

# 遠隔講義受講者のための アクティブな講義映像生成システムの開発

田代 直之<sup>†</sup> 島田 敬士<sup>†</sup> 菅沼 明<sup>†</sup>

<sup>†</sup>九州大学大学院システム情報科学府 〒816-8580 福岡県春日市春日公園 6-1

E-mail: †{tashiro,atsushi,suga}@limu.is.kyushu-u.ac.jp

**概要** 近年、リアルタイム映像配信を用いた遠隔講義が一部の大学や予備校などで行われている。現在の遠隔講義で配信されている映像には、固定カメラによって撮影されたもの、カメラマンが撮影したもの、計算機でカメラを制御することにより撮影されたものがある。しかし、受講者側に提供される映像の多くは配信者側が提供しているため、必ずしも受講者が望んでいる箇所が撮影されているとは限らない。そこで、本システムでは、配信者側では固定カメラによって講義風景全体を撮影した映像を配信するだけにし、受信側の計算機上で配信されてきた講義映像に対して処理を施すことにより動きのある映像を提供する。さらに、受講者が注目したいと思う部分を講義映像中から選択し、その領域を拡大表示する機能などを付加することで従来の遠隔講義支援システムに比べ、受講者のニーズに応えるシステムを実装した。

## System for Generating an Active Video for Participants in a Distant Lecture

Naoyuki TASHIRO<sup>†</sup>, Atsushi SHIMADA<sup>†</sup>, and Akira SUGANUMA<sup>†</sup>

<sup>†</sup> Department of Intelligent Systems, Kyushu University  
6-1, Kasuga-koen, Fukuoka, 816-8580, Japan

E-mail: †{tashiro,atsushi,suga}@limu.is.kyushu-u.ac.jp

**Abstract** We propose a supporting system for participants in a distant lecture. The system extracts an important scene from image sequence by image processing, and automatically makes a similar video (call the video an active video) which a camera-person took. Although systems with such functions already exist, a video which participants in a distant lecture want is not always captured. The video is processed at a distributing side on existing supporting systems in distant lecture. Therefore, we propose a new approach in order to solve this problem. We assume that a video is distributed without any processing at a distributing side and our system makes an active video on each participant's computer. In addition, we implement a function which enables enlarge an area selected by participant from a displayed image on an interface. In this paper, we describe our system and our experimental results.

### 1. はじめに

近年、計算機の普及とネットワーク技術の発達が目覚ましいものがあり、高速なネットワーク環境を個人が利用できる時代になってきている。高速なネットワークを教育の分野に利用した例として、リアルタイム映像配信技術を用いた E-learning が挙げられる。これにより、離れた場所で開講されている講義風景を自宅にいながらにして閲覧することができる (図 1)。また、一部の大学や予備校

などでは、遠隔講義として実際に取り入れている。現在の遠隔講義で配信されている映像は、撮影範囲を固定して撮影したものやカメラマンが撮影したものがある。前者は、コストが安いという利点があるが、撮影される映像が変化のない退屈なものになってしまうという問題がある。一方後者は、教師の説明に合わせてカメラマンが適切な映像を撮影するので、退屈せずに講義を受けることはできる。しかし、カメラマンの雇用やそのコスト、カメラマンが重要と判断したところと受講者が重要と判断

