

# 板書講義映像コンテンツにおける モバイル環境向けカメラワークの提案

佐藤 雄<sup>1,a)</sup> 森澤 竜<sup>1</sup> 市村 哲<sup>1,b)</sup>

**概要:** e-ラーニングの増加や、タブレット、スマートフォンなどのモバイル端末の普及とともに、モバイルラーニングが注目を集めている。筆者らは、PC 向けの学習用講義映像と板書の静止画を作成する自動講義収録システム ChalkTalk の開発を行ってきた。本稿では、モバイル環境向けの学習用講義映像向けのカメラワークとして、板書中心のカメラワークと講師のピクチャインピクチャ表示を行う手法を提案する。実際に、このカメラワークに基づく講義動画を出力するシステムを実装し、講義動画の評価を行った。評価の結果、利用者が見たい板書を重点的に見せながら講師の様子も分かりやすく表示できることを確認した。

**キーワード:** 講義映像、板書認識、カメラワーク、ピクチャインピクチャ、モバイルラーニング

## A Proposal of Camerawork in Lecture Video for Mobile Devices

YUICHI SATO<sup>1,a)</sup> RYO MORISAWA<sup>1</sup> SATOSHI ICHIMURA<sup>1,b)</sup>

**Abstract:** Mobile-learning has become popular as mobile devices, such as tablet or Smartphone, become widely used. In the past, we had developed ChalkTalk, which automatically produces E-learning materials including a lecturer-tracking video and entire blackboard image from videotaped chalk talks. In this paper we propose a new method for mobile devices to improve the ChalkTalk system. The methods include blackboard centric camerawork and picture-in-picture camerawork. As a result of the evaluation, it was found that the new camerawork can better display the area that users want to see.

**Keywords:** lecture video, blackboard recognition, camerawork, picture-in-picture, Mobile-learning

### 1. はじめに

コンピュータやインターネットを利用した学習形態に e-ラーニングがある。学習者は映像や音声、文章などから構成される学習用コンテンツを視聴、閲覧することができる。e-ラーニングには、講師と学習者が一カ所に集まる必要がない、自分のペースで学習できるといった利点があり、大学や企業などで多く採用されている。また近年、Android や iPad などのスマートフォン、タブレット端末が流行し、これらのモバイル端末を利用した学習形態であるモバイル

ラーニングが注目されている。モバイルラーニングでは、e-ラーニングの利点に加え、学習者が自宅や通学中などいつでもどこでも学習ができるという利点がある。ただしこれらには欠点もあり、学習用コンテンツの作成や配信にコストがかかる点などが挙げられる。例えば、講義映像コンテンツの作成を考えた場合、通常、専属カメラマンによる講義の撮影が必要となる。また、コンテンツを配信するには配信用サーバが必要であり、大容量のコンテンツを配信するには十分な設備も必要である。

学習用コンテンツの自動作成を行うシステムとして、筆者らが過去に開発した ChalkTalk がある [1]。ChalkTalk は、板書を用いた講義を教室後方に設置したハイビジョンビデオカメラで撮影し、その映像から講師の追尾動画（講師付近のみを拡大した動画）と黒板全体の静止画を自動作

<sup>1</sup> 東京工科大学  
School of Computer Science, Tokyo University of Technol-  
ogy, Katakura 1404-1, Hachioji, Tokyo, 192-0982, Japan

a) yuichi.sato.public@gmail.com

b) ichimura@stf.teu.ac.jp

成する講義収録システムである。作成したコンテンツはデスクトップ PC やノート PC で視聴でき、利用者は動画と静止画で、講師と板書の両方を見ながら学習することができる。しかし、このような講義映像コンテンツをモバイル端末で視聴しようとした場合、モバイル端末ゆえの問題が発生する。モバイル端末は画面が小さく、例えば、動画で黒板全体を映した場合、文字が小さくて読めないことがある。動画で講師付近をズームして映した場合は、板書が十分に映らないこともある。ChalkTalk が生成したコンテンツのような動画と静止画の複合コンテンツの表示を考えた場合もやはり画面が小さく両方を十分に表示することができない。

筆者らは過去に、ChalkTalk が出力したコンテンツのモバイル環境向け様式に関する実験を行った [2]。実験は、横方向にスクロール可能な動画プレイヤーを用いて、利用者が講義映像の中で何を見たいのか調査するものである。実験の結果、以下のような知見を得た。

- 講師の姿より、書かれた板書を重点的に見たい。また、書いている途中および書かれた直後の板書を良く見たい
- 講師が画面上に映っていない時、講師がどこにいて何をしているのか気になる

