

# 講義映像の自動生成とその効果

藤吉 弘巨

中部大学 工学部情報工学科

## 1 はじめに

近年，知のアーカイブとオープン化という理念の下に，講義映像の収録や配信，さらにはその関連情報のインターネット上における公開が活発化している．特に，コンテンツ配信サイト iTunes U（2007年5月に開設）では，マサチューセッツ工科大学やカリフォルニア大学バークレー校など，多くの有名大学の講義が無料で一般公開されている．講義映像の配信に関しては，映像を PC や iPod 等の携帯型映像端末へ転送することで，利用者は時間や場所に制約されずに学習できるといった利点があり，また，視聴者に講義の雰囲気も伝えることができるという点からもその需要は高い．このような講義映像の収録と配信は多くの大学で取り組む世界的な潮流であり，ICT を活用した教育の導入が活発に行われているのだが，一方では，メディア教育開発センターの報告 [1] によると，e-learning に関するシステム開発やコンテンツ作成，それらの維持のための人員不足やコスト高などの課題も指摘されている．実際のところ，講義収録や配信の専門スタッフを抱えられる大学は世界的にも少数で，人的，予算的にリソースを得ることは難しく，教員に大きな負担がかかっているのが現状である．講義アーカイブとオープン化の重要性は認識されていても，撮影，編集，配信等の負荷を大きく低減しない限り継続は困難である．

我々の開発した講義映像配信のための自動編集システムは，より見やすい映像を生成しながらも，これら問題点を解決している．開発したシステムは，固定のハイビジョン（HD:High Definition）カメラを用いて撮影し，カメラマンは不用である．講義映像から講師や板書等の注目領域に対してトリミングを自動的に行い，不要な映像区間をカットすることで，配信に適したファイルサイズの講義映像を自動生成することが可能となる．さらに，トリミングを行う際に，放送カメラマンのカメラワーク特徴を模倣することにより，学習者が見やすくかつ臨場感あふれる映像を生成することが可能となる．本稿では講義配信を目的とした自動編集処理の技術として，仮想カメラワーク技術と時間短縮処理技術について述べ，システム例と講義配信の効果について紹介する．

## 2 講義映像の自動編集

数学や工学等の理工系の講義では大きな黒板を最大限に生かした講義が多く，黒板の板書文字を判読できるように講義映像を収録する必要がある．講義映像の収録に三脚で固定した HDV カメラを利用すると，幅 6 メートルある黒板でも板書文字を判読することが可能となる．しかし，フル HD 解像度 (1,920 × 1,080 画素) で撮影した講義映像は，90 分の場合そのファイルサイズは約 100GB となり，ダウンロードに時間を要する等の問題で講義配信に向いていない．ファイルサイズを小さくするために画面サイズを縮小することが考えられるが，板書文字を判読することができない．さらに，固定で撮影したカメラ映像は，カメラワークが発生しないため，

