

## 統計処理を用いた打楽器のフォーム分析

辻 靖彦<sup>†</sup>      西方 敦博<sup>‡</sup>  
{tsuji,nisikata}@ns.cradle.titech.ac.jp

### 概要

本研究は、打楽器演奏者の演奏能力をフォーム改善の観点から向上させることを目的としている。そのために、簡易モーションキャプチャシステムを用い、打楽器演奏時の手首の位置情報を取得した。被験者には一定打拍を一定の音の大きさで行わせた。そして得られた位置情報波形データから3種類の特徴量を抽出し、打楽器経験者と未経験者に対して波形に違いがあるのかを統計処理によって分析した。その結果、3種類のテンポにおいて、有意な違いが発見できた。また、その特徴量を独立変数として判別分析を行った結果、打楽器経験者と未経験者を高い判別率で判別できた。これより、本論文で用いた特徴量の有効性が示され、学習支援への応用の可能性が示された。

キーワード: 打楽器演奏, フォーム分析, モーションキャプチャ, t検定, 判別分析

## Statistical Analysis of Drumming Form

Yasuhiko TSUJI<sup>†</sup>      Atsuhiko NISHIKATA<sup>‡</sup>

### Abstract

This paper describes an attempt of statistical analysis for the drumming form, aiming at the development of one's drumming skill by improving their drumming form. A motion capture system was utilized to obtain the motion data of player's wrist of right hand. Subjects were instructed to beat the drum in a constant tempo and strength. Three kinds of feature values were extracted from the motion waveform, and the difference between skilled and non-skilled players were statistically investigated. As a result, we found significant difference between them in three different tempos. the discriminant analysis using these feature values achieved high discrimination ratio. The effectiveness of the feature values were shown, as well as the possibility of its application to the drumming training aid.

**Keywords:** Drumming, Form Analysis, Motion Capture System, t-Test, Discriminant Analysis

### 1 研究背景

近年、小、中、高等学校の音楽の授業において、ギターやドラムなどの、楽器演奏を用いた授業が増

加してきている。楽器演奏を音楽の授業で行なうことは、生徒が日常聴く楽曲を演奏できるため、生徒にとって親しみやすい授業内容と考えられる。また、楽器の演奏を音楽の授業で用いることは、音楽観賞、音楽の歴史の学習などの、座学中心の授業とは異なり、能動的な活動が求められるので、生徒に興味を持たせやすいと考えられる。ところで、楽器演奏においては、リズム能力が重要であると一般に言われ

<sup>†</sup>東京工業大学大学院 社会理工学研究科  
Graduate School of Decision Science and Technology,  
Tokyo Institute of Technology

<sup>‡</sup>東京工業大学 教育工学開発センター  
The Center for Research and Development of  
Educational Technology(CRADLE),  
Tokyo Institute of Technology

ている。楽器演奏の授業において、その指導方法に対してさまざまな実践的研究が行なわれている [1]。

その一方で、計算機を用いた音楽学習の支援を目的とする研究が行なわれている。例えば、音楽ジャンルの機械分類 [2]、和音の学習を目的とした研究 [3] などがある。楽器演奏の支援を行なう研究の一つとしては、ドラムパッドによる人の演奏情報を処理しフィードバックする研究がある [4]。一般に打楽器の演奏においては、良い音を出すためにも、またリズムにおいて重要視されるいわゆるノリやグルーブ感を出すためにも、演奏の際のストロークフォームが重要である。

本研究は、打楽器の演奏支援を、フォーム改善の観点から試みる。フォームの改善により、間接的に演奏者の演奏能力を向上させることを目的とする。そのためには、熟練者の演奏にどのような特徴が見られるのかを調べる必要がある。

