

講義状況の取得方式と取得映像の 教育応用に関する考察

美濃 導彦 亀田 能成

京都大学 総合情報メディアセンター

〒 606-01 京都市左京区吉田本町

E-mail: { minoh, kameda } @media.kyoto-u.ac.jp

あらまし 大学で行われている講義を、情報メディアにより記録したり、遠隔地に伝送するためには、講義の状況を把握し、カメラ制御を行って講義の映像を作成しなければならない。これを計算機が自動的に行うことができれば、簡単に講義映像が取得できるようになり、さまざまな利用が考えられる。本稿では、実際の教育の場で利用可能なシステムを目指して我々の研究室で行っている講義映像取得およびその処理に関する研究を紹介する。このような技術は、これまでの講義に代るものではなく、新しい講義の形態が一つ増え、学生に選択枝を与えるととらえることで、現在の講義と共存する形態を模索している。情報メディアを使った新たな教育環境をどのように利用していくかは教育現場での議論を待つ必要がある。

キーワード 講義状況, コンテキスト, カメラ制御, 映像処理, 教育環境

A method for Taking and Processing Several Videos in a Classroom

Michihiko MINOH Yoshinari KAMEDA

Center for Information and Multimedia Studies

Kyoto University, Kyoto 606-01 Japan

E-mail: { minoh, kameda } @media.kyoto-u.ac.jp

Abstract Lectures are one of the most important functions of the university. To record lectures automatically by a computer will become a key technology for distance learning systems and making course materials. To record lectures in videos by a computer, the context of the lecture has to be considered. Based on the context, the way of controlling the camera has to be changed. Developing this kind of techniques provides the students a new option to take a lecture in the university. The effectiveness of such technology will be evaluated in real educational environment.

Key words Lecture status, Context, Camera control, Video processing, Educational environment

1 はじめに

計算機とコンピュータネットワークの性能向上は、マルチメディアデータを手軽に扱えるマルチメディア環境の構築を可能にした。このようなマルチメディア環境の一つの大きな応用分野は教育であると考えられる。

教育においては、実物の画像、3次元形状やその動き、実験の過程と発生する音、抽象的な概念の可視化映像、エッセンスだけを抽出したわかりやすい図や表などを画像や映像として提示することが生徒に疑似経験をさせるという意味からも重要である。特に、危険が伴う実験や大型設備が必要な実験などに対しては、その効果は絶大である。講義において、このようなマルチメディアデータを十分に活用することは、講義を興味深くするだけでなく、講義をわかりやすくするので、教育効果が上がると予想される。

このためには、講義室に画像や映像を提示できるディスプレイやモニタが設置されている必要がある。このような設備は、マルチメディアPCが広く普及してきた結果、製品コストが下がり性能もよくなって、手に入りやすくなってきている。今後は、大学の多くの教室でこのような設備が導入されて、技術的な情報インフラは徐々に整ってくる。

次の問題はこのような情報インフラをどう活用するか、である。実物の3次元データや実験のビデオなどを作成することは、コンテンツ作りと言われている。コンテンツがないと、情報インフラは役に立たないので、最近では、コンテンツの重要性が強調され始めている。コンテンツは個別に作成する必要がある、本質的に手間のかかる問題である。教育を実際に行う先生が積極的に取り組まなければならないが、残念ながら、現状では、大学における教材作成に対する支援環境は貧弱で、教員側にもこのような意識はあまりない。

技術的な観点から、教育現場で実際に情報メディアを使った遠隔講義、会議などさまざまな試みが始まっている [2]。これらの実験を通して、さまざまな知見が得られつつある。本稿で

は、計算機による講義状況取得方法および講義状況を記録した映像データを利用する方法に関するさまざまな試みについて議論する。

2 情報メディアを利用した教育の支援

大学における教育も、小学校や中学校における教育と同様、先生と生徒による人間対人間のコミュニケーションが必須であることは言うまでもない。計算機を利用した情報メディアの発達により、ある程度の制約はあるが、人間対人間のコミュニケーションを映像というメディアを通して、時間、空間の制約を弱めた形で行えるようになりつつある。