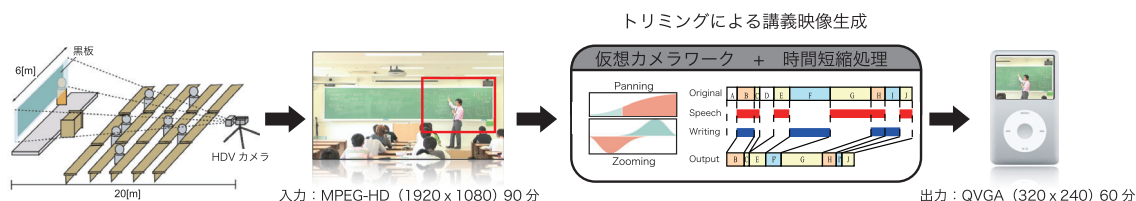


図 1: 講義映像の自動生成の流れ

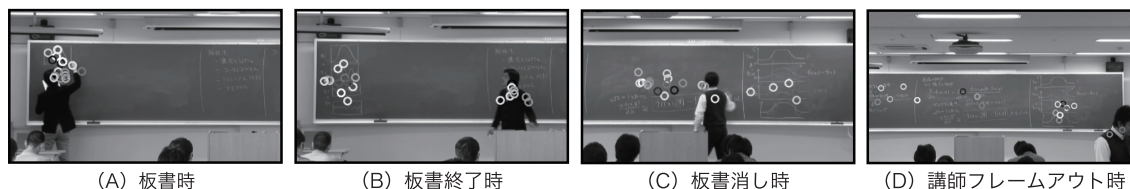


図 2: 視聴者の注目点の分布

単調な講義映像となってしまう。我々の開発した仮想カメラワーク技術はこれら問題点を解決している。

また、配信した講義映像内に講義要素が存在しない区間が多いと、視聴する学生にとっては退屈である。これに対処するために、学習者にとって不必要な区間をカットする時間短縮処理技術を開発した。この処理により、ファイルサイズも大幅に縮小される。

以下では、図 1 に示す講義配信のための自動編集システムで用いる仮想カメラワーク技術と時間短縮処理技術について述べる。

## 2.1 仮想カメラワーク技術

### 2.1.1 講義における受講者の注目点

講義に適した映像を自動生成するために、我々は、受講者が映像中のどこに注目して講義を視聴するかを調査した。20名の学習者の注目点の分布を図 2 に示す。(A) は板書時、(B) は板書終了時、(C) は板書消去時、(D) は講師のフレームアウト時の注目点の分布であり、以下に示す特徴があることを確認した。

- 視聴者は基本的に講師に注目
- 講師が板書時に視聴者は板書領域に注目し、板書終了後には板書領域に注目点が滞在
- 板書消し時や講師のフレームアウト時は、視聴者の注目は黒板上に広く分散

上記の特徴より、講師の動きに合わせて追従するカメラパンニングと、注目点の分散が大きい板書消し時や講師フレームアウト時にはズームアウトする必要があることがわかる。

次に、板書量と注目点の滞在の関係について調査した。板書後に注目点が板書領域から離れるまでの時間を注目固定時間と定義し、学習者の注目固定時間と板書文字ピクセル数を計測した結果を図 3 に示す。図 3 より、学習者の注目点は、板書量が多いほど板書終了後も長い時間に亘り、板書領域に滞在することがわかる。この分布に対して最小 2 乗法により 2 次関数をフィッティングすることで、板書量に対する注目時間のモデル化を行う。

### 2.1.2 放送カメラマンを模倣した仮想カメラワーク

2.1.1 の調査結果に基づいて、HD 映像からフレーム間差分により検出した講師位置と、背景差分により算出した板書量に応じて各フレームにおけるトリミング位置を決定する。その際に、

