

## 動画像処理によるドラム演奏動作の抽出

山本真司<sup>†1</sup> 子安大士<sup>†1</sup> 前川 仁<sup>†1</sup>

本研究では、ドラムの演奏動作を撮影した動画像からの動作抽出手法を提案する。ドラムの学習支援の研究として電子ドラムやモーションキャプチャを用いて演奏動作を解析する研究があるが、普段の練習や演奏とは異なる状況で演奏しなければならない制約がある。そこで、我々は演奏動作を動画像からマーカーレスで腕とスティックの動作を抽出し、動作点の動きから奏法を解析する手法を提案している。しかし、見掛けの変化が生じる正面からの撮影では追跡ができないことや、動作点を見失うとその後のフレームの追跡ができないといった動作抽出での問題がある。そこで本研究では、複数のオプティカルフローの情報を確率的に統合することで領域の移動を推定し、その位置関係から奏法の解析に必要な動作点の検出を行う。動作を点ではなく領域により追跡することで追跡のロバスト性が向上し、同時に見え方の変化への対応が可能となる。高速度カメラで撮影された動画像に対して提案手法を適用し、手法の有効性を確かめた。

### Extraction of Stick Motion from Video Image of Drumming Performance

SHINJI YAMAMOTO,<sup>†1</sup> HIROSHI KOYASU<sup>†1</sup>  
and HITOSHI MAEKAWA<sup>†1</sup>

Motion capturing during actual drumming is important both to support of learning and to analyze the performance. Methods based on specialized equipments such as electronic drums or dedicated capturing system with markers may affect the motion to be performed. Therefore we present a marker-less method of motion extraction of drumming stick from video image. In the method, we first track regions of each body part and stick. To track regions robustly, we use a combination of M-estimation and Bayesian approach. Then feature points such as stick-tip and the grip of basic drumming scene are extracted by the boundary of the regions.

### 1. はじめに

打楽器の演奏は、曲のリズムを決定する重要な役割を担うことが多い。正確なリズムや曲のノリに合ったリズムによって、バンド全体のテンポを保ち、曲の雰囲気表現することができ、激しい曲では動作を大きくし激しく叩き、静かな曲では動作を小さくしゆっくり叩くことで、曲想を視覚的に相手に伝えることができる。さらに、演奏する打楽器によっては、その楽器独特の奏法があり、その奏法に習った動作をすることで、より自然なリズムと音色で演奏することが可能となる。そのため、打楽器におけるリズムや演奏動作は打楽器演奏において重要となる。

しかし、演奏動作を学習する機会は少なく、独学では学習することが難しいため、挫折することも多い。したがって、演奏動作のより良い理解や習得のための支援が必要となる。最近では、DVD やテレビ、動画投稿サイトで経験者による演奏の動画を視聴することができ、さらに、学習用の動画として奏法の説明を行いながら演奏を実際に行ってみせる動画もある。このような動画からおおまかな演奏動作を学ぶことはできるが、経験者と自分の演奏の違いを客観的に比較する手段はない。動作を取得し、双方の比較を行うことができれば経験者と自分の演奏の違いをはっきりと見ることができ、学習に大いに役立つと考えられる。

打楽器学習支援や演奏動作解析に関する研究は従来から行われている。学習支援における分野では電子ドラムを用いることで打拍を取得することで演奏のリズムを解析し、理想のリズムとの差異を学習支援として提示するもの<sup>2)3)4)</sup>、簡易化したモーションキャプチャを用いることで腕の演奏動作を解析するもの<sup>3)4)</sup>、多くのマーカーを使用し、上半身全体の演奏動作の取得を行うもの<sup>5)</sup>など、様々な研究がなされている。

しかし、岩見ら<sup>2)</sup>の手法では電子ドラムによってリズムに対しての演奏支援はしても、演奏動作に関して触れていなかったり、水戸ら<sup>5)</sup>の手法では、演奏動作の解析でモーションキャプチャを使用することで、本来の演奏環境と違う上に、演奏動作に影響を与えてしまったり、辻ら<sup>3)4)</sup>の手法では、手首の動きしか着目していない問題がある。よって、演奏動作に関する支援を重視し、本来の演奏環境を崩さずに演奏動作の取得を行う必要がある。

このような考え方で、上遠野ら<sup>1)</sup>により 50fps のカメラで撮影された動画像のみから動作を取得し、解析する研究が提案された。この研究では、複数の動作点を追跡することによ

