

## 講義映像における板書領域のブロック分割とその応用

大西 正輝<sup>†</sup>      泉 正夫<sup>†</sup>      福永 邦雄<sup>†</sup>

Blackboard Segmentation Using Video Image of Lecture and Its Applications

Masaki ONISHI<sup>†</sup>, Masao IZUMI<sup>†</sup>, and Kunio FUKUNAGA<sup>†</sup>

あらまし ネットワーク技術の発展は、従来の講義形態からネットワークを介して距離的に離れた複数の講義室で同時に講義を行う遠隔講義形態へ移行する可能性を示唆している。遠隔講義では、講義に使用する資料等はネットワークを介して配布する形になる。しかし、あらかじめ電子教材を作成するのは手間がかかることなどから、すぐに普及する傾向にはない。一方、黒板を用いる講義は OHP やビデオが普及した現在でも講義内容を強く印象づけたり、受講者が順次理解していく過程にマッチしている手法として根強い支持を得ており、大学・予備校などでも広く行われている。本論文では板書主体の講義を撮影する時や、ネットワークで利用するための電子教材を作成するときに必要となる板書をブロック分割する手法を提案する。板書のブロックは小さな矩形の集合であると考え、一定時間書き込みのない場合に板書文字領域を囲う小さな矩形を生成する。次に講義のレイアウトから生成したファジールールに従い、これらの矩形を統合・分割することでブロック分割を行う。また、本手法の応用例として、抽出したブロックを順につなぎ合わせて HTML 形式のファイルを作成し、WWW 配信するシステムを構築した。

キーワード 講義映像、板書のブロック分割、時空間画像、ファジールール

### 1. まえがき

遠隔講義の撮影自動化の手法として、受講者にとって見やすい映像を映し出すためのカメラワークを動画画像処理によって決定する手法が提案されている。それらは、講義者の位置・姿勢推定結果からカメラワークを決定するもの [1] や、映像の変化を情報としてとらえて、空間と時間に共通する情報の発生量からカメラワークを決定するもの [2] などである。しかし、これらの手法は講義者と板書文字の明確な区別を行っていないため 1 次元的な情報をもとに機械的なスイッチングを行っているにすぎず、実際のカメラマンのように柔軟な撮影知識をカメラワークに取り入れることは困難であった。

カメラマンはシーンを撮影する際にあらゆるシーン状況を理解し、経験から最適なカメラポジションやカメラワークを決定する。このようなカメラマンの柔軟な撮影知識を計算機に組み込み、自動で最適な映像を撮影していくことを、ここでは知的自動撮影と呼ぶこ

とにする。知的自動撮影を画像処理によって実現するためには、撮影目的に応じてシーン中の様々な状況理解を行う必要がある。講義映像を撮影するための状況理解とは、板書領域や講義者の抽出、更には講義者の行動理解などである。中でも、板書領域の抽出と板書のブロック分割はカメラワークを決定する上で特に重要な要素となる。これらのことより、講義映像から板書領域を抽出し、それらの領域をブロック分割する処理が望まれている。

また、一方でネットワーク技術の発展に伴い、教育のあり方が大きく変わろうとしている [3]。ネットワークを用いた遠隔講義などの教育は、マルチメディア教材の呈示による講義へ移り変わる可能性を示唆している。一方、従来の黒板などの板書を用いた講義はリアルタイムで講義者の説明や雰囲気伝えることができる利点があり、またインターネット用の教材開発の手間が不要などの観点からも根強い支持がある。したがって、自動的に板書主体の講義映像から板書領域を抽出・分割できれば、それらを閲覧できるような HTML 形式のファイルを作成することで、実際の講義映像とともに板書領域を見やすい形で呈示可能となり、マルチメディア教材を開発する手助けにもなり得る。

<sup>†</sup> 大阪府立大学大学院工学研究科，堺市  
Graduate School of Engineering, Osaka Prefecture University 1-1, Gakuen-cho, Sakai-shi, 599-8531 Japan

本論文では、板書主体の講義を撮影する時や、電子教材化するときに必要となる講義映像から講義者並びに板書領域を抽出し、抽出した板書領域をブロック分割する手法を提案する。

この種の研究と関連するものとしては、文書画像のブロック抽出が挙げられる。その代表的な手法として一般図形ポロノイ図を用いた手法[4]や、文字のない白矩形をプリミティブとして処理する手法[5]が挙げられるが、これらの手法は文書画像を対象としているため、高解像度の入力画像を想定している場合や、大きい空白領域をブロックの境界線と考える手法が多い。しかし、講義映像における板書領域のブロック抽出を考えると、リアルタイムで処理を行うためには計算負荷の大きい高解像度の入力画像を望むことはできない。また、板書は講義中に講義者が書き加えていくものであるため、時々刻々と変化する板書領域の時間的な変化をブロック抽出に利用することができるが、必ずしもブロック間に大きな空白領域が境界線として存在するとは限らない。また、板書文字領域の他にも講義者の領域が存在し、板書と撮影カメラの間に割り込むといった難点も生じる。そこで、講義映像から板書文字のブロック抽出を行うには、空間的な情報だけではなく時間的な情報も利用する必要があると考えられる。