空間的な制約を弱めるために行われている映像を使った実時間コミュニケーション（遠隔講義）と実空間を共有するコミュニケーションの違いについては、心理学的、技術的に評価していかねばならないが、映像を使った教育が新たな局面を迎えていることは疑う余地はない。これは今までの講義を置き換えるものではなく、新たな講義の形態を追加するものである。

大学においては、講義室や時間割などの制約から、同じ講義を2度以上繰り返す必要が多く存在する。これは、講義を行っている側からすれば、大学の設備の制約のために生じる不合理な問題であり、何らかの解決が望まれる。このような問題は演習科目に多く起こる。特に、計算機を利用して行う演習では、多くの端末を用意しなければならないので、どうしても小人数のクラスになる。

情報メディアを利用した教育支援は、実時間性を保ったままで空間的制約を弱める方向と時間、空間的制約を同時に弱める方向が考えられる。具体的な講義の支援形態としては、次の二つになる。

- 講義の状況を遠隔地へ伝送する遠隔講義
- 講義の状況を記録し、自習用教材として活用する教材作成

どちらの方向でも、技術的には、まず、最初に講義の状況を取得しなければならない。このためには、講義状況に応じたカメラの制御が必須であり、現在は、カメラマンやディレクタなどが行っていることを、ある程度、自動で行えるようにしなければならない。講義は、スポーツなどと違って動きが激しいものではなく、映像化対象も先生を中心に考えればよいので、計算機による自動化が可能な程度の適当な複雑さであると考えられる。ただ、実時間で実空間で起こっていることを記録するのであるから、カメラ制御を間違えて重要な映像を見逃すことがないように、人間ならば1台のカメラで順次撮影できることを、カメラを複数台設置して冗長性をもたせるような枠組を考えなければならない。

講義状況の取得においては、講義に対して何らかのコンテキストを定義する必要があり、定義したコンテキストを検出し、それに応じて映像化方法を変化させることを考えなければならない。どのようなコンテキストを定義するのか、そのコンテキストに対してどのようにカメラを制御するのが講義状況を取得する上で重要な問題となる。

講義の中心は先生であるので、先生か、先生の指し示しているものを撮影することが基本である。したがって、コンテキストは先生の位置、及び先生の動作で決定できる。

複数台のカメラが設置されている状況では、先生が指し示す可能性のあるところを撮影し続けるカメラを設けることが可能であるので、映像を実時間で取得するだけでなく、実時間でさまざまなデータを集めることにより、実際の講義とは違った形で情報を提供することも可能になる。

3 講義状況の取得方法

3.1 講義の観察

講義は、固定された空間で行われる視覚と聴覚を中心にしたマルチメディアコミュニケーションである。講義状況は、これらのマルチメ

ディアコミュニケーションを総合的にとらえたもので、動画像と音声同期した映像として記録できる。これは、普通のTVカメラで録画することと同等であるが、複数台のカメラを講義室に設置するので、カメラ間で同期がとれていることが必要である。

講義の主役は、先生と生徒であり、人間は全てどちらかに分類できるとする。一般的な講義には、先生が一人、生徒が多数であり、まれに助手などが存在することもあるが、考慮しないことにする。

先生が講義室で存在する可能性のある空間的な位置は、

- 教壇
- 通路
- 黒板
- OHPなどを提示するスクリーン

などである。

また、先生の行動を視覚的に分類すると

- 歩いている
- 止まっている
- 黒板に字を書いている
- OHPなどを指し示している

などが考えられる。

講義は先生の話が中心であるので、先生の行動の聴覚的な分類として

- 喋っている
- 黙っている

が考えられる。

コンテキストは、本来、どのような処理をするかにしたがって定義しなければならない。カメラ制御を例にとると、先生の存在する場所と行動の関係で表3.1に示すようなコンテキストを考えることができる。

複数台のカメラが設置されている状況では、先生が教壇のそばを歩いていると先生を追跡

表 1: カメラ制御のためのコンテキスト

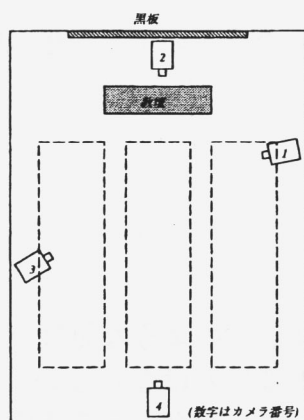
	教壇	通路	黒板の前	スクリーンの前
歩く	○	○		
止まる	○			
字を書く			○	
指示する			○	○

