

ハイビジョンカメラと仮想カメラワークを用いた 簡易な講義ビデオ撮影方式について

永井孝幸^{1*}

¹ 鳥取環境大学 環境情報学部 情報システム学科

概要: 著者の所属する情報システム学科では学科の全講義を学生スタッフにより手動で撮影している。しかし、学生スタッフを主体とした撮影では講義と各学生スタッフのスケジュール調整を常に行う必要があり、カメラ操作要員の不要な撮影方式が望まれていた。

本報告では、ハイビジョンカメラによって講義室全体を無人で撮影し、撮影後に高解像度画像から注目領域だけを抽出する仮想カメラワークによって講義ビデオを生成する方式について述べる。本方式では高解像度カメラの利点を生かし、撮影時にマイクロQRコードをホワイトボード・プロジェクタスクリーンに提示することで素材管理・インデックス付与も撮影と同時に行う。自動生成したビデオを撮影スタッフが評価した結果、復習用途に十分使えるとの反応を得た。

Simple lecture recording with HDV camera and virtual camera work

Takayuki Nagai¹

¹ Department of Information Systems, Tottori University of Environmental Studies (TUES)

Abstract: Since 2004, our department has been recording all lectures by camcorders and streaming them through a campus network. A unique aspect of our activity is that all videos are manually recorded by student staff without expensive auto-recording equipment. However, it needs a daily effort to arrange recording staff under sudden changes of schedules of lectures and students.

In this paper, we study a simplified lecture recording method that uses a fixed HDV (high-definition video) camera and a virtual camera work technique. First, we record an entire classroom with a fixed HDV camera. Then, regions of interest are determined by voting among key frames. Finally, a standard definition video that contains the regions is generated.

1 はじめに

近年、学習支援の一環として講義の様子をビデオに収録して配信する動きが盛んになっている。Open Course Ware や iTunes U をはじめとして、日本国内でも日本オープンコースウェア・コンソーシアムや WIDE プロジェクトの School of Internet による配信が行われており、教育コンテンツを蓄積することが高等教育機関の重要な機能になりつつある。

著者の所属する情報システム学科では、講義の復習機会を提供することを目的に 2004 年度より学科の全講義を学生スタッフにより手動でビデオ撮影し、独自に開発したコース管理システム [1] を通じて学内配信を行っている [2]。この取り組みの特徴は、市販のハンディカムを用いて学生スタッフが撮影を行う方式で学科の全

講義を継続的に撮影していることである [3]。講義を自動的に撮影・アーカイブすることは既に技術的には可能であるが [4]、特定の講義室での撮影ではなく全講義を撮影するには設備の設置コストがかかる。このため、本学のような小規模大学での導入は難しい。市販カメラを用いた有人撮影方式では撮影場所にとらわれず少ない機材で講義のビデオ撮影が可能であり、実績としてわずか 4 台のビデオカメラで学科の 20 科目以上の講義を全て収録できている。

一方で撮影を学生スタッフに頼るという有人撮影方式自体の課題も浮かび上がってきた。一つは撮影スタッフの入れ替わりである。作業に習熟したスタッフほど適切に撮影・編集作業を行うことができるが、学年が上がるにつれて多忙になる、あるいは卒業するなどして引退してしまう。このため、年度初めは作業に不慣れなスタッフが習熟するまでの間、撮影・編集がうまくいかないケースが生じてしまう。もう一つの課題は撮影スケジュールの調整である。学生スタッフには撮

*連絡先: 鳥取環境大学環境情報学部情報システム学科
〒689-1111 鳥取県鳥取市若葉台北 1-1-1
E-mail: nagai@kankyo-u.ac.jp

影だけでなく、自身の科目履修やアルバイト等のスケジュールがあり、休講・補講、病気などの様々な事情に応じて撮影対象科目と学生のスケジュール調整を随時行う必要がある。

この課題を根本的に解決するには無人撮影方式が理想であるが、必ずしも完全自動収録システムが必要となるわけではない。講義室には講義の担当教員と受講生がおり、講義の開始前・終了後に機材の設置や操作を手で行うことは可能だからである。上記の課題を解決するには講義撮影中のカメラ操作だけを自動化できればよい。

カメラ操作を自動化する方法として、パン・チルト・ズームの制御が可能な三脚を用いて被写体を追跡する方式が従来から研究されている [5]。この方式ではリアルタイムに被写体を追跡するために、カメラ・三脚の他に画像解析・カメラ制御を行うためのコンピュータも合わせて必要となる。このため、機材の運搬・設置に手間がかかる。また、長期間の運用において機器の可動部分が故障する点にも留意する必要がある。