映像の時間的な変化を調べたオンラインでの板書記録手法として、背景との差分画像を用いる手法[6],[7]では、自動で板書を記録していく手法が提案されているが、板書のブロック分割までは行っていない。また、板書の記録ではないが研究発表でのスライドを記録する手法[8]も報告されている。これは、映像中からキーフレーム(スライドの変わり目)を探し、発表者の動作を解析することで発表を要約する手法を提案しており、スライドの空間的な分割(ブロック抽出)は行っていない。

本手法は、まず固定したカメラ上に現れる静止する物体の作るエッジと動物体の作るエッジを区別して抽出し、静止物体によるエッジを構成するものを板書文字と考える。次に、一定時間板書文字領域に大きな変化がなければ、それらを一つの矩形で囲む。そして、構成された矩形の時間的な配置と空間的な配置から最終的なブロックを抽出する手法を明らかにしている。

以下、2.で、板書文字と講義者領域を分離して抽出する手法について説明し、3.で板書領域からブロックを抽出する手法を明らかにする。次に、4.で本手法の応用例として、抽出したブロックを順番に記録し

てWWW配信する手法を説明し、5.で実験と考察を行う。

## 2. 板書文字と人物領域の分離

講義の映像に現れる対象物として、大きく二つの対象物が考えられる。一つは講義者であり、もう一つは板書文字である。板書文字を抽出するためには、これらの二つの対象物を分離して抽出する必要がある。黒板に記録された板書文字を抽出する手法として差分を用いる手法[6],[7]が報告されているが、黒板は一様な色分布をしており板書文字は線が細いなどの理由から、差分による抽出法では解像度の高さが要求される。また、差分は色変化や輝度変化の大きい画素を抽出する処理であり、差分値のみから講義者の動きによる色変化か板書文字の出現や消失による色変化かを特定するのは困難である。

そこで、本手法では時空間画像にエッジ処理を施すことで講義者と板書文字を分離して抽出する。エッジ抽出は差分のように注目画素のみを観測するのではなく、注目画素とその周りの画素の色情報も効果的に利用することができるため、差分に比べてロバスト性が高い。また、時空間画像のエッジ方向を検出することで、移動している物体によるエッジが静止している物体によるエッジかを区別することができるため、講義者によるエッジか板書文字によるエッジかを特定することができると考えられる。

時空間画像は、映像中の動きを“流れ”として可視化することができるため、それらの“流れ”の方向から映像中のカメラワークを調べる研究などに利用されている[9],[10]。本手法では、固定したカメラにより得られた時空間画像からエッジ検出を行うため、動きのある物体のエッジと動きのない物体のエッジに分けることが可能になる。

### 2.1 時空間画像

入力画像を時系列に並べた画像を時空間画像と呼び、 $F(x, y, t)$  で表す。また、時空間画像  $F(x, y, t)$  に対して、

$$f(y, t | x_i) = F(x_i, y, t) \quad (1)$$

を定義する。 $f(y, t | x_i)$  は  $x$  を  $x_i$  に固定した時の  $F(x, y, t)$  の値の変化を表す  $y$  と  $t$  の関数であり、 $x = x_i$  における  $F$  の断面を表す。また、 $y = y_i$  における断面は、同様に  $f(x, t | y_i)$  で表す(図1)。そして、この時空間画像を  $x$  軸に垂直な面で切った断面

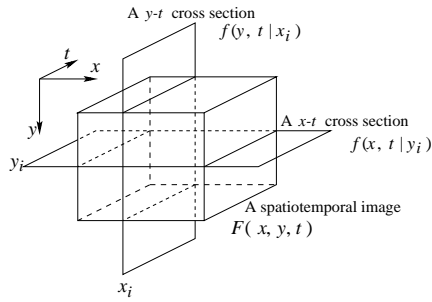


図1 時空間画像の断面図  
Fig.1 Cross sections of spatiotemporal image.

$f(y, t | x_i)$  を  $y - t$  断面図,  $y$  軸に垂直な面で切った断面  $f(x, t | y_i)$  を  $x - t$  断面図と呼ぶ.  $y - t$  断面図及び  $x - t$  断面図の一例を図2に示す. それぞれの断面図は, 入力画像の白線部分の断面図を示しており, 黒板左側で板書していた講義者は右側へ移動した後に, 左へ戻ろうとしているところであることが確認できる.

これらの断面図は, 動き情報が“流れ”として表現されていることから, カメラが固定されていれば撮影対象物の“流れ”が抽出できる. 本手法ではこれらの断面図のエッジ方向を調べることで, 動物体によるエッジと静止物体によるエッジを区別する.

### 2.2 動エッジと静止エッジ

それぞれの断面図に対して, Sobel オペレータを用いてエッジ抽出を行う. 一般に,  $x - t$  断面図においては, 静止物体のエッジは  $t$  軸方向の線として現れ, 動物体エッジは  $x$  軸方向の線として現れる. 一方,  $y - t$  断面図においては, 静止物体のエッジは  $t$  軸方向の線として現れ, 動物体エッジは  $y$  軸方向の線として現れる. ここでは, 動物体により生成されるエッジを動エッジ, 静止物体により生成されるエッジを静止エッジと呼ぶことにする.

