

領域分割によるドラム演奏動作の人体部位追跡

Human Body Parts Tracking in Drum Performance Using Region Segmentation

山本真司[†]
Shinji Yamamoto

子安大士[†]
Hiroshi Koyasu

前川仁[†]
Hitoshi Maekawa

1. まえがき

コンピュータビジョンの応用として、人体の動作推定が広く研究されている。その技術は、ジェスチャインターフェース、CG作成のための動作抽出、監視システムのための人物追跡などに応用されている。一方、スポーツや音楽といった道具を用いる動作を抽出することは重要である。

動作推定には、モーションキャプチャシステム [3, 4] を用いることが多い。しかし、マーカーを身体に取り付けることで動作の妨げとなり、実際の動作環境を保つことができない問題がある。そのため、画像特徴を用いて動作を推定する研究が行われている。Leeら [5] は推定した各動作点位置と人体モデルとのマッチングを行い、Fassatiら [6] は人体シルエットと事前に作成した姿勢の辞書とのマッチングを行うことで、全身の動作を推定している。しかし、マッチングの精度や既知の動作にしか対応できない問題がある。

上遠野ら [2] は、ドラムの演奏動作に着目し、腕の関節点をテンプレートマッチングを用いて追跡することで、演奏動作の解析を可能にしている。しかし、局所的な追跡のため誤差が発生しやすく、その誤差が蓄積することで追跡精度が悪いという問題がある。そこで本研究では、領域追跡を用いる手法を提案する。これまで我々は、腕領域の抽出と追跡によって腕の関節点を検出し、ロバスト性の向上を確認した [1]。今回、領域分割を用いることで全身を追跡するとともに、最適な分割法を検討する。

2. 人体部位追跡手法

2.1. 領域分割

分割領域は人体部位に沿った形状になる事が望ましい。本研究ではそのような手法として、領域拡張法と Maximally Stable External Regions(MSER)[7] を比較する。本研究では、領域分割を行う対象は人体とドラムスティックのみとし、背景差分で得た領域に処理を適応する。その処理例を図1に示す。

領域拡張法

領域拡張法は、初期領域を基準として、エッジの強度が閾値以上の画素を検出するまで、同じ領域として広げていく手法である。

Maximally Stable External Regions

MSER は、画素値が類似している近傍領域を連結する手法である。画素値の極小点から連結する過程で、変動割合が最小の領域を最も安定するものとして採用する。計算量は画素数 n の画像に対して $O(n \log \log n)$ である。

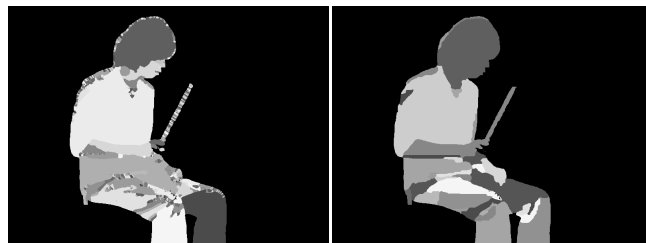


図1: 領域分割例:(左) 領域拡張法,(右)MSER

2.2. 分割領域の移動量

ブロックマッチングによるオプティカルフローを用いて、分割領域の移動量 p_i を求める。 i 番目の分割領域 R_i に対し、 R_i 内を始点とするフローの集合 v_i に対して M 推定を行う。外れ値を除いたフロー v'_i の平均値を、領域 R_i の移動量 p_i とする。このとき、 p_i は式 (1) で求める。

$$p_i = \frac{1}{k} \sum_{j=1}^k v'_i(j) \quad (\text{ここで, } k \text{ は } v'_i \text{ の総数}) \quad (1)$$

2.3. 分割領域の追跡と部位判定

フレーム t において、前フレームの移動量 p_i を用いて領域を追跡し、人体部位を判定する。前フレームで検出した分割領域 $R_i(t-1)$ を p_i だけ平行移動した領域を $R'_i(t)$ とする。そして、現在のフレームで検出した $R_j(t)$ のうち、 $R'_i(t)$ と重なる画素の数が最大となるもので対応付ける。さらに、 $R_i(t-1)$ の部位情報を $R_j(t)$ に与える。補足情報として、 $R_j(t)$ 、 $R'_i(t)$ の重心の2点間の距離が最小となる組み合わせを第二候補とする。前者の方法で対応付けがされなかった $R_j(t)$ は第二候補に対応付ける。

3. 実験

本手法を用いて人体部位を追跡する。入力動画は、200fps の高速度カメラで撮影した経験者によるドラム演奏動作である。

表1: 領域分割法の比較

分割手法	領域拡張法	MSER
分割数	760 - 818	104 - 141
計算時間 [s]	0.071	0.156
安定性	低	高
部位との対応	良	悪

まず、領域分割法の比較を表1に示す。分割数は全身

[†]埼玉大学大学院理工学研究科

に対しての処理をした場合の最大値と最小値, 計算時間は1フレームにかかる平均値を示す. また, 安定性は見えの変化がある場合に領域を同じ形状にする性能, 部位との対応は人体部位を区別して分割する性能を示す. 領域拡張法は, 過剰分割となり安定性が低いが, 複数の分割領域でひとつの人体部位を表すことができる. 一方, MSER は各部位を少ない分割領域で表すことができ, 安定性が高い. しかし, 腕と身体といった複数の部位を一つの領域に結合してしまう場合がある. また, 安定領域を探索する処理が含まれるため, 領域拡張法よりも計算時間が長い.

