

# 画像によるドラムスティックの動作解析

## Motion Analysis of Drumming by Image Processing

上遠野 優†  
Yu Katono

子安 大士‡  
Hiroshi Koyasu

前川 仁‡  
Hitoshi Maekawa

### 1 まえがき

打楽器の演奏において、曲に合ったリズムや音色、感情などを表現するために、演奏動作の習得は重要である。しかし、リズムや曲想に合った演奏動作を身につけるのは、熟練者による直接指導がなければ困難である。そこで演奏動作の学習支援のために、音に直接関係した演奏動作であるストロークに関する動作解析 [1][3][4] や、測定した学習者のストロークに対して、熟練者との違いを教示するシステムの開発 [2] が行われている。これらは主に、一定の間隔で特定のストロークを行う場合の解析であり、打拍毎にストロークが異なるリズムについてはほとんど触れていない。そのため、実際の演奏時に必要となるリズムと演奏動作の関連性の理解を得るのが難しいという問題がある。

そこで本研究では、打楽器演奏を撮影した動画から、どのようなリズムを演奏し、それに対して演奏者がどのようなストロークで演奏したかを判定するシステムの構築を目指す。熟練者の演奏について判定を行うことにより、リズムとストロークの関連性を理解する助けになり、打楽器の学習支援に役立つと考える。今回は基本的な演奏動作である、Down, Up, Tap, Full の4つのストロークに着目する。

### 2 画像からの演奏動作抽出

#### 2.1 演奏動作の撮影

演奏中のスティックの先端であるチップとグリップが常に画面内に収まるように撮影を行う。今回は片手のみの演奏に着目する。使用するカメラは Point Gray Reserch社の DragonflyExpress3.1, サンプリング周期は 120Hz である。

各フレームのチップとグリップの位置をテンプレートマッチングによって検出し、それを観測値として KalmanFilter [5] を適用し、得られた位置からスティックの動きを推定する。テンプレートマッチングのみで位置検出した場合には誤った位置を検出する可能性があるが、KalmanFilter を用いることで誤検出などのノイズの影響を抑え、より真の演奏動作に近い動作を推定できる。

打面付近でチップの速度方向が変化する時刻を打点時刻とし、得られた打点間隔から演奏しているリズムを推定する。

#### 2.2 KalmanFilter によるスティックの運動推定

フレーム間のごく短い時間では、チップとグリップの動きは線形に遷移するシステムとみなせる。そこで、KF を用いてスティックの運動の時系列予測と推定を行う。線形に状態遷移するシステムの状態方程式は以下のように記述される。

$$\mathbf{x}_{t+1} = \mathbf{F}\mathbf{x}_t + \mathbf{w}_t \quad (1)$$

$$\mathbf{y}_t = \mathbf{H}\mathbf{x}_t + \mathbf{u}_t \quad (2)$$

ここで、 $\mathbf{x}_t$  は時刻  $t$  におけるシステムの内部状態を表すベクトル、 $\mathbf{y}_t$  は時刻  $t$  における観測値を表すベクトルである。また、 $\mathbf{F}$  は内部状態の遷移を表す行列、 $\mathbf{H}$  は内部状態を観測値に写像する行列である。 $\mathbf{w}_t$ ,  $\mathbf{u}_t$  はそれぞれプロセスノイズと観測ノイズである。ここで、チップの動きを平常時には等速運動モデル、運動方向が反転する打点時には完全弾性衝突モデルに近似する。位置  $(x, y)$ , 速度  $(\dot{x}, \dot{y})$  を用いて  $\mathbf{x}_t$ ,  $\mathbf{F}$ ,  $\mathbf{H}$  を以下のように定義する。

$$\mathbf{x}_t = \begin{pmatrix} x_t \\ y_t \\ \dot{x}_t \\ \dot{y}_t \end{pmatrix}, \mathbf{F} = \begin{pmatrix} 1 & 0 & \alpha\Delta t & 0 \\ 0 & 1 & 0 & \alpha\Delta t \\ 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \alpha \end{pmatrix} \quad (3)$$

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad (4)$$

等速運動モデルでは  $\alpha = 1$  となり、完全弾性衝突モデルでは  $\alpha = -1$  となる。各時刻においてどちらのモデルを選択するかは、それぞれのモデルを用いて予測した状態と、得られた観測値のマハラノビス距離を計算し、その値が小さいモデルを用いることで決定する。

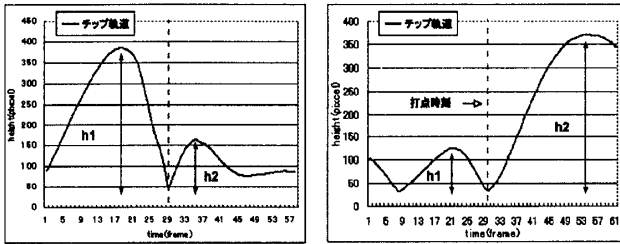
また、求めたチップの位置  $(x_t, y_t)$  とグリップの位置  $(X_t, Y_t)$  から以下の式を用いて各時刻のスティック角度を算出する。

$$\theta_t = \cos^{-1} \left( \frac{X_t - x_t}{\sqrt{(X_t - x_t)^2 + (Y_t - y_t)^2}} \right) \quad (5)$$