まず, 動エッジの検出法について説明する. 前記のように, 動エッジは  $x - t$  断面図による  $x$  軸方向の線,  $y - t$  断面図による  $y$  軸方向の線であるため, 画素  $(x_i, y_i)$  における動エッジの強度は  $t$  方向の1次微分  $f_t$  を用いて次式で定義する.

$$|\nabla f^d| = \sqrt{f_t(x, t | y_i)^2 + f_t(y, t | x_i)^2} \quad (2)$$

ただし, 動物体は動いているため, 現在の時刻での動エッジを求める.

次に, 静止エッジを検出する手法を述べる. 注目画素の静止エッジの強度は,  $x - t$  断面図に対するエッジ

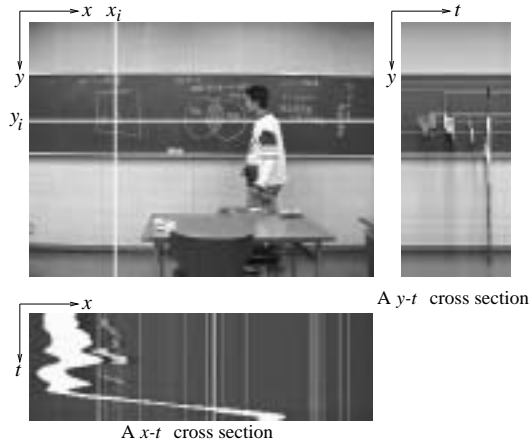


図2  $x - t$  断面図と  $y - t$  断面図の例  
Fig.2 Example of  $x - t$ ,  $y - t$  cross sections.

の垂直成分と  $y - t$  断面図に対するエッジの水平成分の局所積和演算を用いて,

$$|\nabla f^s| = \sqrt{f_x(x, t | y_i)^2 + f_y(y, t | x_i)^2} \quad (3)$$

で定義することができるが, 静止エッジは静止し続けている物体から生成されるエッジであるため, 式(3)をしきい値処理したものを時間軸である  $t$  軸方向に投影することで一定時間静止しているエッジのみを安定して取り出す. つまり, 静止エッジは2値化のしきい値を  $\theta$  として,

$$\sum_t bin(|\nabla f^s|) \quad (4)$$

ただし,

$$bin(x) = \begin{cases} 1 & ; \text{ if } x \geq \theta \\ 0 & ; \text{ others} \end{cases} \quad (5)$$

により算出する.

図3に講義映像を対象としたエッジ画像の例を示す. 動エッジを黒色, 静止エッジを灰色で表している. 動エッジを生成する講義者が黒色, 静止エッジを生成する板書文字が灰色で表されていることが確認できる.

### 3. 板書のブロック分割

本研究の目的は, 講義映像中の講義者と板書文字を抽出し, 板書文字領域から板書ブロックを抽出することである. 板書ブロックとは意味的に一つのかたまりを構成する板書文字や図の集合を指し, 板書ブロック

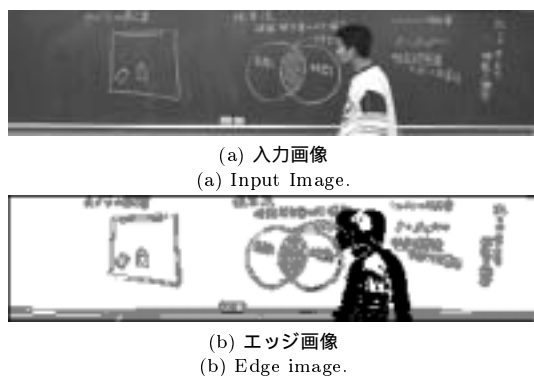


図3 エッジ画像の例

Fig. 3 Example of edge image.

を順に並べたものが受講者の書き写したノートに相当する。一般に、黒板内の領域で静止エッジを生成するものが板書文字であり、動エッジを生成するものが講義者であると考えられる。そこで、板書文字と考えられる静止エッジを用いて板書ブロックの抽出を行う。

### 3.1 板書矩形

板書ブロックは、いくつかの矩形で構成されているものと考え、板書ブロックを構成する矩形の抽出を行う。ここでは、この矩形を板書矩形と呼ぶ。

まず、2. で抽出した黒板内のすべての静止エッジを外接矩形で囲む。入力画像ごとに矩形を求めていき、一定時間矩形の位置と面積に変化がない場合には、その矩形をひとまとまりと考え、板書矩形として記憶する。新しく板書文字が書かれた場合には、まだ板書矩形に囲まれていない静止エッジについて同様の処理を繰り返す。ただし、時間的に連続して板書の下方から上方へ書くことが考えられるため、板書矩形を記憶する際に、静止エッジを  $y$  座標の降順にソートし、連続する静止エッジの  $y$  座標の値がしきい値よりも離れている場合には、強制的に矩形を分割して記憶する。また、矩形を記憶する際には矩形の大きさや中心座標だけでなく、矩形内に含まれる静止エッジの数と矩形のできた時間も同時に記憶しておく。図4に抽出した板書矩形例を示す。time は板書矩形として囲まれた時間をフレーム番号で表し、no は矩形内の静止エッジのドット数を表す。②③④の板書矩形は、消されてしまっているために図中には存在しない。また、矩形内の静止エッジのドット数が増加した場合には順次更新していく。

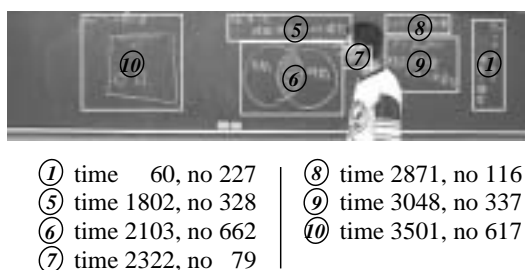


図4 板書矩形と矩形情報

Fig. 4 Extracted rectangles on blackboard and their information.

