Vol.2017-MBL-82 No.16
Vol.2017-UBI-53 No.16
2017/3/9

複数受講者間の頭部動作・瞬きの同期に基づく講義コンテンツの重要シーン検出

笠波昌昭[1]　寺田 努[1,2]　塚本昌彦[1]

概要：近年，情報通信技術の向上と普及に伴い，e-ラーニングやオープンコースウェアのような，講義映像をインターネット上に配信・アーカイブする取組みが広く行われている．しかし，このような講義映像は録画時間が長時間に及ぶことが多く，視聴者や編集者にとって大きな負担となっている可能性が高い．そのような講義映像内での重要なシーンを自動で検出できれば，視聴者は重要なシーンのみを手短に見直すことができ，また編集者はより少ない労力で編集作業を行えると考えられる．そこで本研究では，複数受講者間の頭部動作等のデータから講義コンテンツ内の重要シーンを検出することを目的として，学生による進捗報告ミーティングおよび研究発表会にて複数受講者の頭部動作等を計測し，録画した動画と比較しながらデータの解析を行った．さらに，それを踏まえて複数受講者の頭部動作・瞬きの同期に基づく重要シーンの検出手法を提案し，人手によるダイジェストと比較する評価を行った．評価の結果，あるパラメータ設定において検出した重要シーンのうち 73.6%が人手によるダイジェスト化の際に採用されたシーンと合致した．

1. はじめに

長年の間，講義は講師が演台に立ち受講者がその前に机を並べて授業を受けるというスタイルで行われてきた．しかし，1990 年代に端を発する情報通信技術の向上と普及に伴い，e-ラーニングやオープンコースウェアのような，講義映像や講義資料をインターネット上に配信・アーカイブする取組みが広く行われるようになった．これにより，受講者は教室という物理的，時間的な制限に縛られることなく受講，復習でき，また講師も自身の授業方法や内容の改善 (FD:Faculty Development) に利用できる．

一方，このような実際の教室に出席しない遠隔講義にはいくつか問題点も存在する．

まず，講義コンテンツ編集の巧拙が遠隔受講者の学習効率に影響を与えうる点である．インターネット上で配信される講義映像にはデータ量や通信環境といった仕様上の制約があり，実際の教室で行われる講義において講師から伝えられる情報を必ずしも網羅できるわけではない．その上で，遠隔受講者にとってより良い学習効率を生むためには，編集者による適切な情報の取捨選択が求められる．

次に，講義コンテンツ編集作業に大きな負荷がかかる点である．実際の講義を録画録音したデータは 1 時間を超えることが多く，講義コンテンツの編集者にとって，長時間のデータの中から遠隔受講者にとって必要なシーンを選びぬく編集作業は大きな負担であると考えられる．

もし，講義コンテンツ内から重要なシーンを自動で検出できれば，上述の問題点を改善できると考えられる．そこで本研究では，講義中の複数受講者の頭部動作をセンシングし，そのデータから重要なシーンを検出する手法を提案，評価することを目的とする．このような動機の先行研究として，ヒェウハンらは講義講演動画からスライド中の単語の出現状況などを調べ，それらの情報に基づいて重要シーン抽出する手法を提案し評価している [1]．また，山根らは長時間に及ぶ講義アーカイブの視聴負荷軽減のために，講義映像内から講師と受講者の間のインタラクションを検出する手法を提案している [2]．しかし，複数の受講者にセンサを装着させ，頭部加速度，頭部の向き，瞬きなどを調査したものは筆者の知る限り無い．

本研究では「講義コンテンツ」を，「何らかの発表を行う講師と，それを聴講する複数の受講者が存在する場でもたらされる情報」と定義する．予備調査として学生による進捗報告ミーティングでの複数参加者の頭部動作をセンシングし，その結果を踏まえて，学生による研究発表会での複数受講者の頭部動作および瞬きを調査した．そして研究発表会において得られたデータから，講義コンテンツ内での

[1] 神戸大学大学院工学研究科
Graduate School of Engineering, Kobe University
[2] 科学技術振興機構さきがけ
PRESTO, Japan Science and Technology Agency

重要シーンに特有の複数受講者の動作を検出する手法を検討し，実際に人の手による講義コンテンツのダイジェストと比較を行った．

本稿は以下のように構成される．2 章で関連研究について述べ，3 章ではデータの収集手法と収集したデータについての考察を述べる．さらに 4 章では重要シーン検出手法を説明し，5 章では検出手法の評価を行い，最後に 6 章でまとめを行う．