<sup>†1</sup> 埼玉大学 理工学研究科

Graduate School of Science and Engineering, Saitama University

るリズム推定や、normal stroke と Moller stroke というストロークの判定を可能としている。その結果、リズム推定の正答率は約 8 割、ストロークの判定の正答率は約 9 割という成果を示している。しかし、演奏動画の取得は側面からに限定されており、撮影方向の自由度は限定されてしまっている。現在公開されている色々な演奏動画の撮影方向はそれぞれ違うため、撮影方向が限定されると解析を行うことができる動画も限定されてしまう。また、チップ点の追跡精度が他の追跡点よりも悪いため、リズムとストローク判定の精度が落ちてしまっている。

そこで本研究では、撮影された演奏動作から、上遠野ら<sup>1)</sup>と同じ複数動作点の追跡において撮影方向の限定をなくし、チップ点の追跡精度の向上を目的とし、さらに、動作を点で追跡するのではなく、領域で追跡する手法の提案を行う。

## 2. 関連研究

### 2.1 打楽器演奏の学習支援

打楽器演奏の学習支援の分野で、水戸ら<sup>5)</sup>は、スティックの端 2 点と打面に 2 点を含めた 30 点以上のマーカーを使用したモーションキャプチャによって、実際の打楽器演奏におけるスティック速度の変化を計測し、計測結果からバーチャル環境内でインタラクションモデルを生成することによるバーチャルシミュレーションを行っている。また、バーチャル環境内でモーションキャプチャによるリアルタイムとの同期が可能になれば、教育システムへの構築に有効であると述べている。岩見ら<sup>2)</sup>は、MIDI ドラムを用いて実際のドラムセットに配置されたすべての楽器(シンバル、スネアドラム、バスドラム等)に対して独立にリズムを測定し、理想のタイミングとどれくらい異なっているかをグラフで視覚化することにより、リアルタイムにリズムの習得を支援するシステムの開発を行っている。音符毎の打点強度やズレ、欠落音頻度などの経歴を測定し、それらをディスプレイ中の楽譜の該当音符に色やグラフでリアルタイムに学習に提示することが可能である。辻ら<sup>3)4)</sup>は、片手手首に 1 点のモーションキャプチャセンサと電子ドラムを用いることで、打点毎の音価や打点強度、手首などのブレを測定し、それらをグラフと数値で学習者へ提示することで、打点時の速度、ストロークの安定化の学習支援が可能となっている。しかし、測定点が 1 点であるため、スティックを含めた動作を考慮がされていない。

特に、これらの研究では、ドラム演奏の学習支援情報を使用者に提示することは可能であるが、支援情報がリズム重視であるため演奏動作に関する支援がなかったり、演奏動作を取得する際にマーカーを演奏者に装着しなければならないため、演奏の障害になってしまう。

このような問題点を解決するため、上遠野ら<sup>1)</sup>の研究は、モーションキャプチャを用いず、テンプレートマッチングを用いて、肘、手首、グリップ、チップの位置を検出し動作点追跡を行うことで演奏動作を取得している。また、グリップ、チップ位置の関係からリズム推定やストローク判定を可能としている。

しかし、演奏動画の撮影方向は側面からのみに限定され、対象動画が限定されてしまうことや、初期テンプレートの作成に手動での初期入力が必要であること、初期入力値にその後の追跡結果が大きく変わってしまうことが問題となる。また、動作点探索を行う際に、各動作点を点でとらえて追跡するため、追跡精度が安定しないという問題がある。

### 2.2 人体の姿勢推定

マーカーを用いない人の身体の姿勢推定の分野で、Lee ら<sup>6)7)</sup>は、2 台のカメラで撮影された人体領域の輪郭形状とのマッチングと Particle Filter から各動作点位置を推定し、推定された位置と人体モデルをマッチングすることで姿勢を推定を行っている。また、田中ら<sup>8)</sup>は 3 次元取得をした人体のボリュウムデータに対して細線化を行い、得られたスケルトンと人体モデルとのマッチングを行うことで人の全身の姿勢を推定している。さらに萩本ら<sup>9)</sup>は、人体領域について色情報のヒストグラムを作成し、ヒストグラム投射によって得られた領域に人体モデルを当てはめることで腕の姿勢の推定を行っている。このように、マーカーを用いない人体の姿勢推定の分野では様々な研究がなされている。