### 3.2 板書文字を消した場合への対応

次に、板書文字を消した場合について考える。講義者は矩形ごとに矩形内のすべての文字を消すとは限らず、矩形の一部の領域を消すことも多い。ここでは、板書矩形の情報を用いて板書文字が消されたことを検出する。矩形の一部の板書文字を消すと、その板書矩形内の静止エッジ数が減少する。そこで、板書矩形内の静止エッジ数が減少した場合には、矩形を解放し矩形内に残った静止エッジのみで、もう一度新しく板書矩形を作り直す。ただし、板書文字が講義者で隠れたために静止エッジが減少することも多い。そこで、板書文字を消すときには黒板消しを使用することから、黒板消しが黒板下部のレール上にあるか否かを調べる。まず、黒板消しのエッジ画像をあらかじめモデルとして作成しておき、そのモデルを入力画像から抽出した静止エッジに対してレール左から右へと順にずらしながら各画素値についての残差2乗和を求めていく。もし、レール上から黒板消しがなくなった時(すべての場所での残差2乗和がしきい値より大きい時)には板書文字を消している可能性があると考え、すべての板書矩形の静止エッジ数を調べて減少数の大きい矩形を再構成する。ただし、映像中での黒板下部のレールの位置は既知とする。

### 3.3 板書文字のブロック分割

講義者は板書はある程度のレイアウト規則に従って書いていく。例えば横書きの板書を考えた場合には、上から下へ、左から右へ書く傾向が見受けられる。そこで、板書矩形の時間的・空間的配置をもとに複数の板書矩形をひとまとめにする作業であるブロック分割を行う。一般に、板書は上から下へ向かって書いていき、一番下まで書き終わると新しく右側の上から書き始める。そこで、板書矩形のできた時点で、その板書

矩形と一番新しい板書ブロックの位置関係を調べ、経験的に考えた以下の二つ分割条件を満たす場合にはブロックの分割を行い、その板書矩形から新たな板書ブロックが始まるものとする。一方、満たさない場合には板書ブロックはまだ続いているものとみなして、板書矩形を板書ブロックに含める。また、一番始めに構成される板書矩形は板書ブロックの始まりとする。

分割条件

(1) 現在の板書矩形は黒板の上の方にあり、板書ブロックは黒板の下の方まで続いている。

(2) 現在の板書矩形は板書ブロックと横方向に大きく離れている。

条件 (1) は、一度下まで板書した後に新たに上から新しい板書を始める場合に分割を行う条件であり、条件 (2) は黒板の余白が少なくなってきた場合などに黒板の空いている部分に板書した領域を分割する条件である。多くの場合について条件 (1) で分割することができたが、一部当てはまらない場合があったために条件 (2) を付け加えた。これらの経験的な分割ルールをファジールールにより作成する。ファジールールによるファジー推論は、知識表現を明確に与えることができるため、様々なエキスパートシステムに利用されることが多く [11]、今回のルール作成にも適していると思われる。

条件中の「上(下)の方」「大きく離れている」を表すメンバシップ関数は、図 5 に示す台形を用いる。ただし板書矩形間の横方向の距離は矩形が重なっている場合には 0、離れている場合には  $x$  軸方向の距離を用いる。メンバシップ関数の形や各パラメータは、実際の講義を観察し経験によって決めた。条件 (1) では  $\mu_{\text{top}}(\cdot) \times \mu_{\text{bottom}}(\cdot)$  を、条件 (2) では  $\mu_{\text{away}}(\cdot)$  を活性値とし、これらの活性値がしきい値よりも大きくなった場合に現在の板書矩形を新しい板書ブロックの始まりとする。ただし、 $\mu_{\text{top}}(\cdot)$  は板書矩形の一番上の  $y$  座標を入力とし、 $\mu_{\text{bottom}}(\cdot)$  は板書ブロックの一番下

の  $y$  座標を入力とする。また、 $\mu_{\text{away}}(\cdot)$  は板書矩形と板書ブロックの最も短い距離を入力とする。一方、活性値がしきい値を超えない場合には、その板書矩形はブロック内の矩形であることとして統合していく。

#### 4. 板書の記録と WWW 配信

ここでは、抽出した板書ブロックごとに順番に板書を撮影していき、講義の終わった後にノートとして記録・保存できるシステムの構築を考える。固定カメラにより抽出した板書のブロック分割結果をもとに、パン・チルト・ズーム制御可能なカメラを板書ブロックができるだけ大きいサイズで映るようにセットして画像を取り込む。あらかじめ、2 台のカメラの位置関係やブロックのサイズとズームレベルの関係は調べておく。抽出した板書ブロックを記録するタイミングは、講義者による動エッジがブロック内に含まれない時点として、板書ブロックを 1 枚の画像として取り込む。ただし、講義者が過去に書いていたブロックに追加して書き込むこともあるため、ブロック内の静止エッジ数が増加した場合には、もう一度そのブロックを撮影し直す。また、取り込んだ板書画像は 2 値化処理を行い GIF 形式のファイルで保存しておき、それに合わせた HTML 形式のファイルを自動作成し、インターネットブラウザでの閲覧を可能にする。アーカイブ化した板書はマルチメディア教材としても利用することができる。