そこで、簡易モーションキャプチャシステム [5] を用い、打楽器演奏時の腕の位置情報を取得した。そして演奏時の手首の位置情報波形データから 3 種類の特徴量を抽出し、打楽器経験者と未経験者によって演奏の際の右腕のストローク波形に違いがあるのかを統計処理によって分析した。分析には、始めに特徴量ごとに、経験者と未経験者とのグループ間での t 検定を行ない、特徴量ごとの有意差を定量的に求めた。続いて、ステップワイズ法による判別分析を行ない、両グループ間での線形的な判別を行ない、正準判別関数係数を調べ、判別の際の特徴量の組合せについて検討し、両グループにどのような波形の違いが見られるかを検討した。

## 2 ストロークフォームデータの測定

### 2.1 測定装置と実験課題

図 1 に、本実験で用いたフォームデータの測定装置を示す。CCD カメラが被験者の右手の甲に付いているマーカの 3 次元位置情報を取得し、打拍している間の位置情報がコンピュータへ常に送られる。ここで用いたモーションキャプチャシステム DigitEye3D [5] は、15sps(samples per second) のサンプリングレートを持つ 4 つのマーカで構成されているが、本研究ではそれを改造し、60sps のサンプリングレートのマーカの一つ用いている。

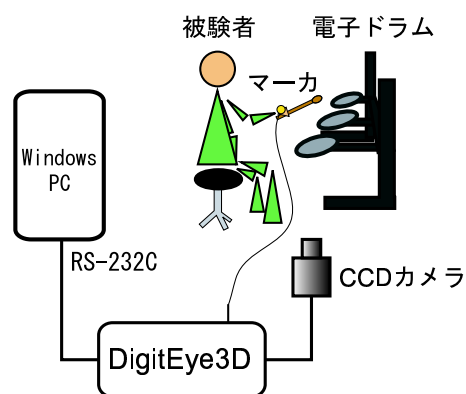


図 1: 実験風景

被験者には、一定の基準テンポで、スティックを用いてドラムパッドを等間隔に打拍する課題を行なわせ、その際の右手の甲に付けたマーカの位置情報を時系列波形として測定した。基準テンポとしては、90BPM (Beats Per Minutes)、120BPM、180BPM の 3 種類のテンポを用いた。測定手順は以下の通りである。まず基準テンポで等間隔なメトロノーム音を被験者に提示し、基準テンポを覚えさせた。その後、120 秒間、提示されたテンポの記憶を頼りに等間隔、等音量でスティックで打拍させた。この測定より、各被験者の各試行につきそれぞれ 120 秒間の打拍ストローク時の右手の甲の 3 次元位置情報を得た。それを各テンポに対して 3 回ずつ、つまり被験者一人あたり 9 回ずつ試行した。そして各テンポにおいて最後の試行の後に、自分の演奏に対する自己評価アンケートに回答させた。さらに全試行が終了後、楽器演奏経験、持っている CD の枚数、好きなジャンルなどの音楽経験アンケートに回答させた。被験者は、一般大学生及び大学院生の、1 週間に 8 時間で 5 年間以上の練習経験を持つ打楽器経験者 4 名、未経験者 10 名である。

### 2.2 測定データ

測定したストローク波形データは、

$$\vec{f}(t), t = 0, \Delta t, 2\Delta t, \dots, 120 \text{ sec}, \Delta t = \frac{1}{60} \text{ sec}$$

で表される 3 次元座標ベクトルデータである。図 2 に、取得した波形データの例の中で、顕著な例を示す。縦軸はマーカの垂直座標、横軸は時間である。測定した波形に見られる特徴は、大よそ以下の 4 種類であった。(a) は周期ごとの大きさが一定しない波形である。(b) は左右対称で正弦波に比較的近い