したがって、モバイル環境向けの講義映像コンテンツでは、利用者が講師と板書の両方を見られることが必要である。

この課題のため、本稿では、ChalkTalk が作成する講義動画をモバイル端末で視聴する際のカメラワークについて検討および提案を行う。提案するカメラワークは、講義動画において板書を中心にズームして映しながら、講師が映らないときはピクチャインピクチャで講師を表示するものである（以後、ピクチャインピクチャのことを PinP と表す）。板書を中心に映しながら端に PinP で講師を映すことで、利用者は講師と板書を確実に見ることができ、このカメラワークに基づく講義動画を作成するシステムを実装し、評価実験を行った。評価の結果、利用者の見たい板書を見せることができるとことや、PinP を用いた板書と講師の両方を見せることができ満足度が高いことを確認した。

## 2. 既存研究

### 2.1 モバイル環境向け講義受講システム

モバイル環境での講義受講システムとしては iPod, iPhone とポッドキャストを利用した大学の講義映像視聴システムが存在する [3][4]。撮影方法としては、専属カメラマンが講師をビデオ撮影するものや、固定カメラや首振りカメラを用いて板書全体を撮影するもの [5][6][7] などがある。撮影した映像を iPod 向けにエンコードし、ポッドキャストコンテンツとして Web サーバに配置している。学生は iTunes 等のポッドキャスト受信ツールにコンテンツを登録するこ

とで、講義映像を自動的に iPod にダウンロードし、通学中等の空いた時間に視聴することができる。しかし、使用するモバイル端末のディスプレイが PC のディスプレイに比べて極めて小さく、そのため、黒板全体を撮影すると板書の文字が小さくなり読み取れないという問題や、講師を中心にズームして撮影すると講師付近の板書しか見えないという問題がある。モバイル端末のディスプレイ面積が小さいため、講師が板書しながら黒板の前を歩き回るような映像を収録するためには、通常、専属カメラマンがビデオカメラを操作し、講師や板書を適宜ズームしながら撮影しなければならなかった。一方、専属カメラマンを導入しない場合には、黒板全体が映るようにビデオカメラを固定設置して撮影し、それをモバイル端末で表示することが考えられるが、板書の文字が小さくなり読み取れないという問題や、講師の表情が分からないという問題が生じる。

### 2.2 ChalkTalk

ChalkTalk はハイビジョン撮影した講義映像から講師の自動追尾動画と黒板全体の静止画アニメーションを出力するツールである。まず、教室の後方にハイビジョンビデオカメラを黒板全体が映るように固定設置し、講義を撮影する。次に撮影した映像を ChalkTalk で処理する。ハイビジョン撮影した講義動画から黒板領域内の移動物体を検出し講師領域とする。その際、移動物体の検出にはフレーム間差分法を用い、移動物体領域として特定されたマクロブロックの数を黒板の横軸座標毎に集計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しカメラワークを行っている。

- (1) 移動物体領域が横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
- (2) 講師の胴体部分から左右に特定のピクセル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) 映像酔いを防ぐため、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ず、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
- (4) 人間が行うカメラワークに似せるため、講師が移動したタイミングから意図的に 0.5 秒遅らせてパンをする

これらの処理を経て、講師近傍領域の小さな動画を出力すると同時に H.264/MPEG4 AVC 技術を用いた動画圧縮を行う。また黒板全体の静止画アニメーションは、講師領域が変化したときに黒板領域の講師がいない部分を更新することにより作成される。元ハイビジョン動画に比べ、これらのコンテンツのサイズは約 1/20 に圧縮されている。また、利用者はデスクトップ PC やノート PC などのブラウザ上でこれらのコンテンツを視聴することができる。

PC などの環境ではディスプレイが大きく、動画と黒板

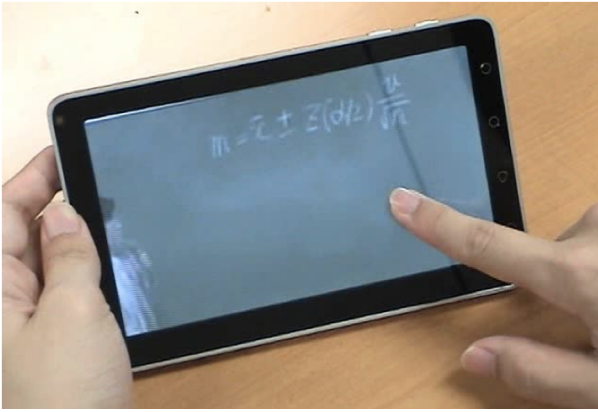


図 1 スクロール操作の様子  
Fig. 1 Scroll action by a user

全体の静止画アニメーションの両方を容易に表示することができたが、スマートフォンなどのモバイル環境ではディスプレイが小さく、黒板全体を表示するのは適切ではない。したがって、モバイル環境においては、ChalkTalkが出力した動画のみを表示することが考えられるが、この場合、講師付近の板書しか見えないという問題が発生する。

### 2.3 ChalkTalk コンテンツのモバイル向け様式の検討

筆者らは過去に、ChalkTalk コンテンツのモバイル環境向け様式について実験を行った [2]。実験は、図 1 のような横方向にスクロール可能な動画プレイヤーを用いてハイビジョンの講義動画を被験者に見てもらい、利用者が講義動画中の何をみたいのか評価するものである。講義動画の縦方向はスマートフォンの画面いっぱいに表示されるが横方向は画面からはみ出ており、はみ出ている部分を見るには画面を横方向に指でなぞりスクロールさせる必要がある。実験の結果、利用者の傾向について以下の知見を得た。

