

## 講義映像における移動物体動線の自動検出 Automatic Flow Detection of Moving Persons in Lecture Scene

西野博貴<sup>†</sup> 齊藤 剛史<sup>†</sup>  
Hiroki Nishino Takeshi Saitoh

### 1. はじめに

近年、教育機関においてアクティブラーニングが注目されている。アクティブラーニングは、従来の講師から学生への一方的な授業と異なり、学生が能動的に授業に参加することが重要である。しかし、アクティブラーニングを効率良く実現できるかは、現場で教育を行う教育担当者やTAの経験や勘に頼る部分が大きく、その効果について、現状では録画された大規模講義シーンを目視で解析している。そのため、解析には多大な労力を要する。そこで本研究では、アクティブラーニングの効果解析を補助するシステムの開発を目的とし、本稿では講義室に固定されたカメラを用いて撮影された講義映像より、混雑環境下において移動物体の動線を自動的に検出する問題に取り組む。

動線の自動検出に関する研究は、駅のホームや改札、電車内に設置されたカメラ映像[1]、歩道や交差点に設置されたカメラ映像[2]が主な対象となっている。動線の自動検出という観点では同じ目的であるが、これらは本稿で対象とする講義シーンとは異なる。基本的に講義シーンに映る人物の多くは着席した学生であり、遅刻者やトイレ等で入退出する学生、TAや講師が移動物体として観測できる。本研究ではこれらの動線を自動検出する手法を提案する。ただし、本研究では図1に示すような固定カメラで撮影された講義映像を用いてオフライン処理による解析を想定している。

### 2. 移動物体の検出

本稿で提案する移動物体動線の自動検出の処理の流れは以下のとおりである。最初に各フレームにおいて特徴点を検出する。次に検出された特徴点を用いて、連続するフレーム画像間で特徴点の対応付けを行い、仮フローを検出する。仮フローは多数検出されるため、フローの連結・統合処理を行い、移動物体のフローを検出する。

#### 2.1 特徴点の検出

多くの特徴点検出手法が提案されているが、本研究ではFAST[3]を適用する。FASTは注目画素とその周囲の円周上における16個の画素を観測する。16個の画素のうち連続して $n$ 個以上の画素値が注目画素と比較して明るい、もしくは暗いときに注目画素をコーナーとして検出する手法である。



図1 講義シーンとフローの検出結果

#### 2.2 仮フローの生成

FASTではフレーム毎に特徴点が検出される。本研究では移動物体のフローを検出することを目的とするため、フレーム間で対応する特徴点を検出する必要がある。ここでFASTは特徴点を検出する手法であり、一般的にSIFT[4]で特徴を記述する。本研究でも特徴量記述にSIFTを用いる。

SIFTで得られた特徴量を用いて、連続するフレーム画像間でBrute Force Matcherを用いて特徴点を対応付ける。本研究では、前フレームから後フレームと後フレームから前フレームの両方向のマッチング処理を行い、対応付けの組み合わせが等しい特徴点ペアを2フレーム間のフローとして検出する。ただし、移動物体の特徴点のフローを求めため、対応付けされた二つの特徴点の座標が各フレーム画像上で等しい場合は移動していないと判断して検出しない。

特徴点ペアとして求めた2フレーム間のフローの中には、正しく対応付けされたフローと誤って対応付けされたフローがある。誤ったフローには、特徴量空間においてマッチング距離が大きいものと、画像空間において距離が離れているものがある。そこで両者にしきい値を設ける。つまり、特徴量空間におけるマッチング距離がしきい値 $T_f$ 以下、かつ画像空間において距離がしきい値 $T_s$ 以下のフローのみを検出する。

上記までの処理はある連続する2フレーム $f$ と $f+1$ 間のフロー検出である。一方、異なる2フレーム $f+1$ と $f+2$ 間や $f+2$ と $f+3$ にも同様にフローが検出される。そのため、3フレーム以上連続する複数フレーム間にまたがるフローが検出される。このフローを本研究では仮フローと定義する。

#### 2.3 仮フローの連結と統合

仮フローを3次元空間上に描画した結果を図2左に示す。図より多くの仮フローが検出されていることが確認できる。

<sup>†</sup>九州工業大学 Kyushu Institute of Technology

これは FAST により検出される特徴点は人物単位で検出されておらず、ある人物に対して複数の仮フローが検出されていることによる。例えば歩いている人物の場合、手の振りや足の動きなど厳密に観測すると異なる動きをしている。同一人物から複数のフローが検出されると動線解析作業に負担を強いる。そこで類似した仮フロー、つまり同一移動物体の仮フローを統合することでフローを整理する。

