

赤外線タグ付きマイクロホンを用いた講義における発言者の追跡撮影

川野 晃寛[†] 吉高 淳夫[†] 平嶋 宗[†]

[†]広島大学大学院工学研究科
〒739-8527 広島県東広島市鏡山 1-4-1

E-mail: [†]{kawano, yoshi, tsukasa}@isl.hiroshima-u.ac.jp

概要 大学などにおいて講義をアーカイブ化して活用する動きが盛んになっている。講義のアーカイブ化のためには講師と受講者が発する情報を映像として記録することが重要である。講師は多くの場合一人であり、行動範囲が限られているため撮影が容易である。しかし、受講者に関しては不特定位置に複数存在するため、撮影対象を特定し撮影することはより困難である。そこで、本研究では位置が不特定な発言中の受講者をミドルショットで撮影する手法を提案する。提案手法では、受講者はマイクロホンを持って発言を行うという条件の下で、マイクロホンに取り付けた赤外線タグの位置を検出し追跡することで発言中の受講者を撮影する。

The person track shooting system using the microphone with the IR-tag in lectures

Akihiro KAWANO[†] Atsuo YOSHITAKA[†] and Tsukasa HIRASHIMA[†]

[†] Graduate School of Engineering, Hiroshima University
1-4-1 Kagamiyama, Higashi-Hiroshima, Hiroshima, 739-8527 Japan

E-mail: [†]{kawano, yoshi, tsukasa}@isl.hiroshima-u.ac.jp

Abstract There are many cases where lectures are recorded to build video archive for improving reusability. It is important to record utterance by the lecturer or students. In many cases, a lecture is given by a single lecturer and the possible area of his/her movement is relatively limited. Therefore, it is easy to shoot the lecturer. However, shooting one of the students who is asking a question is more difficult problem since the position of the student is located at an arbitrary position in the wider area in the classroom. This paper focuses on the issue of shooting a student in utterance by the middle shot size. We propose a novel method of shooting a student in utterance using the IR-tag installed on the microphone. By detecting the IR-tag, a student in utterance is kept tracked.

1. はじめに

大学などにおいて講義の内容を撮影し、Web上に公開する講義映像のアーカイブ化が盛んになってきている。講義を記録する際に必要となる情報としては「講師が発する情報」、「受講者が発する情報」、また電子スライドや板書内容などの「教材」の3つが挙げられる。電子スライドは元々電子データとして記録されており、また、板書内容は電子ホワイトボードなど板書内容を電子データとして記録するものがある。このように、電子スライドや板書内容などの教材に関しては撮影して記録する必要は必ずしもないと考えられる。ここでは、講師と受講者が発する情報を映像として記録することに着目する。

講義を撮影するための方法の一つとして、固定カメラを用いた方法が挙げられる。しかし、固定カメラでは撮影する範囲が固定されており、広範囲を撮影するためには人物の写る大きさであるショットサイズが小さくなる。そのため、講師や受講者の表情などのノンバーバル情報が伝わらず、講義の様子を捉えるには不十分な映像となってしまう。講師や受講者を表情が伝わる程度のショットサイズであるミドルショット（人物の上半身が写っている程度の大きさ）で撮影することで、講師や受講者の様子を詳細に捉えることのできる質の高い映像を撮影することができる。ミドルショットで撮影する方法として、撮影者を配置して撮影することが挙げられる。撮影者を配置して撮影すること

によって、講義中の講師の移動や受講者の発言など、講義の状況に合わせて撮影対象を切り替えながら、適切なズームを行うことでミドルショットで撮影することができる。撮影される映像の質を考えると、撮影者を配置して撮影することが理想的であると考えられるが、大学などで毎日行われている多くの講義を撮影者を配置して撮影するのは、人的コストがかかってしまうため現実的でない。そこで、人的コストを抑えながらも、講義の様子が十分に伝わるような質の高い映像を撮影するため、講義の撮影を自動化する研究が多く行われている。

講師の撮影を対象とした研究として、大西ら[1]は固定カメラ3台を用いて講師と板書内容の自動撮影を行っている。撮影した映像から講師の行動を検出した上で、適切なアングルの映像にスイッチングすることで講義の撮影を行っている。このように、講師の撮影においては、撮影範囲が限られるため固定カメラによって撮影範囲をカバーすることが可能である。また、西口ら[2]の研究では講師に超音波ビーコンを装着することで、講師の位置を検出しパンチルトカメラによって追跡撮影を行っている。講師は一人であることがほとんどであり、センサを取り付けるといったことも可能である。これらのことより、講師の撮影は容易に行うことができると言える。