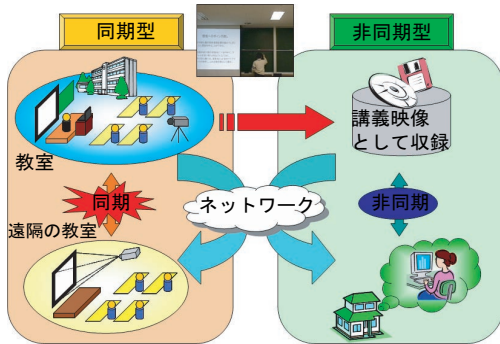


図1 遠隔講義の形態

するところが必ずしも一致するとは限らないという問題などがある。これらの問題点を解決するために様々な遠隔講義支援システムが研究されており、それらは大きく分けて配信者向けの支援システムと受信者向けの支援システムとに分類できる。前者のシステムは、計算機が撮影対象を検出し、カメラマンが撮影したような映像を提供できるという利点がある [1]~[3]。講義風景から撮影対象を推定し、その撮影対象を切り出してカメラで撮影する。一般に、対象を光学ズームで拡大して撮影しているので、文字が小さくて見難いなどの問題は解決できる。遠隔地では、スクリーン上に投影した映像を見ることになる。しかし、導入には配信側に多くの機材が必要となり莫大なコストを要するシステムもある。一方、後者のシステムは、受講者の計算機上に表示するものを指す。受講者側に計算機がある環境を使用するので、そこに講義映像を表示するだけでなく、電子教材を表示したりメモ機能を用意したりして、受講者にとって便利な機能も備えている [4], [5]。しかし、表示される講義映像は、配信側で撮影箇所を決定しているため、必ずしも受講者が注目したい箇所とは限らないという問題がある。そこで、我々は受信者向けの遠隔講義支援システムでかつ、そこに表示する映像はカメラマンが撮影したようなものを自動的に生成できるシステムを提案する。

本システムは、配信側では固定カメラによって講義風景全体を撮影した映像を配信するだけでなく、導入コストも安価なものとなる。また、遠隔講義受信者の計算機上で配信されてきた遠隔講義映像を画像処理技術を用いて解析し、アクティブな映像を生成する。さらに、受講者が注目したいと思う部分を選択し、その領域を拡大表示する機能や講義中にシステムが重要と判断したシーンを自動的に保存する機能などを付加することで、従来の受講者向け遠隔講義支援システムに比べ、受講者のニーズにさらに応えることができるシステムの実現を目指す。

## 2. 設 計

提案するシステムは、4つのモジュールとユーザインタフェースで構成されている。システムの設計概略図を図2

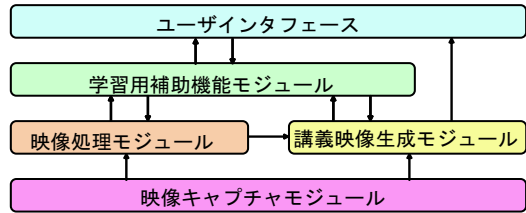


図2 システムの設計概略図

に示す。

映像の拡大表示や画像保存などの映像処理の要求は、学習用補助機能モジュールで解釈され、要求に応じた処理を行っている。映像処理モジュールは、映像キャプチャモジュールの出力を受け取り、画像処理を施している。また、講義映像生成モジュールは、映像処理モジュールや学習用補助機能モジュールからデータを受け取り、そのデータに応じた映像をユーザに提供している。以下では、それぞれのモジュールの処理内容について述べる。

### 2.1 映像キャプチャモジュール

このモジュールでは、下に述べるような形式で入力される遠隔講義の映像をキャプチャし、映像処理用のデータに変換する処理を行う。これ以降のモジュールでは映像処理用のデータだけ扱うように設計できるので、このモジュールで入力メディアの違いを吸収している。

- ・同期型 CometDV<sup>(注1)</sup>を用いることで、リアルタイムに配信されてきた映像をキャプチャし、映像処理用のデータに変換することが可能である。
- ・非同期型 AVI や MPEG などのメディアファイル形式、ダウンロード再生での映像形式、あるいはカメラデバイスから IEEE1394 経由で直接映像を取得する形式に対応している。

### 2.2 映像処理モジュール

このモジュールでは、映像キャプチャモジュールの出力である映像処理用データに対して、画像処理を施すことにより、映像内から重要な箇所を推定する処理を行う。機能の概要としては、講義映像から教師の領域および板書の領域をそれぞれ抽出する。それらの情報を統合して、映像中の重要箇所を推定する。