するカメラからの映像を重視し、先生が黒板に字を書き出すと黒板カメラからの映像を選択する、という形で、表に示したカメラ制御のためのコンテキストを利用する。すなわち、実時間で講義状況を送出している状況では、計算機によるディレクターがこのコンテキストを参照して、送出すべき映像を決定する。

3.2 カメラ機能の構築

計算機によるカメラ制御、すなわち、バーチャルカメラマンを構築する事を考えるとき、制御のエラーを最小限にするためには機能を単純にする必要がある。そのためには、一台のカメラで色々なことをするのは得策ではなく、複数台のカメラを講義室に設置し、個々のカメラに単純な機能を割り当てることを考える。すなわち、単純な機能のカメラを協調させる分散協調視覚の枠組 [5] で考える。

講義室におけるカメラ配置図、及び首振りカメラの写真を図 1 に示す。これらのカメラで講義室を撮影した例を図 2 に示す。



首振りカメラの配置



首振りカメラ

図 1: 実験に供した講義室

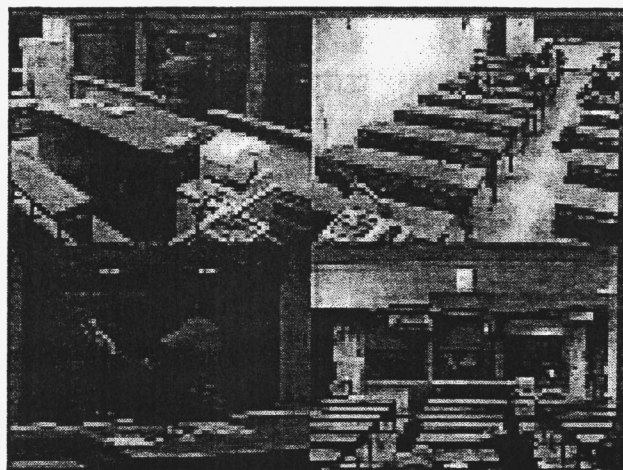


図 2: 講義室撮影画像の例

講義室の空間的対象物で映像を取得する必要があるものとしては、先生、生徒、黒板、スクリーンが考えられる。これらの対象物に対してはその映像を取得するカメラを割り付ける。先生の映像を撮影するカメラは先生を追跡することが必須となる。これに対して、生徒、黒板を撮影するカメラは、映像をよくするために、黒板の場合は字が書かれた部分を、生徒の場合は適当に首を振る程度の処理は必要となる。スクリーンを撮影するカメラは実時間で制御する必要はない。

3.2.1 先生追跡機能

2 台のカメラで 3 角測量の原理により先生の講義室内での位置を計測する。講義においては、講義室の空間的構造は既知であること、動いている物体は先生であること、を利用して、背景差分の原理に基づいて動物体を検出し、それに基づいて先生を追跡する。厳密に言えば、動物体追跡システムで、動物体が先生であるという仮定のもとで先生を追跡する。原理図を図 3 に示す。