もう一つの方法として、高解像度で収録した映像から必要な部分のみをトリミングによって切り出す仮想カメラワークの手法 [6] がある。ここ数年でフルハイビジョン解像度 (1920 × 1080) での撮影に対応した民生品カメラの低価格化が進み、講義室全体を固定カメラを用いて高解像度で撮影することは容易になりつつある。そこで今回、学科の全講義を撮影・配信するという観点から、固定ハイビジョンカメラを用いた講義ビデオ撮影方式と仮想カメラワークを用いた講義ビデオの生成について検討を行った。

まず第2章でこれまでの有人撮影方式で判明した課題について述べる。第3章で今回検討を行った固定ハイビジョンカメラを用いた撮影方式について述べる。第4章では仮想カメラワークを用いた講義ビデオの生成方法について述べる。第5章では実際の講義を対象に今回考案した方法で講義ビデオの無人撮影・自動生成を行った結果について述べる。

2 有人撮影方式の課題

著者の所属する情報システム学科では、2004年度より学科の全講義を学生スタッフによって収録している。この章では、学生スタッフによる講義の有人撮影の取り組みでこれまでに判明した課題について述べる。

2.1 撮影スタッフの確保

有人撮影方式で学科の全講義を集録するには、撮影スタッフの確保が最初の課題となる。これまでの撮影実績から、撮影を円滑に行うには撮影科目数の1.5倍

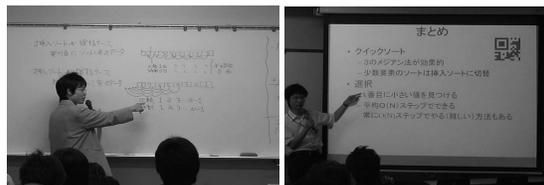


図 1: 講師と説明対象を見やすく撮影した場合

～2倍程度の撮影スタッフが必要なことが分かっている。2008年度前期の場合、全19科目の撮影を26名(2年生5名、3年生7名、4年生6名、修士1年2名、修士2年6名)の学生で手分けして行っている。

撮影科目数以上のスタッフが必要になるのは、急に撮影スタッフ・講義のスケジュールが変更になった場合に備えて予備の撮影スタッフが不可欠なためである。急なスケジュール変更の事例としては、交通ダイヤの乱れによる遅刻、就職活動、補講日程の連絡行き違い、忌引き、体調不良等が挙げられる。

また、時間割によっては特定の時間帯に撮影スタッフのほぼ全員が自分の講義に出席しており、講義撮影を頼めるスタッフがいないという科目が生じることもある。このような場合、過去に撮影スタッフをやっていた学生に臨時に撮影を依頼するなどして、毎週撮影スタッフの確保を続けながら綱渡りの運用を行うことになる。一年度を通してみた場合、特に前期の撮影スタッフの確保が難しい。前年度までの熟練スタッフの卒業に加えて4年生、修士2年の学生が就職活動で忙しくなるためである。

2.2 撮影スタッフ間のスキルのばらつき

復習に使うことを目的として講義ビデオを撮影する場合、図1に示すように講師が説明している対象(板書・スライド)を見やすいように撮影することが重要である。しかしながら、講義の進行に合わせて適切な対象にカメラを合わせるには講義の内容をその場で理解する必要があるため、撮影担当者によってビデオの質にばらつきが生じる。なかでも教室を左右に歩き回りながら説明する講師の科目では、撮影者が講師の動きをうまく予測できるかどうかによってビデオの質が大きく左右される。

全講義を継続的に集録するには一定数の撮影スタッフが必要であること、また、撮影スタッフの後継者を計画的に育成する必要があることから1,2年生も含めて撮影スタッフに採用しているが、全体的な傾向としてやはり低学年の撮影スタッフよりも高学年の撮影スタッフのほうが適切にカメラ操作を行っている。

また、新学期が始まり新規に撮影スタッフになった学生が撮影に慣れるまでの間は、機材の操作ミスにより撮影に失敗するケースや、撮影対象の選び方が不適

切で見にくいビデオになるといったケースがあった。なかでも、画面が水平になっていない、あるいは、ズームしすぎて説明対象が見にくいなど、カメラの映り具合を理解していないと思われる講義ビデオが学期当初に見受けられる。このため、学期が始まってしばらくの間は新人スタッフの撮影した講義ビデオを重点的にチェックし、撮影指導を行う必要がある。

更に、ビデオ配信時のインデックス情報として撮影時に10分毎の講義内容をメモしてもらっているが、メモの内容・量についてもスタッフ間のばらつきは大きい。

3 固定ハイビジョンカメラを用いた撮影

2章で述べたように、学生スタッフによる有人撮影で学科の全講義を集録するにはスタッフの確保とスタッフ間のスキルのばらつきが問題となる。この問題に対処するため、HDD記録型の固定ハイビジョンカメラ一台で講義室全体を撮影する方式について検討を行った。

この方式では講義中のカメラ操作が不要だけでなく、HDD型カメラを用いることでテープ交換も不要になる。そのため録画の開始・終了ボタンを押すだけの操作で済み、講義の講師・受講生自身での収録が可能になる。