波形である。(c)は左右非対称なノコギリ波に近い波形である。(d)は振幅が大きい波形である。

### 2.3 測定データからの支援の考察

前節の4種類の波形に対して学習支援を考える。システムが自動的に(a)の波形を認識することができれば、「ストロークフォームが一定していません。」等のフィードバックをユーザへ返すことができる。また、システムが自動的に(b)、(c)の波形を認識できれば、打楽器経験者のストローク波形を予め調べることにより、フィードバックが可能となる。同様に、システムが自動的に(d)の波形を認識できれば、つまり振幅を計算できれば、打楽器経験者のストローク波形の振幅に合うようなフィードバックをユーザへ提供することが可能となる。さらに、今回の実験では打拍時の音の強さは一定という条件だけであったが、「*mf* (メゾフォルテ) で叩く」などの音の強さの条件を付与した実験環境においては、振幅値がより重要なパラメータとなることが予想できる。

従って、測定データから学習支援を行なうには、上記のような波形を判別できること、もしくは定量的にその特徴を数値化することが必要となる。

そのために本論文では、

- 周期ごとのストロークフォームのずれ (誤差)
- 波の形
- 周期ごとの振幅の大きさ

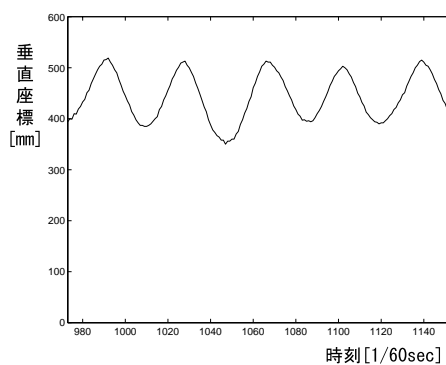
の3種類の要素から特徴量を取得することを考えた。

## 3 特徴量の取得

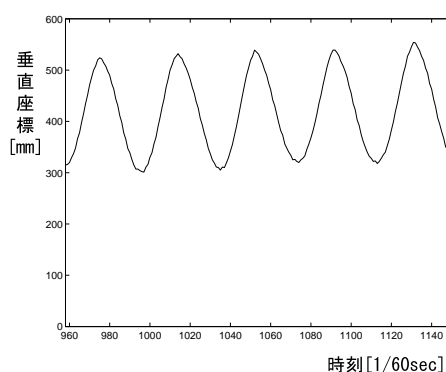
### 3.1 打拍点の推定

前節において、周期ごとに値を計算するためには、周期の境目を特定しなければならない。しかし、測定したデータは被験者の手首の位置情報のみで、演奏情報は含まれていない。従って、ストローク1周期ごとに処理を行ないたい場合には、打拍点を推定する必要がある。

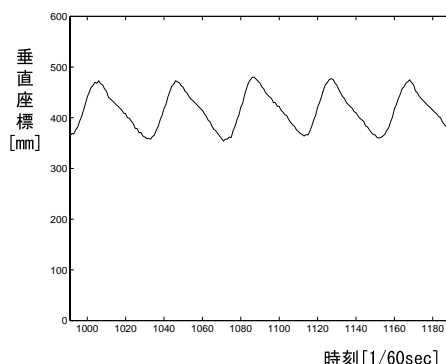
そこで、各周期ごとの垂直座標の最小点を求めるアルゴリズムにより、打拍点を推定した。