本モジュールで行う画像処理には、背景画像や画像中の黒板領域の座標情報が必要である。これらを行ってくる映像から抽出するのは効率的でない。特に、背景画像は講義開始前の画像であるため、講義開始後の映像から作り出すのは容易ではない。そこで本システムでは、配信側が予めこれらの情報を格納した初期設定ファイルを用意しておき、受信側は映像ファイルと初期設定ファイルをダウンロードして動作させるものとして設計している。

(注1)：ネットワーク上で DV バケットをそのまま転送できるハードウェア

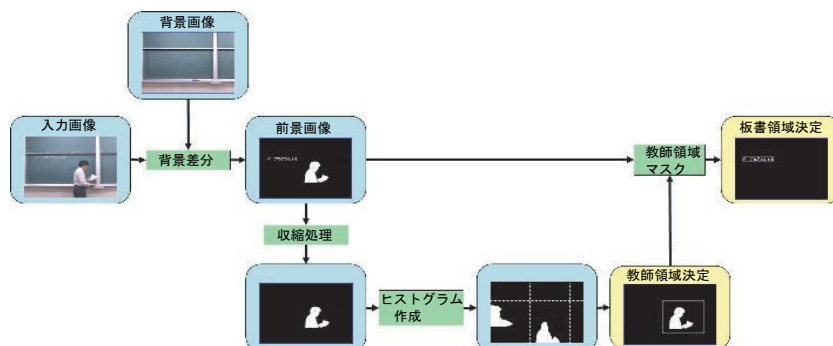


図3 教師領域、板書領域抽出の流れ

## 2.3 講義映像生成モジュール

このモジュールでは、カメラマンが撮影しているかのような滑らかな映像を生成する処理を行う。映像処理モジュールの出力である拡大表示すべき領域を講義映像内から切り出し、デジタルズームを行う。図5(b)は矩形領域をデジタルズームしたものである。

## 2.4 学習用補助機能モジュール

このモジュールでは、遠隔講義を受けるに当たってあれば便利と思われる機能を提供することが目的である。実装した機能としては、電子教材の表示、メモ帳といった従来の受信者向け遠隔講義支援システムが持つ機能に加え、スナップショット、ユーザが選択した領域を拡大表示する機能や選択領域のみを切り抜いて保存する機能がある。さらに、板書領域を記憶しておきある一定量に達したら自動で画像を保存する機能を実装しているの、復習する際には講義の概要を見直すことが容易になる。

## 2.5 ユーザインタフェース

ユーザインタフェースには、計算機が推定した拡大表示すべき領域の映像を出力しておき、ユーザが講義映像内から選択した領域を拡大表示したいという要求をすれば、その映像に切り替わるようにしている。また、学習用補助機能モジュールで作成した機能を実行するボタンを表示しておくことで、ユーザが使いやすいインタフェースを実装している。

# 3. 実装

## 3.1 映像キャプチャモジュール

実装した映像形式のキャプチャの方法を以下に示す。映像キャプチャの開発には、Microsoft社が提供するDirectXを用いた。AVIやMPEGなどのメディアファイル形式、ダウンロード再生での映像形式、あるいはカメラデバイスからIEEE1394経由で直接映像を取得する形式のそれぞれを、DirectXで提供されている関数を利用して映像処理用のデータに変換する。

同期型のメディアでは、送信側と受信側を考える必要がある。送信側では、講義を撮影しているDVカメラとCometDVをケーブルで接続するだけで、映像をDVパ

ケットとして送信できる。受信側では、送られてきたDVパケットをデコードし、映像処理用のデータに変換するように実装した。

## 3.2 映像処理モジュール

映像処理モジュールでは、講義映像内の1枚のフレームを解析することで、教師領域と板書領域とを抽出する。その後、それらの領域を用いて、講義映像中の重要箇所を判断する。講義映像から教師領域及び板書領域の抽出法の処理の流れを図3に示す。