3.1 撮影方法

これまでの講義撮影の経験から、長期間の講義ビデオ撮影においてはテープ記録型カメラよりもHDD記録型カメラの方が記録メディアのコスト・故障発生率・省力化の面で優れていることが分かっている。そこで、フルハイビジョン解像度(1920×1080)での撮影が可能なHDD記録型カメラとして、SONY製HDR-SR12を撮影カメラに用いることとした。

このカメラを用いて、ホワイトボードとスクリーンが一度に撮影領域内に入るように図2に示す構図で講義撮影を行う。この構図ではホワイトボードが約1300×320、プロジェクタスクリーンが約570×440の解像度で撮影される。機材の配置・実寸は図3の通りである。

この状態で撮影した映像から標準ビデオ画質の画面サイズに対応する領域(720×480)を切り出し、シャープ化処理を施したのが図4である。板書・プロジェクタスクリーン上の文字を読めるだけの解像度があることが分かる。なお、撮影レンズの性質としてレンズ中心部の解像度に比べて周辺部の解像度が低くなるため、この撮影構図ではホワイトボード両端付近の文字がにじみのため少し読みづらくなる。

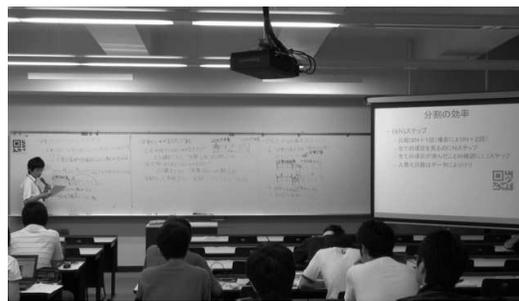


図 2: フルハイビジョンカメラ撮影による教室の全景

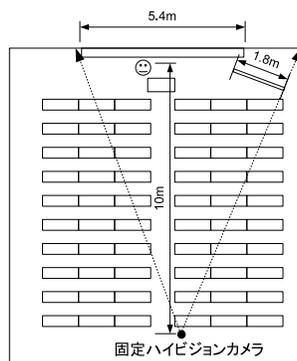


図 3: 教室レイアウトと撮影機材の配置

3.2 QRコードを用いた素材管理

本学科の講義撮影では、撮影したビデオ素材の管理を効率よく行うために講義の冒頭に科目コードに対応するQRコードを撮影し、ビデオ加工時にQRコードを検出することで素材の自動分別を行ってきた[2]。

この作業を無人で行うために、講義開始前に講師自身がホワイトボード上に科目を表すマイクロQRコードを貼り付ける方式を今回考案した。図3に示す撮影機材の配置において、ホワイトボード上に提示した17cm四方のマイクロQRコード(1モジュールあたりのサイズは1.5cm)を撮影後の画像から問題なく検出できることを確認している。

また、プロジェクタスクリーンに投影したマイクロQRコードも検出が可能であり、図5のようなQRコード付きの講義スライドを用意しておくことで講義ビデオと講義スライドの時間的対応をとることができる。これにより、撮影スタッフが講義内容のメモを取らなくても時間毎の話題を特定できる。

撮影に用いるマイクロQRコードをすぐに作成できるように、自作の授業支援システム[1]にマイクロQRコード作成ツールを統合している(図6)。このツールでは連番コードの一括作成・一括ダウンロードが可能であり、教員は講義スライドの余白にQRコード画像を貼り付ける作業だけを行えばよい。20枚の講義スライドにQRコードを貼り付けるのにかかる時間は手作業で5分程度である。

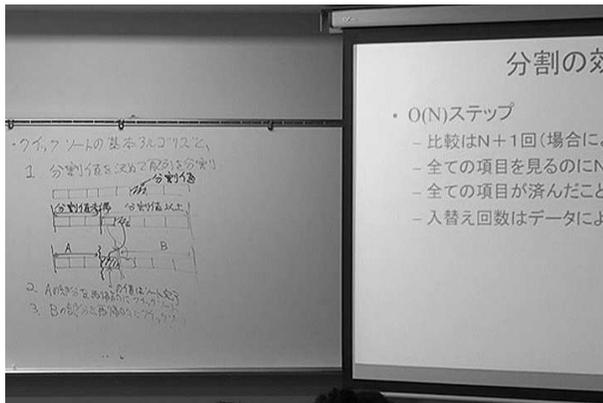


図 4: シャープ化を行ったトリミング画像 (720 × 480)