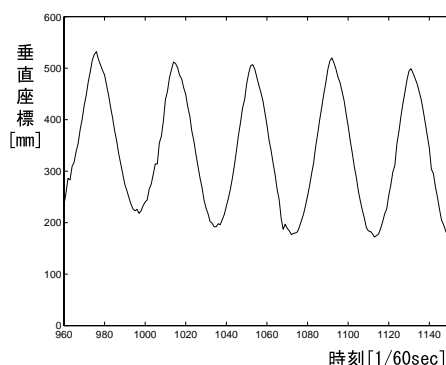
(a) 周期不安定型



(b) 正弦波型



(c) ノコギリ波型



(d) 大振幅型

図 2: ストローク波形の例

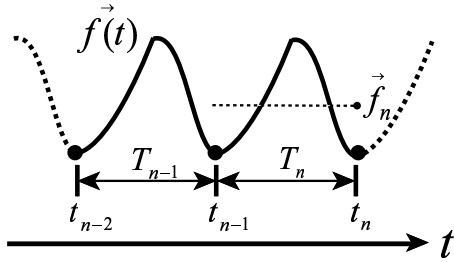


図 3: ストローク誤差の取得

### 3.2 ストローク誤差の定義

本論文では、打拍の再現性を一評価指標とする。周期的な打拍（図3）における繰り返しのストローク誤差  $E_n$  を次式で定義する。

$$E_n = \frac{\frac{1}{T_n} \int_{t_{n-1}}^{t_n} |\vec{f}(t) - \vec{f}(t')|^2 dt}{\frac{1}{T_n} \int_{t_{n-1}}^{t_n} |\vec{f}(t) - \vec{f}_n|^2 dt}, n = 1, 2, \dots$$

ここで  $\vec{f}(t)$  は複数マーカの座標を並べたベクトル、 $T_n$  は  $n$  番目の打拍の周期で  $T_n = t_n - t_{n-1}$ 、 $\vec{f}_n$  は  $n$  番目のフォームの平均値、つまり、 $\vec{f}_n = \frac{1}{T_n} \int_{t_{n-1}}^{t_n} \vec{f}(t) dt$  である。

また、 $t' = t_{n-2} \frac{t_n - t}{T_n} + t_{n-1} \frac{t - t_{n-1}}{T_n}$  であり、 $t_{n-2}$ 、 $t_{n-1}$ 、 $t_n$  は連続する打拍の時刻とする。このストローク誤差  $E_n$  を求めることにより、周期ごとのストロークがどのくらいずれたかが定量的に分かる事が期待できる。特徴量として、被験者の一回の演奏におけるストローク誤差の平均値  $\bar{E}$  及び標準偏差  $SD(E)$  を統計的分析に用いた。

### 3.3 パワースペクトルピーク値

測定データから、ストロークにおける高周波数成分を含めた特徴量を抽出するために、パワースペクトルのピーク値を求めた。測定した3次元位置情報の垂直成分波形に対してパワースペクトル密度関数  $psd(f)$  を求め、そのピーク値  $P_1, P_2, \dots, P_{10}$  を求めた（図4）。さらに、その各ピーク値を  $1[\text{Hz}]$  の一定幅で積分値  $\Delta P_1, \Delta P_2, \dots, \Delta P_{10}$  を求めた。 $\Delta P_n (n = 1, 2, \dots, 10)$  は以下の式で定義できる。

$$\Delta P_n = \int_{f_n - 0.5[\text{Hz}]}^{f_n + 0.5[\text{Hz}]} psd(f) df$$

その比  $\frac{\Delta P_2}{\Delta P_1}, \frac{\Delta P_3}{\Delta P_1}, \dots, \frac{\Delta P_{10}}{\Delta P_1}$  の9つの値を特徴量として後の統計的分析に用いた。

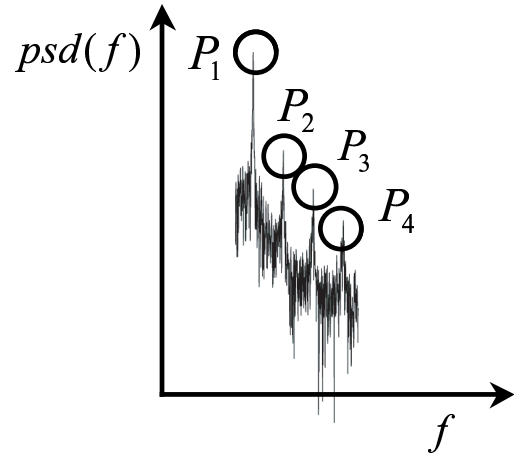


図 4: パワースペクトルピーク値の測定法

この値を求めることにより、2章4節で述べた波の形を特徴量として取得できると考えられる。

### 3.4 振幅のRMS値

振幅のRMS (Root Mean Square: 2乗平均) 値とは、測定した3次元位置情報の垂直成分波形に対して、遮断周波数が  $0.1[\text{Hz}]$  のHPF(High Pass Filter) を通し、周期ごとに振幅の2乗平均値を求めたものである。ストローク波形における、 $n$  周期目の振幅のRMS値  $R_n$  は、以下の式で定義できる。