3.2.2 生徒撮影機能

講義室の前に設置したカメラより、生徒の映像をとるカメラを構築する。撮影機能としては適当に首を振る程度で良い。

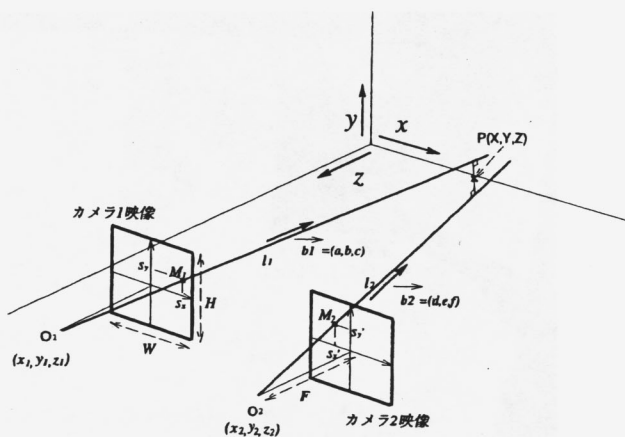


図 3: 三角測量による先生の追跡

生徒を撮影するカメラに付加する機能としては、生徒の着席場所、そこでの生徒の顔画像を集めておけば、あとで出席状況の把握に役立つ。このために、人間の標準顔のテンプレートを作成し、これをもとに画像から顔を検出する。講義中、学生は下を向いていることも多いので、顔の良さを評価し出きるだけ正面顔を記録するようにする [4]。

図 4 に入力された映像の一部と記録した生徒の顔の一覧を示す。

3.2.3 黒板撮影機能

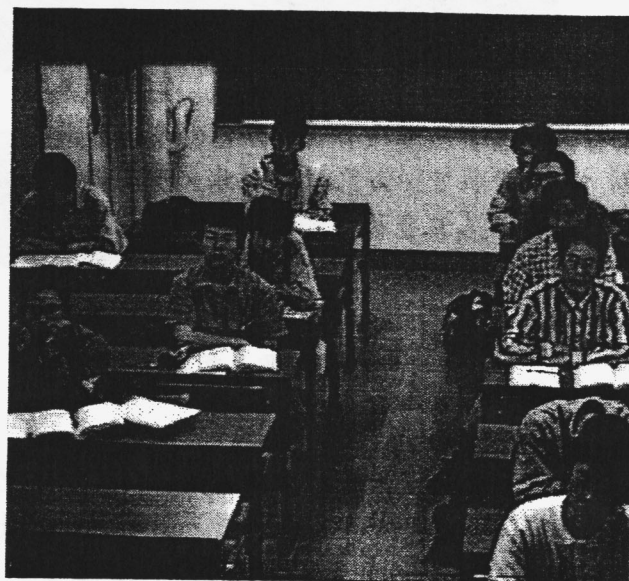
講義中、黒板をとり続けるカメラ機能を考える。カメラを講義室の後方に設置し、先生の板書を取り続けるだけでよいが、黒板の字が読める程度の解像度が必要となる。

先生が黒板を書いているということを検出し、書かれた文字がある場合に文字が読めるようにズームアップして、黒板を撮影し、取得した映像を繋ぎ合わせることにより、講義の期間中の板書を時間順に記録する。

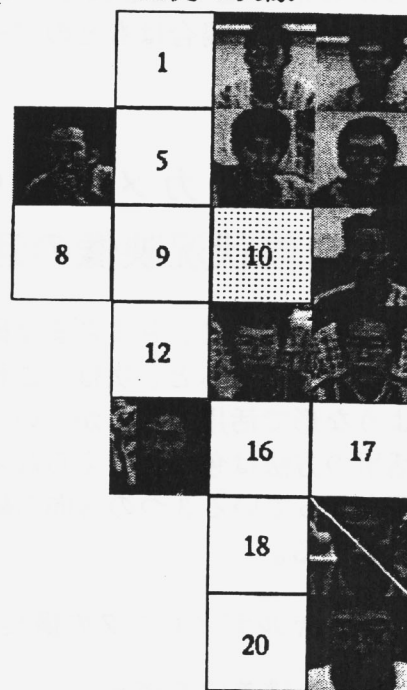
これは、この映像を蓄積し、あとで教材を作成する場合に活用できる [3]。図 5 に黒板映像と検出された文字列を示す。