2. 関連研究

講義コンテンツに関する研究として，篠木らは，黒板や講師を映した 1 台の高解像度カメラから，ディジタルカメラワークを用いて講義映像を自動で生成する手法を提案している [3]．この手法では，講師を追跡するカメラワークを基本としつつ，多くの受講者の注目を集める領域を映像に収めることを目指している．受講者の注目を集める領域の検出手法としては，実際に受講者に視線検出機器を装着させるのが理想としながらも，数などの問題から講義映像を PC で視聴して視聴者の注目点をポインタで示す手法を採っている．上田らは，学習者の視点に合わせた講義アーカイブ映像を作成するための講師・受講者の行動獲得について検討している [4]．この研究では，カメラ 6 台，超音波位置センサ 4 台，電子白板 1 台，後方 3 台のカメラに加えて，講師の胸ポケット・指示棒に取り付けたセンサで講師の動きと位置を観測，教室中央天井の魚眼カメラで受講者の位置や動きを観測している．カメラ 6 台の内，教室前方に受講者へ向けて設置された 3 つのカメラの映像から受講者の顔を認識する．そしてより多く受講者の顔が認識されたカメラの方向に受講者の興味を惹く情報があるとみなし，一定以上の受講者が向いている方向にある情報を講義映像に採用しアーカイブする．服部らは，教室に設置されたセンサによって得られた講師・受講者の観測データから自動的に時系列コンテキストを獲得するシステムを構築している [5]．このシステムでは，講師の行動としてスライド情報・板書情報・3 次元位置情報，受講者の行動として顔上げ動作を検出することで時系列コンテキストを取得している．大山らは，講義映像を視聴している受講者の瞳孔径を測定しスペクトル分析を行うことで講義の特徴を調査すると同時に，視線データを扱う上でのプライバシーといった倫理的問題についても検討を行っている [6]．大西らは，会議参加者の頭部に加速度・角速度センサを装着し，発話動作，うなずき，首かしげ等の動作を認識し，会議映像に自動でタグ付けを行うシステムを提案している [7]．

これらの先行研究より，受講者の顔を上げる動作や注目対象の情報を用いて講義コンテンツ内から重要な情報が推定出来る可能性が高い．さらに受講者にセンサを直接装着することで，顔上げ動作に限らない，より詳細な受講者動作を検出できると考えられる．

図 1　ミーティング俯瞰画像

図 2　ミーティング参加者

3. データ収集

講義コンテンツから重要シーンを検出する手法を提案するにあたり，発表者の発表を複数の受講者が聞くという状況において，複数の受講者がどのような頭部動作を行うのか調査するために，学生による進捗報告ミーティングおよび研究発表会にてデータ収集を行い，得られたデータに対して考察を行った．

3.1 進捗報告ミーティングでのデータ収集

研究室内での進捗報告ミーティングにおいて，教員および複数学生の頭部の動きのデータを収集し，同時にその様子を映像に記録した．このデータ収集では，ミーティングの内容に応じて行われる，頷く，身振りを行う，発言者を見るといった参加者の行動が参加者頭部の加速度・角速度データにどのような特徴を与えるか検証することを目的とした．

データ収集環境

ミーティングの様子を俯瞰した画像を図 1 に示す．教員を含めた各進捗報告ミーティング参加者は，頭部の動きを取得するために，加速度・角速度センサを右耳側に固定したヘッドホンを装着した．ヘッドホンを装着した参加者の様子を図 2 に示す．これらの加速度・角速度センサは全てデータ収集用の PC に接続されており，ミーティングと同時にデータの収集を行った．さらに映像データとして，各参加者の PC のフロントカメラを用いて各自の頭部の動きを収めた映像，ビデオカメラを用いてミーティング全体をとらえた俯瞰映像を録画した．このとき，PC フロントカメラ映像にはその PC の現在時刻を焼き込み，また俯瞰映像にはテーブルに置いたタブレット端末の時刻を映し込んでおくことで，各センサデータとの時刻同期の助けとした．確実な映像とセンサの時刻同期を図るために，各参加者はデータ収集の開始直後に大きく数回頷くという特徴的な動作を行った．ミーティングでは参加学生が 1 人ずつ順番に進捗報告を行い，報告中の学生は教員と一対一で会話するという形をとる．