### a) 教師領域の抽出

まず、講義映像から教師を含む矩形領域を下のようにして抽出する。ここで、 $I_{x,y}$ を座標 $(x,y)$ の画素の値を表すものとし、 $M \times N$ の画像を $\mathbf{I} = (I_{1,1}, I_{1,2}, \dots, I_{1,N}, \dots, I_{x,y}, \dots, I_{M,N})$ と表現する。カラー画像のときは、それぞれの画素値 $I_{x,y}$ にはRGBの成分があり、赤成分を $I_{x,y}(r)$ 、緑成分を $I_{x,y}(g)$ 、青成分を $I_{x,y}(b)$ と表す。それぞれの画素値は0~255の値を取る。

講義を開始する前の教壇の映像を背景画像 $\mathbf{B}$ とすると、講義中の映像に写る教師は前景となる。教師を抽出するには、講義中の入力画像 $\mathbf{I}$ から背景画像を取り除くことで抽出できる。しかし、映像中の各フレームで対応する画素の値がまったく同じ値になることはまれなので、背景画像 $\mathbf{B}$ と入力画像 $\mathbf{I}$ との画素値の差を下のようにして求めて、前景画像 $\mathbf{F}$ を得る。

$$\begin{aligned} R &= |I_{x,y}(r) - B_{x,y}(r)| \\ G &= |I_{x,y}(g) - B_{x,y}(g)| \\ B &= |I_{x,y}(b) - B_{x,y}(b)| \end{aligned} \quad (1)$$

$$F_{x,y} = \begin{cases} 255 & (R > TH \text{ または } G > TH, B > TH) \\ 0 & (\text{それ以外}) \end{cases} \quad (2)$$

ここで、 $TH$ は予め決めた閾値である。式(2)により2値画像 $\mathbf{F}$ を得る。式で代入する値255はその画素が白であることを意味し、値0は黒であることを意味する。

背景差分によって得られた前景画像 $\mathbf{F}$ には教師以外に光の影響によるノイズや板書が含まれる。前景画像に含まれる教師領域は大きい領域として抽出される。一方、ノ

イズなどは細かな点として現れる．そのため，前景画像に対して収縮処理を施すことでノイズなどの教師以外の点を除去する．収縮処理は，白色の注目画素の近傍をみて注目画素の値を変化させるもので，隣接8近傍のうち1つでも黒画素があると，注目画素は黒に変える．点として現れる白画素や小さな連結成分として現れる白画素はこの処理で消えてしまう．しかし，ある程度大きな白画素の塊は外側の画素では黒に変化するが，内側の画素は白画素として残る．これを利用して，教師だけを残す．

収縮処理によって得られた白色の画素の分布を調べて，教師を囲む矩形領域を検出する．その際，画像の縦方向と横方向で白画素のヒストグラムを計算し，ある閾値を超える区間を検出する．両方向ともその区間に入る領域を教師領域として抽出する．

#### b) 板書領域の抽出

教師が黒板上に書いた板書を囲む矩形領域を抽出する．黒板に書かれた文字は前景であるので，背景差分を行うことにより黒板の文字を抽出することはできる．しかし，背景差分で得られた前景には教師も含まれてしまうので黒板に書かれた文字だけを抽出するためには，教師の領域をマスクする必要がある．

$$F'_{x,y} = \begin{cases} 0 & (\text{教師領域内}) \\ F_{x,y} & (\text{それ以外}) \end{cases} \quad (3)$$

上式を用いることで，前景画像  $F$  から処理 (a) で推定した教師領域をマスクした画像  $F'$  が得られる．マスクした画像  $F'$  に対して  $3 \times 3$  のメディアンフィルタをかけることによりノイズ処理を行う． $3 \times 3$  のメディアンフィルタの処理は，下の式で表される計算を行うことになる．

$$M_{x,y} = \text{median}\{F_{x-1,y-1}, F_{x,y-1}, F_{x+1,y-1}, F_{x-1,y}, F_{x,y}, F_{x+1,y}, F_{x-1,y+1}, F_{x,y+1}, F_{x+1,y+1}\} \quad (4)$$

