 \leq H < 150$ を黒板色と定めることが適当であることが分かった。このHの範囲は12色相環¹²⁾の黄緑から青緑に相当する領域で

ある。

一方、黒板の形状に関しては長方形であることを仮定した。撮影画像の上下の端、または、左右の端からそれぞれ水平、および、垂直方向に画素を走査して、上記 HSV 系の色特性が黒板である領域を絞り込むという方法によって行った。水平方向に黒板色と認識されるマクロブロックがいくつ存在するかを調べ、閾値を超えたところを黒板領域の上端と定めた。

黒板の大きさや設置されている建物が異なる 10 教室の黒板について上記方法によって黒板領域自動検出を試みた結果、黒板領域が正しく認識されないケースは存在しなかった。

3.2 講師領域自動検出

黒板領域が特定されると、次は、その領域に対し講師領域検出処理が行われる。講師領域は、その後の、講師の姿を消去した板書静止画を作成する処理、および、ハイビジョン映像から講師近傍動画を切り出す処理に使用されるため、本システムにおいては、講師領域を正確に検出することがきわめて重要である。しかしながら、以前の試作システムでは、比較的単純な背景差分法とフレーム間差分法の組合せによって講師領域を検出していたためノイズ等に弱く、講師領域を誤って認識してしまう状況が少なからず発生していた。

ここでは、今回新たに実装した講師領域自動検出方式について述べる。本方式は、a. 移動物体領域検出処理、b. 板書撮影時に発生するノイズの除去、および、c. 講師の胴体と手の関係を考慮した講師領域検出、から構成されている。以下、それぞれについて述べる。

a. 移動物体領域検出処理

まず、移動物体領域検出処理の基本となるフレーム間差分法¹³⁾について説明する。フレーム間差分法は、撮影された時間の近い 2 つの画像で差分を計算する方法であり、画素変化のあった領域を移動物体が映っている領域と定める方法である。輝度によって 2 値化した画像が処理対象とされることが多い。ただし、一般的なフレーム間差分法では、動いている物体の領域として、時刻 t における位置と、時刻 $t-1$ における位置の両方が抽出されてしまうという問題がある。実際に、講師が速く移動したような場合に、時刻 t における位置と時刻 $t-1$ における位置の両方の領域が移動領域として抽出されてしまう不具合が観察された。また、講師の動きが遅い場合には、移動物体が部分的に重なってしまうため、移動物体の輪郭が抽出されにくいという問題点があることが観察された。

そこで今回、時間的に連続した 3 フレームの画像（輝度によって 2 値化した画像）に対し、2 段階のフ

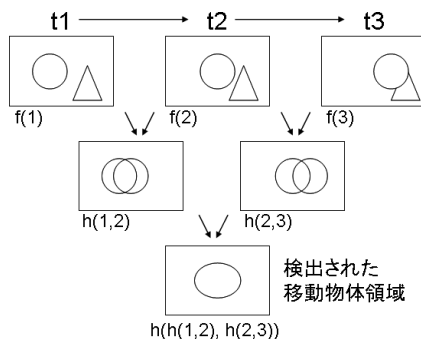


図 2 移動物体検出

Fig. 2 Detection of moving objects.

レーム間差分法を施すことによって移動物体領域を検出するようにした（本実装において、フレーム間差分抽出処理は 8×8 ピクセルのマクロブロックごとに行われる）。図 2 に示すように連続したフレーム画像を $f(1)$, $f(2)$, $f(3)$ とするとき、最初に、 $f(1)$ と $f(2)$ の差分画像 $h(1,2)$ と、 $f(2)$ と $f(3)$ の差分画像 $h(2,3)$ とを作成する。次に、差分画像 $h(1,2)$ と差分画像 $h(2,3)$ の差分画像 $h(h(1,2), h(2,3))$ を作成するようにした。そして、この差分画像 $h(h(1,2), h(2,3))$ に消えずに存在している画像領域を、画像 $f(1)$, $f(2)$, $f(3)$ から得られた移動物体領域として定めた。これによって、時刻 t における位置と時刻 $t-1$ における位置の両方が移動物体として抽出されてしまうという問題や、移動物体の動きが遅い場合に移動物体の輪郭が抽出されにくいという問題を大幅に緩和することができた。