Vol.2017-MBL-82 No.16
Vol.2017-UBI-53 No.16
2017/3/9

図 3　データビューワアプリケーション

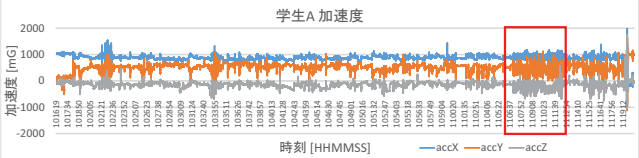

図 4　学生 A の加速度

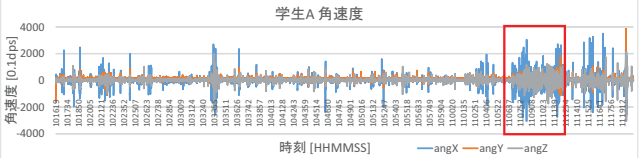

図 5　学生 A の角速度

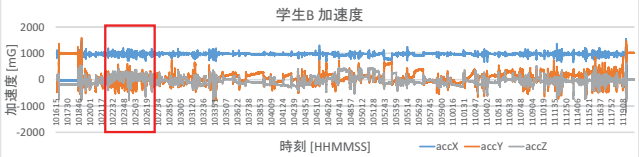

図 6　学生 B の加速度

データビューワアプリケーション

次に，取得した各種映像データとセンサデータを比較するためのアプリケーションを作成した．アプリケーションの概観を図 3 に示す．本アプリケーションは左側に動画表示部，右側にグラフ表示部をもち，再生/停止ボタンを押すことで動画およびグラフを同時に再生/停止でき，グラフ表示部下のシークバーを操作することで再生位置を調整できる．さらに，記録開始時刻にずれがある映像データとセンサデータを同期するための機能を実装した．この機能は，映像とグラフの同期がずれている場合，再生を一時停止したのち，動画をセンサデータと同期している位置までコマ送りすることで同期位置を調整できる．この映像とセンサデータの同期の情報は XML ファイルに保存され，次回からはこの XML ファイルを読み込むことで同期が取れた状態で両データを比較できる．

結果と考察

データ収集の結果として，ミーティングに参加した学生 6 名のうち 2 名の加速度・角速度データを図 4〜図 7 に示す．図中の赤枠は，当該学生が進捗報告を行っていた時間帯を表す．学生 A，学生 B ともに報告中はそうでない時に比べ頭部が頻繁に動いた事がわかる．特に，自身の PC 画面に表示された資料と教員の顔を交互に見るため水平方向 (角速度 X 軸) の頭部の動きが大きく，かつ多くなった．10 時 33 分 50 秒付近で各参加者で頭部の動きが同期している箇所があった．動画で確認したところ，話題が部屋の隅に居たミーティングに参加していない第 3 者へ逸れ，皆がその方向を向いたためであるとわかった．大きな頭部の動きが検出されたタイミングにもかかわらず，報告中ではなく，かつ他の参加者と動きが同期していないタイミングも幾つか見つかった．動画を確認してみたところ，進捗報告者や教員の方向を向くといったミーティングに直接関係のある動きもあったが，椅子に座り直す，首を倒してストレッチするなど，ミーティングの内容に全く関係のない動きも散見された．

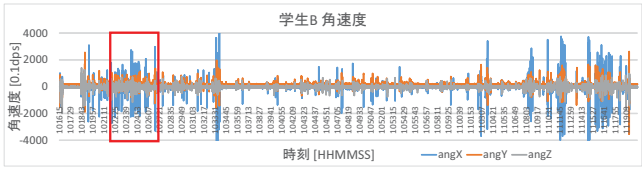

図 7　学生 B の角速度

これらから，ミーティングにおいて何らかの情報発信/受信を行っている参加者は頭部の動きがそうでないときに比べ頻繁に発生することがわかった．さらに，複数の参加者で頭部の動きが同期しているタイミングでは，多くの参加者の興味を引く何らかの出来事が発生しており，それが講義コンテンツにおいて重要なシーンになる可能性が高い．また，頭部の加速度・角速度センサのデータだけでは，上の空状態での何気ない仕草や椅子に座り直すなどのミーティングに関係ない動作も，ミーティング内容への反応と誤認識されうるので，その行動がミーティング内容に応じたものか推定するためには，参加者間の同期をみる，他のセンサも同時に用いるなどの対策を行う必要があると考えられる．