#### 5. 実験・考察

##### 5.1 実験

本手法を用いて、模擬的に行った講義六つ (3 人、合計約 70 分)、大学生を対象とした実際の講義三つ (3 人、合計約 170 分) について、ブロック抽出の実験を行った。模擬的な講義は受講者がいない以外は実際の講義と同じ形態で行った。固定カメラを用いて撮影した入力画像の大きさは、横長の黒板がちょうど映像内に収まる大きさとして、 $320 \times 80$  画素とした。処理速度は毎秒 7 フレーム程度であり、時間方向に連続した 60 フレームの画像のうち 30 フレーム以上において式 (3) の値が大きい画素を静止エッジとした。また、静止エッジに外接する矩形の面積と位置が 30 フレームの間変化しない場合に板書矩形として記憶した。板書のブロック分割に用いるメンバシップ関数の各値は、図 5 の値を用いており、活性値を 2 値化するしきい値は 0.5 とした。実験結果の例を図 6 左段に示す。それ

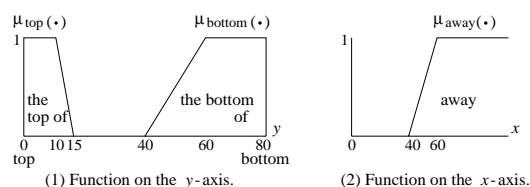


図 5 メンバシップ関数

Fig. 5 Membership functions.

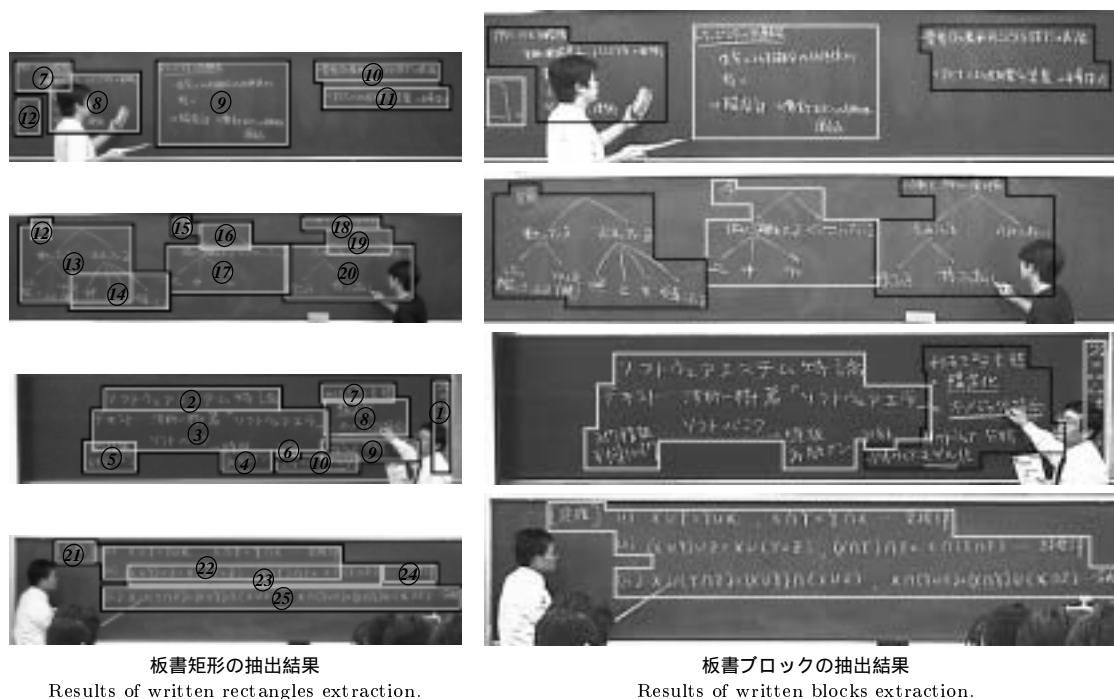


図 6 実験結果 (ブロック分割)  
Fig. 6 Experimental results (block extraction).

表 1 実験結果 (ブロック抽出率)  
Table 1 Experimental results (correct ratio of block extraction).