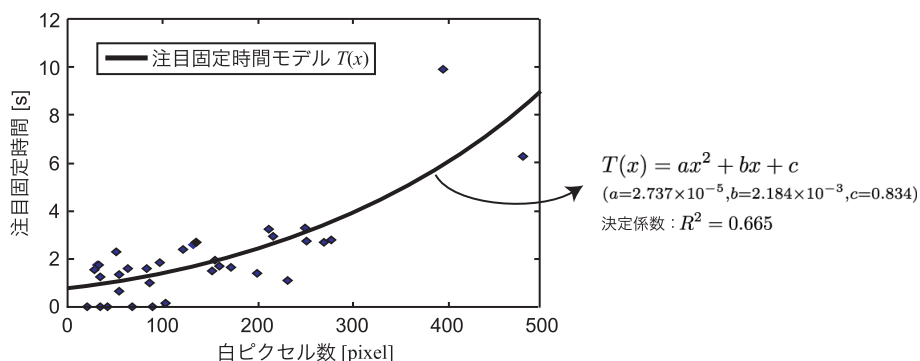


図 3: 板書文字との関連

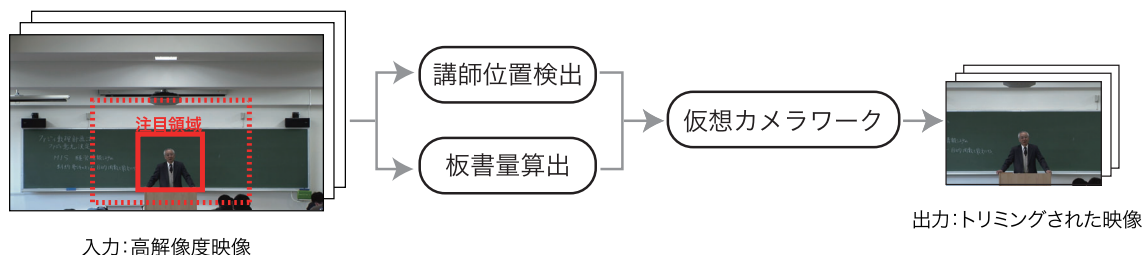


図 4: 仮想カメラワーク技術の流れ

滑らかなカメラワークの講義映像を自動生成するために、放送カメラマンのカメラワーク特徴を模倣する。NHK 放送技研による報告 [4] では、放送カメラマンのカメラワークに以下の特徴があることが報告されている。

- パンニング速度曲線は非対称で減速時間が加速時間に比べて 6 割程度長い
- ズーミングの加速時間と減速時間の平均値は、加速時間がズーミング時間の 47～60% に対し、減速時間は 40～53% である。ズームアウトの場合は、加速時間はズーミング時間の 38～63% で、減速時間は 37～61% の範囲である。

我々は上記のカメラワーク特徴を数式でモデル化 [3] し、講師位置と板書量で決定した位置から仮想カメラワークに対応するトリミング位置を最終決定する。以下に、仮想カメラワークの流れを示す。

- Step1 HD 映像からフレーム間差分により講師位置を検出
- Step2 背景差分により板書量を算出
- Step3 講師位置と板書量を基にパンニングやズーミングの決定
- Step4 仮想カメラワークによりトリミングを行い講義映像を自動生成

図 5 に、仮想カメラワーク技術により生成した講義映像の例を示す。仮想カメラワーク技術を用いることで、人が撮影したような滑らかで臨場感溢れるカメラワーク映像を得ることができた。また、トリミングを行うことで、板書文字の判読性を損なうことなく、ファイルサイズを大幅に減少した講義映像の生成が可能となった。

## 2.2 講義映像の時間短縮処理技術

講義映像を再生する際、学習者にとっては重要な区間だけを視聴できることが望ましい。我々は、講義映像から講師の発話、動き、板書区間を自動検出し、講義に無関係な区間をカットし、板書のみを早回しして短縮映像を生成する。

