

音楽音響信号の楽曲特徴量に基づいた照明パターン自動生成システムの提案

的場 達矢

Tatsuya MATOBA

1 はじめに

ライブやコンサートにおいて、演奏者のリアクションや舞台照明など多くの視覚的要素が、聴衆の感動を増幅させる効果がある。そのため、大掛かりなコンサートなどでは、照明デザイナーなどの専門家によって楽曲ごとの雰囲気に合わせて照明パターンがデザインされる。以前は、ライトの数に比例する規模の照明制御機器をリアルタイムにて操作していたが、近年ではデジタル化と規格化が進み DMX(Digital MultipleX) 規格のコントローラを用いることで、照明パターンを予めプログラムしておくことが可能になった。

照明装置に関しても、LED が開発され、光量だけでなく光色もプログラムで制御可能な照明システム¹⁾が市販されるようになった。また、LED の低価格化により、音響信号を入力として複数のフルカラー LED の点灯パターンを変化させる製品が市販されるようになってきた^{2, 3)}。これらを利用し、様々な舞台芸術に照明が利用され始めてきた⁴⁾。さらに、音楽と同期して点灯パターンが変化する照明装置(単色、光量変化無し)の開発に関する研究も報告されている⁵⁾。

これらはいずれも、楽曲の情報というよりも、音響情報を何らかの手段により照明パターンに変換することで、楽曲に対応する照明パターンの自動生成を実現している。そのため、楽曲の特徴(曲調)に基づく一定の規則の下に照明パターンを生成することが非常に困難である。これに対して、SMF 形式で記述された楽曲から抽出した特徴を利用し、照明パターンを自動的に生成するシステムの研究もなされている⁶⁾。しかしながら、全ての楽曲が SMF 形式で記述されておらず、より汎用性を持たせるためには音響信号から照明パターンを生成することが必要となってくる。

この問題を解決する為に、音響信号から楽曲の特徴量を抽出し、照明パターンを自動的に生成する手段が提供できれば、様々な応用先が期待できる。本講では、楽曲の繰り返し区間、拍情報及び、楽曲の RMS 情報を音響信号から取り出し、照明切り替えタイミング及び照明点灯パターンを自動的に生成するシステムを提案する。

2 提案システムの概要

2.1 前提条件

適応対象をポピュラーミュージックとし、楽曲信号が記録された Windows 標準音声フォーマットの wav(RIFF Waveform Audio Format) ファイルを利用する。また、使用するものは楽曲は全てのビートが 4 分の 4 拍子であ

るとする。ポピュラーミュージックとはジャズや演歌などの専門的なジャンルを除いた、ヒットチャートの主流の音楽である。適応対象をポピュラーミュージックとする理由は専門的なジャンルの楽曲において特殊な楽曲構造が原因により今回利用する特徴の抽出が困難なためである。wav ファイルは、市販の CD と同じ 16bit, 44.1kHz であるものを利用する。

2.2 システムの全体構成

本システムは楽曲(wav ファイル)を読み込ませることで自動的に照明パターンを生成し、音楽と同期して照明を制御するシステムである。このシステムを利用することで、楽曲信号から音楽の曲調に合わせた照明パターンを生成し、音楽と同期させて再生及び点灯させることが出来る。Fig. 1 に本システムの構成を示す。本システムは、三つのレイヤで構成する。三つのレイヤは上位レイヤから順に音楽分析レイヤ、照明パターン生成レイヤ及び音楽照明再生レイヤの三部分から構成される。音楽分析部では再生事前に楽曲ファイルから音楽の感性情報を抽出し、感性情報ファイルを生成する。照明パターン生成レイヤでは生成された感性情報ファイルを利用し、照明パターンを記述した照明ファイルを作成する。音楽照明再生部では生成された照明ファイルと照明ファイルの生成に利用した音楽ファイルを利用し、音楽と同期して照明を再生する。全てのレイヤは独立して動作し、下位レイヤは上位レイヤの処理が終了している前提で処理を行う。