今回のデータ収集において，各種データの時刻同期に参加者の PC の内部時刻を用いたが，後に PC ごとにこの時刻のズレが存在すると判明した．複数参加者の動作の同期が重要な意味を含みうることから，多くのデータを収集する際には確実に同期の取れたタイムスタンプをデータに付与できるような仕組みが求められる．

3.2　学生発表会でのデータ収集

学生による研究発表会において，スライドを用いた発表を受講する際の複数受講者の頭部動作および瞬きのデータを収集した．このデータ収集では，進捗報告ミーティングでのデータ収集における考察を踏まえて，計測する複数受

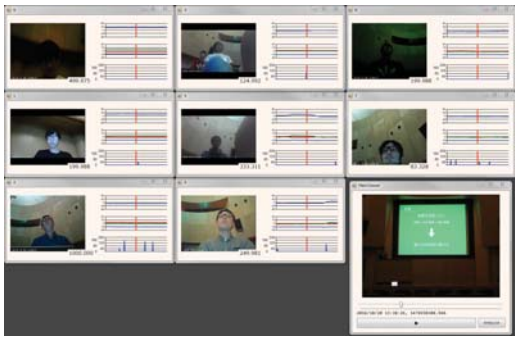

図 8　データビューワアプリケーション 2

講者の動作データの種類を増やすと同時に受講者間の動作の同期を確認し，重要シーン検出手法を検討することを目的とする．

データ収集環境

本研究発表会は神戸大学百年記念館六甲ホールにて行われた．研究発表会は，発表者 1 人あたり発表 15 分質疑 10 分の計 25 分が与えられ，各発表者はステージ上の演台に立ち，動画コンテンツを含む PowerPoint スライドを用いて研究発表を行った．参加した学生のうち 9 人に，センシングデバイスとして 3 軸加速度・頭部姿勢・瞬きなどが計測できる JIN 社のアイウェア，JINS MEME[8] を装着させた．この JINS MEME はそれぞれ対となる iPod touch と接続されており，データはそこに逐次蓄えられる．加えて，JINS MEME から得られたデータと実際の受講者の行動を比較するために，各受講者の受講中の様子を膝上に置いた PC のインナーカメラで録画した．以降この動画を個人動画と呼ぶ．さらに，発表内容と講師の様子を記録するために会場中央にビデオカメラを設置し，発表中の学生とスクリーンを録画した．以降この動画を前方動画と呼ぶ．進捗報告ミーティングでのデータ収集と同様，今回も動画とデータの同期のために，各種動画に時刻を焼き込み，受講者はデータ収集開始直後に数回頷く動作を行った．

データビューワアプリケーション 2

実験後，センシングした全受講者のデータのグラフと個人動画，前方動画を同時に再生するアプリケーションを用いてデータの解析を行った．アプリケーションの概観を図 8 に示す．本アプリケーションは前節のデータビューワアプリケーションと同様に個人動画とセンサデータを時刻同期でき，複数受講者の個人動画・センサデータと前方動画を同時に再生できる．

結果と考察

センシングした 9 人の受講者のうち 1 人が個人動画の撮影に失敗したため，残りの 8 人分のデータを用いて，どのようなシーンに動作等が同期していたのか解析した．4 人