- 講師の姿より、書かれた板書を重点的に見たい。また、書いている途中および書かれた直後の板書を良く見たい
- 講師が画面上に映っていない時、講師がどこにいて何をしているのか気になる

以上より、本研究において解決すべき課題は以下の通りである。

**課題 1** 見たい板書を重点的に見せる必要がある

**課題 2** 見たい板書に加え、講師の様子も分かるようにする必要がある

## 3. 提案

前章最後に挙げた課題を解決するため、講師と板書の両方を十分に見せるカメラワークを提案する。板書を中央に映しながら、端に PinP で講師を表示するものである。以下にその詳細を示す。

ChalkTalk の講師追尾動画は、講師を中心にカメラワー

クを行っており、板書を考慮していなかった。本稿では、モバイル端末において動画だけで板書も十分に見せるため、板書中心のカメラワークを提案する。提案するカメラワークを以下に示す。

- (1) 講師が板書を書いているときは、講師を中心にズームする。これは ChalkTalk のカメラワークと同じである
- (2) 講師が板書を書き始めるとき、これから書かれるひとまとまりの板書が中央に映るようにカメラをパンし、板書終了後、一定時間経過するまで固定する。
- (3) 一定時間経過し、新たな板書が検出されれば、(2)と同様にその板書が映るようにカメラをパンし一定時間映す。その後、また新たな板書があれば、これを繰り返す。新たな板書が無ければ、(1)のカメラワークに戻る
- (4) カメラを板書に固定しているときに講師が板書から離れ、画面内に映らなくなった場合、講師付近の様子を PinP で動画内に表示し、一定時間後に (1) のカメラワークに戻る。

(1) は課題 1 に対応し、板書がまだ書かれていない時は主に先生が映っていることが望ましいと考えられるためである。(2)、(3) は課題 2 に対応し、書いている途中および書かれた直後の板書をよく見たいという要望があったためである。(4) は課題 1 に対応し、講師が画面から消えると、どこにて何をしているのか気になるという意見のためである。これらを合わせることで、利用者は講師と板書の両方を見ることができる。

(2) における、ひとまとまりの板書とは、講師が黒板上に距離的、および時間的に連続して書いた文字や記号の集合のことである。例えば、黒板上に書かれた文字列や数式等がひとまとまりの板書である。(2) では、このひとまとまりの板書領域が画面中央に映るようにカメラをパンする。しかし、板書を書き始めた時には、これから書かれる板書がどこにどれだけ書かれるかは分からない。したがって、あらかじめ、その板書を書き終わるところまで動画を読み込み、板書領域の認識を行っておくようにした。(4) では、板書を中央に映しながら、端に PinP で講師を表示する。PinP の表示位置は画面四隅のうち板書がないところに表示するのが望ましい。図 2 は PinP の例で、画面中央に書き終わった板書を一定時間経過するまで表示しながら、画面右端に PinP で講師の様子を表示するようにした。

## 4. 実装

筆者らは、提案したカメラワークに基づく講義動画収録システムを実装した。これには、まず講師の位置や板書の位置、講師の板書の書き始め、書き終わりタイミングの認識が必要となった。

板書の位置に関しては、背景モデルを用いる手法 [8] や、

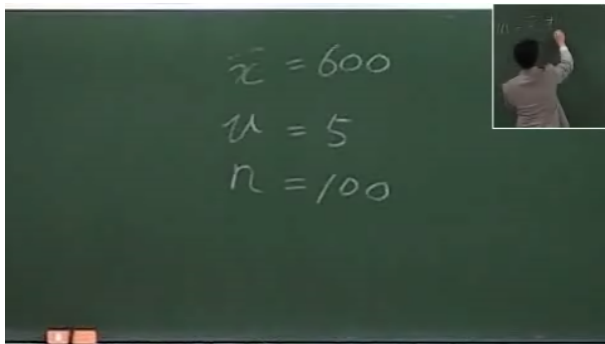


図 2 PinP の例  
 Fig. 2 Example of PinP

時空間画像からエッジ抽出を行い分析を行う手法 [9] がある。前者は背景差分法、後者はフレーム間差分法にそれぞれ基づく手法である。講師の板書の書き始め、書き終わりタイミングの認識については、フレーム間差分法やチョーク音を用いる手法 [10] がある。

今回の実装では、講師の位置、板書の位置、板書の書き始め、書き終わりの認識について、時空間画像からエッジ抽出を行い分析を行う手法を用いた。時空間画像やエッジ抽出を用いる既存研究 [9] と類似しているが、板書の書き始めと書き終わりの認識を行うための処理を変更および追加している点が異なる。また講義動画作成時にカメラワークを決定するには、3章で述べたように事前に板書位置の認識が必要である。そこで、動画作成前に一度、元のハイビジョン動画を読み込んで、板書位置、板書の書き始め、書き終わりなど必要な情報の認識を行うようにした。