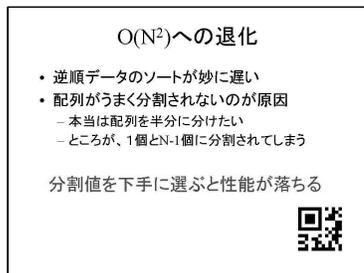


図 5: マイクロ QR コード付き講義スライド

4 仮想カメラワークによる講義ビデオ生成

ハイビジョンカメラで収録した講義ビデオは学生がいつでも自由に閲覧できることが望ましいが、本学の学生が主に利用している数年前のノート PC では CPU 性能が不足するためフルハイビジョン画質での視聴は不可能であった。また、高性能な CPU を備えた機種であってもノート PC の液晶サイズではビデオを縮小して再生する必要があり、高解像度で撮影したにもかかわらず小さな見にくい映像として視聴することになってしまう。

そこで今回、固定ハイビジョンカメラを用いて撮影された高解像度 (1920 × 1080) 映像に対して、仮想カメラワークの技法を用いて標準ビデオ解像度 (720 × 480) の講義ビデオを生成する方式について検討を行った。

仮想カメラワークとは、映像をトリミングする際にトリミング位置を時間的に変化させることで擬似的にカメラワークを実現する技法のことである。この技法ではパン・チルト・ズームの速度を任意にコントロールできる他、カメラスイッチングの効果も実現できる。この利点を生かし、大西らは仮想カメラワークを用いてスポーツ選手の演技撮影や講義撮影における自然なカメラワークを作り出す手法について研究を行っている [6]。また、横井らは板書講義に対して、仮想カメラワークを用いて自動的に講義ビデオを生成する手法の

連番QRコード作成 / 連番QRコード作成

余白	16	
ドットサイズ	4	
接頭辞	4207	
開始番号	01	
終了番号	30	



図 6: 連番 QR コード作成ツール

研究を行っている [7][8]。

これに対し、今回考案した手法はホワイトボードとプロジェクタスクリーンを組み合わせた講義のビデオを前提としているため、話者の位置とカメラの向きが一致しないケースを考慮する必要がある。さらに加工したビデオは単独のコンテンツとして視聴するのではなく、学内の授業支援システム上でインデックスを付与した状態で配信する。学生は復習に必要な箇所だけを選んで視聴することから、視聴者を長時間退屈させないためのカメラワークは特に必要としない。そのため、ここでのカメラワークはパン操作のみを考える。

4.1 基本アイデア

これまでの有人撮影においては板書・スライドの文字を見やすく撮影することを原則とし、カメラはできるだけ動かさずに講師が説明している対象 (板書・スライド) を固定アングルで捉えるという方法で撮影してきた。この撮影は (1) 板書の進行に合わせてカメラを時々パンする、(2) 講義スライド説明中はカメラをプロジェクタスクリーンに向ける、という 2 つのカメラ操作から成り立っている。ほとんどの教員がホワイトボードを大まかに 4 つの領域に区切って板書を進めていくため、前もって決められたカメラアングル間をパン操作でつないでいくという撮影方法になる。今回考案した仮想カメラワークにおいても上記の撮影方法を基本とし、「撮影範囲をいくつかのブロックに分けておき、大きな変化のあるブロックにカメラを向ける」という方針でカメラワークの計算を行う。

プロジェクタスクリーンを用いる講義では話者の立ち位置とスクリーンの位置が離れる場面があるため、単純に話者の位置を追跡してもカメラ位置の決定は行えない。そのため Condensation [9] やパーティクルフィルタ等の物体追跡の手法は用いず、単純にフレーム間差分を用いて映像上の変化が大きい領域を見つけることとした。さらにパン操作をできるだけ減らすため、ある時刻でのカメラ位置を決める際に過去と未来の映像フレームも参照し、前後数秒間における映像の変化範

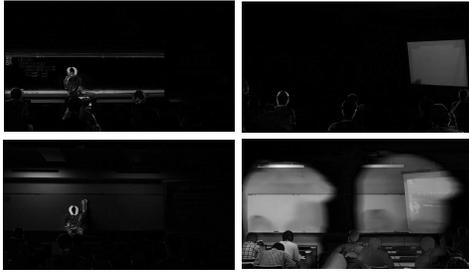


図 7: 外乱を受けた差分フレームの例

囲がカバーされるようにカメラ位置を算出する。

なお、実際の撮影では図7に示すようにホワイトボードの位置調整やプロジェクター操作、照明操作、演習中に TA の学生がカメラを横切ることによる突発的な大きな外乱を考慮する必要がある。この点に関する工夫は 4.3 節で述べる。

4.2 フレーム差分による注目ブロックの検出

撮影環境毎の照明条件の変動を受けないようにするため、フレーム差分を用いて変化の大きい領域の検出を行う。差分フレームを矩形ブロックに分け、最も変化の大きいブロックを注目ブロックとして検出する。