しかし、本研究にこれらの手法を適用するにあたり、各動作点の形状でのマッチングでは、動作点を画素レベルで追跡するため、誤検出が起こった場合、その後の結果に大きく影響する。また、全身のシルエットでのマッチングでは、追跡対象の腕が身体と密接したフレームでは誤検出が起きやすいといった問題点がある。

そこで、各動作領域を追跡し、その位置関係から奏法の解析に必要な動作点の検出を行う手法を提案する。複数のオプティカルフローの情報を確率的に統合することで各領域を追跡するため、画素単位での誤検出は動作点の検出に大きく影響することがなく、腕と身体を区別して追跡するため、密接したフレームでも腕の検出が可能となる。

## 3. 動画像からの演奏動作抽出

撮影した動画像から、図 1 に示した各動作領域(スティック、グリップ、腕)の追跡を行うことで、各領域の境目や先端から、同図に示した上遠野ら<sup>1)</sup>の手法での奏法解析に必要な各動作点(チップ、グリップ)の位置を検出する。

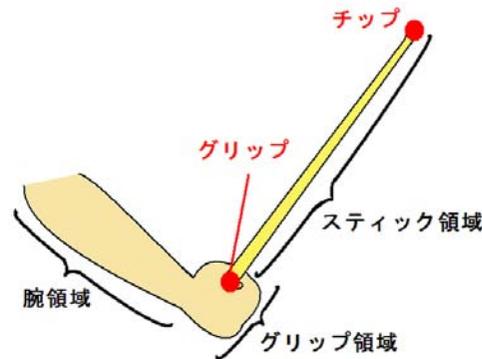


図1 追跡・検出対象

Fig.1 Tracking regions and points to be detected.

### 3.1 オプティカルフローの計算と誤差修正

オプティカルフローの計算には、大別して、輝度勾配を用いる勾配法と画像中の小ブロックをテンプレートとしてマッチングを行うブロックマッチングがある。本研究では、ブロックマッチングを用いる。テンプレートを作成するブロックはフレーム間差分によって抽出された領域内に限定し、計算時間の削減を行う。マッチングに使用するブロックのサイズは  $8 \times 8$  [pixel] とする。図2に計算されたフローの例を示す。大まかな動作方向にフローが観測されていることが分かるが、見た目に特徴の少ないスティックや手におけるフローは誤対応が存在し、中には大きく外れた方向へのフローも存在する。よって、外れ値に強い誤差修正を行う必要がある。

本研究で追跡する対象は、腕やスティックといった関節構造をしているため、同じ動作領域内のオプティカルフローはほぼ同じ方向になると考えられる。そのため、外れ値の影響を受けにくいロバストな手法である M 推定<sup>10)</sup> によって誤差の修正を行う。M 推定を用いた誤差修正の計算の流れを以下に示す。

検出されたフローのうち、まず修正を行う  $i$  番目フローに注目する。そのフローの始点  $s_i$  とその他のフローの始点  $s_j$  との距離  $d_j$  を次式で求める。

$$d_j = \sqrt{(s_i - s_j)^2} \quad (1)$$

注目したフローを含め、 $d_j$  が設定した閾値内にあるフローの集合を  $v_k (k = 0, 1, 2, \dots, n)$



図2 ブロックマッチングにより検出されるオプティカルフローの例  
Fig.2 Obtained optical flow by block matching.

とし、移動ベクトルとする。さらに、 $v_k$  の平均ベクトル  $\bar{v}$  を次式で求める。

$$\bar{v} = \frac{\sum_{k=0}^n v_k}{n+1} \quad (2)$$

平均ベクトル  $\bar{v}$  から平均ベクトル方向  $\bar{\theta}$  を求め、ベクトル  $v_k$  の方向  $\theta_k$  との誤差より、 $W$  を誤差の許容範囲とした重み  $w(\theta_k)$  を次式で求める。

$$w(\theta_k) = \left\{ 1 - \left( \frac{\theta_k - \bar{\theta}}{W} \right)^2 \right\}^2 \quad (3)$$

重み  $w(\theta_k)$  にて重み付けを行ったものの平均を移動ベクトル  $v_i$  とし、注目したフローの始点から  $v_i$  移動した点をそのフローの新しい終点  $p_i$  とする。

$$v_i = \frac{\sum_{k=0}^n w(\theta_k) v_k}{\sum_{k=0}^n w(\theta_k)} \quad (4)$$

$$p_i = s_i + v_i \quad (5)$$

M 推定によって誤差修正を行うことで、方向について、より正確なフローを得ることができる。図3に M 推定によって誤差修正されたフローの例を示す。修正前では乱雑で、同じ画素に複数対応していたフローが、より正確な移動方向を示している。