#### 4.1 時空間画像

認識アルゴリズムの詳細を述べる。本稿における時空間画像とは、図 3 のように動画の各フレーム画像を時系列順に並べた断面画像のことである。図中の  $x-t$  で示されたある  $y$  座標における断面画像を  $x-t$  時空間画像、 $y-t$  で示されたある  $x$  座標における断面画像を  $y-t$  時空間画像とする。図 4 と図 5 はそれぞれ  $x-t$  時空間画像と  $y-t$  時空間画像の例である。図 4 の横方向は図 3 における  $x$  方向、縦方向は図 3 における時間方向  $t$  を表している。図 5 の横方向は図 3 における時間方向  $t$ 、縦方向は図 3 における  $y$  方向を表している。

また、時空間画像の作成においては、長い動画になると多大な記憶領域や処理時間が必要となる。そのため今回の実装では、30fps の動画を想定し、10 フレームにつき 1 フレーム (0.33 秒につき 1 フレーム) ずつ抽出し時空間画像を作成することとした。

#### 4.2 エッジ抽出

時空間画像の作成後、エッジ抽出を行う。エッジ抽出は縦と横の 2 種類の Sobel オペレータを用いて行う。以下に

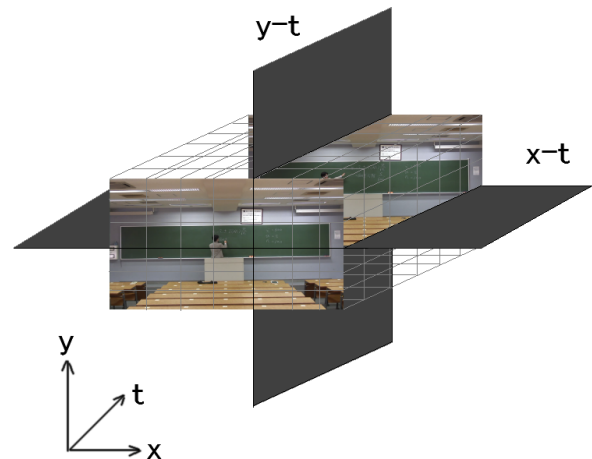


図 3 時空間画像の例  
 Fig. 3 Example of spatiotemporal image

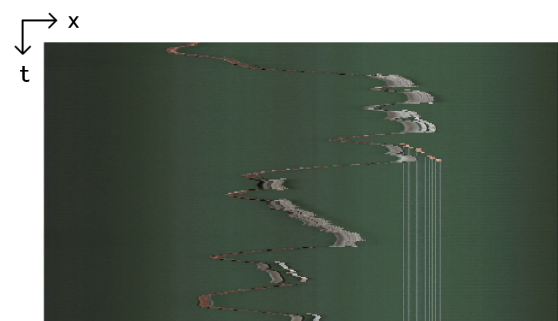


図 4  $x-t$  時空間画像の例  
 Fig. 4 Example of  $x-t$  spatiotemporal image

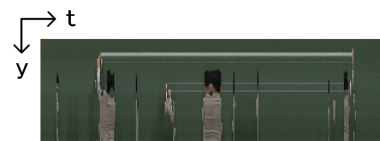


図 5  $y-t$  時空間画像の例  
 Fig. 5 Example of  $y-t$  spatiotemporal image

示す  $s_x$  が縦方向のエッジを、 $s_y$  が横方向のエッジをそれぞれ抽出する Sobel オペレータである。

$$s_x : \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad s_y : \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

時空間画像において、縦横 2 種類のエッジ抽出オペレータを用いるのは、静止物体 (板書) と動体 (講師) を認識するためである。  $x-t$  時空間画像において縦方向のエッジを抽出すると、 $x$  方向における静止物体が濃く現れ、横方向のエッジを抽出すると、 $x$  方向における動体が濃く現れる。以後これらを  $x-t$  静エッジ画像、 $x-t$  動エッジ画像とそれぞれ表す。  $y-t$  時空間画像においては縦方向のエッジを抽出すると、 $y$  方向における動体が濃く現れ、横方向のエッジ