	講義時間 (分'秒")	ブロック 抽出率	誤抽出				板書矩形 (個数)	ブロック 抽出率	誤抽出			
			過分割	過統合	過抽出	無抽出			過分割	過統合	過抽出	無抽出
講義 A	12'34"	96.9%	0.0%	0.0%	0.0%	3.1%	14	100%	0.0%	0.0%	0.0%	—
講義 B	11'30"	100%	0.0%	0.0%	0.0%	0.0%	12	100%	0.0%	0.0%	0.0%	—
講義 C	8'32"	76.4%	0.0%	23.6%	0.0%	0.0%	10	80.0%	0.0%	20.0%	0.0%	—
講義 D	10'52"	88.7%	4.6%	0.0%	0.0%	6.7%	19	94.7%	5.3%	0.0%	0.0%	—
講義 E	10'32"	100%	0.0%	0.0%	0.0%	0.0%	20	100%	0.0%	0.0%	0.0%	—
講義 F	15'29"	100%	0.0%	0.0%	0.0%	0.0%	12	100%	0.0%	0.0%	0.0%	—
小計	69'29"	94.8%	0.7%	3.0%	0.0%	1.6%	87	96.6%	1.1%	2.3%	0.0%	—
講義 1	42'52"	100%	0.0%	0.0%	0.0%	0.0%	21	100%	0.0%	0.0%	0.0%	—
講義 2	67'29"	90.6%	4.8%	4.6%	0.0%	0.0%	42	88.1%	7.1%	4.8%	0.0%	—
講義 3	60'32"	62.8%	0.0%	33.9%	3.3%	28.4%	27	74.1%	0.0%	22.2%	3.7%	—
小計	170'53"	83.1%	1.9%	13.8%	1.2%	10.1%	90	86.7%	3.3%	8.9%	1.1%	—
合計	240'22"	86.5%	1.5%	10.7%	0.8%	7.6%	177	91.5%	2.3%	5.6%	0.6%	—

それぞれの板書矩形を白線で囲み、囲まれた順に番号を付けた。また黒線がブロック分割結果を表す。これらの実験で用いた画像の解像度は 320 × 80 画素と低いため、紙面上で十分に文字などを読み取ることができない。そこで、右段に 640 × 160 画素で取り込んだ画像にそれぞれの板書ブロックの座標を計算してブロックの境界線のみを合成した画像を示す。ここでは、板書ブロックの境界線を黒線と白線で適宜書き分けて表示している。

次にブロック抽出率を表 1 に示す。左半分は、講義時間全体の中で正しく分割が行えた時間をブロック抽出率としてパーセントで示している。抽出ができていないかどうかは、模擬的に行った講義については講義者に判断させ、実際の講義については受講者に判断させた。講義 A ~ F が模擬的に行った講義における抽出率であり、講義 1 ~ 3 が実際の講義における抽出率である。本評価方法は講義時間中で正しくブロック分割が行えている時間を評価しているため、講義中に全く何

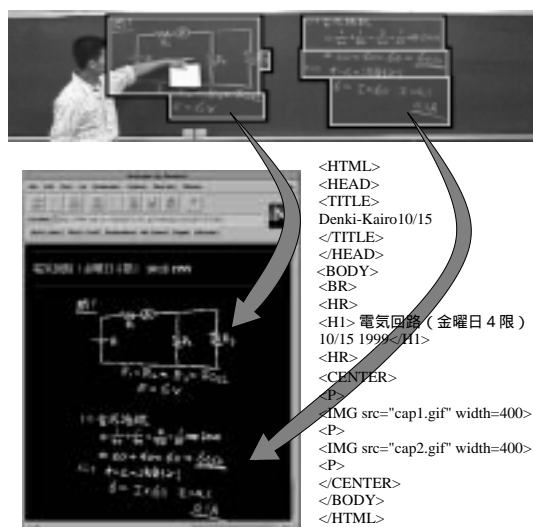


図7 板書の WWW 配信例

Fig. 7 Example of demonstration on WWW.

も書かなかった場合にも抽出率は 100% になってしまう。そこで、同時にブロック数に合わせた評価も行った。表右半分は板書矩形が現れたときにその矩形が正しくブロック分割できている割合をパーセントで示している。

また、4. のシステムを実際に構築した。固定したカメラから得られた講義映像の板書をブロック分割した後、ブロック内に動エッジが含まれない時点で、パン・チルト・ズームの制御可能な首振りカメラを用いて板書ブロックを記録していく。制御可能なカメラで取り込んだ画像の解像度は  $640 \times 480$  画素であり、ブロック領域のみを切り出して記録・保存する。板書を記録した WWW 配信の例を図 7 に示す。

## 5.2 考察

模擬的な講義と実際の講義で板書のブロック分割を行い、ブロックの抽出率が約 90% と高い抽出率を得ることができた。実験における模擬的な講義では小中学生を対象にした講義経験をもつ講師を中心に選出しており、板書は比較的まとまっていて見やすい印象を受けた。このような場合には時間の抽出率が平均 94.8%、矩形ごとの抽出率が平均 96.6% であり、特に高い抽出率を得ることができた。一方、実際の講義は大学生を対象とした講義であり、板書が見やすいとはいえない印象であった。このため、過統合やそれに続いて誘導される未抽出が多く生じた講義もあった。しかし、表 1 からわかるとおり、一般的な講義者が行う講義

であれば十分適用可能である。

本手法は一度誤抽出が起こるとそれ以後の抽出に支障を来たす可能性が高いため、時間の抽出率は矩形数の抽出率より低くなる場合が多い。ブロック分割の典型的な誤り例は以下のとおりである。

[過分割] 一つのブロックであるにもかかわらず、上部に書き加えたために条件 (1) を満たして過分割が生じた。

[過統合] 前のブロックに書き加えた矩形を現在のブロックに含めてしまったために過統合が生じた。

[過抽出] 講義者が長時間動かなかった場合に板書文字と見なされて過抽出が生じた。

[無抽出] 以前にブロックとして囲まれた矩形内に新しいブロックを書き込んだために抽出されない無抽出が生じた。

これらについて詳しく説明するとともに、それ以外の問題点についても言及する。また、本手法の応用例について考察する。

- 本手法では、生成された板書矩形の順に空間的な矩形の配置からブロック分割を行っている。この手法では、容易にブロック分割の規則を生成することができる反面、時間や空間の配置を無視した以前のブロックへの書き加えなどがあつた場合に図 8 に示すような過分割や過統合を引き起こす可能性がある。図 8 (a) の過分割の誤り例では、