(1) フレーム切り出し

フルハイビジョン画質のまま連続フレーム間の差分計算を行うのは時間がかかるため、10フレーム毎にフレームを読み出し 960×540 の 256 階調グレースケール画像に変換する。切り出された各キーフレーム f_1, f_2, \dots に対し、差分フレーム $g_i = |f_i - f_{i-1}|$ を計算する。これは 1/3 秒ごとの映像の変化を調べることに対応する。

(2) 注目ブロック検出

各差分フレーム g_i の上側 960×350 の領域を 72×70 の小ブロックに分け、ブロック内の成分の和 (差分値) が最も大きいブロック (注目ブロック) の添え字を記録していく。図 8 に示す例ではブロック 5-2 が注目ブロックとなる。ここで対象領域をフレーム上方に限定しているのは、座席にいる学生の動きを検出しないようにするためである。

4.3 注目ゾーンの検出

4.2 節で検出した注目ブロックにもとづき、各差分フレームにおけるカメラの横座標を定める。差分フレームを図 8 に示すように横幅 72 の縦長のゾーンに分け、頻繁に注目ブロックが出現するゾーン (注目ゾーン) をカメラが捉えるようにする。この時、カメラ位置をあまり頻繁に変更せずに注目ブロックをカメラが連続して捉え続けられるようにしたい。

ゾーン	ゾーン1	ゾーン2	ゾーン3	ゾーン4	ゾーン5	ゾーン6	ゾーン7	ゾーン8	ゾーン9	ゾーン10	ゾーン11	ゾーン12	ゾーン13
1-1	1-2	1-3	1-4	1-5	1-6	1-7	1-8	1-9	1-10	1-11	1-12	1-13	
2-1	2-2	2-3	2-4	2-5	2-6	2-7	2-8	2-9	2-10	2-11	2-12	2-13	
3-1	3-2	3-3	3-4	3-5	3-6	3-7	3-8	3-9	3-10	3-11	3-12	3-13	
4-1	4-2	4-3	4-4	4-5	4-6	4-7	4-8	4-9	4-10	4-11	4-12	4-13	
5-1	5-2	5-3	5-4	5-5	5-6	5-7	5-8	5-9	5-10	5-11	5-12	5-13	

図 8: 差分フレームのブロック・ゾーンへの分割

そこで差分フレーム g_i だけでなく、前後各 k 枚の差分フレーム $g_{i-k}, g_{i-k+1}, \dots, g_{i-1}, g_{i+1}, g_{i+2}, \dots, g_{i+k}$ も用いて以下のようにして差分フレーム g_i における注目ゾーンを定める。 k の値として現在は経験的に 12 を指定し、前後 4 秒の変化を見るようにしている。

(1) 注目ブロック出現回数の計算

各ゾーン $z_i (i = 1, 2, \dots, 13)$ に対して上記の $2k + 1$ フレームにおいて注目ブロックがゾーン z_i に出現する回数 c_i を数え、 c_i の値が最も大きいゾーンを注目ゾーンとする。現在のフレームだけでなく過去と未来のフレームも含めた多数決によって注目ゾーンを定めるため、図 7 に示したような突発的な外乱によって出現した注目ブロックの影響を取り除くことができる。

(2) カメラ位置選択

注目ゾーンがトリミング範囲に入るようにカメラ位置を定めればよいが、トリミング範囲の幅は 5 ゾーン分あるためカメラ位置 (トリミング範囲の左端のゾーン位置) には任意性がある。

注目ゾーンだけでなくできるだけ多くの注目ブロックがトリミング範囲内に入るようにするため、各カメラ位置 $i (i = 1, 2, \dots, 9)$ に対してスコア $s_i = c_i + c_{i+1} + c_{i+2} + c_{i+3} + c_{i+4}$ を計算し、最もスコアの高いカメラ位置をキーフレーム f_i におけるカメラ位置 p_i として選択する。注目ブロックを多く捉えられるカメラ位置ほどスコアが高くなり、例えば $(c_1, \dots, c_{14}) = (2, 6, 2, 1, 2, 2, 3, 8, 0, 0, 0, 0, 0, 0)$ となるキーフレームにおいてはカメラ位置 4 が最もスコアが高い。

また、スコアが同じになるカメラ位置が複数ある場合は、変化領域の重心が画面中央に近いカメラ位置を採用する。そのために各カメラ位置 c_i における重み付き重心 $g_i = (-2, -1, 0, 1, 2) \cdot (c_i, c_{i+1}, c_{i+2}, c_{i+3}, c_{i+4})$ を求め、重心が 0 に最も近くなるカメラ位置を採用する。例えば、 $(c_1, \dots, c_{14}) = (0, 1, 0, 0, 0, 1, 24, 0, 0, 0, 0, 0, 0, 0)$ となるキーフレームにおいてはスコア s_3, s_4, s_5, s_6 の値はいずれも 25 であるが、重心の値はそれぞれ $g_3 =$



図 9: スクリーンから離れて説明している場面

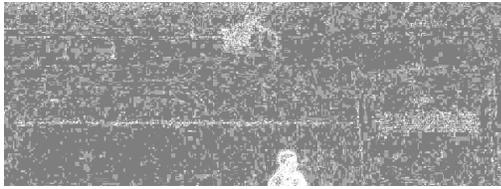


図 10: 図 9 の時点の差分フレーム (強調画像)

49, $g_4 = 24, g_5 = -1, g_6 = -26$ であり, 重心位置が最も 0 に近くなるカメラ位置 5 が選択される. もし重心位置においても優劣がつかない場合は, 残ったカメラ位置の中で最も左の位置を採用する.