しかし、受講者の撮影は講師の撮影に比べて困難であることが言える。なぜなら、受講者は講義室に複数存在するため、受講者全員にビーコン等を取り付けるといったことは現実的でない。また、固定カメラで受講者全体を撮影した場合、撮影範囲が広いためミドルショットで撮影することができない。テレビ会議において発言者の拡大機能を付けた場合と付けない場合とで映像の印象評価を行った調査[3]では、拡大機能を付けた方が発言者の表情が伝わり、映像としての評価が高いという結果が得られている。このように、受講者の撮影を行う際には、複数の受講者の中から撮影対象を特定し、適切なズーム操作を行うことで表情などのノンバーバル情報も記録することが重要であると考えられ、それを十分に捉えることができるミドルショットで撮影する必要があると言える。

受講者の撮影を対象とする手法の一つとして、受講者を広い範囲でカメラにおさめて撮影するものがある[4]。しかし、受講者の中から撮影対象である発言者を特定し、ミドルショットで撮影することは行っていない。そのため、受講者の表情などが捉えられる質の高い映像とはなっていない。西口ら[2]は発言中の受講者の撮影を目的としており、講義室の天井に複数のマイクロホンを設置し、マイクロホンに到達する音声の位相差を求めることで、発言中の受講者の位置を検出し

ている。発言者の正確な位置を検出できるため、発言中の受講者をミドルショットで撮影することができる。その反面、マイクロホンを設置した講義室でないと利用できないといった欠点がある。さらに、個々のマイクロホンには感度差があるため、それらの誤差を調整する作業が必要である。また、設置方法や経年変化の影響も受けるため安定した性能を得るには定期的な調整が必要である[5]。このようなマイクロホンの特性上、マイクロホンを付け替えて他の場所で使用することは容易ではないと言える。

しかし、講義の撮影を行うシステムとしては、複数存在する講義室の中から講義内容や受講者の人数などに合わせて場所の選択をする必要があり、任意の場所に設置し利用できることが好ましい。本稿では、不特定多数の受講者の中から発言中の受講者の位置を検出し、従来手法の問題点である設置コストを抑えつつ、発言者の表情を捉えることのできるミドルショットでの質の高い映像を撮影する手法を提案する。

本稿では、まず2章で提案手法について述べ、3章でシステム構成について説明する。4章では、実験により求めた赤外線タグの動きの特徴について述べ、5章で赤外線タグの動きの特徴に基づいた追跡手法について述べる。最後に、6章でまとめと今後の課題について述べる。

2. 提案手法

本稿では、講義の撮影における発言中の受講者の撮影に着目する。提案手法では、講義中において受講者はマイクロホンを持って発言するという状況を想定する。マイクロホンには赤外線タグを取り付け、赤外線タグの位置を検出することで発言中の受講者の位置を特定し撮影を行う。発言者が常にマイクロホンを持っているという条件については、広い講義室では他の受講者に音声を送達するためにマイクロホンが使用されることも多く、また、講義の記録を目的と考えた際に音声を取得するためにマイクロホンを持って発言してもらうのは自然な前提条件であると考えている。提案手法では、マイクロホンに取り付けた赤外線タグの位置をカメラの映像から検出し、マイクロホンを持って発言を行っている受講者の撮影を行う。人物を追跡する方法としては、カメラの映像による背景差分やオプティカルフローを用いた手法[6]がいくつか提案されてきているが、いずれの手法も複数存在する受講者の中から発言者を特定することは困難であると考えられる。提案手法では、発言時には赤外線タグの付いたマイクロホンを用いるという制約を設けることによって、複数存在する受講者の中から発言者を特定することができる。

また、赤外線タグの位置を検出することでカメラを発言者の方に向けることは可能であるが、カメラと発言者との距離がわからなければミドルショットで発言者を撮影することはできない。そこで、カメラパラメータの一つであるフォーカス情報による被写体までの測定距離に基づいたズーム操作を行い、ミドルショットで発言者を撮影する手法を提案する。カメラはオートフォーカス機能が付いているものを用いる。その際に、フォーカス情報としてフォーカス値というものが得られる。フォーカス値とは、ピントを合わせ際に焦点距離に応じて変化する値であり、カメラから被写体までの距離が短いとフォーカス値は大きい値となり、距離が長いとフォーカス値は小さい値を返す。このフォーカス値によりカメラから被写体までの距離を換算し、距離に応じたズームを行うことでミドルショットで発言者を撮影する。