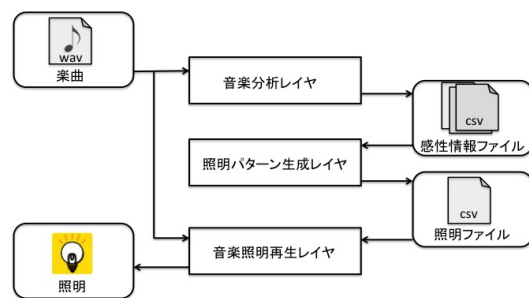


Fig.1 システムの構成

2.3 音楽分析レイヤ

音響信号から感性情報として抽出可能な特徴量はさまざまな情報が考えられるが、本システムで利用可能であ

る特徴量には下記のものが考えられる。

- 楽曲の繰り返し区間

ポピュラーミュージックは一般的にリフレイン形式で構成されている。リフレイン形式とは楽曲の形式の1つであり、サビ (Chorus) と呼ばれる楽曲の中で最も盛り上がる部分の前に A メロ (Verse) 及び B メロ (Bridge) が存在し、それらが繰り返される楽曲の形式である。ここではこのサビ、A メロ、B メロなどを楽曲のセクションと呼び、この繰り返し区間を楽曲における感性情報として抽出する。

- 楽曲の拍情報

ポピュラーミュージックは規則的なリズムで構成されており、ビート (4分音符) が存在する。このビートの位置を楽曲における感性情報として抽出する。

- 楽曲信号の RMS

ポピュラーミュージックにおいて一般に楽曲が盛り上がるほど楽曲信号の RMS は高くなる。これを利用して、楽曲の感性情報として抽出する。

これらの3つはそれぞれ、照明パターンの切り替え、照明切り替えタイミング、照明の盛り上がり度の決定に利用する。以下に3つの抽出手法について説明する。

2.3.1 楽曲の繰り返し区間の抽出手法

楽曲の繰り返し区間は楽曲の繰り返し構造を利用することで検出することが可能である。典型的なポピュラー音楽の楽曲構造は、Fig. 2 のようになっている。ここではそれぞれのブロックをここではセクションと呼ぶ。このセクションの検出に後藤の方法 (RefrainD)⁷⁾ を利用する。以下に楽曲の繰り返し構造の抽出手法を記述する。

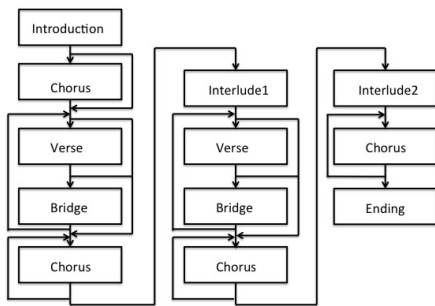


Fig.2 典型的なポピュラーミュージックの構造

まず、最初に入力された音響信号に対して高速フーリエ変換 (FFT) を行い周波数スペクトルを得る。観測区間を時間軸方向に単位時間ずつシフトしながら FFT を適応することで時間軸に対する周波数スペクトル $p(t, f)$ を得る。次に情報量の圧縮の為に音高情報を圧縮した12次元クロマベクトルを生成する。12次元クロマベクトル $v(t)$ とは、人間の可聴領域のうち、主となる約7オクターブの音高情報のパワーを加算したものである。生成した

クロマベクトルの時刻 t と時刻 $t - \text{lag}$ の類似度を算出した後、後藤の方法⁷⁾にてノイズ除去を行った後、ハフ変換を用いることで楽曲の繰り返し構造を検出している。

しかしこの方法では楽曲の繰り返し区間を検出することは可能であるが、その繰り返し区間が楽曲に対し、どのようなセクションであるかを検出することは出来ない。ここでは楽曲中における相関が高い部分、その開始時刻、終了時刻を感性情報として抽出する。今回はこれを照明パターンの切り替えに利用する。