(3) 不必要なパン操作の除去

注目ゾーンを単純に追跡するようにカメラ位置を定めるとカメラ位置が短時間で左右することがあるが, 見やすい映像にするためには一定時間以上 (典型的には 4 秒間) カメラ位置を静止させる必要がある [10]. そこで「動いてもすぐに (4 秒以内に) 元の位置に戻る」カメラのパン操作を取り消すために, 「 $p_i \neq p_{i+1}$ かつ $p_i = p_{i+j}$ ($j < 12$)」ならば「 $p_{i+1}, p_{i+2}, \dots, p_{i+j-1}$ の値を p_i に変更」する.

4.4 プロジェクタスクリーンの増感処理

講義スライドを主体とした講義では講師が必ずしもプロジェクタスクリーンの近くに立っているとは限らない. 講師によってはスクリーンから離れた位置に立つことが多く, 映像上の変化が大きい講師に対してカメラが向けられると説明対象の講義スライドが映らないビデオになってしまう. 本節ではこの問題の解決策について述べる.

図 9 のように講師がスクリーンから離れた位置に立って説明を行っている場面におけるフレーム差分を調べた結果, 2 番目に変化の大きいブロックがスクリーン上 (ゾーン 10~13 の範囲) に現れることが分かった (図 10). これは講師が講義スライドを切替えたり, スクリーンに指示棒を当てた際のわずかな画像の変化が検出されるためである.

そこで, 注目ブロックを検出する際に, 最も変化の大きいブロック (第 1 注目ブロック) と 2 番目に変化の大きいブロック (第 2 注目ブロック) の位置関係を調べ, 「第 1 注目ブロックがスクリーン上にないが, 第 2 注目

ブロックはスクリーン上にある」場合については第 2 注目ブロックを注目ブロックとして採用する. このプロジェクター増感処理により, 講師がプロジェクタスクリーンから離れた位置に立っている場合でも優先的にプロジェクタスクリーンにカメラが向けられるようになる.

講義スライドを使っている場面に対して, プロジェクター増感処理を施した仮想カメラワークの計算結果 (40 秒分) を図 11 に示す. 図中, x_1, x_2 とあるのがそれぞれ第 1 注目ブロックのゾーン番号, 第 2 注目ブロックのゾーン番号であり, 2 本の破線に囲まれた範囲が仮想カメラの撮影範囲を表す. この場面は (1) 講師が教室中央でスライドの説明を行い, その後 (2) 教室左端の教卓で PC を操作してプロジェクタスクリーンの近くまで移動し, (3) 最後にもう一度教卓まで戻り書画装置で資料を投影するという動きで成り立っている.

冒頭の 30 フレームでは, 講師と講義スライドが明瞭に区別して検出されており, 講義スライドにカメラが向いていることが分かる. 続く 30~40 フレームでは講師が教室左端に移動するのに合わせてカメラも左に移動し, 50~80 フレームでは講義スライドの変化と講師の右方への移動に合わせてカメラが教室右側に固定された状態になっている. 100 フレーム以降は x_1 の値が振動し不安定であるが, これは教室左端の講師の動きとスクリーンの変化を両方捉えているためである. この場面でもプロジェクター増感処理によりカメラは教室右側に安定して向けられている.

4.5 カメラ位置補間計算

4.3 節の手法で決定した各キーフレーム f_i でのカメラ位置 p_i を元に, 各フレームにおけるカメラ位置 x_i を計算する. ここで, 収録映像の各フレームを h_1, h_2, \dots で表し, $f_i = h_{10i}$ であるとする. また, 以下の説明では 1 ゾーンの幅を 1 で表すものとする.

(1) カメラ目標位置の算出

フレーム h_i での目標カメラ位置 q_i を, 隣接キーフレームのカメラ位置を線形補間して算出する. $i = 10m + n$ ($0 \leq n < 9$) であるとする, $q_i = \frac{(10-n)p_m + np_{m+1}}{10}$ となる.