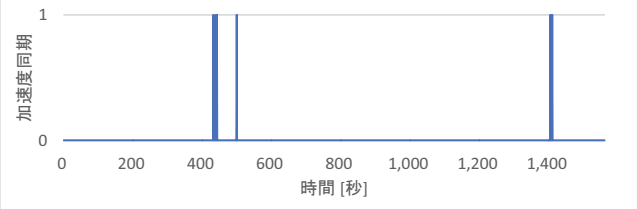

図 9　加速度合成値の増加が同期したタイミング

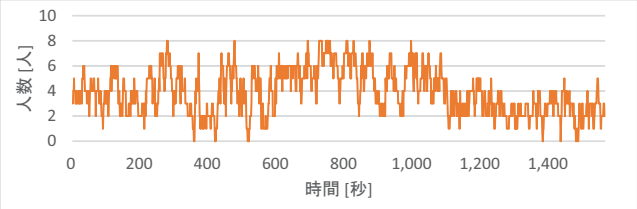

図 10　瞬き頻度が増加した人数

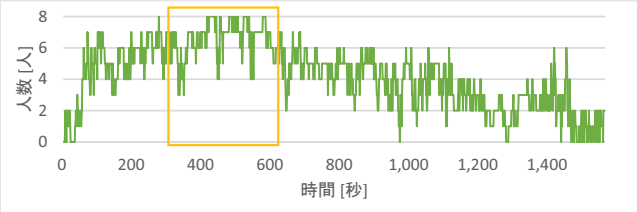

図 11　頭部が 5 [deg] を超えて上を向いている人数

の発表者のうち，ある 1 人が発表中の受講者 8 人のデータを解析したグラフを図 9 から図 11 に示す．それぞれのグラフについて説明する．図 9 のグラフは，1 秒間隔で区切ったウィンドウ内での各受講者の加速度合成値の標準偏差を計算し，その値が閾値を超えた受講者の人数が 7 人を上回ったタイミングを表す．図 10 は，10 秒間隔で区切ったウィンドウ内での瞬き頻度を各受講者について計算し，その値が発表会全体での瞬き頻度の平均を超えていた受講者の人数を表す．図 11 は，水平を 0 [deg] とし，頭部が 5 [deg] を超えて上を向いている受講者の人数を表す．図中の黄枠は，スクリーンに動画が表示されていた時間帯を表す．

図 9 について，加速度標準偏差が閾値を超えて増加したタイミングが同期した際の個人動画，前方動画を確認したところ，その多くの場面で会場内で笑いが巻き起こったり，拍手が行われていたことがわかった．これは，複数の受講者の笑い，拍手による比較的大きな振動を同時に検知したためである．この特徴は，あとで長時間の映像の中で盛り上がった場面を見返す際などに有用であり，重要シーン検出においても大きな意味を持つと考えられる．図 10 および図 11 について，多くの受講者が上を向いており，かつ瞬き頻度が平均より低下した時間帯において，個人動画・前方動画を確認したところ，その多くの場面でスクリーンに動画，画像のような注目性の高いコンテンツが表示されており，受講者はそれを注視していたことがわかった．これは，スクリーン上に複数の受講者にとって重要な価値を持

つ情報が存在していることを意味しており，遠隔受講者，講義映像編集者にとっても価値を持つと考えられる．反対に，多くの受講者が下を向いている時間帯では，発表者交代のための準備中である，発表者がスライドの文章を読み上げる，スライドとは無関係な質疑が行われているなど，受講者にとってスライドの重要性が薄まっていたことがわかった．

4. 検出手法

研究発表会にて取得したデータより，重要シーンを検出する手法について提案する．3 章の考察より，本手法では以下の条件を満たすシーンを抽出する．

- 3 軸加速度合成値の標準偏差が α を超えた人数が n_a 人以上となる．
- 頭部の角度が θ を超えた人数が n_{b1} 人以上となり，かつ瞬き頻度が発表会全体の平均頻度を下回った人数が n_{b2} 人以下となる．
- 頭部の角度が θ を超えた人数が n_c 人以上となる．

(ただし，頭部角度について水平を 0[deg] とし，上向きを正，下向きを負とする．)

加えて，重要シーンとして検出されたタイミングから 10 秒以内に再び重要シーンを検出した場合，その間の区間も重要なシーンが連続していると仮定し重要シーンとする．

5. 評価

4 章で検討した重要シーン検出手法について評価を行った．本評価では，人手による研究発表会映像のダイジェスト化の際に採用されたシーンを重要シーンの正解データとし，提案手法で検出された重要シーンと比較するという形で行った．

評価に先立ち，発表会映像の人手によるダイジェスト化を行った．作成したダイジェストは，およそ 100 分 (発表・質疑 25 分 ×4 人) の研究発表会映像 (前方動画) を，20 分 (1 人あたり 5 分 ×4 人) に縮小したもので，大学院修士課程の学生 3 人が行った．この作業では動画データや音声データに注釈をつけるソフトウェア ELAN[9] を用いた．ダイジェストで採用されたシーンを図 13 に示す．図 13 において，横軸は動画開始からの秒数を表し，縦軸は当該秒数のシーンがダイジェストを作成した 3 人のうち何人に採用されたかを表す．動画全体 6312 秒のうち，少なくとも 1 人以上に採用されたシーンは合計 2298 秒 (36.4%)，全員に採用されたシーンは合計 313 秒 (5.0%) であった．

さらに，提案手法の各種パラメータを調整しながら，人手によるダイジェスト採用シーンと比較できるアプリケーションを作成した．アプリケーションの概観を図 12 に示す．