b. 板書撮影時に発生するノイズの除去

しかしながら、移動物体領域検出処理を終えた段階で、講師が存在しない領域を移動物体領域と誤認する例が少なからず見受けられた。そこで原因を確認するために、移動物体領域検出処理を 1 ピクセルごとに施し、移動物体と認識されたピクセルを着色して画像表示する実験を行った。複数の講義収録ビデオを対象に実験を行った結果、黒板にチョークで書かれた文字が移動物体として誤認識されることが多いこと、さらに、同じ文字でも白色チョークで書かれた文字の誤認識率が高いことが分かった。次に、白色文字の誤認識率が高い理由を調べるために、屋外で撮影したビデオや室内で撮影したビデオ等、様々なビデオを ChalkTalk-OSS に入力して移動物体検出実験を行った。しかしながらこれらの実験においては、白色の部分の誤認識率が高いと判定されたビデオは少なく、むしろ、そのような傾向は見られない場合がほとんどであった。

そこで、ハイスピードカメラを用いてビデオ映像の分析を行うとともに、講義収録ビデオに立ち返って再

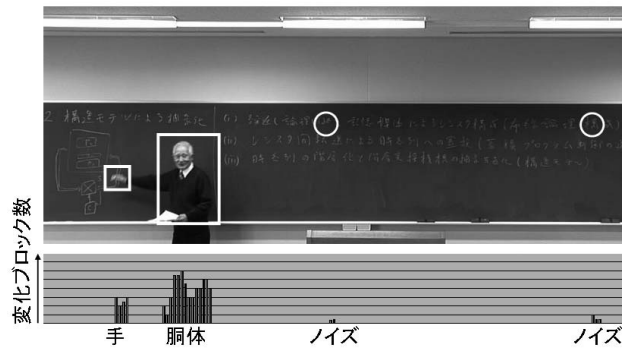


図3 講師領域検出

Fig. 3 Detection of lecturer's area.

実験をすることにした．200 フレーム/秒で撮影できるハイスピードカメラを用いて映像を分析した結果，肉眼では分からなかった蛍光灯によるフリッカが，板書の明度を大きく変化させていることが分かった．実際に，講義収録ビデオに対する再実験の結果，黑板に書かれた白色文字の中でも，黑板灯（黑板を照らす目的で設置された蛍光灯）に近い黑板上部に書かれた白色文字の誤認識率が高いことが確認できた（図3の丸枠部分）．白色は他の色よりも蛍光灯の光を反射する率が高いため，その影響が顕著に現れ，白色文字が移動物体として誤認識されたと推測できる．

このことを検証するために，白色の文字部分と濃緑色の黑板部分とで，黑板灯のフリッカによる影響度合いの違いを比較する実験を行った．講義ビデオ映像を輝度 256 段階のグレースケール映像に変換して実験したところ，黑板上の同じ場所（黑板灯に比較的近い場所）において，濃緑色の黑板部分ではフリッカの影響による輝度変化は平均 9.0 であったが，一方，白色の文字部分の輝度変化は平均 18.0 であり，輝度変化量が平均 2 倍であることが分かった．板書の視認性をあげるために設置されている黑板灯が，移動物体領域検出処理に悪影響を与えていたといえる．

この実験結果に基づき，システムの改良を行った．移動物体を漏れなく検出するためには，輝度変化の閾値は極力小さい方が好ましいが，一方で，白色文字部分に関して閾値を小さくすると，板書文字が移動物体として誤認識される可能性が高くなる．これらの条件を鑑み，フリッカによる影響を受けやすい白色部分（RGB 値がほぼ均等に大きい値（輝度 256 段階の 210 以上）を示す部分）に関しては，他の色の部分に比べて約 2 倍の輝度変化を許容するようにシステムを変更した．白色文字部分を複数講義映像から無作為に 900 カ所選り出して実験したところ，システム変更後は変