### 3.2 領域の移動と推定による動作点検出

#### 3.2.1 ベイズ推定による動作領域の追跡

領域の追跡にはベイズ推定を用いる。オプティカルフローにより領域の移動と事前確率を計算し、すべての画素に対してフローの移動距離、方向、その画素の色情報から尤度を求め、ベイズの式から事後確率を計算し、領域の追跡を行う。スティック、グリップ、腕の各動作

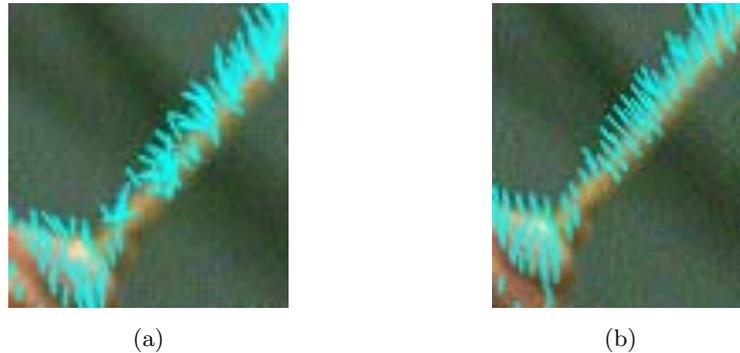


図 3 M 推定による誤差修正 (a):M 推定適用前 (b):M 推定適用後

Fig.3 Error correction of the optical flow by M-estimation. (a)Original flows, and (b) corrected by M-estimation.

領域を  $x$  とし、領域の判定に用いる移動距離、フロー方向、色情報をそれぞれ  $z_1, z_2, z_3$  としたときのベイズの式を式 (6) に示す。

$$p(x|z) = \frac{P(z|x)P(x)}{P(z)}$$

$$p(x|z_1, z_2, z_3) = \frac{P(z_1|x)P(z_2|x)P(z_3|x)P(x)}{P(z)} \quad (6)$$

また、ベイズ推定による追跡の手順は以下ようになる。

**step-1** 背景差分値  $b_i$ , フレーム間差分値  $m_i$ , 背景差分閾値  $B$ , フレーム間差分閾値  $M$  から図 4 の状態遷移図より、動作領域と静止領域を求める。

**step-2** 初期状態  $p_0(x)$  または前フレームの事後確率  $p(x|z)$  から、各動作領域  $x$  に対する事前確率  $p(x)$  を求める。初期状態  $p_0(x)$  は手で設定した色情報を基にスティック、グリッ、腕の 3 つの領域に分け、それぞれの領域の確率に 1 を与える。確率(領域)の移動はフロー情報を利用する。図 5 のように、始点から前フレームの確率を取得し、終点画素から半径  $r$  内の領域へ確率分布的に値を加算する。すべてのフローについて計算を行い、最後にすべての画素についてすべての確率の和が 1 となるように正規化を行う。静止領域として抽出された領域にどの動作領域についても確率が与えられていない場合、前フレームの事後確率をそのまま与える。また、グリッの確率について、動作中のグリッ領域は大きさ、形状の変化しない剛体であると考えられるため、重心位

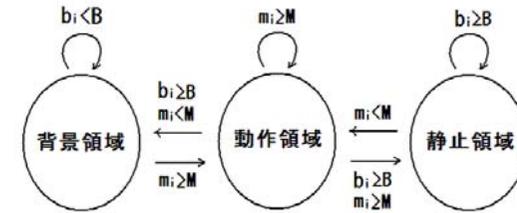


図 4 動作領域と静止領域の状態遷移図

Fig.4 State transition between background-, moving-, and stationary-regions.

置から一定の範囲外の確率は 0 として計算から除外することで、必要以上にグリッ領域が広がることをなくす。

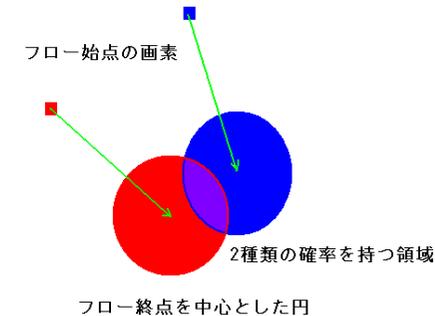


図 5 フロー情報を用いた事前確率の計算例

Fig.5 Calculation of the prior of the flow.