ここで，画像  $F'$  に対して， $3 \times 3$  のメディアンフィルタをかけた結果得られた画像を  $M$  とする．また，関数  $\text{median}\{X_1, X_2, \dots, X_n\}$  は， $n$  個の値から中央値を返すものとする．つまり，マスクした画像  $F'$  内の注目画素の値及び隣接8近傍の画素値の中央値を注目画素の新たな値とするものである．フィルタをかけることによって得られた画像  $M$  の外接矩形を板書領域として抽出する．常に最新の板書情報を獲得するために，抽出した板書領域を背景画像に埋め込み背景を更新する．さらに，板書が書かれた黒板上の位置と板書領域として抽出された時刻を記録する．

#### c) 拡大表示すべき領域の決定

処理 (a) で抽出した教師領域と処理 (b) で抽出した板書領域を利用して，重要箇所を抽出する．教師が書いた板書は複数の板書領域として検出される．そのため，1つの板書領域として登録されている範囲では，板書の内容は分からないことが多い．そのため，数秒前から現在ま

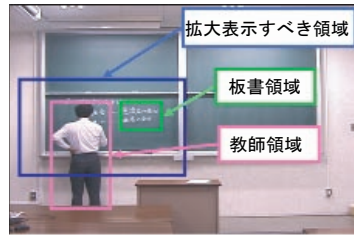
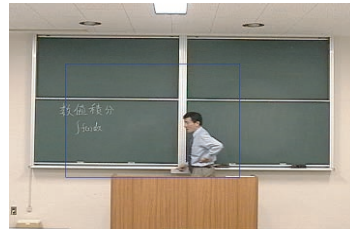
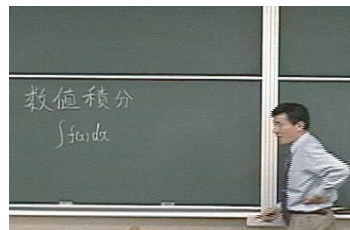


図 4 拡大表示領域の推定



(a) 講義全体の映像



(b) デジタルズームを行った映像

図 5 拡大表示映像

でに検出した板書領域と教師領域を含む外接矩形領域を算出し，その領域を含む矩形領域を重要箇所として抽出する (図 4)．

### 3.3 講義映像生成モジュール

デジタルズームには，最近傍補間法を用いる．最近傍補間法は推定する画素にもっとも近い画素の値をそのまま与える方法である．図 5(b) は講義映像 (図 5(a)) から拡大表示すべき領域を切り出し，デジタルズームを行なったものである．ここで，図 5(a) 内の矩形領域は，映像処理モジュールで推定された重要な箇所を表している．

映像処理モジュールで推定される重要箇所は滑らかに変化しないため，領域の移動やズーム率の変化を滑らかに補完することで滑らかな映像を提供する．また，領域の移動やズーム率の変化が大きいつきは，別のカメラ映像に切り替えたかのように瞬時にその領域を表示するように実装している．さらに，映像処理モジュールで推定される重要箇所が頻繁に変化すると，受講者にとって見難い映像となってしまふ．そこで，現在の表示領域と映



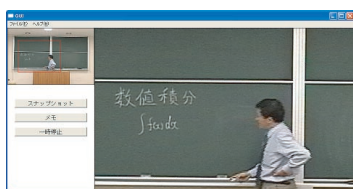


図 6 ユーザインタフェース

像処理モジュールで推定された重要箇所を比較し、細かな変化であれば、表示領域を変えないようにすることで見やすい映像を提供する。

### 3.4 学習用補助機能モジュール

現時点で実装した機能は、以下のようなものがある。

**ユーザ選択領域の拡大表示機能** 本システムで最も重要なものとして、講義全体の映像からユーザーが注目したい場所をマウスで指定すれば、クリックした点を中心として2.0倍にデジタルズームした映像を表示することができる。

**メモ帳** メモ帳のボタンをクリックすれば、メモ帳機能を持った新しいウィンドウが生成され、メモを取ることができる。

**スナップショット** スナップショットのボタンをクリックすれば、現在表示されているフレームを静止画像として保存することができる。

**一時停止** 一時停止ボタンを用意しており、それをクリックすれば、現在表示されている映像を一時停止して見ることができる。

### 3.5 ユーザインタフェース

本システムの動作例を図6に示す。図中の大きな画像の部分拡大表示画面と呼ぶ。そこには、普段はシステムが重要箇所と判断した範囲が映像として表示される。左上の小さな映像は教壇の全体の映像が表示されていて、受講者が注目したい箇所をクリックすれば、その点を中心とした映像が拡大されて拡大表示画面に現れる。その他の学習用補助機能を起動するためのボタンを左下に用意している。