2.1 提案手法と従来手法との比較

マイクロホンによる音声位相差を用いた従来手法では、正確に発言者の位置を検出し、ミドルショットで質の高い映像の撮影が可能である。しかし、マイクロホンを正確な位置に設置する必要がある、任意の場所に容易に付け替えて使用することはできず、設置コストの面に問題があると言える。本手法では、赤外線タグ付きマイクロホンによって発言者の位置を検出し、フォーカス情報によるズームを行うことで発言者を従来手法と同等なミドルショットで撮影する。正確な設置や調整などが不要であるため設置コストが抑えられ、任意の場所に持ち運んで使用することが可能である。

また、音源による位置検出では、発言を開始した時点から発言者の位置の検出を行うため、発言を開始した直後の発言者の映像の取得が困難であるといった問題や発言者の音声を拡声するためにマイクロホンを使用するような状況では発言者の位置検出ができないといった問題が挙げられる。本手法では、発言を行う前の状態から赤外線タグの追跡を行っているため、発言を開始する時点から発言者を撮影することが可能であり、音声をマイクロホンで拡声する状況でも利用可能である。一方、赤外線タグを用いた場合に起こる問題としては、赤外線タグの以外の光として太陽光や蛍光灯の光にも赤外領域の光が含まれ、誤検出の原因となることが挙げられる。本手法では、蛍光灯が点灯した状態であっても、赤外線タグを一定の周期で点滅させることで蛍光灯の光と赤外線タグの光を識別する処理を行う。また、太陽光に関しては太陽光の光が赤外線タグの光よりも強く、点滅周期などの方法で識別することが困難である。そこで、講義室のカーテンを閉めることで太陽光を遮断する。蛍光灯を点けているため、

講義室は明るい状態でありカーテンを閉めることは講義を行うのに影響がないと考えている。

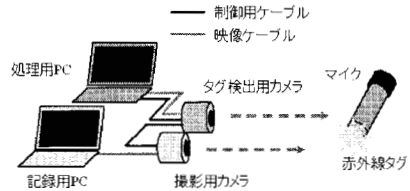


図1 システム構成

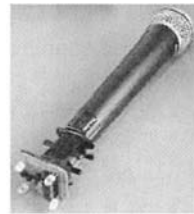


図2 赤外線タグ付きマイクロホン

3. システム構成

システム構成を図1に示す。本システムは赤外線タグ付きマイクロホンとそれを追跡するタグ検出用カメラ、タグ検出用カメラと同期して講義の様子を撮影する撮影用カメラ、カメラを制御する処理用PCと映像を記録するための記録用PCで構成されている。カメラは水平、垂直方向に旋回動作とズームの制御が可能なパンチルトズームカメラを使用している。処理用PCはタグ検出用カメラからの映像を入力とし、赤外線タグの位置を検出しタグ検出用カメラを赤外線タグの方に向ける制御を行っている。この際、タグ検出用カメラと撮影用カメラを同位置に設置し動きを同期させることで、撮影用カメラも同時に赤外線タグの方に向ける。記録用PCには撮影カメラからの映像を入力し、撮影を行っている映像を記録する。

また、作成した赤外線タグ付きマイクロホンを図2に示す。赤外線タグには4個の赤外LEDを配置しPICを用いた回路で点滅制御しており、電源にはリチウム電池2本を使用している。LEDが0.1秒ごとに点灯と消灯を繰り返すようにPICをプログラムしている。それをマイクロホンの先端に取り付けている。以下の節では、赤外線タグの位置検出手法とフォーカス情報によるズーム撮影手法について述べる。

処理に使用したPCは、CPUがPentium IV、2.66GHzのノート型PCである。タグ検出用カメラからの入力映像のフレームサイズは320×240[pixel]である。カメラはキヤノン製のVC-C50iを使用した。