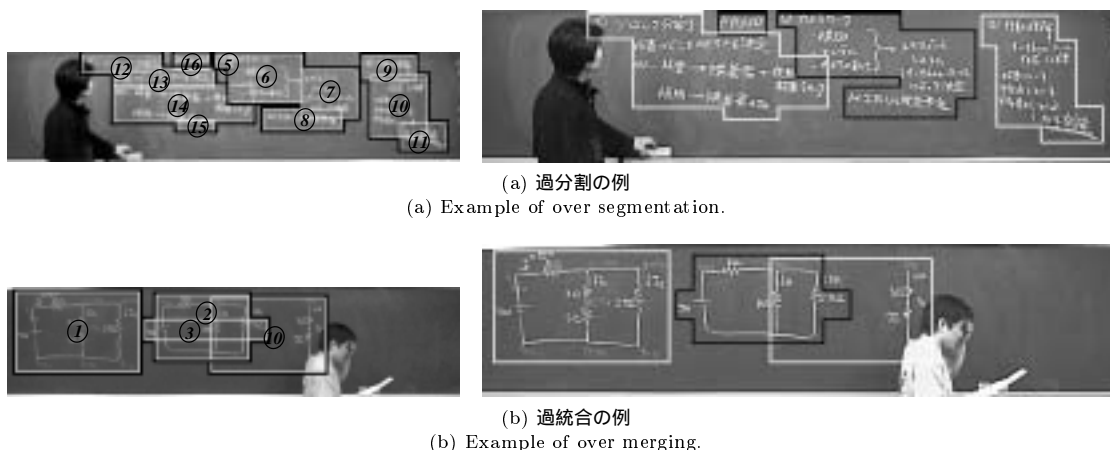
$\{5\} \{6\} \{7\} \{8\} \{9\} \{10\} \{11\} \{12\} \{13\} \{14\} \{15\} \{16\}$

のようにブロック分割されているが、正しい分割結果は、

$\{5\} \{6\} \{7\} \{8\} \{9\} \{10\} \{11\} \{12\} \{13\} \{14\} \{15\} \{16\}$

である。しかし、矩形の位置関係のみからでは、 $\{16\}$  の矩形が、独立した一つのブロックなのか、 $\{5\} \{6\} \{7\} \{8\}$  のブロックに属する矩形なのか、 $\{12\} \{13\} \{14\} \{15\}$  のブロックに属する矩形なのかを判断することは難しい。また、図 8 (b) の過統合例において  $\{16\}$  の矩形は  $\{2\} \{3\}$  のブロックに書き加えた板書を新しいブロックに統合してしまっている。空間・時間の配置を同時に考慮した分割規則を用いることで、このような過分割や過統合をある程度は回避することができると考えられるが、更に抽出精度を向上させるためには、矩形内に書かれた文字の意味や講義の内容まで理解する必要があると考えられ、本手法のように矩形の時間的・空間的な位置関係のみで分割する手法には限界があると思われる。

- 本分割手法は、過去の矩形抽出やブロック分割の結果に大きく依存する。そのため一度誤抽出が起こ



板書矩形の抽出結果  
Results of written rectangles extraction.

板書ブロックの抽出結果  
Results of written blocks extraction.

図8 誤りの例  
Fig. 8 Examples of miss extraction.

ると、後々まで誤抽出を誘導する可能性がある。これらの問題を回避するためには、矩形やブロックの候補を複数抽出し、その時々によって最も評価の高い結果を選出するなどといった誤抽出から回復する手法が必要になると考えられる。

- ここで述べた手法では、板書文字を消した可能性を黑板下のレールに黑板消しがあるかないかをもとにして調べている。しかし、講義者がカメラと黑板消しの間に割り込むことで黑板消しが見えなくなる場合もあり、その結果として矩形を解放してしまい誤抽出を招くこともある。更には、少し消して少し書くといった場合に矩形の解放に失敗することも考えられる。これに関しては、人物の行動を認識する手法が検討されており [12]、講義者が黑板を消すという行動を起こしたときに消した領域を解放することで、このような誤抽出を回避することができると思われる。

- 今回実験に用いた黑板の横幅は 350 cm であり、制御用のカメラから得られた  $640 \times 480$  画素の記録用の板書画像は、一つの板書ブロックがちょうど画像中に収まる程度に拡大して撮影した場合には、人が文字認識するのに十分な解像度を得ることができた。しかし、更に大きな黑板に大きなブロックで板書した場合には記録した板書文字を認識できるほどの解像度が得られるとは限らない。この場合には、適切なズームパラメータを決定する手法 [13] などを併用し、分割して板書を撮影するといった処理が必要となる。