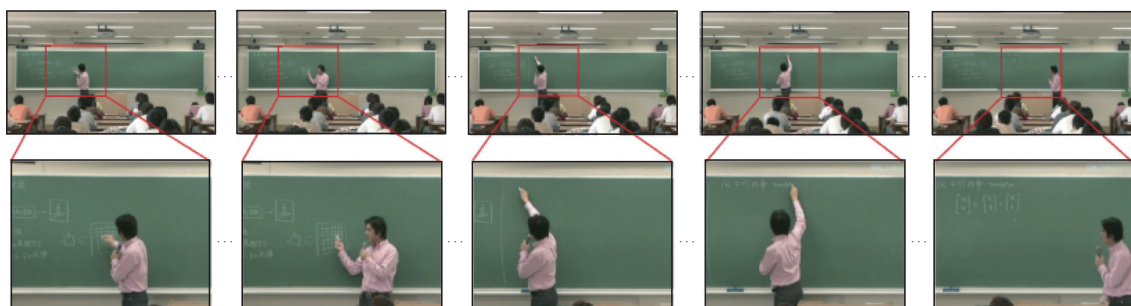


図 5: 仮想カメラワークによる講義映像の例

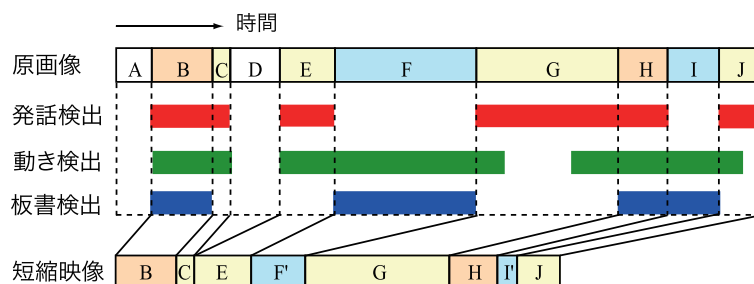


図 6: 講義映像の時間短縮

例えば、図 6 に示すブロック A や D (無発話、動きなし、板書なし) のような何も講義要素が存在しない区間は、講義の内容に関する情報量が少ないとして削除する。この時、映像が急激に変化し違和感のある映像となる場合があるため、切り替わり前後数フレームでクロスフェード処理を施し、自然な映像の切り替わりを行う。また、ブロック F のように板書中で無発話の区間は、講師の板書動作が確認できれば良いため、3 倍速で再生を行う。我々の調査では、講義時間に対して無発話で板書なしの割合は 4% ~ 10%、板書のみは 10% ~ 30% であり、90 分の講義は 60 ~ 80 分に短縮できることになる。

### 2.3 仮想カメラワーク技術と時間短縮処理技術による効果

仮想カメラワーク技術と時間短縮処理技術により、90 分の講義映像 (容量 100GB) は 60 ~ 80 分程度の映像 (容量 500MB) となり、約 1/200 にファイルサイズを減少させることができる。これにより、1 つの講義が 500MB になると、例えば、iPod 等の携帯型映像端末に 15 回分の講義を全て保存しておくことが可能となる。本技術を用いた講義映像配信システムは、収録や編集にかかる人員を必要とせず、また、学習者にとって利便性が高く、滑らかで見やすい映像であるため、大学や学習塾等の教育機関で需要が増えつつある。

## 3 システム運用例

### 3.1 追尾型講義コンテンツ自動作成ソフトウェアとポッドキャストリング

本手法を基にしたソフトウェア「i-Collabo.AutoRec」が NEC より製品化されている [5]。HDV カメラで黒板全体が映るように撮影し、ソフトウェアに取り込み、開始ボタンを押すと 2 章で述べた技術により配信に適した講義映像が自動生成される。現在は講義 1 つの自動生成に約 3 時間かかるが、人手による編集に比べ開始ボタンを押すだけでよく、編集作業は一切必要としないため手間がかからない。公開者は生成した講義映像を web サーバ上にアップロードし、RSS2.0



図 7: 講義コンテンツ自動作成ソフトを用いたポッドキャスト

を通して公開する．RSS を iTunes に登録した学習者は，ポッドキャストにより最新の講義映像を自動でダウンロードすることが可能となる．我々は，本ソフトウェアを利用して 2006 年度から講義配信を開始し，講義の補助教材の一つとして活用している [6]．

講義配信を利用した学生 20 人に，自由記述型のアンケート調査を行ったところ，以下の回答が得られた．

- 復習に使える（何度も見られることが良い）
- 欠席や試験勉強の際に手助けになった
- 板書にはなかった先生の説明を再度聞けることが良い
- 書き逃した板書を確認することができるため助かった
- 知識情報を共有できることが良い

また，期末試験の結果，講義映像を視聴した学生の平均点は視聴しなかった学生に比べ 8 点高く，視聴回数が多い学生ほど高得点であった．講義配信を利用することで自主学習の助けとなり，復習の効果も上がったものと考えられる．