仮フローを統合するために本手法では時間軸において途切れたフローを接続する連結処理と並列した動きのフローを一つのフローにまとめる統合処理を行う。

連結処理は、時間軸において重なりのない二つの仮フローに対して、特徴量空間および画像空間においてしきい値以下の場合、両仮フロー間を線形で補間する。

統合処理は、重なりのある二つの仮フローに対して、重なるフレーム間の平均画像空間距離がしきい値以下の場合、一つのフローにまとめる。

図2右は図2左に対して連結および統合処理を適用した結果である。同じ動きのフローが一つのフローに統合されていることが確認できる。

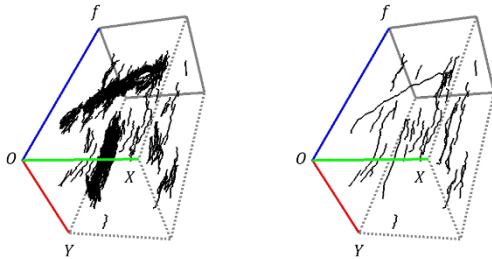


図2 フローの連結と統合 (左: 処理前, 右: 処理後)

## 2.4 移動物体フローの検出

前節で統合されたフローに対して、移動物体により生じたフローを検出する。移動物体か否かを判定するため、検出されたフローの画像空間における範囲に着目する。つまり着目フローの X 軸および Y 軸における最小と最大を求めてフローの範囲を計算し、その範囲がしきい値  $T_0$  以上であれば移動物体のフローとして検出する。

図1のシーンにおいて一つの動線を描画している。この動線は画像の左側からテーブルの間を通過して椅子に着席している人物まで描かれている。この動線は検出された移動物体のフローにおける末尾フレーム画像に、フローの座標を折れ線で描画している。この図より、学生が講義中にトイレに出て戻ってきた動線、あるいは遅刻して入ってきた動線と推測できる。前者の場合、トイレに行くための動線も検出されるため、両者は容易に判別できる。

## 3. 評価実験

九州工業大学情報工学部にはグループワーク向けのインタラクティブ教室 MILAiS がある。MILAiS の教室面積は  $260\text{m}^2$  であり、収容人数は 90 名である。同教室には、パン・チルト可能な部屋全体あるいは部屋の一部を視野とする 5 台の周囲カメラ、各テーブルを頭上から撮影する 11 台のテーブルカメラを合わせた 16 台のカメラが天井に設置されている。全てのカメラを同期して撮影でき、取得される

動画の画像サイズは  $640 \times 480$  画素、フレームレートは 7.5fps である。

本実験では、MILAiS で撮影された一つの講義シーンを用いる。録画時間は約 1 時間であるが、ここでは 5 台の周囲カメラで撮影されたシーンの中より人物の動きが観測できるシーンを六つ手動で切り出して入力データとして利用する。6 シーンの平均フレーム数は 398 であり、おおよそ 1 分弱のシーンを対象とする。

提案手法の有効性を調べるため、検出および統合された全てのフローに対して正しく検出すべきフローと検出すべきでないフローのラベルを割り当てた。次に、しきい値  $T_0$  を変化させて各フローを判別し、しきい値と処理結果に対する適合率  $P$ 、再現率  $R$  と F 値  $F$  を求めた。その結果を図 3 に示す。  $T_0 = 100$  で  $F = 0.77$  の最大値を得た。

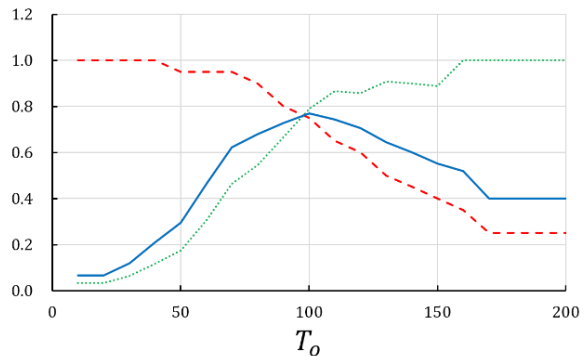


図3 実験結果

## 4. おわりに

本稿ではアクティブラーニングの効果解析を補助するシステムの開発を目的とし、講義室に固定されたカメラを用いて撮影された講義映像に対して、特徴点のマッチング処理およびフローの連結・統合処理を適用することで、移動物体の動線を自動的に検出する手法を提案した。

本稿では一つの講義映像の短時間のシーンのみを対象とした。今後は講義全体のシーンを対象とした実験や複数講義のシーンを対象とした実験に取り組む。

### 謝辞

本研究を遂行するにあたり、ご助言をいただいた九州工業大学学習教育センター近藤秀樹先生に感謝致します。

### 参考文献

- [1] 鶴飼正人, 長峰望, “異常行動検知のための画像処理手法の開発”, 鉄道総研報告, Vol.21, No.11 (2007).
- [2] 杉村大輔, 木谷クリス真実, 岡部孝弘, 佐藤洋一, 杉本晃宏, “歩容特徴と局所的見えを考慮した特徴点軌跡のクラスタリングによる混雑環境下人物追跡”, 信学論, Vol.J93-D, No.8, (2010).
- [3] E. Rosten and T. Drummond, “Machine learning for high speed corner detection,” Proc. of ECCV (2006).
- [4] D.G. Lowe, “Object recognition from local scale-invariant features,” Proc. of ICCV (1999).