3.1 赤外線タグの位置検出手法

赤外線タグの検出手法について述べる。位置検出にはタグ検出用カメラから入力された映像を用いて、点滅している赤外線タグを検出する。タグ検出用カメラには赤外線タグの波長に合わせた可視光カットフィルムを取り付けており、可視領域を除去するようにしている。タグ検出用カメラからの赤外領域のみの画像を図3に示す。タグ検出用カメラから入力された映像に対し、二値化を行い赤外線タグの光のみを抽出する。二値化後の画像を図4(a)に示す。図4(a)の上の方を見ると蛍光灯の光が残っているのがわかる。蛍光灯の光には可視領域の光だけでなく赤外領域の光も含まれており、蛍光灯の光が赤外線タグ検出の際のノイズとなる。このように、蛍光灯等の光を赤外線タグの光と誤認識しないように、赤外線タグの光と蛍光灯の光を判別する必要がある。そこで、赤外線タグのLEDが0.1秒ごとに点灯と消灯を繰り返すようにPICをプログラムし、点滅を識別することで赤外線タグの光と蛍光灯の光の判別を行う。0.1秒ごとに点灯と消灯を繰り返す赤外線タグの発光を検出するために0.1秒ごとにタグ検出用カメラから入力された映像を処理し、0.3秒間で点灯、消灯、点灯、もしくは消灯、消灯を繰り返している画素を赤外線タグの発光と見なして赤外線タグの位置と定義している。0.1秒ごとの点滅を識別する処理を行った際の映像を図4(b)に示す。点滅の識別を行うことで、赤外線タグの光のみを検出することが可能である。



図3 赤外線カメラの映像



図4 赤外線カメラの映像の二値化画像(a)と点滅を識別し赤外線タグの位置を検出した画像(b)

3.2 フォーカス情報によるズーム操作

カメラが発言中の受講者を撮影しているとき、表情

を捉えられるようなミドルショットでの撮影をするために、適切なズーム操作を行う必要がある。本手法では、ズーム操作を行うために、リアルタイムで処理が可能で処理の負担が少ないフォーカス情報によるズーム操作を行う。ここで、フォーカス情報とはピントを合わす際の焦点距離に応じて決まる値である。被写体が近いときフォーカス値は大きくなり、遠いときフォーカス値は小さくなる。このフォーカス値を距離に換算することでズームパラメータを適切に制御する。本システムで用いたカメラのフォーカス値と被写体までの距離との関係を実験により求め、フォーカス値から適切なズームパラメータを決定するための変換テーブルを作成し、それに基づきズームパラメータを決定した。フォーカス情報を正確に取得した際には、確実にミドルショットで撮影できることを確認した。ただし、フォーカス値から距離を求めるには被写体に対してミドルショットよりもさらにズームインした状態にしなければならない。そこで、距離を求めるために一度被写体に向かって大きくズームインを行い、フォーカス値を取得し距離を求めた後に被写体が適切な大ききで映るようにズームアウトするという処理を行うことで、フォーカス情報によるズーム操作を可能にした。

また、今回使用したパンチルトカメラではズーム操作を開始してからミドルショットで撮影する状態になるまでの時間に6~12秒程度を要する。発言者にマイクが渡ってから発言を開始するまでに多少時間があると考えられるが、マイクが渡ってから発言を開始するまでにズーム操作を終えることができるかについて検証する必要があると考えている。ズーム操作に要する時間は、ズームイン・アウトをする時間とフォーカスを合わせる時間の和となる。これは、本システムで用いたカメラでの時間であり、ズーム速度やフォーカス機能の性能が高いカメラを用いることで時間を縮められる可能性は十分にあると考えている。フォーカス情報から取得可能な距離は約2~11mであり、これは一般的な講義室の大部分をカバーできる距離であると言える。11m以上の距離については11m離れた人物をミドルショットで撮影できるズームパラメータに設定する。

4. 赤外線タグの動きの特徴

赤外線タグの動きと受講者の動きは完全に一致している訳ではないため、赤外線タグが中心となるようにカメラを制御しても適切に受講者を撮影できるとは言えない。赤外線タグの動きから受講者を追跡するためには、講義中のマイクロホンの動きを想定し、マイクロホンに取り付けた赤外線タグの動きの特徴を知る必要がある。講義中に起こりえる状況としては、マイクロホンを持ち運ぶ人物が発言の際に発言者のもとにマ

