

音楽音響信号を対象とした 時間周波数領域での スペクトル変動による 感情価の変化

宮本 真治

Shinji MIYAMOTO

情報メディア学専攻修士課程 2015 年度修了

1. はじめに

2015 年 10 月 10 日に筑波大学で開催された日本音響学会音楽音響研究会（2015 年 10 月研究会）に参加し、ポスターセッションで発表を行なった。発表題目は「音楽音響信号を対象とした時間周波数領域でのスペクトル変動による感情価の変化」である。本報告では、当該研究会で発表した内容及び感想を述べる。

2. 研究内容

2.1 背景

楽曲演奏における感情の伝達について、演奏者はテンポや音圧レベルなどの演奏の特徴を使い分けて演奏を行なうことで感情を表現し、聴取者は演奏の特徴から奏者の意図する感情を読み取ることができる。また、計算機によって楽曲の印象や雰囲気、すなわち感情を自動的に制御することが可能な MIDI を対象とした感情ハンドリングシステム（以降、ERS）もある。この機能を音響信号に適用し感情表現を行なわせる「ERSaudio」を実現するには、感情表現に寄与する音響パラメタが必要となる。Justin (2001) によると、「Timber は Valence 軸、articulation は Activity 軸との密接な関係がある」とある。これはオンセットや発声の明瞭性が Valence 軸と Activity 軸に与える影響があると解釈されるので、これを音響信号で実現するために、時間軸における音色の変化に強弱を与えるという特徴を実現させる方法が考えられる。よって、ここでは Spectrum Centroid Variability（以降、SCV）を提案する。

2.2 Spectrum Centroid Variability とは

音楽音響信号における SCV とは、オンセットや発声などが生じた際に、大きく変化する音色の明るさの変動量を意味する。例えばクラシックやピアノのソロ演奏など、比較的緩やかな演奏が行なわれる楽曲においては SCV が小さいといえる。これに対して、ドラムスやエレキギターなどの複数の楽器を用いて演奏される、比較的激しい曲調の楽曲においては SCV が大きいといえる。もし変動が大きい区間をより強調（または抑制）し、変動が小さい区間をより抑制（または強調）することが可能となれば、心理的には「LowActivity/HighActivity」を制御できると期待される。

2.3 提案する SCV の変更方法

音色の「明るい」や「暗い」などの印象は、音響信号のパワースペクトル上でのエネルギーの分布と関係している。そのため、このエネルギーの分布を変化させる事ができれば、音色を相対的に変化させることが出来ると考えられる。そのため、入力された音楽音響信号の Spectrum Centroid（以降、SC）の値を増加または減少させる必要がある。これを実現するためには、音色を明るく、すなわち SC の値を増加する場合は、入力された音楽音響信号のパワースペクトルにおける低周波数成分を抑制させ、高周波数成分を強調させる必要があると考えられる。そこで、本研究では、「明るい音色」や「暗い音色」などの音色の変更を実現するために、音響信号の振幅スペクトルに対して、高い周波数または低い周波数にパワーの分布が偏る様な関数（以降、補正関数）を用いて周波数の補正を行なうことで音色を変化させる。具体的には、STFT によって得られた振幅スペクトルに対して、補正関数を乗算することでスペクトルの補正を行なう。周波数特性を変更するために用いる補正関数を式 (1) 及び (2) に示す。ここで、式 (1) 及び (2) における $n(n = 1, 2, \dots, N)$ は周波数ビンの ID を表している。

$$\alpha_n = 0.95 + 0.1 \frac{n}{N} \quad (1)$$

$$\beta_n = 1.05 - 0.1 \frac{n}{N} \quad (2)$$

各補正関数は入力されたスペクトルの中心周波数を軸に低周波数領域及び高周波数領域に分け、強調または抑制の処理を行なう。式 (1) に示される補正関数は n の単調増加関数、式 (2) に示される補正関数は n の単調減少関数であり、これを入力された音響波形の振幅スペクトルに対して乗算することで、高周波成分または低周波数成分を強調し、低周波数成分または高周波数成分の抑制を行ない、音色を変化させる。この式 (1) および (2) を切り替えることによって、時間軸での変化を再現する。

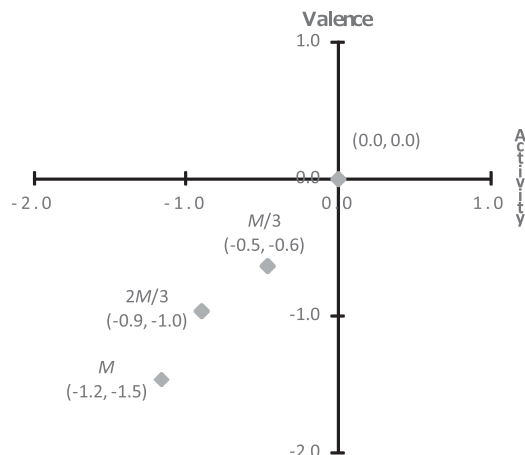


図1 SCV が感情価に与える影響

2.4 SCV が感情価に与える影響の調査

SCV の程度が感情価にどのような影響を与えるかの調査を行なう。そこで、ある楽曲の音響信号と、その SCV を複数のパターンで変更した音楽音響信号を聴取させ、その感情価を被験者に回答させた。使用した楽曲は Pops, Classic, Instrumental の 3 ジャンルから 2 曲ずつの計 6 曲を用いた。提示刺激は元楽曲 1 パターンと、その SCV を $M/3 \sim M$ 倍した 3 パターンの計 4 パターンとした。ここで、SCV における M とは、予備実験によって求められた、各元楽曲の SCV を強調した際に不自然とならない最大の変化率とする。SCV を強調する際、一定の値を最大の変化率とすると元楽曲の SCV の大きさによっては、その変化量が小さすぎて変化を認識できないことや変化が大きすぎてクリッピングを起こしてしまうことがあり、正確な感情価が認識できない場合がある。そのため、各元楽曲における不自然に聴こえない最大の変化率 M を予備実験により個別に設定する。また、回答方法は SAM 法を用いることとした。被験者は計 5 名 (平均年齢 21.8 歳, SD 1.1) であった。

2.5 結果と考察

SCV が感情価に与える影響を図 1 に示す。図 1 に示されるように、SCV を強調 ($M/3 \sim M$ 倍) するにつれて Activity 及び Valence がより小さく、す

なわち感情価がより沈静的かつ否定的に変化することが確認された。これは、SCV の特徴である時間軸における音色の変化に強弱を与えることにより、音高が変わるたびに音色の変動が大きくなるため、音と音の間に断続感が発生していることに起因していると考えられる。つまり、当初の目的であったオンセットや発声が明瞭に聞き取れるようにするのではなく、なめらかな演奏が粗い演奏に聞こえているためであると見受けられる。しかし、図 1 に示されるように SCV は悲しみの感情制御に有効なパラメータであるため、今後はシステムへの実装が望まれる。

3. 発表について

本研究会は、非常に有意義な発表ができたと感じた。他大学の学生や教授の方からの質疑応答ではこれまで考え付かなかったようなご指摘をいただき、また発表後においてもディスカッションが絶えず行なわれた。今後は自身の研究にこの経験を生かしていきたいと考えている。

4. おわりに

最後に、今回の発表を行なうにあたりご指導いただいた三浦雅展講師に深く感謝致します。また、多方面にわたりご支援いただいた多くの方々に感謝致します。