検出結果として，各種パラメータを表 1 に設定した場合のものを図 14 に示す．これをダイジェストデータと比較

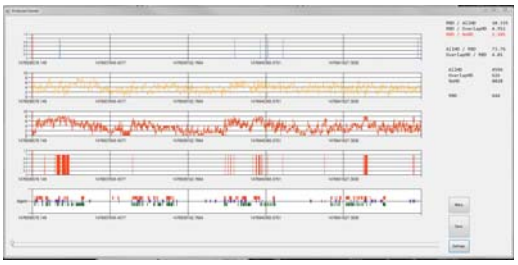

図 12 人手ダイジェストと比較できるアプリケーション

表 1 各種パラメータ設定

パラメータ	値
α	0.245 [m/s^2]
n_a	7 [人]
θ	2.0 [deg]
n_{b1}	7 [人]
n_{b2}	2 [人]
n_c	8 [人]

したところ，提案手法によって検出された全シーンのうち 73.6%のシーンが人手によるダイジェストに採用されたシーンであることがわかった．以下，この割合を正解率と呼ぶ．しかし，検出されたシーンは計 323 秒と短い．そのため，ダイジェストに採用されたシーンのうち，提案手法によって検出されたシーンの割合は，少なくとも 1 人に採用されたシーン全体の 10.3%，3 人全員に採用されたシーン全体の 5.0%にとどまった．以下，これらの割合をそれぞれ，被覆率，全員採用シーンにおける被覆率と呼ぶ．

より検出シーンが増加した際の評価を行うために，表 1 のパラメータのうち α および θ を，検出シーン長が 2 倍を超えるまで変化させた際の各種割合について調査した．パラメータ α を 0.061 [m/s^2] まで低下させたところ，検出シーンは 728 秒まで増加した．正解率は 52.3%まで低下し，被覆率は 16%，全員採用シーンにおける被覆率は 11.5%に増加した．パラメータ θ を-4.0 [deg] まで低下させたところ，検出シーンが 664 秒と 2 倍以上に伸びた．この時，正解率は 64.9%まで低下したものの，被覆率は 18.5%，全員採用シーンにおける被覆率は 7.3%に増加した．両パラメータともに，値を下げた場合，正解率が低下し，被覆率・全員採用シーンにおける被覆率は増加した．これは，パラメータの値を下げることで検出シーンが増加し，被覆率・全員採用シーンにおける被覆率が増加した一方で，ダイジェストで採用されなかったシーンを重要と検出する回数も増えたためである．この評価より，α の値を下げた場合に比べ，θ の値を下げた場合のほうが正解率，被覆率の面から良いといえる．しかし，全員採用シーンにおける被覆率については α の値を下げた場合のほうが高い．この理由として，θ の値を下げることで，笑い等が発生したシーンを θ を下げたときより多く検出でき，3 人がダイジェストに採用したシーンに合致するシーンが増えたものの，講義コンテン