2.3.2 楽曲の拍情報の抽出手法

楽曲の拍情報を抽出する為にビートトラッキングを行う。ビートトラッキングとは、人間が音楽に合わせて手拍子を打つように、曲のビート (4分音符) の位置を認識する技術である。ビートトラッキングにおいては BeatTrackingSystem (BTS)⁸⁾ を利用する。BTS を用いた結果、楽曲の全てのビートの表拍と裏拍の発音時刻を検出し、それを楽曲の特徴量として抽出する。今回はこれを照明切り替えタイミングに利用する。

2.3.3 楽曲の RMS の抽出手法

次に楽曲の RMS の抽出手法について述べる。RMS は変化する値の強度を統計系的に示す値であり式 (1) で表される。

$$N \text{ 個の数 } \{x_{t+1} + x_{t+2} + \dots x_{t+N}\} \text{ に対して}$$

$$x_{RMS}(t) = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2(1)}$$

(t : 時刻, N : 1 フレームに含まれるサンプル数)

今回は1フレーム4096点、約93msとし楽曲中の全てのフレームにおけるRMSを算出し、楽曲の特徴量として抽出する。Fig. 3に楽曲から抽出したRMSの例を示す。

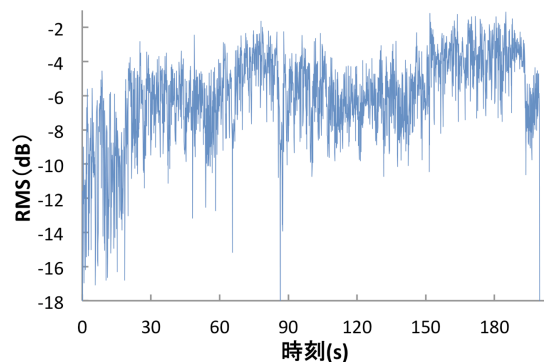


Fig.3 楽曲から抽出した RMS の例

今回示した楽曲におけるサビ区間は68sから85s、152sから194sである。Fig. 3から、サビにおけるRMSの平均値が楽曲において最も高い値を示していることがわかる。今回はこれを楽曲の盛り上がり度の情報として利用する。

2.4 照明パターン生成レイヤ:照明変更頻度の決定手法

音楽分析レイヤで抽出した感性情報は、楽曲における繰り返し区間の開始時刻、及び終了時刻、楽曲の全てのビートタイム、楽曲の RMS 情報である。これを利用して照明パターンを生成する。一般的に楽曲が盛り上がる区間では音数が増加する特徴が見られる。これを情報として利用するために音楽分析レイヤで抽出した各繰り返し区間に対して、RMS の平均値を求める。しかし各区間における RMS の平均値は楽曲のジャンルや作成された時代などにより分布が異なるため、一定の閾値を用いて楽曲の盛り上がりを判断することは出来ない。そこで各繰り返し区間に対して求めた RMS の平均値の上下関係からその区間における楽曲中での相対的な盛り上がり具合を判断する。ここで各区間における RMS の平均値について述べた先行研究が見当たらないため下記の予備実験を行った。

2.4.1 予備実験

下記に示す 20 曲のポピュラーミュージックの 8 小節以上の繰り返し区間における RMS の平均値を求め、それらを比較する。RMS は 2.3.3 の手法で抽出を行う。

曲名	アーティスト
蕾	コブクロ
波乗りジョニー	桑田佳祐
ガラスの少年	kinki kids
BREAK OUT!	東方神起
Believe	嵐
嘘	シド
世界に一つだけの花	SMAP
Butterfly	木村カエラ
keep the faith	KAT-TUN
ふたつの唇	EXILE
Progress	浜崎あゆみ
SAKURA	いきものがかり
カブトムシ	aiko
Squall	福山雅治
愛をこめて花束を	superfly
キセキ	GreeeeN
ギミギミック	RADWIMPS
さあ	surface
等身大のラブソング	Aqua Timez

