

単音を対象としたリズム練習支援システム

松尾 章弘[†] 原 佑輔[†] 土江田 織枝[†] 山田 昌尚[†]

概要: 音楽を演奏する上で音高とリズムは基礎的で重要な要素である。音高に関してはチューナーなどが普及しているのに対し、リズムに関してはそのように数値化、可視化するようなものはない。現在、我々は発音検出を用いて正確に自分の演奏したリズムを確認できるシステムを開発中である。これまで発音検出の精度の評価、発音検出に用いるしきい値の自動化、ユーザインタフェイスの作成を行った。

1. はじめに

本研究はリアルタイム発音検出での動的しきい値のパラメータを自動的に最適化する方法を検討するものである。現在、我々はリズム感の向上を目的として、正確に自分の演奏したリズムを確認できるように発音検出を用いたリズム練習支援システムを開発中である[1]。発音検出に関しては一定の精度が得られたが、これまでは発音検出のための関数に用いるしきい値のパラメータを手動で設定してきた。今回は、ユーザインタフェイスの作成し、しきい値パラメータを自動で最適化する方法について検討した。

2. 発音検出としきい値

音響信号から発音のタイミングを得ることを発音検出という。発音検出は一般的に、前処理、検出関数、ピーク抽出の3段階からなる[2]。第1段階の前処理としては正規化を施す。第2段階の検出関数として、周波数ごとの信号スペクトル強度を利用するスペクトルフラックスや HFC (High Frequency Content), 位相の変化を利用する方法など、各種が提案されている。本研究では、周波数ごとの信号スペクトル強度の変化が大きき場合に発音となるピークが現れるスペクトルフラックスを使用する。スペクトルフラックスは次式で表される。

$$SF(n) = \sum_{k=1}^{\frac{N}{2}-1} \max(0, |X(n, k)| - |X(n-1, k)|)$$

ここで N はFFTフレームのデータ数、 n は時間、 k は周波数であり、 $X(n, k)$ はスペクトログラム(時間周波数信号)を表す。0とのmaxをとることでスペクトル強度が増加する場合のみを対象としている。第3段階のピーク抽出では、しきい値を超える局所最大値(local maxima)を検出し、その時刻を発音時刻とする。しきい値として次式による動的しきい値を用いる。

$$TH(n) = \delta + \lambda \cdot \text{median}(SF(n - v_1 : n + v_2)) \\ + \alpha \cdot \text{mean}(SF(n - v_1 : n + v_2))$$

ここで δ はしきい値の定数項、 λ および α はそれぞれ中央値、平均値に対する重みであり、 v_1, v_2 は動的しきい値の対象幅を表す。今回はこの δ, λ, α の最適値を求める。このしきい値を用いて、検出関数 $DF(n)$ および発音時刻 $OD(n)$ は次のように求められる。

$$DF(n) = SF(n) - TH(n)$$

$$OD(n) = \begin{cases} 1, DF(n) > 0 \text{ and } \operatorname{argmax}_{w_1 < m < w_2} DF(m) = n \\ 0, \text{otherwise} \end{cases}$$

3. システム構成

図1にシステム構成を示す。メトロノーム音はChuckで生成し、スピーカから出力する。強拍と弱拍は周波数で区別する。強拍は880Hz、弱拍は440Hzとした。システムのユーザはこのメトロノーム音を聴きながら楽器を演奏する。ChuckからProcessingへメトロノーム音と発音検出結果のタイミングを送るためにOSC(Open Sound Control)を用いる。OSCは電子楽器およびコンピュータ間で音楽演奏データ等を送受信するための通信プロトコルである。OSCではURL形式でデータを送るため、メトロノームのタイミングとオンセットを容易に区別して情報伝達をできる。図2にユーザインタフェイスを示す。画面右に上からしきい値を調整するための感度調整スイッチ、メトロノームのオンオフを切り替えるスイッチ、テンポと拍子を指定するテキストボックスがある。左側にメトロノームが表示される。メトロノームは1小節1段で表され、図2では4拍子の小節が4小節分表示されている。縦の青い線の間隔が1拍の長さを表している。黒い棒が左から右に流れていき、入力信号で発音だと判断されたときに赤い点を画面上に描画する。

4. 検出精度の評価

発音検出の精度の評価としてF値(F-measure)を用いる。そのF値を求めるのに必要な発音検出の結果の状態を分けるために、実際に発音されている場合で発音だと判断したものをTrue Positive(TP)、発音だと判断できなかったものをFalse Negative(FN)、実際に発音されていない場合で発音だと判断しなかったものをTrue Negative(TN)、発音だと判断