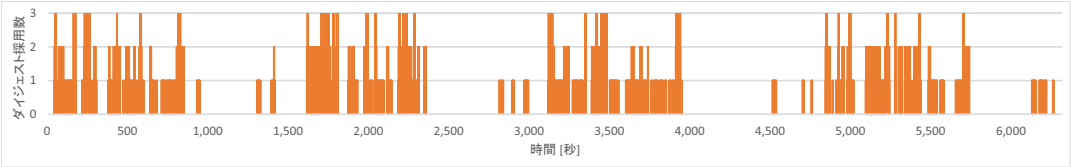

図 13 ダイジェスト採用シーン

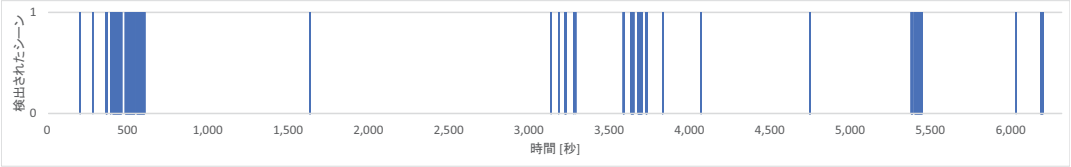

図 14 検出シーン

ツのダイジェストにおいて笑い等が発生したシーンが必ずしも採用されるとは限らないためと考えられる．

また，検出シーンが増加した際に正解率が下がる理由として，提案手法と人手によるダイジェストとのシーン採用基準の差異が挙げられる．人手によるダイジェストでは，短い時間で発表のストーリーを伝えるために，重要なスライドが表示されたタイミングを採用シーンにした箇所が多く見られた．対して，提案した検出手法では，主に受講者がスライドを見ているかどうかを基準にして，重要なシーンか否かの判断を行っている．そのため，文字が多く読み上げるだけ，または内容が少なくすぐに切り替わるスライドなどが，発表において重要な意味をもっていても，複数受講者の注目が集まらないために重要シーンと判定されないことがあった．さらに，スライドに画像のような注目性の高いコンテンツが存在するにもかかわらず，発表者が口頭で受講者の視線を誘導しなかったために注目が集まらないこともあった．

これらの考察より，本提案手法が検出する重要シーンは，発表の流れを重視するようなダイジェストの生成に用いるよりも，発表への受講者の反応を定量的に評価する，例えば遠隔受講者に対してコンテンツの注目度を提示したり，講師自身の発表への振り返りなどに用いるほうが有用性が高いと考えられる．

6. まとめ

本研究では，複数受講者間の頭部動作等のデータから講義コンテンツ内の重要シーンを検出することを目的として，学生による進捗報告ミーティングおよび研究発表会にて複数受講者の頭部動作等を計測し，各種動画と比較しながらデータの解析を行った．さらに，それを踏まえて複数受講者の頭部動作・瞬きの同期に基づく重要シーンの検出手法を提案し，人手によるダイジェストと比較する評価を行った．

評価の結果，あるパラメータ設定において提案手法によって検出した重要シーンのうち 73.6%が人手によるダイジェスト化の際に採用されたシーンと合致した．しかし，検出シーンは 323 秒と短く，検出シーンを長くしようと各種パラメータを調整すると，被覆率などは向上するものの正解率が低下することがわかった．

今後は，機械学習を用いた重要シーン検出を提案し評価すること，実際の授業でデータ収集を行うこと，講義コンテンツ以外の分野へ応用することを予定している．

謝辞 本研究の一部は，科学技術振興機構戦略的創造研究推進事業 (さきがけ) および文部科学省科学研究費補助金挑戦的萌芽研究 (25540084) によるものである．ここに記して謝意を表す．

参考文献

[1] レー ヒェウハン, ティティポーン ルートラットデーチャークン, 渡部徹太郎, 横田治夫: 講義講演ビデオからダイジェスト自動作成のための重要シーン抽出手法の評価, 第 19 回電子情報通信学会データ工学ワークショップ (DEWS2008) 論文集, pp. E4-1 (2008).
[2] 山根卓也, 中村和晃, 上田真由美, 椋木雅之, 美濃導彦: 講義中の行動分析に基づく講師受講者間インタラクションの検出, 先進的学習科学と工学研究会, Vol. 60, pp. 7–14 (2010).
[3] 篠木雄大, 藤吉弘亘: 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, 映像情報メディア学会誌, Vol. 62, No. 2, pp. 240–246 (2008).
[4] 上田真由美, 服部博憲, 森村吉貴, 丸谷宜史, 角所 考, 美濃導彦: 学習者の視点にあわせた講義アーカイブ作成のための講義室内インタラクション獲得に関する検討, 先進的学習科学と工学研究会, Vol. 54, pp. 13–18 (Nov. 2008).
[5] 服部博憲, 正司哲朗, 丸谷宜史, 森村吉貴, 角所 考, 美濃導彦: 講義時における講師・受講者の行動に基づく時系列コンテキストの獲得, 電子情報通信学会総合大会講演論文集 2008 年情報・システム (1), p. 200 (Mar. 2008).
[6] 大山貴紀, 金子 格, 小野文孝, 曽根順治, 花村 剛: 瞳孔径による授業評価, 情報科学技術フォーラム講演論文集, Vol. 30, No. 3, pp. 781–782 (Sep. 2011).
[7] 大西鮎美, 村尾和哉, 寺田 努, 塚本昌彦: 装着型センサを用いた会議ログの構造化システム, マルチメディア，分散，協調とモバイル (DICOMO2014) 論文集, Vol. 2014, pp. 1860–1868 (Jul. 2014).
[8] JINS MEME: https://jins-meme.com/ja/.
[9] ELAN: https://tla.mpi.nl/tools/tla-tools/elan/.