

DIVA: 画像の印象に合わせた音楽自動アレンジの一手法の提案

大山喜冴 伊藤貴之
お茶の水女子大学 理学部情報科学科

1. 概要

映画やテレビCMの制作において、画像と音楽は密接な関係にある。また例えば個人でも、「自分で制作したホームページに印象の合った音楽を載せたい」という感想は、多くの人が一度は持つような感想だろう。しかし音楽に精通していない人は、画像に印象の合う音楽を選べない場合があり、画像と音楽の相関性のない一見不釣り合いなホームページをつくる場合が多い。

最近では画像に合った音楽を検索できるシステムの研究[1]や、画像に合った音楽の自動作曲システムの研究[2]が盛んに行われている。しかし、入力画像の雰囲気マッチする音楽が、必ずしもシステムに登録されているとは限らない。また、このシステムにおいて選曲された楽曲や自動作曲された音楽が、ユーザーの好きな作曲者・演奏家による楽曲、ユーザーの好みのメロディである保証はない。

この問題を解消するために我々は、ユーザーが任意の音楽と画像を入力した際に、画像の印象に合わせて音楽を自動アレンジする手法の研究を進めている。本研究は現時点では、画像から色分布を算出し、その色分布から連想されるリズムパターンを導き、そのリズムパターンを音楽に適用して自動アレンジする。

本手法ではまず前処理として、被験者に質問を課し、その被験者が所定の色から所定のリズムパターンを連想する度合いを「連想度」として数値化する。続いて任意の画像を入力し、その画像中における所定の色的重要性を算出し、それにリズムパターン連想度を乗じ、これを累算する。この処理により、所定のリズムパターンの入力画像からの連想度を算出する。この連想度が最大であるリズムパターンを用いて、入力音楽を自動アレンジする。

2. 関連研究

音楽の持つ感性と画像の持つ感性をマッチングにする方法には、多様なメディアデータに対し書誌情報による通常の検索に加え感性検索も可能とするマルチメディア感性データベース管理システムの研究[3]や言語情報と画像情報のマッチングし、自由な表現が可能な自然言語を検索キーとした言語・画像型のシステムの研究がある[3]。楽曲からリズム認識をする研究[4]、楽

曲からメロディを抽出する研究[5]も本研究の楽曲を解析する際に関連付けられそうである。

しかし、画像の印象に基づいて音楽を自動アレンジする研究は今のところ見つかっていない。

3. 提案内容

論文[6]では、音楽の印象と色彩の印象に相関性があることを実証する実験結果が示されている。例えば色相と音楽の印象の相関性には、「赤には迫力のある音楽が似合う」「緑には明るい音楽が似合う」という回答が多いとの結果が出ている。また、明度や彩度と音楽の印象の相関性には、「明度が高くなるに連れて音楽を明るく感じる傾向にある」「彩度の高低と音楽の力強さの度合いが対応する傾向にある」という回答が多いとの結果が出ている。このことから画像の持つ「色」が音楽に与える影響は大きいと考えられる。

また文献[7]においても、画像の持つ印象は主に配色が重要である、という実験結果が報告されている。このことから本研究では、画像の特徴のうち配色に着目する。画像からの他の特徴(例えば画像中の物体形状)の抽出は、現時点では未着手である。

一方、楽曲の構成要素には「調(長調/単調)」「テンポ」「旋律(上昇/下降)」「音高(高/低)」「和声(単純/複雑)」「リズム(固定/流動)」などがある。文献[4]では、この構成要素の中でも「リズム」「旋律」「和声」「音高」の順で印象に残りやすく、楽曲を特徴付ける大きな要素となっていると述べている。このことから本研究では、音楽の構成要素のうちリズムパターンに着目してアレンジを行う。他の構成要素は、現時点ではアレンジには用いていない。

以上の背景により本研究では、入力画像中の色分布を用いて、画像からの連想度が高いと推定されるリズムパターンを特定し、これを音楽に適用させて自動アレンジすることを考える。

まず本研究では準備段階として、被験者に十分に多くの色を提示して、各々の色に対して印象の近いリズムを選択させる。ここで回答対象となる色の制定方法について述べる。まず、色特徴の解釈手法はHSV色空間を用いることにする[8]。本研究では現段階では68色を被験者に提示している。68色の内訳は以下の通りである。有彩色については、太陽光をプリズムで虹色7色に分割すること[9]から7段階に、彩度および明度を3段階に分類する。無彩色は5段階に分類する。以上の処理により本研究では、被験者に提示する68色を得る。本研究ではこの68色と同時に、複数のリズムパ

“Digital Image Varies Arrangement”

Kisa Ohyama, Takayuki Itoh

Ochanomizu University

{kisa, itot}@itolab.is.ocha.ac.jp

ターンを被験者に提示し、各色に対するリズムパターンの連想度を回答させる。本研究では、 i 番目の色に対する j 番目のリズムの印象の近さを R_{ij} と記述する。