$$R_n = \sqrt{\frac{1}{T_n} \int_{t_{n-1}}^{t_n} f_y^2(t) dt}, n = 1, 2, \dots$$

振幅のRMS値を求めることにより、被験者のストロークの大きさが求められる。特徴量として、RMS値の平均値  $\bar{R}$ 、標準偏差  $SD(R)$  を統計的分析に用いた。

## 4 t検定による分析

測定した被験者のフォームデータから得られた各特徴量  $\frac{\Delta P_2}{\Delta P_1}, \frac{\Delta P_3}{\Delta P_1}, \dots, \frac{\Delta P_{10}}{\Delta P_1}, \bar{E}, SD(E), \bar{R}, SD(R)$  に対して、経験者、未経験者のグループ間におけるt検定を行なった。t検定は3種類のテンポごとに別々に行なった。その理由は、テンポが異なると演奏におけるストロークの動作も異なると仮定しているからである。

被験者合計14名の90bpmのテンポによる演奏フォーム波形データより抽出した、合計40サンプルの特徴量において、経験者、未経験者のグループ

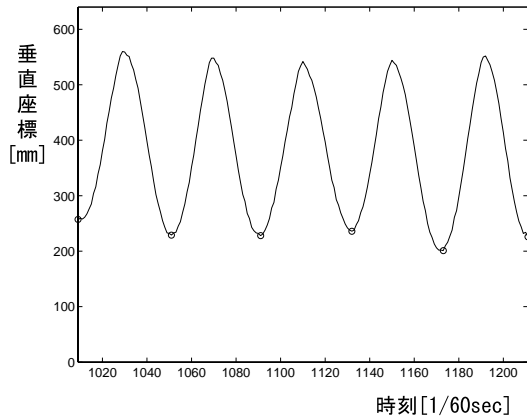


図 5: 90bpm における打楽器未経験者のストローク波形の例

間における  $t$  検定を行なった。その結果、 $\frac{\Delta P_2}{\Delta P_1}$  の特徴量において、経験者と未経験者で有意な差が見られた ( $t(38) = -3.352, p < 0.01$ )。これにより、打楽器経験者の方が未経験者よりも基本周波数に対して 2 倍の高調波の割合が有意に大きくなっていることが分かる。

また、 $\bar{R}, SD(R)$  の特徴量においても、経験者と未経験者では有意差が見られた ( $\bar{R}: t(38) = 3.649, p < 0.01, SD(R): t(38) = 3.447, p < 0.01$ )。これより、打楽器経験者は未経験者に比べて、リズムキープすることを最重要目的に打拍した場合、有意にストロークの振幅が小さく、また振幅の分布が狭いことが分かった。顕著な例として未経験者のストローク波形 (図 5) と経験者の波形 (図 6) を比較すると、図 5 の方は観察すると波形が左右対称で、sin 波に近い波形であるのに対し、図 6 の方は若干、左右非対称な形で 2 倍高調波を多く含むノコギリ波に近い波形であることが分かる。

120bpm と 180bpm のテンポに関しても同様に、被験者計 14 名による演奏フォーム波形より抽出したそれぞれ合計 40、42 サンプルの特徴量において、経験者、未経験者のグループ間における  $t$  検定を行なった。

その結果、120bpm の演奏に関しては、 $\frac{\Delta P_3}{\Delta P_1}$  の特徴量において、経験者と未経験者で有意な差が見られた ( $t(38) = 2.112, p < 0.05$ )。これより、打楽器経験者の方が未経験者よりも基本周波数に対して 3 倍の高調波の割合が有意に小さくなっていることが分かる。さらに、 $\bar{E}, SD(E), \bar{R}, SD(R)$  の特徴量においても、経験者と未経験者では有意差が