3.2.4 スクリーン撮影機能

スクリーンを撮影するカメラの映像を図 6 に示す。撮影のためにはこれで十分であるが、あ

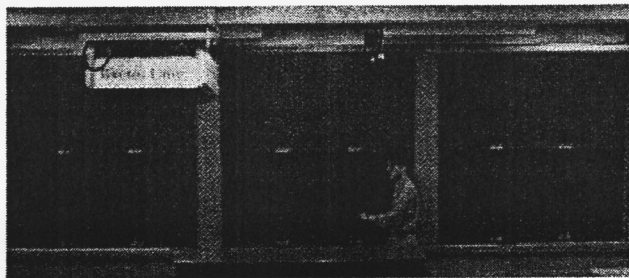


生徒の映像

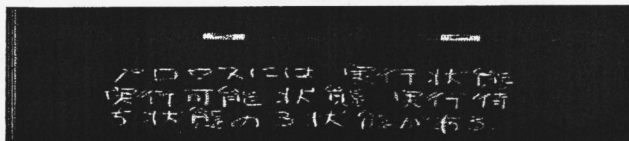


処理結果

図 4: 生徒映像の取得と処理結果



黒板映像



ズームアップして処理した結果

図 5: 黒板映像と板書取得結果

とで利用できる形にするためには、OHP が代ったことを検出する必要があり、画像の変化を検出することによりこれを行う [7]。もとの OHP が何らかの形で電子化されていると仮定すれば、後で利用する場合はもとのデータを活用できる。

4 複数台のカメラから得られる講義状況映像の活用

複数台のカメラで、さまざまな撮影機能を実現できるようになると、次は、これらの映像をどのような形で活用できるかという問題が生じる。活用の方法は色々と考えられるが、ここでは、現在行っている 3つの方向の研究について簡単に述べる。

- バーチャルディレクタの構築
- 3次元の講義室を復元
- 自習教材の作成

4.1 複数カメラ映像を選択するディレクタの構築

遠隔講義のようにオンラインで講義を送出する場合は、複数台のカメラから得られる映像

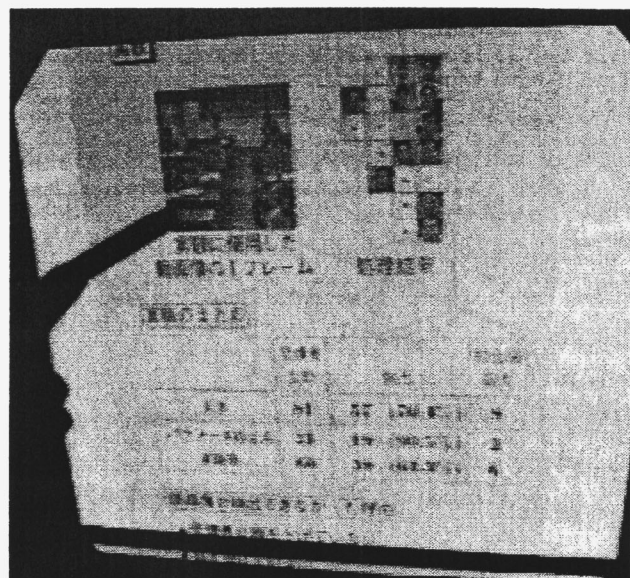


図 6: OHP 映像の例

を何らかの形で取捨選択する基準を設け、1つの映像を選択しなければならない。人間社会では、この仕事はディレクタの仕事であるので、この機能をバーチャルディレクタと名付ける。

図 7 にシステム全体の構成図を示す。カメラは講義室の特定の場所に固定されているので、カメラ機能を講義のコンテキストに応じて割り付ける。各カメラは自己の取得している映像の評価を評価値として表現し、その値をディレクタに送る。ディレクタは、評価値と、講義を見ているであろうユーザの見やすさを考慮して、映像を切替える [6]。

このシステムは、遠隔講義におけるカメラ制御として利用できる。遠隔講義をするためにカメラマンを雇うことは不可能なので、このようなシステムを利用すれば、受講者に退屈させない映像を送出できる可能性がある。

4.2 モデルに基づく動的 3次元空間の再構成

複数台のカメラから得られる同期した映像を利用して講義室の 3次元空間を再構成し、この空間に実映像を重ねあわせることにより、遠隔から講義に参加しているユーザに仮想空間とインタラクションを許すシステムの構築を目指し