- 本手法の応用例として、講義映像から板書を自動的に記録するシステムを構築した。これ以外の応用として、遠隔講義などに用いる臨場感のある講義映像を自動で撮影する知的自動撮影システムが考えられる。単純な手法として、現在使用中の板書ブロックを中心にカメラワークを決定する手法 [14] などが挙げられるが、カメラマンのような柔軟な撮影技法を実現するために講義者の行動なども同時に推定し、それらの多次元的な情報を統合してカメラワークを決定するができれば、より臨場感のある映像作成が可能であろう。

## 6. むすび

時空間画像の断面図からエッジを抽出することで、静止しているものと動いているものを分離し、静止しているエッジである板書文字を囲む矩形の位置関係からファジールールを用いて板書を分割する手法を提案した。また、抽出した板書ブロックを順に撮影していくことで、板書を自動で記録するシステムを構築した。本システムにより、講義室内のみに閉鎖された講義を WWW などを用いて外部へ公開することが可能となる。また、他の応用例として、遠隔講義などで用いる講義映像を自動で見やすいアングルから撮影するシステムなどが考えられる。

謝辞 実際の講義を撮影させて頂いた本学工学部の松本啓之亮教授、模擬的な講義を撮影させて頂いた本学大学院生の渡辺浩司氏、浅沼克紀氏に深く感謝致し



ます。

### 文 献

- [1] 大西正輝, 松本昌紀, 福永邦雄, “動画像処理による遠隔講義映像の自動生成とその評価,” 信学技報, ET98-62, Oct. 1998.
- [2] 大西正輝, 泉 正夫, 福永邦雄, “情報発生量の分布に基づく遠隔講義撮影の自動化,” 信学論 (D-II), vol.J82-D-II, no.10, pp.1590-1597, Oct. 1999.
- [3] 前田香織, 相原玲二, 川本佳代, 寺内睦博, 河野英太郎, 西村浩二, “遠隔講義のためのマルチメディア通信環境,” 信学論 (B-I), vol.J80-B-I, no.6, pp.348-354, June 1997.
- [4] K.Kise, A.Sato, and M.Iwata, “Segmentation of page images using the area Voronoi diagram,” Computer Vision and Image Understanding, vol.70, no.3, pp.370-382, June 1998.
- [5] T.Pavlidis, and J.Zhou, “Page segmentation and classification,” GVGIP, vol.54, no.6, pp.484-496, Nov. 1992.
- [6] 西口敏司, 仙田修司, 美濃導彦, 池田克夫, “首振りカメラによる黒板の記録手法,” 画像の認識・理解シンポジウム (MIRU'96), vol.1, pp.I-37-I-42, July 1996.
- [7] 先山卓朗, 椋木雅之, 池田克夫, 飯塚重善, 辻本雅彦, “アーカイブのための黒板映像編集,” 情処学第 58 回全国大会, 4U-2, pp.3-181-182, March 1999.
- [8] S.X.Ju, M.J.Black, S.Minneman, and D.Kimber, “Summarization of videotaped presentations: automatic analysis of motion and gesture,” IEEE Trans. Circuits and Systems for Video Technology, vol.8, no.5, pp.686-696, Sept. 1998.
- [9] 阿久津明人, 外村佳伸, “投影法を用いた映像の解析方法と映像ハンドリングへの応用,” 信学論 (D-II), vol.J70-D-II, no.5, pp.675-686, May 1996.
- [10] F.M.Idris, S.Panchanathan, “Spatio-Temporal Indexing of Vector Quantized Video Sequences,” IEEE Trans. on Circuits and Systems for Video Technology, vol.7, no.5, pp.728-740, Oct. 1997.
- [11] 匹田志朗, 駒谷喜代俊, “ファジィ・ルールベースを用いた新しいエレベータ群管理システム,” 計測自動制御学会論文誌, vol.25, no.1, pp.99-104, Jan. 1989.
- [12] 小島篤博, 田原典枝, 田村武志, 福永邦雄, “動画像における人物行動の自然言語による説明の生成,” 信学論 (D-II), vol.J81-D-II, no.8, pp.1867-1875, Aug. 1998.
- [13] 戸田真志, 川嶋稔夫, 青木由直, “画像情報分布に基づく視覚ズームの適応的制御,” 信学論 (D-II), vol.J81-D-II, no.1, pp.84-92, Jan. 1998.
- [14] 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抽出とその自動撮影,” 信学技報, PRMU99-128, Nov. 1999.

(平成 12 年 2 月 28 日受付, 6 月 12 日再受付)

大西 正輝 (学生員)

平 9 阪府大・工・情報卒・平 11 同大学院博士前期課程了。現在, 同大学院博士後期課程在学中。コンピュータビジョンに関する研究に従事。映像の自動撮影・自動編集に興味をもつ。電気学会, 日本ロボット学会, 映像情報メディア学会各学生員。

泉 正夫 (正員)

昭 58 阪府大・工・電気卒。昭 60 同大学院博士前期課程了。同年シャープ(株)入社。平 2 より阪府大・工・電気助手。現在, 同大・工・情報助教授。画像認識・理解に関する研究に従事。情報処理学会, システム制御情報学会, IEEE 各会員。工博。

福永 邦雄 (正員)

昭 42 阪府大・工・電気卒。昭 44 同大学院修士課程了。同年同大・工・電気助手。現在, 同大・工・情報教授。コンピュータビジョン, グラフ理論とその応用などの研究に従事。情報処理学会, システム制御情報学会, IEEE 各会員。工博。