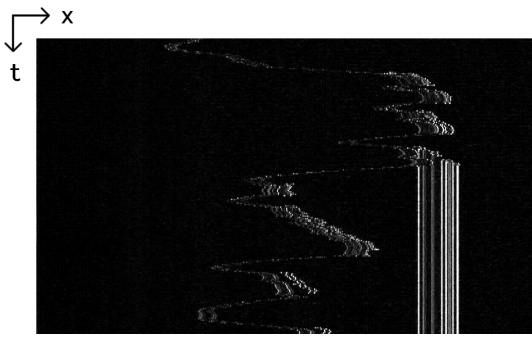


図 6 x-t 静エッジ画像

Fig. 6 Example of x-t static edge image

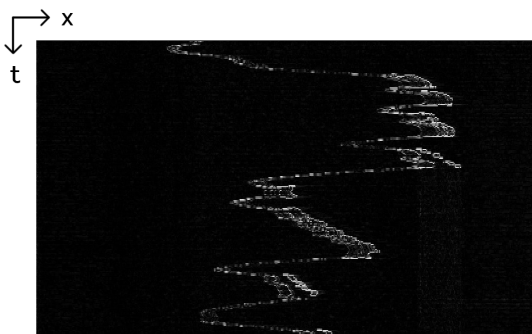


図 7 x-t 動エッジ画像

Fig. 7 Example of x-t dynamic image

を抽出すると、y 方向における静止物体が濃く現れる。以後これらを y-t 動エッジ画像、y-t 静エッジ画像とそれぞれ表す。図 6 は x-t 静エッジ画像、図 7 は x-t 動エッジ画像である。図 6 の右部分には、縦方向にまっすぐ伸びるエッジが濃く現れている。これが板書の文字の一部を表している。図 7 では、図 7 で現れていた縦方向にまっすぐ伸びるエッジが消え、講師位置と思われるエッジが残っており、濃く現れている。y-t 時空間画像からも同様のエッジ画像が作成でき、これらのエッジ画像をもとに講師の位置、板書の位置、板書の書き始め、書き終わりタイミングを認識する。

#### 4.3 講師位置の認識

講師位置は、図 7 のような、x-t 動エッジ画像を用いて認識する。x-t 時空間画像は、動画の縦の長さの数だけ作成されるが、作成時に y 座標を指定し、その y 座標の x-t 時空間画像 1 つだけを用いる。これは、経験的に 1 つだけで十分に講師位置を認識できたためである。また y-t 動エッジ画像を用いないのは、講師は縦方向には移動せず、縦方向の動きを見る必要がないためである。x-t 動エッジ画像においては、講師はある時間 t において濃いエッジが多く出ている x 座標付近にいると考えられる。これより、以下のように各時間における講師位置を決定する。

- (1) 全ての x 座標について、その x 座標とその付近の x 座標のエッジ量の合計値を計算する

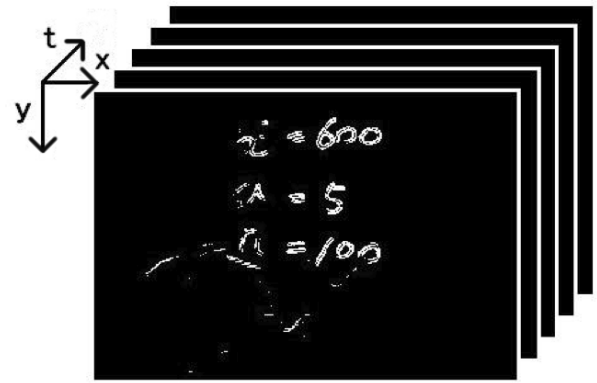


図 8 静エッジ直方体の例

Fig. 8 Example of static edge cuboid

- (2) 最大の合計値を探し、その値が閾値以上ならば、その x 座標を講師位置とする
  - (3) 閾値以下ならば、前後の時間から線形補間する。
- (2), (3) で閾値を用いるのは、講師が動いていない場合エッジ量の合計値が小さくなる傾向があり、他に偶発的に動画内に入る受講者の頭などが講師位置として認識されることがあったためである。閾値を用いることで、講師の移動のようなより大きな動きのみを認識することができる。また閾値以下のときは、講師はほとんど動いておらず、線形補間で問題ないと考えられる。

#### 4.4 板書位置の認識

板書位置は、静エッジ画像を用いて認識する。まず x-t 静エッジ画像と y-t 静エッジ画像を組合せ、静エッジ直方体を作成する。x-t 静エッジ画像の各ピクセル値  $f_x(x, t)$ 、y-t 静エッジ画像の各ピクセル値  $f_y(y, t)$  から、以下の値を計算する。

$$f(x, y, t) = \sqrt{f_x(x, t)^2 + f_y(y, t)^2}$$

この  $f(x, y, t)$  を 2 値化したものが静エッジ直方体である。図 8 は、静エッジ直方体の例である。これは静止物体をエッジとして抽出した画像を時系列順に並べたもの、すなわち静止物体のみを表示する動画のようなものである。図 8 では、板書文字がエッジとして現れているのが分かる。板書の下には講師の肩と腕の輪郭と思われるエッジも現れている。時系列的に次の画像を見ると、多くの場合、板書文字エッジの位置はそのまま、講師の肩と腕のエッジの位置は移動すると考えられる。したがって、板書文字を認識するには、時系列的に見て一定時間同じ位置に存在するエッジを探せばよい。以下に、板書位置の認識手順を示す。