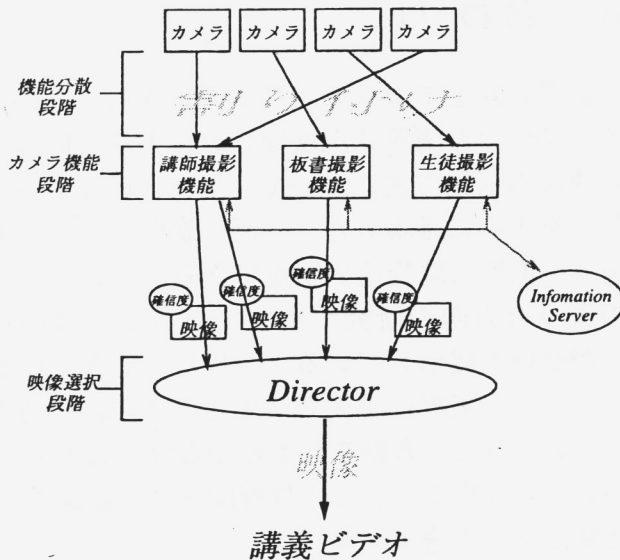


図 7: 講義送出システムの概念図

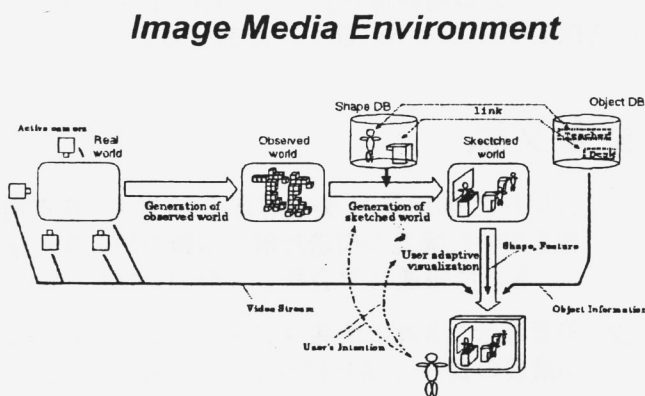


図 8: システム構成図

た研究を行っている [5]。

このシステムでは、実時間で講義室の 3 次元状況を再構築し、そこにさまざまな情報をハイパーリンクという形で付加することにより、現実の講義室よりもリッチな情報環境を提供することを目的としている。現在のところ、かなり粗い形ではあるが、実時間で講義室の 3 次元復元ができています。動いているのが先生しかないという仮定のもとでの結果の一部を図 8 に示す。

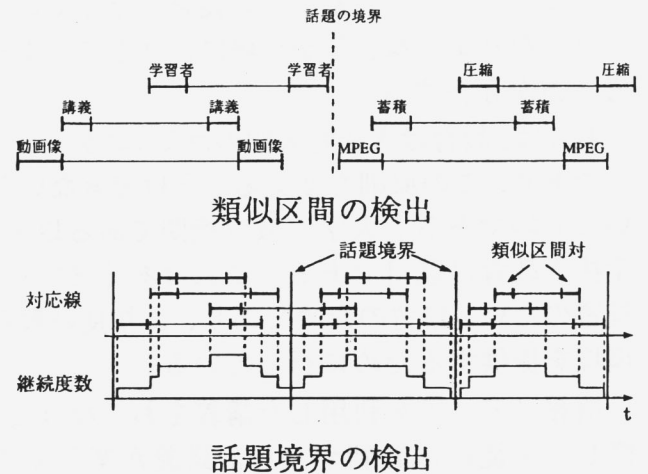


図 9: 話題境界の検出の例

4.3 講義映像をもとにした自習教材作成

これまでの応用では、音声はそのまま処理する必要はなかったが、講義の映像をもとに自習教材を作成しようとする、講義をある意味のある単位に分割し、分割した単位に対応する音声を張り付けることが必要になる。分割は、音声の切れ目、OHP の変更、板書のはじまりなど、さまざまなイベントを定義することで行うことができる。講義状況のコンテキストに基づいて、それぞれの分割単位に適当な映像を付加した後に、意味的な関係を利用して、それぞれの分割にハイパーリンクを張ることを目的としている。

講義音声の中での同じ単語が多く出現する区間を類似度区間とし、類似度区間を蓄積したヒストグラムをもとに話題境界を検出する。模擬講義を行って実験した結果を図 9 に示す [1]。このような手法で、ある程度、講義をセマンティックスに基づいて分割できる可能性がある。