更前に比べて誤認識回数が 55%減少する結果となった．また有意水準 5%で Z 検定（片側）した結果からも，システム変更後は変更前と比べて誤認識する率が有意に低下したことが分かった（P 値 0.001 以下）．

c. 講師の胴体と手の関係を考慮した講師領域検出

板書撮影時に発生するノイズの除去が実行された後，講師の胴体と手の位置関係を考慮した講師領域検出処理が行われる（図3の長方形枠部分）．具体的には，移動物体検出処理によって移動物体領域と特定されたマクロブロックの数（単位時間あたりの度数）を黑板の横軸座標で整理してヒストグラム化し，以下のルールに従って講師領域を特定するようになっている．

- (1) 移動物体領域が X 軸方向に最も広い範囲で存在する領域を講師の胴体部分と見なす．胴体部分の中心付近を撮影の中心点とする．
- (2) 講師の胴体部分から左右に 120 ピクセル（15 マクロブロック）以内に存在する動きの大きい移動物体領域は講師の手部分と見なして，胴体部分と領域をつなげて講師領域とし，講師領域の中心付近を撮影の中心点とする．
- (3) X 軸方向に狭い移動物体領域が単独で存在，または，講師の胴体部分から左右に 120 ピクセル以上離れて存在するとき，ノイズと見なして講師領域とはしない．

図3に，黑板の横軸座標方向に移動物体領域のマクロブロック数をプロットした実際のデータを示す．これは，黑板の色と比較的似た色の服を着ている講師を撮影した際のデータであるが，図に示されるように，胴体と手の部分の間に移動物体領域がまったく存在しない部分があることが分かる．講師が何を指し示したか，または，どのようなジェスチャをしたかは講義の中で非常に重要な情報であり，講師の手部分と予測される領域と胴体部分と予測される領域とをつなげて保

存することができれば、講義内容がより分かりやすくなると考えた。

くわえて、重要箇所を指し示しているときや板書をしているときは、必然的に手の方向に顔が向いているため、手と顔（胴体）の中心付近を撮影の中心点とすることで、講師の顔が撮影画像の内側を向いている場面が多くなる。映像撮影や写真撮影のテクニックとして、顔の向いている方向に空間を開けた方が画面に奥行き感がでて見やすくなる効果があるといわれている^{14),15)}。この観点からも、講義中の講師の手と顔の中心付近を撮影の中心点とすることは好ましいと考えた。

過去において開発した試作システムと、今回開発した ChalkTalk-OSS とで講師領域の誤認識回数を比較する実験を行った。講師 3 名による 12 講義からそれぞれ 10 分間ずつの映像を抜き出して実験したところ、試作システムでは誤認識回数が 10 分あたり平均 14.3 回であったのに対し ChalkTalk-OSS では平均 0.7 回であり、誤認識回数が大きく低下したことが分かった。

3.3 板書静止画と講師動画の作成

講師領域が特定されると、次に、板書静止画と講師動画の作成が行われる。またこれと同時に、講師領域として特定された領域は、ハイビジョン映像から切り出されて講師動画としてストリーミングビデオ形式で出力される。

3.3.1 板書静止画からの講師の姿の除去

特定された講師領域の画像が、講師が存在しなかった時点の画像（講師がその領域に入る直前の画像）によって置き換えられ、あたかもその時点で講師が存在しなかったような画像が作成される。具体的には、特定された講師領域の画像を、講師が存在しなかった時点の画像（本実装の場合は、2 秒前の講師映像除去画像）によって置き換えるようになっている。したがって、作成された板書静止画は、見かけ上は 1 枚の写真画像のように見えるが、実際は、それぞれの部分が異なる時間に撮影された画像の貼り合わせであるといえる。さらに、板書静止画が保存されるのは板書内容に変更があったと認識された場合のみであり、板書が書き加えられたり消されたりしない間は新たな静止画は保存されない。これによってファイル保存容量を極力小さくしている。