- (1) 静エッジ直方体の全て xy 座標について、t 方向にエッジを見ていき、エッジが一定割合以上存在する一定長さ以上の区間を探す。区間は x 座標、y 座標、エッジ開始時間、エッジ終了時間で表せ、静エッジ直方体内の t 方向の線分と見なせる。以後、これをエッジ線分と表す

(2) (1) で求めたエッジ線分のうち近くにあるものを連結し仮板書とする。連結するエッジ線分はエッジ開始時間が近く、かつ、 $xy$  座標における距離が近いものである。また、3つ以上連結する際には、それぞれのエッジ線分が互いに近い必要はなく、どれか2つのエッジ線分が近ければ連結する

(3) (2) で求めた各仮板書の領域を求める。連結したエッジ線分のうち、 $xy$  座標において左端と右端にあるエッジ線分の  $x$  座標をそれぞれ板書領域の左端と右端、上端と下端にあるエッジ線分の  $y$  座標をそれぞれ板書領域の上端と下端とする

(4) (3) で求めた仮板書の領域の大きさが一定以上のものを真の板書とする

(1) でエッジが割合以上存在する区間を探しているのは、板書が講師の体によって隠れることがあり、途中隠れることがあっても板書を構成する文字を認識できるようにするためである。(2) によって距離的に時間的に近いエッジ線分群は連結され1つの仮板書となり、遠いエッジ線分群は、別の仮板書として連結される。(4) で板書領域の大きさを調べ、小さいものはノイズとして板書と認めない。

#### 4.5 板書タイミングの認識

板書の書き始め、書き終わりタイミングは、板書位置の認識で用いたエッジ線分を用いて認識する。板書位置の認識においては、板書は複数のエッジ線分から構成され、各エッジ線分を表す情報の1つにエッジ開始時間があった。これはエッジが出始めた時間、すなわち、講師がそのエッジ線分の  $xy$  座標に板書しそれが初めてカメラに映った時間である。そこで、板書を構成するエッジ線分のエッジ開始時間で、最も早いものを板書の書き始め、最も遅いものを板書の書き終わりタイミングとした。これは講師の体によって板書が隠れてエッジ開始時間が実際の講師の動作より遅くなることがあり、厳密ではないと考えられる。しかし、提案したカメラワークを行うことを考えると、板書が現れた時にカメラワークを行うことが重要であり、この誤差は逆に望ましいと考えられる。

### 5. 評価

提案したカメラワークに基づく講義動画を作成するシステムを実装した。このシステムが作成する講義動画が、本研究の課題を解決できているか確かめるため2種類の評価実験を行った。実験の目的は、以下の通りである。

**実験目的 1** 見たい板書を重点的に見せられているか確認する

**実験目的 2** 板書と同時に講師の様子を確認できるか評価する

実験目的 1 は課題 1 に、実験目的 2 は課題 2 にそれぞれ対

表 1 実験に用いた元の講義動画

Table 1 Original HD lecture video used in experiment

	サイズ [ピクセル]	長さ [sec]	講義内容
動画 1	1920x1080	180	プログラミング
動画 2	1920x1080	180	統計学
動画 3	1920x1080	180	物理学

応している。以後、ChalkTalk のことを旧 ChalkTalk、今回実装したシステムのことを新 ChalkTalk とそれぞれ表すことにする。また ChalkTalk で変換する前の動画である元のハイビジョンの動画のことを元動画と表す。各実験に用いた、元動画の詳細を表 1 に示す。

#### 5.1 実験 1

実験 1 は実験目的 1 に対応しており、新 ChalkTalk の講義動画が講義動画中における被験者が見たい板書を映せているか評価する実験である。まず被験者が見たいものを調査するため、被験者に過去の実験で用いた横方向にスクロール可能な動画プレイヤーと同様のソフトウェアを用いて、被験者に元動画を視聴してもらった。この時のスクロールの履歴は自動的に保存されており、各フレームにおいて被験者が何を見ていたか分析できる。次に動画の各フレームにおいて、新および旧 ChalkTalk が作成した講義動画のカメラ位置と、スクロール履歴から分析した被験者のスクロール位置との距離を求め、全フレームにおける平均値を求めた。この距離の平均値が小さいほど被験者が見たい板書を見せられていると考えられる。

##### 5.1.1 実験条件

実験の被験者は大学生 10 人である。実験に用いた元の講義動画は、表 1 で示した 3 種類の動画である。動画プレイヤーおよび新旧 ChalkTalk が作成する講義動画のサイズは、ともに 480x272 ピクセルとした。

##### 5.1.2 実験結果

被験者のスクロールを分析した結果を表 2 に示す。いずれの被験者も板書と講師両方を同時に見ようとするスクロールが見られた。加えて、過去の板書を見直すようなスクロールも被験者 C, D, E, F の 4 人で見られた。また、各被験者のスクロール位置と講師動画のカメラ位置の平均距離を表 3 に示す。いずれの被験者のスクロールと比較しても、旧 ChalkTalk の講義動画に比べ、新 ChalkTalk 講義動画のカメラ位置は被験者のスクロール位置に近い結果となった。また平均距離の平均について、片側有意水準 5% の  $t$  検定を行ったところ有意差が認められた。