(2) カメラ移動量の計算

次に, 先頭フレーム h_1 におけるカメラの初期位置 x_1 をキーフレーム f_1 におけるカメラ位置 p_1 とし, 目標カメラ位置をなぞるように後続フレーム h_i のカメラ位置を定めていく. カメラ位置 x_{i-1} が目標位置 q_i より右 (左) にずれていたら, カメラ位置 x_i を目標位置に近くよう x_{i-1} より左 (右) に ϵ_i ゾーン分だけ動かす. ϵ_i の値はフレーム h_i における目標カメラ位置 q_i とフレー

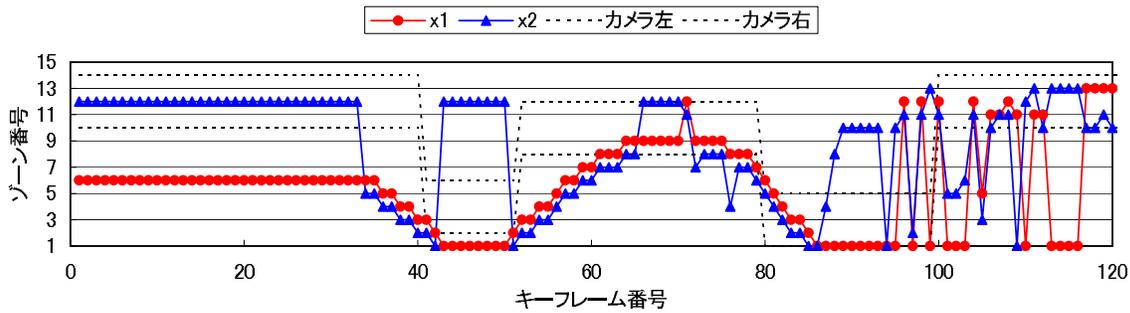


図 11: 仮想カメラワークの計算結果 (40 秒分)



図 12: 加工済み講義ビデオ

ム h_{i-1} におけるカメラ位置 x_{i-1} とのずれ $|q_i - x_{i-1}|$ に基づき、次式で計算される。

$$\epsilon_i = \begin{cases} 0.1 & (0 \leq |q_i - x_{i-1}| < 1) \\ 0.1 \times |q_i - x_{i-1}| & (1 \leq |q_i - x_{i-1}| < 3) \\ |q_i - x_{i-1}| & (|q_i - x_{i-1}| \geq 3) \end{cases}$$

基本的には目標カメラ位置とのずれに比例して移動量を大きくしているが、位置のずれが3ゾーン以上ある場合はカメラ位置を目標位置にジャンプさせることでカメラスイッチを行っている。

5 試行結果

2008 年度前期の複数の講義に対して、今回考案した手法を用いてビデオ撮影・加工実験を行った。ここでは著者自身の担当講義「データ構造とアルゴリズム第 12 回 クイックソート (2008/07/07)」での試行結果について述べる。収録は講義の直前に著者自身が図 3 に示すレイアウトで機材を設置し、マイクロ QR コード付きの講義スライドと板書を併用して講義を行った。

収録したフルハイビジョン映像に対して仮想カメラワークの計算・マイクロ QR コードの検出を行い、スライドタイトル付きの標準ビデオ映像に変換した結果を図 12 に示す。

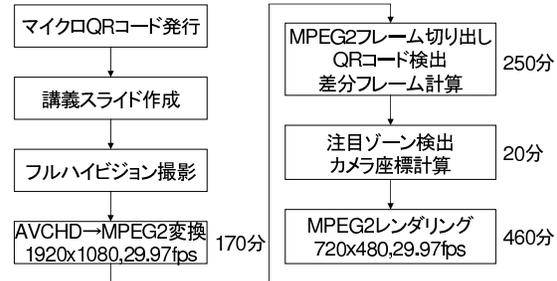


図 13: 全体の作業工程

5.1 作業工程

講義の撮影準備から加工済みビデオが出力されるまでの作業工程を図 13 に示す。AVCHD 形式はデコード処理に時間がかかるため、取り扱いやすい MPEG2 形式に変換してから処理を行うようにしている。AVCHD 形式から MPEG2 形式への変換・MPEG2 レンダリングには Windows XP の計算機 (Core2 Duo 2.2GHz, 3GBMem) を、その他の処理にはもう一台の Windows XP の計算機 (Xeon 3.6GHz, 2GBMem) を用いた。ビデオの加工には ffmpeg, HC encoder, AVISynth といった既存ツールの他に自作の画像加工ツールを組み合わせさせており、バッチ処理が可能のように全てコマンドラインから操作できるように実装している。

現在の簡易実装では 90 分のフルハイビジョン映像を加工するのに、約 15 時間かかった。最も計算時間のかかる最終工程の MPEG2 レンダリングでは、フレーム読み出し・トリミング・シャープ化・テロップ追加の処理を Java で行っており、エンコーダ・デコーダとのプロセス間通信がボトルネックになっている。この処理を C 言語で実装し高速化することが今後の課題である。