なお、2 秒というタイミングは、講師が漢字 1 文字を板書するために費やす時間のほぼ平均時間であり、講義撮影ビデオを観察しこの値を定めた。当然、記憶容量に余裕がある場合には、2 秒より短い間隔で上記動作を繰り返すようにしてもかまわない。

3.3.2 自然なカメラワークを有する講師動画の作成

特定された講師領域は、その後、講師近傍動画を作成するためにも利用される。具体的には、ハイビジョン映像から講師領域部分が切り出され、講師動画としてストリーミングビデオ形式で出力される。しかしながら以前の試作システムでは、一定間隔（4 秒間隔）で講師位置を中心とした映像に強制的に切り替えていたために、映像に不連続感があり見にくいという問題があった。

この問題を解決するため、今回、2 パス方式で講師動画を作成する方式を導入し、カメラがパンしているような映像効果を加えるようにした。1 パス目で、講義ビデオ全体を通しての 1 秒ごとの講師の位置をインデックスファイルに書き出し、2 パス目で、そのインデックスファイルを参照して、撮影の中心点を線形補完によって滑らかに移動させるようにした。たとえば、1 パス目で、 t_1 時点の講師座標が x_1 であり、その 1 秒後の t_2 時点の講師座標が x_2 であるとする。このとき、2 パス目では、 t_1 から t_2 の間の任意時刻 t における撮影の中心点 x は、 x_1 と x_2 の座標を線形補完した座標となり次式で求まる。

$$x = x_1 + \frac{x_2 - x_1}{t_2 - t_1}(t - t_1) \quad (\text{ただし, } t_1 \leq t < t_2)$$

以上のように、任意時点の撮影の中心点を決定するために、最大 1 秒先の講師座標を利用する構成となっている。なお 1 秒という値は、講師が比較的素早く動いた場合に、約 1 秒間で撮影範囲から出てしまうという実験結果から求められた値である。この変更によりカメラワークはスムーズになり、映像に不連続感があるという問題は解決された。

しかしながら、生成映像を長時間学生に視聴させたところ、「撮影の中心点がつねに小刻みに変化する（パンが止まらない）ため、映像に酔ってしまうような感覚がある」という指摘を受けた。そこでシステムを改良し、講師領域の移動が少ない場合には、あえてカメラワークを行わないように変更した。具体的には、講師が左右に 100 ピクセル分以上動いた場合のみ、1 秒先の講師位置に（パンしながら）撮影の中心点を移動させるようにした。100 ピクセルという値は、「講師の胴体や手が画面からはみ出しそうになったらパンをする」という意味合いを有している。

くわえて、生成映像を学生に視聴させた際、「講師が動く先にカメラが先回りするカメラワークは、やや不自然であり違和感がある」との指摘があった。実際、最大 1 秒先の講師座標を利用して作成された講師動画



図 4 利用画面

Fig. 4 The use of the system.

のカメラワークは、人間がカメラワークを行ったときより平均 0.5 秒早いタイミングでスタートしていた。対策として、あえてカメラワークを 0.5 秒間遅らせるようにシステムに変更を施し、より人間のカメラワークに近い映像を作成するようにした。

以上の方法によって自動生成した講師動画（大学講義の収録映像）を、大学生被験者 11 名に視聴させ、見やすさに関して 5 段階評価（1：とても見にくい～5：とても見やすい）する実験を行った。具体的には、人が撮影した講師動画 Mh と、ChalkTalk-OSS が生成した講師動画 Mc とを、被験者にはどの動画がどの方法で作成された動画かを知らせずに約 5 分間視聴させた。なお、Mh は、三脚に固定した DV カメラを撮影者（ビデオ撮影に慣れている 30 歳代男性）が手動でパンさせて講師を撮影したものである。