マイクロホンを持っていく状況や、マイクロホンを手渡しするような状況が考えられる。そこで、教室の前方にカメラを固定して設置し、マイクロホンの持ち運び、発言中等の状況における赤外線タグの撮影を行った。この際、発言者には立って発言を行ってもらった。実験で得られた映像は、マイクロホンを発言者のもとに持ち運んで発言を行ってもらった映像と手渡しを行った映像の合計5分程度の映像である。実験は、10×15m程度の大きさで約80人の受講者が座れる講義室を使用し、カメラから赤外線タグまでの距離は2~15m程度であった。

4.1 赤外線タグの動きの実験結果

赤外線タグの動きを求めるために、実験によって得られた映像中の赤外線タグの位置を画面上の座標として求めた。映像のフレームサイズは320×240[pixel]であり、フレームレートが30[fps]の映像に対して10[frame]ごとに座標として調査した。解析したデータについて表1に示す。7790[frame]の映像に対して、10[frame]ごとに調査したので全データ数は779個となる。実験では、取得した映像から10[frame]ごとに二値化した画像を表示し、画面上にタグが映っている場合と隠れている場合を確認しながら赤外線タグの位置の調査を行った。ただし、これらの実験結果については実験で用いた特定のカメラや講義室を用いた環境でのデータを基にした結果であり、講義室等の環境が異なる場合の影響を今後考慮する必要があると言える。以下では、現段階で取得できているデータを基に述べる。

講義中における赤外線タグの動きのパターンとしては、実験データからマイクロホンを持ち運んでいる、またはマイクロホンを手渡ししている移動中と、マイクロホンを持って発言している発言中の2つに分けられることがわかった。移動中と発言中における赤外線タグの動きの特徴の違いから、移動中と発言中の識別を行う。撮影中に赤外線タグが物陰や人の体の陰になり隠れる状況が起こるが、このような状況に対処する方法を考察するためにも、赤外線タグが隠れた場合についても述べる。

表1 実験の解析データ

全データ数	779個	
タグが映っているデータ数	645個	移動中 312個
		発言中 333個
タグが映っていないデータ数	134個	移動中 126個
		発言中 8個

4.1.1 赤外線タグが映っている場合

赤外線タグが映っている際の、赤外線タグの移動量について述べる。まず、10[frame]中の赤外線タグの移

動量を原点(0,0)を基準にプロットしたものを移動中と発言中とに分けて、それぞれ図5、図6に示す。例えば、赤外線タグが座標(100,100)から10[frame]後に座標(121,77)に移動した際の移動量は(21,-23)であり、グラフ中の(21,-23)の位置にプロットしている。さらに、10[frame]中に移動した量をまとめたものを表2に示す。表2では、座標データを距離に換算している。移動量が(21,-23)の場合の赤外線タグの移動量は $\sqrt{21^2 + (-23)^2} = 31.1$ となる。これらの実験結果から、移動中と発言中を識別するために、発言中における赤外線タグの移動量の範囲に関する閾値を定め、範囲内であれば発言中、範囲外であれば移動中と識別する。移動中と発言中を識別するための移動量の範囲に関する閾値を定めるために、実験データから傾向を調べた。その結果、30[frame]中での移動量が9[pixel]以上であれば移動中、9[pixel]以下であれば発言中とすることで、赤外線タグの動きから約86%の精度で移動中と発言中を識別できることがわかった。

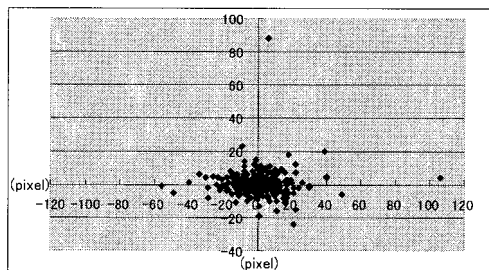


図5 赤外線タグの移動量 (移動中)

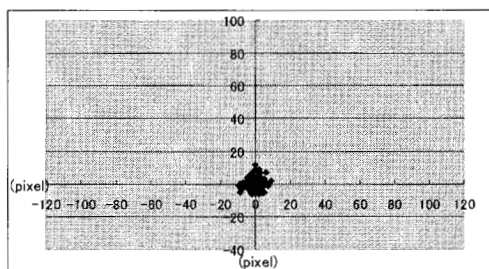


図6 赤外線タグの移動量 (発言中)