5.2 評価

加工後のビデオを撮影・編集スタッフ 3 名、教員 1 名に見てもらい、画質・カメラワーク・見やすさについてアンケートを行った。

Q1. 板書・スライドは支障なく読めますか？

A. 「十分に読める」… 3名, 「読める」… 1名

画質に関しては今回考案した収録方式で基本的に問題ないと思われる。ただし, 「小さい字は若干ぼやけている感じがする」「字の小さい先生は少し板書が見にくいかもしれない」との意見があった。

Q2. カメラワークは自然ですか？

A. 「自然」… 2名, 「やや不自然」… 1名, 「不自然」… 1名

教員がプロジェクタスクリーンの横で説明している場面で, カメラを動かさなくて良い場面でもカメラが動くことがあるのが不自然との指摘であった。この点はまだ改良の必要があると考えている。

Q3. 復習用ビデオとして使えますか？

A. 「十分に使える」… 4名

この結果から, 復習用講義ビデオの撮影を省力化するという当初の目標は達成できたと考えられる。「下手に人が撮るより見やすい」とのコメントもあった。その他に「カメラを設置した後は, 講義内容の記録に専念できる点が良い」「ビデオにコメント(タイトル)がついていて良い」とのコメントもあり, カメラワークに一部不自然さはあるものの概ね好意的な評価であった。

6 まとめ

本報告では, 固定フルハイビジョンカメラによる講義撮影と仮想カメラワークを組み合わせた講義ビデオ生成方式について検討を行った。この方式では録画の開始と終了の操作だけを行えばよいため, 撮影中にカメラを操作する人員が不要である。また, 撮影時にマイクロ QR コードをホワイトボード・プロジェクタスクリーンに提示することで素材管理・インデックス付与も自動的に行うことができる。

仮想カメラワークの計算はフレーム間差分に基づいて変化の大きい領域を求めることで行っている。この時, 過去・未来両方のフレーム情報を活用することで撮影時の突発的な外乱の影響を受けにくいようにしている。またプロジェクタスクリーンに優先的にカメラを向けるようにすることで, 板書だけでなく講義スライドを併用する講義にも対応している。

板書・スライドを併用する実際の講義において今回考案した方式でビデオの作成を行い, 講義ビデオの撮影・編集スタッフに対してアンケートを行った。その結果, 板書・スライドを読むのに支障のない画質であり, また, 復習用ビデオとして利用可能であるとの反応を得た。

本方式の最大のメリットは講義ビデオを収録したい教員が自分1人で手軽に収録から編集まで行えること

であり, より多くの教育機関での活用が期待される。

謝辞

マイクロ QR コードのエンコードルーチンを作成した新野哲平君(現 NEC システムテクノロジー), 講義ビデオ撮影・編集スタッフの皆さん, ならびに講義撮影スケジュールの調整にご尽力いただいた亀井清華先生(現広島大学), 天寄聡介先生(現岡山県立大学)にこの場を借りてお礼申し上げます。

参考文献

- [1] 永井孝幸: 鳥取環境大学次期授業支援システムの開発, 第 82 回 情報処理学会研究報告 2005-CE-82, pp.7-14 (2005-12)
- [2] Takayuki Nagai, "Laboursaving Video Archive System for Daily Manual Recording", 8th International Conference on Information Technology Based Higher Education and Training
- [3] 長瀧寛之, 永井孝幸, 都倉信樹, 「学生が作業スタッフとして参加する学科内全科目の講義撮影・配信への取り組み」, 社団法人日本工学教育協会論文誌工学教育 2007 年 9 月号, vol.55 no.5 pp.8-14
- [4] 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, 電子情報通信学会論文誌 D-II, vol.J88-D-II, no.8, pp.530-540, 2005
- [5] Y.Rui, L.He, A. Gupta and Q.Liu: "Building an Intelligent Camera Management System", Proceedings of the ninth ACM international conference on Multimedia, pp. 2 - 11, 2001.
- [6] 大西正輝, 泉正夫, 福永邦雄: "デジタルカメラワークを用いた自動映像生成", 画像の認識・理解シンポジウム (MIUR2000), vol.I, pp.331-336, 2000.
- [7] 横井隆雄, 藤吉弘亘: "高解像度映像からの自動講義ビデオ生成-仮想カメラワークの実現-" 第 11 回画像センシングシンポジウム, pp.73-76, 2005.
- [8] 横井隆雄, 桐井孝嘉, 藤吉弘亘: "講義イベント検出に基づく短縮講義ビデオの自動生成", 第 12 回画像センシングシンポジウム予稿集, pp.535-530, 2006.
- [9] M. Isard and A. Blake, "CONDENSATION - conditional density propagation for visual tracking", International Journal of Computer Vision, 1998.
- [10] Q.Liu, Y.Rui, A.Gupta and J.Cadiz: "Automating Camera Management for Lecture Room Environments", Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 442 - 449, 2001.