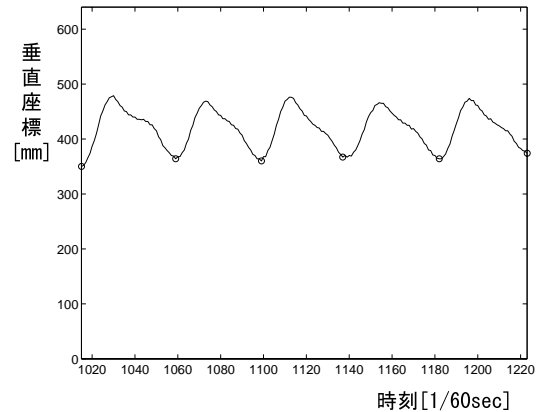


図 6: 90bpm における打楽器経験者のストローク波形の例

見られた ( $\bar{E}: t(38) = 2.749, p < 0.01, SD(E): t(38) = 2.029, p < 0.10, \bar{R}: t(38) = 3.182, p < 0.01, SD(R): t(38) = 4.306, p < 0.01$ )。これより、120bpm の演奏において打楽器経験者は、未経験者に比べて有意にストローク誤差及びその標準偏差が小さく、また有意にストロークの振幅が小さく、その分布も狭いことが分かった。従って、打楽器経験者の方がコンパクトで、正確なストロークフォームであったことが言える。

続いて、180bpm の演奏に関しては、フォーム波形データより得られたパワースペクトル密度関数から抽出した特徴量に関して、多少偏りはあったものの、有意な差は見当たらなかった。しかし、 $\bar{R}, SD(R)$  の特徴量においては経験者と未経験者で有意な差が見られた ( $\bar{R}: t(40) = 4.039, p < 0.001, SD(R): t(40) = 3.537, p < 0.01$ )。これより、180bpm の演奏において打楽器経験者は、有意にストロークの振幅が小さく、その分布も狭いことが分かった。従って、打楽器経験者の方がコンパクトなストロークフォームで演奏していたことが言える。

## 5 打楽器経験者と未経験者の判別分析

前節の  $t$  検定の結果を基に、打楽器経験者か未経験者かを目的変数、特徴量を独立変数とした判別分析を行なった。分析にはステップワイズ法を用い、2つのグループ間のマハラノビスの距離が最大となるような独立変数を順次、選定し判別させた。

その結果、90bpm のテンポにおいて、必要な特徴量は  $\frac{\Delta P_2}{\Delta P_1}, \frac{\Delta P_4}{\Delta P_1}, \bar{R}$  の 3 種類の特徴量が選定され、90%

表 1: 判別率集計表

単位はすべて[人] 判別率のみ[%]	予測グループ			判別率
	未経験者	経験者	合計	
元のデータ (90bpm) 未経験者	28	2	30	90.0
経験者	2	8	10	
元のデータ (120bpm) 未経験者	29	1	30	95.0
経験者	2	8	10	
元のデータ (180bpm) 未経験者	24	6	30	78.6
経験者	3	9	12	

の判別率が得られた。また、120bpm のテンポにおいて、必要な特徴量は  $\frac{\Delta P_2}{\Delta P_1}, \frac{\Delta P_3}{\Delta P_1}, \frac{\Delta P_4}{\Delta P_1}, \frac{\Delta P_5}{\Delta P_1}, SD(R)$  の 4 種類の特徴量が選定され、95.0% の判別率が得られた。180bpm のテンポにおいては、必要な特徴量は  $\frac{\Delta P_3}{\Delta P_1}, \bar{R}$  の 2 種類の特徴量が選定され、78.6% の判別率に留まった。(表 1、表 2)。この判別率が他の 2 種類のテンポよりも小さい原因としては、テンポが速いため、被験者のストロークフォームの自由度が少なくなり、経験者と未経験者の違いが現れにくくなったためと推察できる。