予備実験の結果、各繰り返し区間における RMS の平均値を求め、他の区間と比較を行った所、最大の値は 95 % でサビ区間を示していた。よって、各繰り返し区間における RMS の平均値を求め、その最大値を含む区間をサビと判断し、楽曲において盛り上がる区間 (Up Section) と判断する。楽曲においては盛り下がる区間も存在し、盛り下がる区間では音数が他の区間よりも減少し RMS の平均値が低くなる。しかし、楽曲において盛り下がる区間が繰り返し区間に含まれていない特殊な区間である可能性も存在する。そこで検出した繰り返し区間の間に存在す

る繰り返されていない区間それぞれにラベルを付け、そのそれぞれの区間に対する RMS の平均値を算出し、繰り返し区間における RMS の平均値と比較することで楽曲中において相対的に盛り下がる区間を判定する。楽曲中に盛り下がる区間が複数存在することは望ましくないため、楽曲中に盛り下がる区間は1か所とし、最も RMS の平均値が低い区間を盛り下がる区間 (Down Section) として判断する。ここまでの処理で楽曲における全ての区間は、盛り上がる区間 (Up Section)、盛り下がる区間 (Down Section)、その他の区間 (Other Section) の3つに判別される。その3つの区間に別々の照明変更パターンを割り当てる。

2.5 各区間における照明パターンの割り当て手法

2.5.1 Up Section

Up Section は盛り上がる区間であると同時に聞かせる区間であるため、使用する光色が複数になると、楽曲を妨害することが考えられる。そのため、使用する光色は一色とする。しかし、照明変化によりビートを感じることで照明点灯が魅力的に感じる事が考えられるため、予め用意した照明点灯パターンをビートに合わせて変更する。今回は予め Up Section 用の2つの点灯パターンを用意した。この2つの点灯パターンをビートの表拍として交互に点灯させる。照明変更頻度は少なすぎると盛り上がりを感じられないが多すぎると不快と感じる為、ここでは2分音符で照明を変更する。

2.5.2 Down Section

Down Section は聞かせる区間であるため、Up Section と同様に使用する光色は一色とする。また、Down Section において照明は Down Section の開始時の最初の表拍で点灯する。

2.5.3 Other Section

Other Section は Up Section, Down Section に対し楽曲の中で比較的聞かせる区間ではないため、使用する光色はメインとなる光色一色及び、アクセントとして用いる修飾色一色とする。Other Section では Other Section の開始後、最初の表拍で照明はメインの光色で格子状に点灯し、その後、表拍ごとに点灯していない照明が1つずつ修飾色で点灯していく。メインの光色が点灯してから4小節後、つまりメインの光色が点灯してから8つ目の表拍でメインの光色が格子状に点灯する配置を変更し、修飾色を全て消灯する。その後、再び表拍ごとに点灯していない照明が1つずつ修飾色で点灯していく。修飾色で点灯する照明は前回点灯した修飾色の照明から遠い位置にあるものが点灯するように割り当てておく。

2.6 照明パターン生成レイヤ:光色の決定手法

照明パターン生成レイヤでは音楽分析レイヤで抽出した感性情報から節 2.4 で分類した各繰り返し区間に対してそれぞれに異なる光色を割り当てる。音楽の感性と色の間には関係性があることが考えられる。先行研究⁵⁾では、事前アンケートにより音楽の印象に感性的名を調査し、

感性名に対して、マンセルの色相環 (Fig. 4) に対して滑らかに変化させている。また先行研究から圧迫感や軽快さを感じる音楽はテンポが速く、華やかさや安らぎを感じる音楽はテンポが遅いことがわかっている。これを利用し、音楽のテンポにより中心色を Table 1 のように決定する。楽曲をマンセルの色相環を中心色を中心に点灯させる。Up Section では中心色が点灯し、Down Section では中心色に対し、マンセルの色相環において、最も遠い色が点灯させる。また、Other Section における光色は Up Section と Down Section で使用した光色から遠い 2 色をメインの光色と修飾色に割り当てる。Other Section において、メインの光色と修飾色は区間ごとに入れ替わるものとする。