**step-3** 移動距離、フロー方向、色情報について、すべての画素において事前確率  $p(x)$  を重みとし、各平均値  $\mu$  を求める。その後、各動作領域について事前確率  $p(x)$  を持つ画素の情報  $a$  より、正規分布の式 (7) から尤度  $p(z|x)$  を求める。

$$p(z|x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(a-\mu)^2}{2\sigma^2}\right) \quad (7)$$

**step-4** 各尤度が求められたら、ベイズの式 (6) より事後確率  $p(x|z)$  を求める。本研究で

は、分子の計算のみを行い、全体の確率の和が1となるように正規化を行うことで事後確率とする。

### 3.2.2 各動作領域からの動作点検出

各動作領域である確率が、閾値以上であるものを各動作領域とし、スティック領域については、連続している領域を一つの領域ととらえ、その中で領域の面積が最大となるものをスティック領域とする。次に、スティック領域とグリップ領域に注目し、その2つの領域が接している画素の中間点をグリップ点とする。また、スティック領域内で、グリップ点から一番離れている画素をチップ点とする。

## 4. 実 験

撮影された動画像に本手法を適用した結果について、提案手法と先行研究<sup>1)</sup>との結果の比較を行う。また、それぞれの結果から追跡の成功例や失敗例、検出数、誤検出数、検出率を示し、内容の考察を述べる。

実験動画の撮影は、PCに接続したカメラの正面に奏者(ドラム経験者)、スティック、そして打面が収まるようにし、奏者の側面と正面の2方向より行う。また、カメラは図6のVISION RESEARCH社高速度カメラ Miro eX4<sup>11)</sup>を用い、カメラの仕様は表1に示す。撮影にはフレームレート200fpsの設定で約7秒間の撮影を行う。この環境で撮影した画像の例を図7に示す。



図6 撮影に使用したカメラ (a):レンズ装着前 (b):レンズ装着時  
Fig. 6 Used high-speed camera. (a)Camera body, and (b)the camera with lens.

### 4.1 結 果

各結果の図中での緑色の点が検出したグリップ点、赤色の点が検出したチップ点を表して

表1 カメラの仕様

Table 1 Specifications of camera used.

使用したカメラ	VISION RESEARCH 社製 Miro eX4
フレームレート	最大 1200[fps]
総画素数	800×600[pixel]
絞り, 明るさ, コントラスト	手動設定
露光時間	最短 2[マイクロ秒]
階調	24[bit]
感度	ISO1200
レンズ	SIGMA 200M f=24-70mm, F=1:2.8



(a) (b)  
図7 撮影画像例 (a):側面から撮影された画像 (b):正面から撮影された画像  
Fig. 7 Captured image frame. (a)Side view, and (b)front view.

いる。図8は側面から撮影された動画に対する検出成功例であり、図9が失敗例である。図8では、いずれも、各動作点位置を正しく検出できていることが分かる。しかし、図9のように、動作点を誤検出している場合もある。また、図10、図11は正面から撮影された動画に対する追跡の成功例と失敗例である。側面と同様、成功例では動作点を正しく検出し、失敗例では動作点を誤検出している。

#### 精度の評価

それぞれの動画像での動作点検出結果を表2に示す。なお、側面撮影をA、正面撮影をBとする。目視によって1フレーム毎の検出位置の正誤を判断し、総フレームでの検出数、誤検出数から精度の評価を行う。

側面撮影(A)と正面撮影(B)に対し、どちらも動作点位置を正しく検出ができています。チップはモーションブラーの影響により、図9の187フレーム目のように、検出の外れが起きる場合はありますが、側面撮影(A)ではモーションブラーで広がってしまったスティック領



$i=211$



$i=233$



$i=269$



$i=295$

図 8 側面から撮影された動画の動作点検出成功例 ( $i$ [フレーム])

Fig. 8 Image frames that feature points detection are succeeded(side view).( $i$ [frame])



$i=137$



$i=150$



$i=185$



$i=196$

図 10 正面から撮影された動画の動作点検出成功例 ( $i$ [フレーム])

Fig. 10 Image frames that feature points detection are succeeded(front view).( $i$ [frame])