表2 10[frame]ごとの赤外線タグの移動量

タグの移動量(pixel)	最大値	最小値	平均値
移動中	106.0	0.0	6.7
発言中	12.0	0.0	1.6

4.1.2 赤外線タグが隠れた場合

赤外線タグが隠れた際に、再び現れるまでに移動した量をまとめたものを図5と同様な形式でプロットし

たものを図7に示す。さらに、表2と同様の形式で表したものを表3に示す。タグが隠れる状況としては、持ち運ぶ際に手の振りによって体の影に隠れる状況が最も多く見られた。その他にも、マイクロホンの向きによって赤外線タグの光がカメラに届かない状況などが見られた。しかし、赤外線タグが隠れる状況はほぼ移動中が占めていた。よって、発言中においては赤外線タグが隠れることはなく、赤外線タグの移動量も小さいことから、確実に撮影することが求められる発言中の受講者の撮影は可能であると言える。移動中においては、タグが隠れることが多く常にタグの位置を検出することが困難であると言える。しかし、移動中の人物の撮影は重要ではなく、追跡ができていればよいと考えられる。つまり、移動中においてはタグが隠れた状態でも再び現れた時に、タグを見失うことなく追跡できるかどうかの問題であり、そこで重要となるのがタグの現れた際の移動量となる。画面の外に赤外線タグが移動してしまうと見失った状態となってしまう。今回のデータでは、隠れてから再び現れるまでの最大移動量は98.4[pixel]であった。映像のフレームサイズが320×240[pixel]であるので、赤外線タグを画面の中央に追えている状況であれば左右に160[pixel]、上下に120[pixel]の移動量内であれば見失うことなく再び追跡が可能であると考えられる。このことから、今回の実験で想定した状況においては赤外線タグが隠れてから再び現れた際にも画面上に赤外線タグが映っており、追跡は可能であると言える。

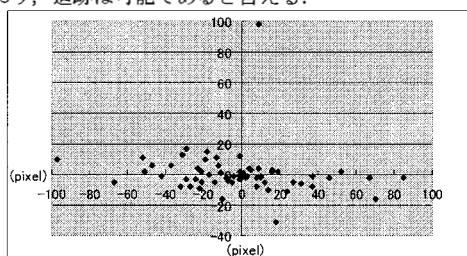


図7 隠れてから現れるまでの赤外線タグの移動量

表3 隠れてから現れるまでの赤外線タグの移動量

	最大値	最小値	平均値
タグの移動量(pixel)	98.4	1.4	17.4

5. 赤外線タグの動きの特徴に基づいた追跡

赤外線タグの動きの特徴に基づいたカメラの追跡手法について述べる。実験のデータから、移動中と発言中の識別が可能であることがわかった。移動中と発言中のそれぞれの状況に合わせた追跡手法を述べる。また、発言中であると判断した場合は、ズーム操作を行い発言者の表情が捉えられるようにズーム制御を行

う。

移動中と発言中の識別については、音声の有無と赤外線タグの動きの2つの情報を用いて行う。マイクロホンに入る音から音声区間を検出することで、確実に移動中と発言中を識別することができる。さらに、赤外線タグの動きから発言を開始する前に発言中であることを認識できる場合があることから、赤外線タグの動きからも移動中と発言中の認識を行う。移動中と発言中の追跡を制御する際のフローチャートを図8に示す。カメラの入力映像から、音声区間を基に移動中と発言中を識別する。さらに、赤外線タグ動きによって発言を開始するよりも先に発言中であると識別された際には、発言を開始する前からミドルショットで撮影が行えるようにズーム操作を開始する。

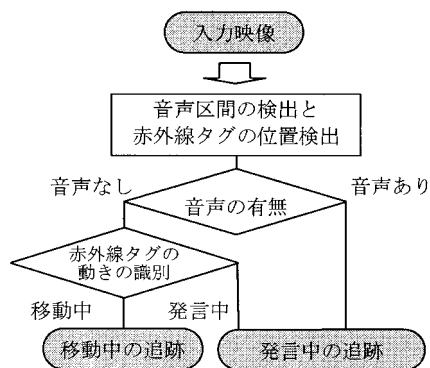


図8 追跡制御フローチャート

5.1 移動中の追跡手法

移動中の追跡手法について述べる。移動中に赤外線タグを追跡するため、常に赤外線タグの位置を把握し、追跡し続ける必要がある。追跡手法としては、赤外線タグの位置を検出し、カメラの中心に赤外線タグが位置するようにカメラの制御を行った。ここで、移動する人物をカメラが追跡できるか検証するために、追跡可能な限界速度を求める実験を行った。その結果、画面上での移動量が110[pixel/sec.]に達するまでは、画面上に人物を捉えたまま追跡撮影することができた。これは、カメラからの距離が4m程度でカメラに対して平行に歩行する人物の速度に相当し、カメラに近い状態でなければ歩行する人物の追跡が行える。また、カメラからの距離が遠ければ画面上の動きが小さくなるため、カメラから離れた位置であれば走る人物の追跡も可能である。