#### 5.2 実験 2

実験 2 は実験目的 2 に対応しており、新 ChalkTalk の講義動画が板書に加えて講師の様子も十分表示できているか評価する実験である。まず被験者に新旧 ChalkTalk が作成

表 2 被験者のスクロールの分析  
Table 2 Anlyzation of users' scroll action

被験者	元動画	スクロール回数	特徴
A	動画 1	7	スクロール回数が少なく、板書が映るように停止している時間が長い
B	動画 1	11	同上
C	動画 1	11	同上 加えて、過去の板書を見直すようなスクロールが見られた
D	動画 1	16	板書が映るように停止している時間が長い。2つの板書を交互に見比べるようなカメラ移動が複数回見られた
E	動画 2	46	板書を主にしながら講師に合わせて少しずらすようなスクロール、および、過去の板書と現在書いている途中の板書を見比べるようなスクロールが複数回見られた
F	動画 2	22	板書を主にしながら講師に合わせて少しずらすようなスクロール、および、過去の板書を見直すようなスクロールが見られた
G	動画 2	20	講師の動きに合わせたスクロールが多く見られた
H	動画 3	29	板書中および板書した直後は板書を、それ以外の身振りでの説明中の時は講師を見るようなスクロールが見られた
I	動画 3	18	同上
J	動画 3	16	板書見ており講師が映っていない時間が長い

表 3 各被験者のスクロール位置と新旧 ChalkTalk 動画のカメラ位置の平均距離 (単位:ピクセル)

Table 3 Difference between users' scroll and system-generated scroll

	新 ChalkTalk	旧 ChalkTalk
平均	93.69	158.81
標準偏差	41.43	82.94

した講義動画の5段階印象評価を行ってもらった。評価項目は「勉強を行うのにこの講義動画は見やすいか」で、5が見やすい、1が見つらいとした。印象評価のあと、動画の見やすさについてどう感じたかインタビューを行った。また評価してもらった講義動画は新旧 ChalkTalk の2種類に、新 ChalkTalk の講義動画の PinP なしのもを作成しこれを加えて3種類とした。

### 5.2.1 実験条件

実験の被験者は大学生9人である。9人それぞれに旧 ChalkTalk と新 ChalkTalk (PinP あり) と新 ChalkTalk (PinP なし) の3種類の講義動画を見せ評価を行ってもらった。見せる順番や元動画の組合せは、極力異なるようにした。実験に用いた元の講義動画は、表1で示した3種類の動画である。講義動画のサイズは480ピクセルx272ピクセルで、新 ChalkTalk の講義動画における PinP のサイズは50ピクセルx50ピクセルとした。また、新 ChalkTalk のカメラワークにおいて各板書を見せる時間は10秒とした。

### 5.2.2 実験結果

実験を行った結果、各講義動画の評価平均および分散は表4のようになった。この結果について、両側有意水準5%でマンホイットニーのU検定を行ったところ、旧 ChalkTalk と新 ChalkTalk (PinP あり)、旧 ChalkTalk と新 ChalkTalk (PinP なし) で、有意差が認められた。この

表 4 印象評価の結果

Table 4 Result of impression evaluation

	平均	分散
旧 ChalkTalk	2.11	1.11
新 ChalkTalk (PinP あり)	3.89	0.61
新 ChalkTalk (PinP なし)	3.33	0.75

ため、板書重視カメラワークの講義動画は、講師追尾動画と比べ見やすいと言える。PinP ありとなしの差について有意差は認められなかった。

インタビューの結果を表5、表6、表7に示す。それぞれ、旧 ChalkTalk、新 ChalkTalk (PinP あり)、新 ChalkTalk (PinP なし) の講義動画に対するインタビュー結果をまとめたものである。旧 ChalkTalk 講義動画に対しては、講師のみを表示するだけで板書が見つらいという否定的な意見が多い一方で、講師が見れて良かったという肯定的な意見もあった。新 ChalkTalk (PinP あり) 講義動画については、PinP で講師が映っていることに対して肯定的な意見が多く、講師を見たいという被験者の欲求に応えられていると考えられる。また PinP によって板書が見つらかったという問題の指摘や、板書さえ見えていればよいという否定的な意見もあった。PinP なしの講義動画に対しては、PinP ありのものとは比べ、やはり講師が見えないことを指摘する意見が多く見られた。また板書重視のカメラワークについてカメラ遷移の問題を指摘する意見もあった。階印象評価においては PinP ありとなしで有意差は認められなかったにも関わらず、このような意見が新 ChalkTalk (PinP あり) のときに出なかったのは、板書を中心に表示し講師が映っていないときでも PinP によって講師の様子が分かり被験者を安心させられたためではないかと考えられる。