この結果より、90bpm 及び 120bpm のテンポにおいては、これらの特徴量を用いることにより大よその分類が可能であることが分かった。

120bpm のテンポに関しては、表 2 を見ると、基本周波数の偶数倍成分のピーク値が大きいほど経験者側、奇数倍成分のピーク値が大きいほど未経験者側に判別される結果になった。これより、打楽器経験者のストロークフォーム波形はノコギリ波などの左右非対称な波形である傾向があり、それに対して未経験者のストロークフォーム波形は、三角波などの左右対称な傾向があると言える。また、ストロークの振幅の標準偏差  $SD(R)$  も、打楽器未経験者の方が、演奏全体を通して、広い分布を持つ傾向があることが分かった。

## 6 むすび

以下、本研究の結果をまとめる。

簡易モーションキャプチャシステムを用い、打楽器演奏時の手首の位置情報を取得した。被験者には一定打拍を一定の音の大きさで行わせた。そして得られた位置情報波形データから 3 種類の特徴量を抽出し、打楽器経験者と未経験者に対して波形に違いがあるのかを t 検定によって分析した。その結果、3 種類のテンポにおいて、有意な違いが発見できた。経験者グループと未経験者グループ間の線形的な判別を行なった結果、3 種類のテンポにおいて、90.0%、

表 2: 正準判別関数係数

tempo	特徴量	係数
90bpm	$\frac{\Delta P_2}{\Delta P_1}$	-2.488
	$\frac{\Delta P_4}{\Delta P_1}$	2.950
	$\bar{R}$	0.830
120bpm	$\frac{\Delta P_2}{\Delta P_1}$	-2.389
	$\frac{\Delta P_3}{\Delta P_1}$	4.170
	$\frac{\Delta P_4}{\Delta P_1}$	-2.516
	$\frac{\Delta P_5}{\Delta P_1}$	1.170
	$SD(R)$	0.916
180bpm	$\frac{\Delta P_3}{\Delta P_1}$	0.947
	$\bar{R}$	1.131

95.0%、78.6% の判別率が得られた。

上記の t 検定の結果より、打楽器経験者及び未経験者のストローク波形の傾向が発見できた。本研究で用いた特徴量を用いることにより、打拍を行なったユーザへ対して、「フォームがずれました。」「もう少しスナップを効かせて叩きましょう。」「もう少し腕の振りを小さくしましょう。」等のフィードバックを行える可能性が示されたと言える。

また判別分析の結果より、本研究で用いた特徴量より被験者の大よその分類が可能であることが示された。

以上より、本研究で提案した特徴量の有効性が示されたと考えられる。

## 参考文献

- [1] Debbie Rohwer, "Effect of Movement Instruction on Steady Beat Perception, Synchronization, and Performance", *Journal of Research in Music Education*, Vol. 46, no. 3, pp. 414-424, 1998
- [2] 辻 靖彦, 赤堀 侃司, ニューラルネットワークを用いた楽曲のジャンル推定とその教育利用, 日本教育工学会論文誌, Vol.24, No.3, pp.183-191, 2000
- [3] 米川孝宏, 西方 敦博, 2 純音和音の再生テストの成績と楽典テスト・自己評価との関係, 電子情報通信学会論文誌, VOL.J85-DI, No.5, pp.484-488, May 2002
- [4] 米川孝宏, 西方 敦博, 中山実, 清水康敬, テンポ変動度に基づく基礎リズム能力評価法の検討, 日本教育工学会論文誌, Vol.24, No.3, pp.153-160, 2000
- [5] 簡易モーションキャプチャセンサ DigitEye3D-M100 <http://www.kansai-ri.co.jp/icr/digiteye/>
- [6] 多変量解析法入門, 永田 靖, 棟近雅彦, サイエンス社