5 教育への応用に関する考察

講義状況を取得すると取得した映像をどう使うのかということがよく問題になる。これは何のためにやっておられるのですか、という質問をよく受ける。学生が講義をサボってもいいよ

うにですか、とか、先生が講義をしなくていいようになりますね、とか立場によりさまざまな考え方が存在する。

我々は、教育は人間対人間のコミュニケーションであり、この原則を変えることはできないという立場である。大学が教育機関である以上、手抜き教育は駄目である。これはあくまでも、補習のため、自習のためであり、より良い教育環境を提供するために行っている。

情報メディアを利用した講義をおこなうと、概して生徒には評判が良い。講義が楽しくてわかりやすいというのがその理由である。しかし、教材の作成、講義の準備にこれまでのノートは使えない。最初から講義を設計し直す必要があり、これが多くの先生方が躊躇される理由である。

しかし、学内を中心としたデータベースを整備すれば、素材、教材、資料などが共有できる。データが蓄積されれば、ノウハウが蓄積されれば、情報メディアを活用した講義や、実際に行った講義を教材として用いる講義などが可能になる。このようになれば、講義の変更やデータの変更が容易になり、多くの先生方が利用して戴けるものと確信している。

このような状況をつくり出すためには、最初が肝心である。熱心な先生方が、魅力のある講義を実施され、それが情報メディアを通して公開されることが重要である。これで、生徒の側に人気が出て、講義に人が集まるようになり、他の先生方も積極的に情報メディアを利用されるようになるというのが理想である。このためには、大学の講義に冗長性をもたせ、競争原理を導入しなければならず、これには異論が多く出る可能性がある。

一つの大きな方向は、情報メディアによる講義の公開であり、これがひいては大学の公開、外部評価、講義の質の向上をもたらすことを期待している。学生にとっては、講義の映像を活用することにより、自習、独学が可能になり、大学にどれだけ在学したかよりも、大学でどのようなことを学んだかが重要になる。そのためにも、教材の作成、収集に努めていくのが、今後、大学の大きな責任になっていくであろう。

6 おわりに

講義に対してコンテキストを定義し、その定義に基づいた講義状況映像の取得方法、取得した映像の利用方法について、現在、行っている研究を中心に議論した。まだ、改良すべき点も多く、実用には程遠い状況である。しかし、このような方向の研究が重要であることは間違いがないと確信している。今後は、できているところから実際に講義で活用し、その教育効果について評価、考察をしてゆかなければならない。その過程で、多くの教育関係者と議論する必要があると考えている。

謝辞 本研究には、京都大学大学院工学研究科池田克夫教授、及び池田研究室の先山卓朗氏の協力を得ました。ここに感謝します。研究費は、科学研究費基盤研究 B-09558034、学術振興会未来開拓プロジェクト「分散協調視覚」(JSPS-RFTF 96P00501)の支援を受けている。記して感謝する。

参考文献

- [1] 吉川将之 仙田修司 美濃導彦 池田克夫：音声情報を用いた講義の構造理解と話題の類似度の検出，電子情報通信学会総合大会，D-517，1996.
- [2] 美濃導彦：講義の情報メディアによる支援－遠隔講義の要件－，情報処理学会研究報告，情報メディア 27-7，p.45～52，1996.
- [3] 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記録手法，MIRU'96，Vol.1，I-301～306，1996.
- [4] 先山卓朗 亀田能成 美濃導彦 池田克夫：テンプレートマッチングによる動画像からの受講者の顔画像検出，MIRU'96，Vol.1，I-301～306，1996.
- [5] 松山隆司 浅田稔 美濃導彦 和田俊和：分散協調視覚プロジェクト－分散協調視覚研究、システム開発の概要－，情報処理学会研究報告，CVIM103-4，1997.
- [6] 亀田能成 吉吉健太郎 美濃導彦：講義室における複数カメラからのビデオ送出法，映像メディアシンポジウム，pp.13-14，1997.
- [7] 先山卓朗 椋木雅之 美濃導彦 池田克夫：講義中のイベントに基づく講義の分割記録システム，第3回知能情報メディアシンポジウム，1997.