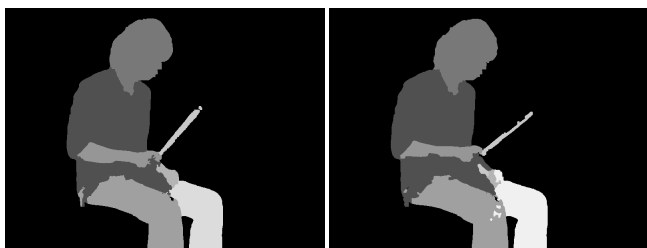


図 2: 領域拡張法を用いたときの追跡結果:(左) 追跡成功例 (10 フレーム目), (右) 追跡失敗例 (38 フレーム目)

領域拡張法を用いて追跡を行った結果が図 2 である. 結果画像では各部位を色分けして表示する. 図 2(左) では各部位の位置を追跡できていることが分かる. しかし, 図 2(右) のように右腕を振り下ろした際に, 誤判定が起きる. 原因は, 分割領域が隠れてしまい, その後の追跡ができないことだと考えられる.

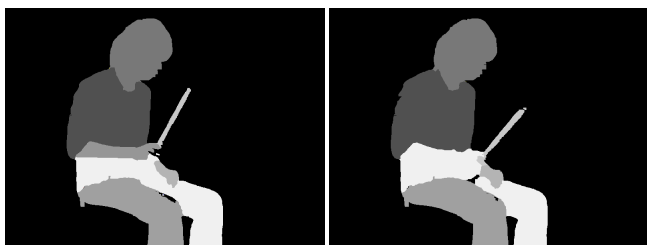


図 3: MSER を用いたときの追跡結果:(左) 追跡成功例 (5 フレーム目), (右) 追跡失敗例 (10 フレーム目)

次に, MSER を用いて追跡を行った結果を図 3 に示す. 図 3(左) では, 領域拡張法に比べ誤判定が多いが, 各部位の大まかな位置を追跡することができている. しかし, 図 3(右) では, 右腕を左足と判定している. 原因は, 分割領域が結合して別々の部位をひとつにしてしまうことである.

提案手法では, 領域拡張法の図 2(右) のように, わずかな誤判定であれば, その後の領域分割で正しい部位領域が結合され修正される. このことから, 領域追跡によるロバスト性の向上が確認できた.

4. まとめ

領域分割を用いて人体部位を追跡する手法を提案した. 演奏動作の領域追跡実験により, 追跡のロバスト性の向

上を確認した.

提案手法の追跡精度は, 各部位に対応した分割ができる領域拡張法が MSER よりも良いという結果になった. 原因は, 別々の部位をひとつの領域に結合してしまうことである. しかし MSER に比べ, 領域拡張法は一つの部位に対しての分割数が多く, 安定性が低い. これらの問題に対しては, 時系列情報を用いることで, より最適な分割ができ, 追跡精度が向上すると考えられる.

また, 追跡する際に隠蔽に対応できない問題がある. そのため, 検出している領域の一部から全体を推定することや, 人体部位の検出位置と姿勢モデルを用いて姿勢を最適化することで, 隠蔽時の部位位置を修正することが必要だと考えられる.

参考文献

- [1] 山本, 子安, 前川, “動画像処理によるドラム演奏動作の抽出”, コンピュータビジョンとイメージメディア研究発表会, CVIM172-25, 2010.
- [2] 上遠野, 子安, 前川, “画像によるドラムスティックの動作解析”, 情報処理学会研究報告, 2007-MUS-72, pp.61-66, 2007.
- [3] 奥田, 岡本, 難波, “モーションキャプチャを用いた野球バッティング動作の三次元解析”, 情報科学技術フォーラム一般講演論文集, pp.13-14, 2006.
- [4] 辻, 西方, “リズムと打拍フォームに基づく打楽器学習支援システムの開発と評価”, 電子情報通信学会論文誌, Vol.J88-D-1, No.2, pp.508-516, 2005.
- [5] M. Lee, I. Cohen, “Human Body Tracking with Auxiliary Measurements”, IEEE International Workshop on Analysis and Modeling of Faces and Gestures, 2003.
- [6] A. Fossati, M. Dimitrijevic, V. Lepetit, P. Fua, “From Canonical Poses to 3D Motion Capture Using a Single Camera”, IEEE Trans. on PAMI, Vol.32, No.7, 2010.
- [7] J. Matas, O. Chum, M. Urban, T. Pajdla, “Robust Wide Baseline Stereo from Maximally Stable External Regions”, in proceedings of the 13th BMVC, Cardiff, pp.384-393, 2002.