## 4. 実験

### 4.1 実験環境

あらかじめDVカメラで撮影しておいた講義映像に対して、オフラインで画像処理を施すことで実験を行う。映像の1フレームの画像サイズは720×480画素である。それぞれ異なった教室で撮影された3本の講義映像に対して以下のような実験を行った。

### 4.2 教師領域及び板書領域の抽出の評価

本システムがどれだけ教師領域及び板書領域を抽出することができるかを評価する実験を行った。表1は、本システムが教師領域及び板書領域をどれだけ正確に抽出できているか適合率と再現率により評価したものである。

表 1 教師領域、板書領域抽出結果

	教師領域	板書領域	
	適合率 (%)	適合率 (%)	再現率 (%)
講義映像 1	98.4	81.3	92.9
講義映像 2	99.0	82.4	87.5
講義映像 3	98.4	83.3	89.2

教師領域の抽出では、適合率のみを用いた評価を行った。これは、実験で使用した講義映像内には常に教師が存在し、システムも必ず教師領域を抽出するため、分母が同じになるので、適合率と再現率は同じ値になるからである。ここで、適合率及び再現率はそれぞれ式(5)、(6)で計算されるものである。

$$\text{適合率} = \frac{\text{抽出した領域のうち正しい領域の総数}}{\text{システムが抽出した領域の総数}} \times 100 \quad (5)$$

$$\text{再現率} = \frac{\text{抽出した領域のうち正しい領域の総数}}{\text{システムに抽出してほしい領域の総数}} \times 100 \quad (6)$$

教師領域及び板書領域について、いづれの映像に対しても高精度で抽出ができていていることがわかる。教師領域及び板書領域の抽出がうまくいかないのは次のような場合である。

- 文字が教師領域内に書かれた場合

板書領域を抽出するには、背景差分によって得られた前景画像から教師領域をマスクする必要がある。そのため教師領域内に文字を書いた場合、黒板の文字もマスクされるため板書領域の抽出がうまくいかず、黒板の文字が教師領域外になったときに板書領域を抽出できるため、教師が連続して文字を書く場合やあまり移動しない場合にはなかなか板書領域が抽出できない。

- 黒板をスライドした場合

実験に用いた映像の講義では、上下にスライドできる黒板を使用しているため、黒板をスライドした際に背景差分を行うと前景として教師領域や板書領域以外のものまで抽出されてしまう(図7)。そのため、教師領域及び板書領域がうまく抽出できていない。

- 教師が指差し動作を行った場合

実験に用いた映像では、教師が指で説明対象を指すという行為がある。教師が指差し動作を行ったとき、教師の指が推定された教師領域内含まれていれば指の領域は教師領域としてマスクされ板書領域もうまく抽出することができる。しかし、前景として抽出された指領域は非常に小さいので収縮処理を行ったときにさらに小さな領域となったり、消えてしまったりすることがある。すると、指領域が教師領域内に含まれなくなり、教師領域がマスクされても前景画像に指領域が残ってしまう。その結果、その領域を板書領域として誤検出してしまう場合がある。

上に挙げたような問題があるが、これらの現象は瞬間的なものである。また、映像処理モジュールで背景画像を更新する処理を行っているため、しばらくすると教師領域及び板書領域をうまく抽出できるようになる。