$i=187$



$i=614$

図 9 側面から撮影された動画の動作点誤検出例 ( $i$ [フレーム])

Fig. 9 Image frames that feature points detection are failed(side view).( $i$ [frame])



$i=397$



$i=406$

図 11 正面から撮影された動画の動作点誤検出例 ( $i$ [フレーム])

Fig. 11 Image frames that feature points detection are failed(front view).( $i$ [frame])

表 2 各動作点検出結果 (A:側面撮影, B:正面撮影)  
Table 2 Result of feature points detection.(A:side, B:front)

	総フレーム	検出	誤検出	検出率
(A) チップ	629	574	55	91.2%
(A) グリップ	629	553	76	87.9%
(B) チップ	700	664	36	94.9%
(B) グリップ	700	585	115	83.6%

域から外れることはなく、微小なずれは起きているものの、ほぼ正しい位置で検出できている。グリップに関しては、同図 614 フレーム目のように右腕の奥に写っている左腕が静止領域として抽出されたことにより、グリップ領域が広がってしまったため誤検出が起きている。

また、正面撮影 (B) でも同じことが言えるが、側面撮影 (A) に比べ、グリップ領域とスティック領域のフローの方向に差が出にくいいため、グリップ位置が安定せず、図 11 の 406 フレーム目のように誤検出となったり、動作、静止領域の抽出時にノイズが残っていたことにより同図 397 フレーム目のように動作点の誤検出が起こる場合があるが、8 割を超えるフレームで各動作点の正しい位置が取得できている。チップの検出精度が上がったのはモーションブレンダーが少なかったためである。

#### 提案手法と先行手法との比較

提案手法と先行手法との比較を行う。表 3 にそれぞれの手法での結果を示す。また、側面撮影を A、正面撮影を B とする。

表 3 提案手法と先行研究 (上遠野<sup>1)</sup>) との結果比較 (A:側面撮影, B:正面撮影)  
Table 3 Comparison between proposed method and previous study by Katono, et.al<sup>1)</sup>.(A:side, B:front)

	総フレーム	チップ			グリップ		
		検出	誤検出	検出率	検出	誤検出	検出率
(A) 先行研究	629	571	58	90.8%	478	151	76.0%
(A) 提案手法	629	574	55	91.2%	553	76	87.9%
(B) 先行研究	700	44	656	6.3%	424	276	60.6%
(B) 提案手法	700	664	36	94.9%	585	115	83.6%

表 3 に示したように、各動作点、各撮影方向において、先行研究よりも精度よく検出できていることがわかる。特に正面撮影 (B) の場合、先行研究ではほぼ追跡不可能だったものが提案手法では 8 割を超える検出率で追跡が行えている。しかし、正面撮影 (B) でのグ

リップの検出精度はチップに比べると低く、2 点の追跡からの奏法解析を行うには改善が必要となる。

#### 4.2 考 察

提案手法についての考察を以下に述べ、手法に対する利点、問題点、改善案を示す。経験者が演奏した動画像に対し、提案手法で動作領域から動作点を検出し、演奏動作を取得することができることが確認できた。先行研究との大きな違いは動作を点で追跡するのではなく領域で追跡することであり、先行研究では、テンプレートマッチングで検出した追跡点の位置の少ない誤差でも、その後の追跡と解析の精度に大きく影響を与えていたが、領域の追跡での誤差ならば、動作点の検出時点での影響は少ないといえる。また、形状特徴に頼らないため、見掛けの変化に対応することで、撮影方向の制約もなくなり、側面方向と正面方向の動画への処理が可能となった。

しかし、未だモーションブレンダーによる影響で、図 9 の 187 フレーム目のように、チップ点の誤検出がおきている。さらに、フローの方向に差が出にくい正面から撮影された動画では、図 11 の 406 フレーム目のように、グリップ点の誤検出が存在する。ベイズ推定の尤度の計算には色情報とフローに関する情報を用いているが、その他にも輪郭の形状やエッジの情報などを用いることで改善できると考えられる。また、背景差分とフレーム間差分で得られた値から状態遷移を用い、動作領域と静止領域の判定を行ったが、手腕領域とスティック領域以外の領域もそれらの領域として抽出してしまい、動作点を誤検出することがあるので、動作領域と静止領域の判定とともに肌色領域の抽出を行うことでノイズを軽減し、動作点の誤検出を減らすことが可能だと考えられる。