### 3.2 マルチスケールに対応した講義配信システム

中部大学の教育研究施設であるデジタルラボでは，マルチスケールに対応した講義配信システム<sup>1</sup>を実現した（図 8 参照）．デジタルラボで開講される全ての講義は教室の後方に設置されているフル HDV カメラで撮影し，HD 映像としてアーカイブしている．アーカイブした HD 映像は，PC 等のストリーミング配信用 640×480 画素や iPod 等の携帯端末用 320×240 画素に自動編集され，コンテンツ管理システム（LMS: Learning Management System）を介して配信される．また，2 台のフル HD プロジェクタの合成表示により，アーカイブした HD 映像を 7.2m × 2.4m の大スクリーンに原寸大で再生することが可能であり，あたかも教室で講義を受けているかのような迫力が得られる．マルチスケールに対応することで，大学，自宅，通学時など時間を選ばず，状況に応じて学習することが可能である．

## 4 おわりに

本稿では，講義映像の自動生成技術とそのシステム例について述べた．本手法による講義映像の生成は自動化されており，人員，予算のリソース不足問題を解決している．今後，より多くの教育機関において講義映像の収録と配信が積極的に行われ，知のオープン化により，知の共有，知の継承が広く社会に浸透することを望んでいる．

<sup>1</sup>マルチスケールに対応した講義配信は世界初と思われる

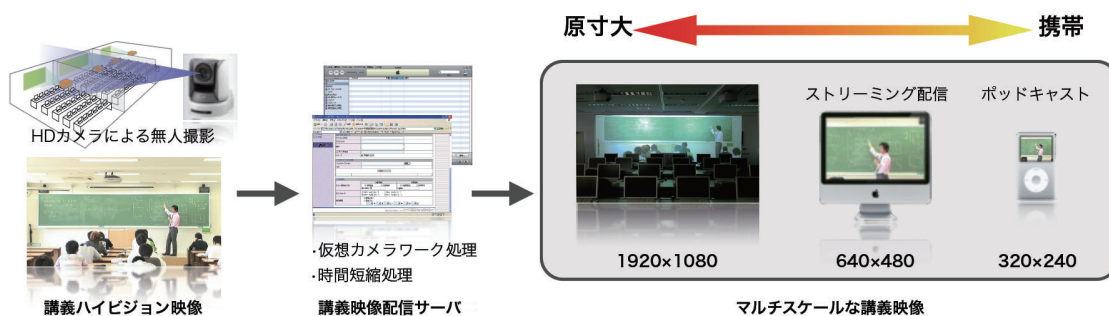


図 8: マルチスケールに対応した講義配信システム

## 参考文献

- [1] NIME (独立行政法人メディア教育開発センター), “e ラーニング等の ICT を活用した教育に関する調査報告書”, <http://www.code.u-air.ac.jp/reports/001/2007/>
- [2] 横井 隆雄, 桐井 孝嘉, 藤吉 弘巨 “講義イベント検出に基づく短縮講義ビデオの自動生成”, SSII06, pp.535-540, (2006)
- [3] 篠木 雄大, 藤吉 弘巨, “高解像度映像からの視聴者の注目点を考慮した講義映像の自動生成”, 映像学誌 2008, Vol.62, No2, pp.1-7 (2008)
- [4] 加藤 大一郎, 山田 光穂, 阿部 一雄, “スタジオ番組における放送カメラマンのカメラワークと視線の動きの分析”, テレビ誌, Vol.49, No.8, pp.1023-1031, (1995)
- [5] NEC, “追尾型講義コンテンツ自動作成ソフト i-Collabo.AutoRec”, <http://www.i-collabo.jp/autorec/>
- [6] 中部大学 工学部 情報工学科「画像情報処理」, <http://www.vision.cs.chubu.ac.jp/CULVA>