(a) 入力画像



(b) 教師領域抽出結果

図 7 黒板をスライドした場合

表 2 拡大表示領域抽出結果

	成功率 (%)
講義映像 1	88.5
講義映像 2	92.1
講義映像 3	91.4

#### 4.3 重要箇所の抽出の評価

本システムがどれだけ正しく重要箇所を抽出できるかを評価する実験を行った。重要箇所と判断した範囲に最新の板書領域が含まれており、かつズーム率が 1.0 より大きい、つまり講義映像全体ではないときを評価の対象とした。

その結果を表 2 に表す。実験は、講義映像を見て、作成者が注目したいと考える領域を本システムが適切に抽出することができたかを検証するという設定で行った。

いずれの映像に対しても高精度で重要箇所が推定できている。板書領域の抽出において、抽出したものの約 80% 強のものしか正しくないと言う結果を得ている。しかし、重要箇所の抽出ではそれ以上の精度で抽出できている。これは、重要箇所は、数秒前から現在までに検出された板書領域と教師領域を含む外接矩形領域を算出し、その領域を含むようにして抽出している。そのため、比較的短時間で起こる板書領域や教師領域の抽出誤りは、重要箇所の抽出にはあまり影響を与えないためである。また、デジタルズーム率を最大で 2.0 倍としているので教師が説明している場所を厳密に限定する必要がなく、拡大表示してほしい板書情報や図形が重要箇所の領域に含まれていればよいということも表 2 のような結果が得られた要

因であると考えられる。ここで、デジタルズーム率を最大で 2.0 倍にしているのは、ズーム率が 2.0 より大きくなると見難い映像になってしまうと著者が判断したためである。

## 5. おわりに

本稿では、カメラマンが撮影したような動きのある映像を受信側で生成する受信者向けの遠隔講義支援システムを提案した。配信側では、固定カメラによって講義風景全体を撮影した映像を配信するだけでなく、受信側では、提案したソフトウェアを導入することでカメラマンが撮影したような動きのある映像を生成することが可能なので、比較的低コストで導入することができる。さらに、今までの遠隔講義支援システムと異なり、受信側で映像を生成するので、講義風景全体からユーザが注目したい領域を表示することが可能となり、従来の支援システムに比べ、より受講者のニーズに応えることができるシステムを実現した。

本システムのユーザインタフェースは開発途中のものである。そのため、操作性などの評価のため、今後、多くのユーザに本ソフトウェアを使用してもらい、遠隔講義受講者にとってより使いやすく改良する必要がある。

本システムは同期型と非同期型の両方をひとつの枠組みで設計してきた。しかし、同期型の遠隔講義を行う際には、教師へのフィードバックが重要である。教師は受講生の態度や質疑などのアクションに応じて講義の進め方を決定している。そのため、同期型の遠隔講義を支援するために、遠隔教室の受講生の受講態度や理解度、質疑の有無などを教師に知らせるようなシステムが必要になると考えられる。

## 文 献

- [1] Atsushi Shimada, Akira Suganuma, Rin-ichiro Taniguchi, "Automatic Camera Control System for a Distant Lecture Based on Estimation of Teacher's Behavior," proceeding of the Seventh IASTED International Conference on Computers and Advanced Technology in Education, pp.106-111, 2005.
- [2] 西口敏司, 亀田能成, 角所考, 美濃導彦, "大学における実運用のための講義自動アーカイブシステムの開発," 電子情報通信学会, Vol.J88-D-II, No.3, pp.530-540, 2005.
- [3] 大西正輝, 村上昌史, 福永邦雄, "状況理解と映像評価に基づく講義の知的自動撮影," 電気情報通信学会, vol.J85-D-II, No.4, pp.594-603, 2002.
- [4] 大西淑雅, 山口真之介, 西野和典, 小林史典, "九州工業大学における e-Learning の実践," Journal of Multimedia Aided Education Research, No.1, pp.45-58, 2004.
- [5] 岩田陽子, 加藤直樹, 中川正樹, "対話型電子白板を用いた電子化授業への遠隔受講者参加方式の試作," 情報処理学会研究報告 (CE-67), Vol.2002, No.119, pp.33-40, 2002.