評価の結果を述べる。実験開始当初、自動生成された映像が手動撮影映像に迫る評価を得ることを期待していたが、結果的には、手動撮影 Mh の平均点が 2.7 に対し、自動生成 Mc の平均点が 3.3 となり、自動生成した映像の方が優位との結果になった。また、講師動画 Mc が自動生成された映像であると気付いた被験者は 1 人もいなかった。さらに自由筆記によるアンケートからは、「Mh は細かいカメラ振れや急速なパンが多く、見ていてつかれる」、「Mh はつねにカメラが動いていて、映像に酔ってしまう」という意見が得られた。実際、Mh と Mc のパンの回数はそれぞれ 38 回と 16 回であった。ChalkTalk-OSS の場合は、講師が次はどこに移動するかを把握しているためにカメラワークが最低限の回数で済んだと考えられる。

3.4 板書静止画と講師動画の配信

ChalkTalk-OSS システムは、板書静止画と講師動画をユーザーに配信するための Web ページをあわせて

生成するため、Web サーバにコピーしてただちにコンテンツ配信ができるようになっている。生成された Web ページには、動画の進行にあわせて板書静止画を自動的に切り替える機能、および、選択された板書静止画に対応した時点から動画を再生する機能が備わっており、動画プレーヤプラグインと JavaScript とを組み合わせることで実装されている（図 4 参照）。

動画の進行にあわせて板書静止画を切り替える仕組みとしては、Web ブラウザが JavaScript の機能によって一定時間間隔で動画プレーヤプラグインの動画再生時点を監視し、必要に応じて、この動画再生時点に対応した板書静止画を Web サーバからダウンロードして表示するようになっている（各板書静止画のファイル名（URL）は、動画撮影開始時刻からの経過時間情報を含んでおり、JavaScript の機能によって選択的にダウンロードおよび表示が可能である）。一方、Web ブラウザ上でユーザが板書静止画を選択すると、JavaScript の機能によってこの板書静止画に対応した動画再生時点を算出し、動画プレーヤプラグインに対しシーク命令を発行するようになっている。

実際の講義を本システムで処理して Web ブラウザに表示する実験を行ったところ、大学生 11 名に対する自由筆記アンケート調査において、板書静止画が高画質であること（ハイビジョンカメラの画質に関し、板書文字が読めないという意見はどの被験者からも得られなかった）、黒板の死角をなくしていること、および、つねに黒板全体を見渡せること等から、試作システムが作成した板書静止画は見やすいとの意見を得ることができた。

また、ハイビジョンカメラで黒板領域全体を動画として保存したファイル容量に比較して、本システムが作成した全静止画ファイル（平均 1,400 × 平均 300 ピ

クセル)と講師近傍動画ファイル(320×240ピクセル300Kbps WMV=WindowsMedia形式)の合計容量は1/50以下であり,有効に帯域を圧縮できることが確認できた.講師と板書を単一動画ファイルに記録したとすると,板書内容が読める解像度を保つ必要があることから圧縮率をあまり高めることはできないが,講師の顔や姿のみであれば圧縮率をかなり高めて記録しても実質的に問題が生じない.このことも,ファイル容量の削減に大きく貢献していると考えられる.

4. おわりに

板書を使用する講義の映像から,マルチメディア Web 教材を自動的に作成する講義収録システム ChalkTalk-OSS を提案した.現状では,講師の胴体部分から左右に所定ピクセル以内に存在する移動物体領域は講師の手部分と見なして,胴体部分と領域をつなげて講師領域としているが,映像から講師の胴体幅を認識して,つなげるべき幅を自動設定する機能が望まれる.また,ChalkTalk-OSS を電子プレゼンテーションの収録にも使いたいというユーザの声もあり,静止画作成の最短時間間隔の最適化を行う等の検討をする必要がある.さらに,今回の評価実験においては,比較的短時間で結果が得られるコンテンツの「見やすさ」に関して評価を行い有効性を判断したが,より望ましくは,様々な教育コンテンツにおいて受講者の理解を助けることができたかどうかによって判断する必要がある.生成されたコンテンツが受講者の理解度に与える影響等については,今後の実運用を通じて評価していきたいと考えている.