これらにより得られた各時刻におけるスティックの位置と角度から、スティックの運動を推定する。

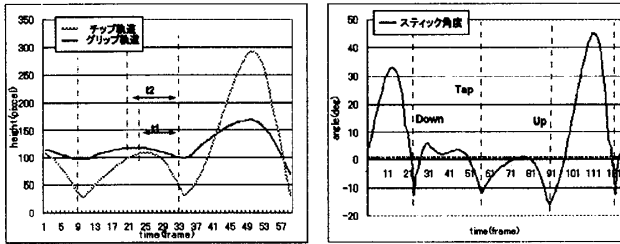
†埼玉大学大学院理工学研究科数理電子情報系専攻

‡埼玉大学大学院理工学研究科数理電子情報部門



Down のストローク比      Up のストローク比

図 1: グリップ先行度 (左), 打拍スティック角度 (右)



スティック軌道遷移      スティック角度推移

図 2: グリップ先行度 (左), 打拍スティック角度 (右)

### 3 ストローク判定

画像から、演奏動作中のチップとグリップの位置と速度、及びスティックの角度、角速度の時系列データが得られる。これらから、各ストロークの性質を表す特徴量であるストローク比  $S$ 、グリップ先行度  $P$ 、打拍スティック角度  $A$  を以下のように定義する。

[ストローク比  $S$ ]

Down, Up は打点前後で大きくチップの高さが変化する。また、Full, Tap は打点前後のチップ高さがほとんど変化しないという傾向がある。そこで、打点前の打面に対するチップの高さ  $h_1$ 、打点後の打面に対するチップの高さ  $h_2$  の比  $h_1/h_2$  をストローク比と定義する。

[グリップ先行度  $P$ ]

アクセントをつける Full, Down は勢いをつけるため、チップに対してグリップが先行する傾向がある。そこで、チップの動き出しから打点までの時間  $t_1$ 、グリップの動き出しから打点までの時間  $t_2$  の比  $t_2/t_1$  をグリップ先行度と定義する。

[打拍スティック角度  $A$ ]

Up は打点後に素早くスティックを上げる準備をするために、打点付近でグリップが先行して上がり始める傾向がある。そのため、打点時のスティックの角度が負の方向に大きい値をとる。そこで、打点時の打面に対するスティック角度を打拍スティック角度と定義する。

各打拍についてこれらの特徴量を求め、ストロークを判定する。 $S$  が設定した閾値より大きければ Down、小さければ Up、1 に近ければ Full, Tap であると判定し、各ストロークらしさを表す評価値を判定結果に基づき加算する。 $P$  が閾値より大きい場合は Full, Down、小さい場

表 1: 実験結果

	判定数	割合 (%)
正しく判定	60	85.7
Tap と Full を誤認識	8	11.4
Down と Tap を誤認識	2	2.9

合は Up, Tap について同様に評価値の加算を行う。また、 $A$  が閾値より大きければ Up の評価値を加算する。これにより決定される評価値が最大となるものをその打拍におけるストロークであると判定する。

### 4 実験と結果

各打拍毎にストロークを指定したリズムパターンを用意し、被験者に指定通りの演奏をしてもらい、その一連の演奏動作を撮影した。今回は 4 つのリズムパターン、計 70 打に対しストローク判定を行った。結果を表 1 に示す。

70 打中 60 打、約 85.7% の精度で指定どおりのストローク判定結果が得られた。誤った結果となった 10 打中 8 打が Tap と Full を誤認識したものである。グリップ先行度において Tap と Full で同じような値をとっている打拍があることが主な原因であると考えられる。また Down 後に Tap を打つ際、リバウンドを得てチップが高い位置から動き出す場合があり、その時ストローク比が Down に近い値となることも原因であると考えられる。

### 5 まとめ

KalmanFilter を用いて、画像から演奏動作を推定した。その結果得られる動作の時系列データから、各ストロークを判定する方法を提案した。今回行った実験では高い確率で各ストロークを判定することができたが、Tap と Full の誤判定が目立った。

今後の課題として、より確実な判定のための特徴の選択が考えられる。また、実際の演奏環境に近づけるため、両手での演奏や 3 次元での動作解析が必要である。

### 参考文献

- [1] 辻, 西方: 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, 信学論, vol.J88-D-1, No.1, pp.99-107, 2005
- [2] 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, 信学論, vol.J88-D-1, No.2, pp.508-516, 2005
- [3] S.Dahl: Measurements of the Motion of the Hand and Drumstick in a Drumming Sequence with Interleaved Accented Strokes, TMH-QPSR, 1997
- [4] 三浦, 北村: スネアドラム演奏のロール奏法に関する熟達者間での共通性と多様性, 日本音響学会秋季研究発表会, 2004
- [5] 片山徹: 新版応用カルマンフィルタ, 朝倉書店, 2000