表 5 旧 ChalkTalk 講義動画に関するインタビュー結果

Table 5 Answer to interview about old ChalkTalk

件数	内容
6	講師を追尾する講義動画で板書が見づらい
1	講師をちゃんと追いかけていて良かった

表 6 新 ChalkTalk 講義動画 (PinP あり) に関するインタビュー結果

Table 6 Answer to interview about new ChalkTalk(with PinP)

件数	内容
6	PinP で講師が映っており見やすかった
2	PinP で板書が見えづらくなるがあった
1	講師が見えていなくても良い

表 7 新 ChalkTalk 講義動画 (PinP なし) の講義動画に関するインタビュー結果

Table 7 Answer to interview about new ChalkTalk(without PinP)

件数	内容
4	講師が見えないのが気になる
2	板書の出現が時間的に連続した場合のカメラ遷移が遅れている・速すぎる

## 6. 考察

### 6.1 実験結果について

課題 1 について評価した実験 1 の結果より、提案したカメラワークは被験者の見たい板書により近づけられたと考えられる。それでも、新 ChalkTalk のカメラ位置と被験者のスクロール位置の平均距離の平均は 93.69 ピクセルという数値であり、改善の余地がある。これは、まず多くの被験者が板書と講師が両方映せるときは、両方映るようにスクロール位置を調整しており板書が画面端の方になることが度々あったためである。実装した板書中心のカメラワークでは、板書をカメラの中心に表示し PinP で講師を同時に表示するため、板書が端になることは少ない。したがって、数値上は改善の余地がありそうだが、実際には板書は十分見せられていると考えられる。

課題 2 について評価した実験 2 の結果より、板書を中心にカメラワークを行いながら、PinP で講師の様子を表示する手法は被験者から高い評価を得られた。特に PinP で講師が映っていて見やすかったという意見が多く、課題は解決できていると考えられる。

### 6.2 今後の課題

インタビューの結果から板書の表示に関する問題点が見つかった。インタビューで得られた意見の「PinP で板書が見えづらくなることもある」というものである。提案したカメラワークでは PinP の表示位置を考慮していなかつ

たためである。そのため、PinP の表示位置と板書位置を考慮したカメラ位置の決定を行えば解決すると考えられる。また、「板書の出現が時間的に連続した場合のカメラ遷移が遅すぎたり早すぎたりする場合がある」というものがあった。現在は、1つの板書を見せる時間は、書き終わってあるいは見せ始めてから一定時間、今回の実験では 10 秒にしている。そのため、書き始めから見せている板書は次の板書を見せ始めるまで長く感じられたり、板書の量が多いものや少ないものは、それぞれ早すぎたり遅すぎたり感じられるものと思われる。改善策としては、板書の量に応じて板書を見せる時間を決定したり、書き始めから見せている板書は書き終わってから見せる時間を短くしたりするなどが考えられる。

## 7. おわりに

本稿では、板書講義映像コンテンツにおけるモバイル環境向けのカメラワークを検討および提案し、その実装方法について論じた。実際に実装を行い提案したカメラワークの評価を行った結果、提案したカメラワークは板書と講師の両方を十分に見せられおり、被験者からは高い評価を得られた。また一方で、板書の表示に関してカメラワーク上の問題点も見つかった。今後は板書を利用者にとって、より見やすく表示することが課題である。

## 参考文献

- [1] 市村, 井上, 宇田, 伊藤, 田胡, 松下: ChalkTalk: 講師動画と板書静止画の同時記録が可能講義自動収録システム, 上昇処理学会論文誌, Vol.47, No.3 pp.924-931 (2006).
- [2] 市村, 中村, 梶並: 講師動画と板書画像を含むモバイル学習用コンテンツ様式の検討, 情報処理学会シンポジウムシリーズ, vol.2011, no.7, pp.41-46 (2011).
- [3] Chandra, S.: Lecture video capture for the masses, Proc. of ACM ITiCSE'07, pp.276-280 (2007).
- [4] Hurst, W., Welte, M., Jung, S.: An evaluation of the mobile usage of e-lecture podcasts, Proc. of ACM Mobility'07, pp.16-23 (2007).
- [5] 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).
- [6] Liu, Q., Rui, Y., Gupta, A., Cadiz, J.: Automating Camera Management for Lecture Room Environments, Proc. Of ACM Conference on Human Factors in Computing System (CHI 2001), Vol.3, pp.442-449 (2001).
- [7] Zhang, C., Crawford, J., Rui, Y., He, L.: An Automated End-to-End Lecture Capturing and Broadcasting System, Proc. Of ACM Multimedia 05, pp.808-809 (2005).
- [8] 錦織, 菅沼, 谷口: 黒板講義を対象とした講義自動撮影システムの構築, 電子情報通信学会技術報告, PRMU, 100(701), pp.79-86 (2001).
- [9] 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-1, No.11, pp.1187-1195(2000).
- [10] 芦川, 菅沼, 谷口: チョーク音による板書動作の開始と終了の判定, 情報技術レターズ 1, pp.245-246(2002).