参考文献

- 市村, 富野, 井上, 松下: 講師映像と板書静止画の記録が可能な講義自動収録システム, 情報処理学会, グループウェアとネットワークサービス研究会報告, GN-56-2 (2005).
- 森田, 井上, 市村, 松下: 講義自動収録システムにおける板書静止画記録法, 情報処理学会全国大会, 3Z-5 (2005).
- 板宮, 林, 千代倉: ワンマン録画可能な講義ビデオ作成システム, 情報処理学会コンピュータと教育研究報告, No.70, pp.17-20 (2003).
- 大西, 村上, 福永: 状況理解と映像評価に基づく講義の知的自動撮影, 電子情報通信学会論文誌 (D-II), Vol.J85-D-II, No.4, pp.594-603 (2002).
- 大西, 泉, 福永: 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.J83-D-I, No.11, pp.1187-1195 (2000).
- 大西, 泉, 福永: 情報発生量の分布に基づく遠隔講義撮影の自動化, 電子情報通信学会論文誌 (D-II), Vol.J82-D-II, No.10, pp.1590-1597 (1999).
- 宮崎, 亀田, 美濃: 複数カメラを用いた複数ユーザに対する講義の実時間映像化法, 電子情報通信学会論文誌 (D-II), Vol.J82-D-II, No.10, pp.1598-1605 (1999).
- <http://www.sony.jp/products/Consumer/handycam/PRODUCTS/HDR-HC1/>
- Cruz, G. and Hill, R.: Capturing and Playing Multimedia Events with STREAMS, *Proc. ACM Multimedia 94*, pp.193-200 (1994).
- Chiu, P., Kapuskar, A., Reitmeier, S. and Wilcox, L.: NoteLook: Taking notes in meetings with digital video and ink, *Proc. ACM Multimedia 99*, pp.149-158 (1999).
- Uchihashi, S.: Improvising camera control for capturing meeting activities using a floor plan, *Proc. ACM Multimedia 01*, pp.12-18 (2001).
- 昌達 K'z: 画像処理を極めるアルゴリズムラボ, C MAGAZINE, 2003年6月号 (2003).
- 宮崎: 動画像処理技術による映像監視の高度化, 沖テクニカルレビュー, Vol.70, No.3, pp.72-75 (2003). http://www.oki.com/jp/Home/JIS/Books/KENKAI/n195/pdf/195_R17.pdf
- ジャレミー・ヴィンヤード: 映画技法完全レファレンス, フィルムアート社 (2002).
- 松本: 図解デジタルビデオ編集のしくみ, ディー・アート社 (2001).

(平成 18 年 1 月 18 日受付)

(平成 18 年 7 月 4 日採録)

推薦文

計算機を駆使した授業支援システムは数多く提案されているが, 黒板を用いた授業が依然として多いのも事実である. 本論文は, 黒板を用いた授業の記録において, 講師の影になって見えない部分を処理によって見やすくした, 現実的な講義自動収録システム ChalkTalk-OSS を提案している. 本システムは以前のシステムと比較して講師位置検出を工夫し, また人間が行うカメラワークに近い教師動画作成処理等を提案して, 有用性に特に優れていると考えられるため, 推薦論文に値すると判断した.

(グループウェアとネットワークサービス研究会主査
宗森 純)



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, 電子情報通信学会会員。



福井登志也

1989年日立製作所半導体事業部入社。1992年日立茨城工業専門学院管理工学科卒業。1997年東京ソフト販売入社。1997~2003年富士ゼロックスに派遣、ソフトウェア開発業務に従事。2003~2005年蝶理情報システムに派遣、官公庁システムの開発業務に従事。2005年より東京工科大学 Linux オープンソースソフトウェアセンターにてソフトウェアの研究開発業務に従事、現在に至る。



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年同大学大学院理工学研究科前期博士課程修了。2005年同大学院理工学研究科後期博士課程修了。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。グループウェア、マルチメディアコンテンツ処理の研究に従事。DICOMO 2006 ヤングリサーチ賞受賞。



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年イリノイ大学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。現在、住宅情報化推進協議会会長。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、バ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 研究会委員長等を歴任。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年および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。