[†] 釧路工業高等専門学校

National Institute of Technology, Kushiro College

してしまったものを False Positive (FP)とする。

次にこの4つの状態から、発音と判断したもののうち正しい時刻で発音と判断できた割合を示す再現率Pと正解の発音をどれだけ発音だと判断できたかの割合を示す適合率Rを求めることができる。

$$P = \frac{TP}{TP + FP}, R = \frac{TP}{TP + FN}$$

そしてこの2つの調和平均であるF値をこのシステムの検出精度の指標とする。F値は以下のように求められる。

$$F = \frac{2 \cdot P \cdot R}{P + R}$$

本システムはFP, FNが多くなるとユーザへ正しい情報のフィードバックを行うことができなくなってしまう。よって、その2つを同時に評価することのできるF値を精度の指標として扱う。

5. しきい値自動最適化

今回は最急降下法[3]によりしきい値の最適化を検討した。最急降下法によるパラメータの更新式は次式で表される。

$$V_{t+1} = V_t + \beta \frac{\partial F(V_t)}{\partial V}$$

ここで $V_t = (\delta, \lambda, \alpha)$ とし、 α は学習率、Fは最大化する目的関数であるF値である。終了条件は $F(V_{t+1}) - F(V_t) < 0.001$ とした。発音検出に用いたデータは、サンプリング周波数44.1kHzで保存された48個の発音があるホルンの演奏の音響データである。正解の発音時刻は3人でアンテーションし、その平均時刻を各発音の発音時刻とした。 $v_1 = 50ms$, $v_2 = 0$ として、このデータに対して最急降下法によるパラメータの最適化を行った。今回はしきい値のパラメータの自動最適化の手法を検討するため、計算時間を短縮する目的でフリーの数値解析ソフトであるScilabを用いて実験を行った。

λ, α をそれぞれ0~1.5の範囲で0.15刻みで11個、 δ を0~0.2の範囲で0.02刻みで11個用意し初期値を変えて1331個の組み合わせで最急降下法によるパラメータの最適化を行った。F値が0.9を越えるものが299個となった。0.8を下回るものはそれぞれ100個以下であり、最急降下法で、F値の高くなるパラメータの組み合わせを得ることができた。F値の高くなる時のパラメータの組み合わせは複数あることが確認できた。

次に最急降下法で初期値を変えて最適化を行った組み合わせの例を例1~4として表1に示す。その結果を縦軸をF値、横軸を繰り返し回数として図3に示す。初期値によって、収束する値が異なることが図3からわかる。

6. まとめ

現在開発中のリズム練習支援システムの構成要素の1つとして今まで手動で行ってきたしきい値の設定を自動化する方法について検討した。結果は、F値の高いパラメータの組み合わせを求めることができた。最適値は複数あることと、最急降下法により最適値を求める場合、初期値の設定が結果に影響することがわかった。今後の課題としては、今回はホルン以外の種類の楽器についても同様の結果が得られるかの確認が必要である。

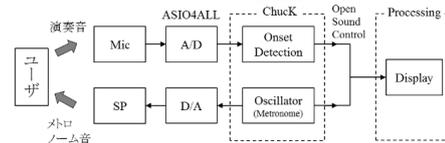


図1 リズム練習支援システム構成図

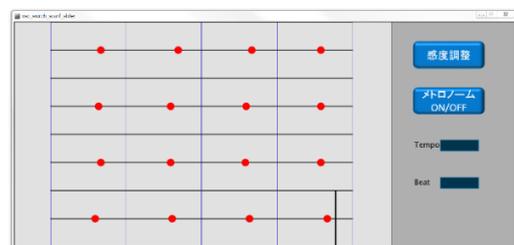


図2 ユーザインタフェース

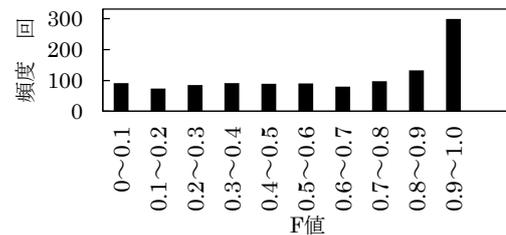


図3 最急降下法によるF値のヒストグラム

表1 最急降下法での初期値の組み合わせ

	α	λ	δ
例1	0.2306	0.5452	0.4151
例2	0.2306	0.2370	0.0607
例3	0.2306	0.1104	0.4151
例4	0.0519	0.1104	0.0607

参考文献

- [1] 松尾章弘, 原佑輔, 土江田織枝, 山田昌尚, “リズム練習支援のためのリアルタイム発音検出手法の検討”, 第14回情報科学技術フォーラム(FIT)論文集, 第2分冊, pp.205-206(2015)
- [2] Bello, Juan Pablo, et al., “A tutorial on onset detection in music signals”, Speech and Audio Processing, IEEE Transactions, Vol.13, No.5, pp.1035-1047 (2005)
- [3] 北村充, “数理計画法による最適化 実際の問題に生かすための考え方と手法”, 森北出版株式会社, p.72(2015)