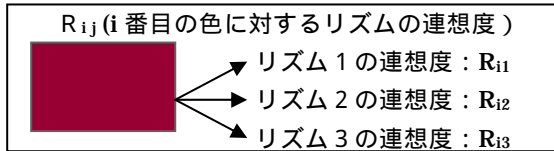


図 1 : 68 色それぞれが持つリズム連想度

続いて、画像を 68 色に分類し、その画像中における 68 色の重要性の割合を算出する。ここで色重要度は

- ・ 画像中に占める面積
- ・ 隣り合う色との差分
- ・ 配置場所
- ・ 色自体の持つ印象

等により算出される、その色の画像中における印象の強さを意味する。

本研究では、ある画像中における i 番目の色の重要度を C_i と記述する。このとき、この画像に対する j 番目のリズムパターンの連想度は、式(1)で表現される。

$$C_1R_{i1}+C_2R_{i2}+\dots+C_{68}R_{i68} \dots(1)$$

本研究では各々のリズムパターンに対して、式(1)を用いて連想度を算出する。そして、この値が最も大きいリズムパターンを、入力画像に最も印象の近いリズムパターンであると判断し、自動アレンジに用いる。

4. 実行結果

本研究ではまず準備段階として、同じメロディ A に対して異なる 7 種類のリズムパターンを適用した 7 曲を、被験者に鑑賞させた。同時に所定の 68 色を被験者に提示し、各々の色からどのリズムパターンを連想したかを回答させ、この回答結果から値 R_{ij} を得た。同様に上記の実験で行った楽曲とは違うメロディ B に対して、メロディ A と同様に 7 種類のリズムパターンを適用した 7 曲と画像 20 枚を提示し、各々の画像からどのリズムパターンを連想したかを選択させた。20 枚の画像は様々な風景画像を用い、メロディ A, B は著者の一人である伊藤によって作曲された楽曲を使用した。アンケート回答者は男女 7 人である。

続いて本研究では、この 20 枚の参照画像について、所定の 68 色の各々が占める重要度 C_i を算出し、式(1)を用いて各々のリズムパターンの連想度を算出し、連想度が最大であるリズムパターンを自動選択した。この自動選択されたリズムパターンと、準備段階で被験者が画像から選択したリズムパターンとを比較した。

この実験では各々の被験者の回答から導いた R_{ij} 値の傾向が被験者ごとにより似通っていないことに対し、画像から選ばれたリズムパターンは非常に似通っていた。また、68 色から選ぶ場合でも画像より選ぶ場合でも、リズムパターンの選ばれる頻度に非常に偏りがあ

った。あまり選ばれないリズムパターンはメロディと印象の合わないものである場合が多かった。画像から被験者が選んだリズムパターンと式(1)を用いて参照画像から連想されたリズムパターンの一致率は現時点で 60%であり、まだ改善の余地があると考えられる。一致率の高い画像は彩度や明度の低い色の面積が多い画像であった。反対に明度・彩度が高い色が多く使われている画像では一致率が極端に低かった。

5. まとめ

本論文では、画像から色分布を求め、その色分布から最も連想されるリズムパターンを用いることで、画像に印象が近くなるように音楽を自動アレンジする手法を提案した。

今後の課題として

- ・ 画像からの配色以外の特徴抽出
- ・ リズムパターン以外の構成要素を変化させる音楽の自動アレンジ
- ・ 楽曲の持つ雰囲気も考慮し、より多彩な楽曲にアレンジができるようなリズムパターンの作成
- ・ 学習アルゴリズムなどの適用により、よりユーザーの好みを反映するシステムの研究開発

なども考慮していきたいと考えている。

謝辞

本研究の被験者の方々に感謝の意を表します。

参考文献

- [1] 古賀、下塩、画像に合った音楽の選定技術、ヒューマンコミュニケーション基礎研究会技報、平 11-9、1999.
- [2] 佐藤、英画像から音楽を自動演奏「ピクチャーメロディー」v1.2
- [3] 坂井、大塚、宮崎、マルチメディア感性データベース YAMAKAN、第 13 回データ工学ワークショップ (DEWS2002)、2002.
- [4] 感性(印象語)語による検索
<http://www.slis.keio.ac.jp/~ueda/semi/99onsei.html>
- [5] 武田、西本、嵯峨山、確率モデルによる多声音楽演奏の MIDI 信号のリズム認識、情報処理学会論文誌, Vol. 45, No. 3, pp.670-679, March, 2004
- [6] 安達、岩宮、色彩と音楽が互いに及ぼす影響--ショパンのエチュードを手がかりに、第 5 回学生のための研究発表会講演論文集(日本音響学会九州支部), pp.13-16、2003
- [7] 北島、土居、画像の構成を手掛かりとした音楽の検索、平成 15 年度情報処理学会関西支部支部大会 講演論文集, pp39-42, 2003.
- [8] 原田、感性語句を用いた自然言語文による画像データベースの対話的検索、静岡大学博士論文、工博甲第 175 号、1995.
- [9] 大林、銀河の道 虹の架け橋、小学館、1999.