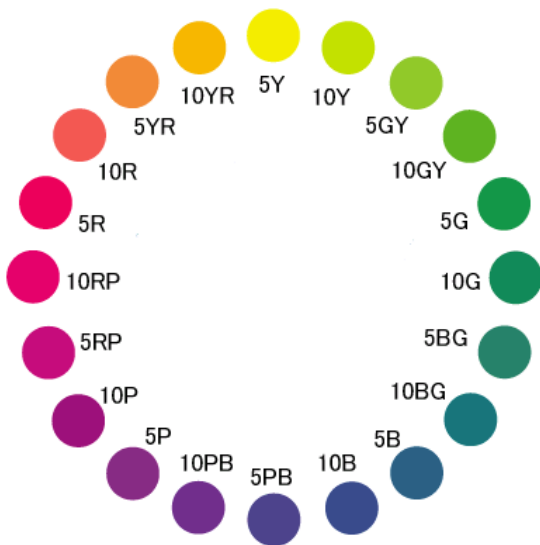


Fig.4 マンセルの色相環

Table1 テンポと中心色

BPM(Beat Per Minute)	中心色
180-	赤 (5R)
150-179	黄 (5Y)
130-149	緑 (5G)
100-129	青 (5B)
-99	紫 (5P)

2.7 音楽照明再生レイヤ

音楽照明再生レイヤでは照明パターン生成レイヤで生成された照明パターンを音楽と同期させて再生させる。照明パターン音楽と同期して照明を点灯させる場合、重要な要素としてあげられるのが照明の点灯タイミングである。照明パターン生成レイヤでは照明点灯場所と照明の光色、及び明るさが照明点灯時刻とともに表記されている。時間 SMPTE タイムコード形式の 30 ノンドロップフレームを点灯タイミングの時間形式に利用する。照明の点灯タイミングの制御に SMPTE タイムコード形式

とは映像、音響機器の同期および編集で用いられる時間形式である。時間、分、秒、フレームの 4 つの要素で構成され、フレームが最小単位である。

本システムは、リアルタイム制御を行うため、動作上のレイテンシは可能な限り小さくする必要がある。そこで、音楽照明再生レイヤの照明点灯タイミングの制御アルゴリズムを次のように定義する。

1. n 回目の照明制御時刻を $t(n)$ とする。 ($n= 0,1,\dots,N$)
2. n 回目の照明制御を行う。
3. $T(n)=t(n+1)-t(n)$ とし、 $T(n)$ 時間停止する。
4. n を 1 増加させる。
5. n が N でない場合、2 に戻る。

上記のアルゴリズムにより、CPU の使用回数を削減し、その結果動作上のレイテンシを削減する。

3 実験

3.1 実験環境

同志社大学香知館 KC111 で実験を行い、実験室における格子状に配置されたフルカラー LED29 灯の照明 (Fig. 5) を利用する。照明の色は Fig. 4 のマンセルの色相環に示す 20 色を予め用意し、本システムを利用し照明パターンを生成する。

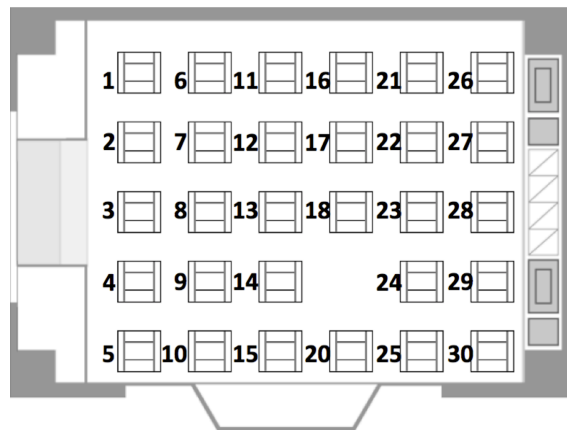


Fig.5 照明配置図

3.2 実験方法

今回は評価のために、室内を白色蛍光灯により 700lx で点灯させた場合 (状況 1)、一定時刻 (今回は 5 秒とする) ごとにランダムに照明の色を変更し点灯させた場合 (状況 2)、本システムを利用して作成した照明パターンで点灯させた場合 (状況 3) の 3 つの状況で被験者に楽曲を聴いてもらう。各状況で試験を行うごとに Table.2 に示す感性語を用いた 7 段階の SD 法のアンケートに答えてもらう。また、被験者が部屋中央後方に着席した状態から実験を開始する。被験者に下記の実験手順を予め説明した状態で実験を開始する。