#### 5. ま と め

本論文では、ドラムの演奏動作の取得や学習についての重要性を述べ、それらに関連する研究や先行研究の問題点を挙げた。また、その問題を解決するために、動画像を用いた動作点検出手法を提案し、映像処理による学習支援の利点を述べた。そして、実際に撮影したドラム演奏の動画像に対して動作点検出を行い、先行研究との検出精度の比較を行った。実験として提案した手法を実装し演奏動画に対して動作点検出を行い、各動作点、各動作領域を追跡した。先行研究と同様に動作点の二次元位置を取得をすることができ、また、様々な方向から撮影された動画に対応して、追跡を行うことが可能となった。しかし、腕とスティックが静止したフレームでの腕とスティックの領域抽出が必要となり、その抽出にノイズが含まれたり、非検出となり、動作領域として検出される領域と、実際の腕やスティックの領域

では差があるため、各動作点を正確に追跡できず誤検出する場合がある。

各動作点の追跡精度を向上させることが今後の課題となる。特に、腕とスティックが静止するフレームや、急激な動きでモーションブラーが発生するフレームでは、各動作点を正確に追跡できず誤検出する場合がある。そこで改善案として、肌色領域の抽出や動的に背景画像を変えることで、静止している身体領域を背景差分のみで除去するというものが挙げられる。また、人体のモデルを作成し、各動作領域へマッチングを行うことも、各動作点の追跡に有効だと考えられる。

本論文では扱っていないが、ドラムの両手演奏に対して、各腕の動作点を検出することが必要である。領域による追跡手法では、特徴点ごとの追跡で完全に追跡点が隠蔽してしまう場合が少なく、領域の一部が隠されても、隠れていない箇所から隠蔽部分を位置推定することにより追跡が可能だと考えられる。よって、両手の動きを区別することができれば、片腕ごとに追跡を行うことが可能となる。そのためには、精度よくオプティカルフローを検出することが必要となり、また、両腕が同じ動作をしたときの各腕を区別するための他の情報が必要となる。これらの問題には2台のカメラを用いて得られる情報や人体モデルを使用することが挙げられる。

また、本研究ではドラムの演奏環境は打面が1面となる環境のみを想定している。実際の演奏環境は打面が複数存在する場合もあり、その機材によって奏者領域には隠蔽が発生することも考えられる。したがって、複数打面における打点の検出、また、隠蔽に強い追跡手法を考える必要がある。

## 参 考 文 献

- 1) 上遠野優, 子安大士, 前川仁: 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.
- 2) 岩見直樹, 三浦雅文: MIDI 楽器をもちいたドラム演奏練習支援システムの提案, 情報処理学会研究報告, 2007-MUS-72, pp.85-90, 2007.
- 3) 辻靖彦, 西方敦博: リズムと打拍フォームの同時測定に基づく打楽器の演奏分析, 電子情報通信学会論文誌, Vol. J88-D-1, No.1, pp.99-107.
- 4) 辻靖彦, 西方敦博: 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.
- 5) 三戸勇氣, 日高徹哉, 川上央, 丸茂菜穂子: モーションキャプチャデータを利用した打楽器演奏のシミュレーション, 音響学会春秋季論文集, pp.819-820, 2007.
- 6) Mun Wai Lee, Isaac Cohen and Soon Ki Jung: Particle Filter with Analytical Inference for Human Body Tracking, IEEE Workshop on Motion and Video Computing

November 2002, Florida.

- 7) Mun Wai Lee, Isaac Cohen: Human Body Tracking with Auxiliary Measurements, IEEE International Workshop on Analysis and Modeling of Faces and Gestures, 2003.
- 8) 田中秀典, 中澤篤志, 竹村治雄: ボリュームデータの細線化とグラフマッチングを用いた事例ベース人体姿勢推定, 電子情報通信学会論文誌, Vol. J91-D, No.6, pp.1580-1591.
- 9) 萩本憲俊, 杉山岳弘, 阿部圭一, 佐治齊, 中谷広正: 色と動き情報を用いた人物の腕の動作追跡, 電子情報通信学会, Technical Report of IEICE, PRMU99-104(1999-11).
- 10) Michael J. Black: The Robust Estimation of Multiple Motions: Parametric and Piecewise-Smooth Flow Fields, Computer Vision and Image Understanding, Vol.63, No.1, January, pp.75-104, 1996.
- 11) Vision Research, Phantom Miro eX4, <http://www.visionresearch.com/>