しかし、実際の移動における赤外線タグの動きは手の振りになどによって左右に振れている。その動きに合わせて追跡を行うとカメラが左右に動作し、追跡が遅れる原因となってしまう。そのため、移動中の追跡

においては赤外線タグの位置を平均し、平均した位置にカメラを向けるように制御することで、不必要なカメラの振りを抑えることができる。さらに、移動中の赤外線タグの動きの速さは常に同じではないため、速さに応じたパンチルト速度を設定し、移動速度に応じた追跡を行う。また、赤外線タグが隠れた際には、カメラを停止し再びタグが現れた場合に追跡を再開する。しかし、広い講義室では赤外線タグがカメラの撮影範囲外に移動することも考えられ、一定時間以上赤外線タグが現れない際には赤外線タグを見失ったと判断して、講義室内で赤外線タグを探るようにカメラを左右に振るパン操作を行う。

5.2 発言中の追跡手法

発言中の追跡手法について述べる。音声の有無、または赤外線タグの動きから発言中であると識別された場合、即座にフォーカス情報によるズーム操作を行うことで発言中の受講者をミドルショットで撮影する。また、発言者を撮影している際に赤外線タグは小さく動いており、その動きに合わせてカメラを制御すると常に細かくカメラが動いてしまい非常に見づらい映像となってしまう。そこで、発言中の赤外線タグの追跡においては赤外線タグの許容移動範囲を定義し、その範囲に赤外線タグが納まっていればパンチルト動作を行わないようにすることで、カメラが細かく動くのを抑える。ここで、発言中における赤外線タグの許容移動範囲については人物の上半身の大きさを基準に定義した。こうすることで、カメラは発言中の受講者を自然な映像として記録することができる。

6. まとめと今後の課題

本稿では、赤外線タグを取り付けたマイクロホンを用いることで位置が不特定な発言中の受講者をミドルショットで撮影する手法について述べた。設置コストを抑えることで、容易に持ち運んで利用することが可能である。また、赤外線タグを適切なカメラ制御によって追跡するために、講義中における赤外線タグの動きを実験により調査し特徴を求めた。そのデータを基に移動中と発言中の識別を行い、移動中と発言中のそれぞれに適した赤外線タグの追跡手法を提案した。

今後の課題として、赤外線タグの動きの特徴を求め実験に関して、より一般性のある実験を行うことが挙げられる。また、本手法を講義の撮影に用いることで、追跡手法の評価を行う。

参考文献

[1] 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”, 電子情報通

信学会論文誌 D-II, Vol. J85-D-II, No. 4, pp. 594-603, 2002.

- [2] 西口敏司, 村上正行, 亀田能成, 角所考, 美濃導彦, “受講者撮影機能を持つ双方向コミュニケーション記録型講義自動アーカイブシステム”, 日本知能情報ファジィ学会誌, Vol. 17, No. 5, pp. 587-598, 2005.
- [3] 富野剛, 井上亮文, 市村哲, 松下温, “多人数参加型テレビ会議システムにおける発言者拡大映像の作成”, 情報処理学会論文誌, Vol. 47, No. 7, pp. 2091-2098, 2006.
- [4] 宮崎英明, 亀田能成, 美能導彦, “複数のカメラを用いた複数ユーザに対する講義の実時間映像化法”, 電気情報通信学会論文誌, Vol. J82-D-II, No. 10, pp. 1598-1605, 1999.
- [5] 宝珠山治, 杉山昭彦, “マイクロホンアレイの研究動向と実現技術”, 電子情報通信学会研究報告.DSP, デジタル信号処理, Vol. 99, No. 504, pp. 75-82, 1999.
- [6] 白井良明, 三浦純, “複雑背景における人の追跡”, 情報処理学会論文誌. コンピュータビジョンとイメージメディア, Vol. 43, No. SIG4, pp. 33-42, 2002.