実験手順を下記に示す。

1. 状況1で楽曲を聴く。
2. アンケート項目を解答する。
3. 状況2で楽曲を聴く。
4. アンケート項目を解答する。
5. 状況3で楽曲を聴く。
6. アンケート項目を解答する。
7. 楽曲を変更し,1に戻る。

なお、アンケート項目は下記に示す。

Table2 評価実験に用いる感性語対

因子名	感性語
軽快感因子	軽快な-重厚な リズムカルな-リズムカルでない
華やか感因子	華やかな-華やかでない 快い-不快
安らぎ感因子	安らぎのある-安らぎの無い 美しい-美しくない
圧迫感因子	圧倒される-圧倒されない 迫力がある-迫力がない
その他	照明が音楽に合っている-合っていない 雰囲気が良い-雰囲気が悪い 楽しい-つまらない 感動する-感動しない 魅力的である-魅力的でない

参考文献

- 1) Philips Color Kinetics, "Color Play1.8". <http://colorkinetics.com/ls/controllers/colorplay/>
- 2) タカラトミー,"Flower Rock 2.0". <http://www.takaratomy.co.jp/products/flower-rock/whats/index.html>
- 3) Qiaohua Electric Co. Ltd, "Strobe Lights". <http://www.qiaohua.com/>
- 4) FOTB2011 MAIN TITLES. http://www.refinedcolors.com/refined_J.html, 2011 Flash On The Beach".
- 5) 合志和洋, 清田公保, 三好正純, 古賀広昭. 音楽の印象に合わせた照明制御システム. Vol. 35, , 2008. 熊本電波工業高等専門学校 研究紀要.
- 6) 小長谷実希, 水上嘉樹, 松田 憲. SMF 形式の楽曲に基づく照明パターン自動生成システムの開発. Vol. 243, , 2009. 画像電子学会.
- 7) 後藤真孝. リアルタイム音楽情景記述システム:サビ区間検出手法. Vol. 47, No. 100, pp. 27-34, 2002. 情報処理学会, 音楽情報科学研究会, 研究報告.
- 8) 後藤真孝. An Audio-based Real-time Beat Tracking System for Music With or Without Drum-sounds. Vol. 30, No. 2, pp. 159-171, 2001. Journal of New Music Research.

4 今後の展望

本講では、音楽音響信号から楽曲の感性情報として3種の特徴量を抽出し、照明パターンを自動的に生成するシステムを提案した。本システムは楽曲の全てのビートが4分の4拍子であるポピュラーミュージックを参照し、楽曲から楽曲の繰り返し区間、ビート情報、RMS情報する。抽出した情報量を用いて楽曲に合わせた照明パターンを生成し、音楽と同期し再生する。今後の展望として本システムを実装し、実験を行い、システムの改良に役立てる。

本システムは繰り返し区間とRMS情報を特徴量として用いることで、楽曲の音楽的な盛り上がり合わせた照明パターンを生成する。また照明の制御時刻にビート情報を利用しているため、楽曲に合わせて照明が点灯する。これにより、音楽と同期させて照明を点灯させることで音楽を楽しむことが出来ると考えられる。しかし、本システムでは各繰り返し区間における照明点灯頻度、光色の中心色を予め決定しているが、これらは筆者が作為的に設定した感性パラメータである。この感性パラメータは楽曲や利用者の感性により最適であると感じられるものが異なり、一意に決定することが困難である。そこで今後の研究課題として、感性パラメータの学習に対話型遺伝的アルゴリズムを導入すること検討している。