

楽師 μ ～画像情報を用いた音楽変換と、 変換音楽からの画像復元システム～

鈴木 正敏 川角 和弘 酒光 真理 濱川 礼

中京大学 情報理工学部

1 概要

本稿では、画像情報とユーザが入力する画像の雰囲気情報から音楽を生成する手法について述べる。

画像と音楽は私たちの身近なものである。しかし、独立した扱いがみられ相互関連がない。そこで、画像情報と画像の雰囲気を取り入れた音楽を生成することにより、画像と音楽との一体化を図ることで、両者を関連づけられるのではないかと考えた。

また、画像情報を保持した音楽を生成することにより、生成された音楽から元の画像へ復元する。それにより、限られたコミュニティ内に向けて公開したい画像を、音楽ファイルで受け渡すことが可能である。

2 関連研究

2.1 既存の研究

画像情報から音楽を生成する手段として、画像に含まれる要素と音楽に含まれる要素の相互関係が導きやすい要素を関連付け、画像の雰囲気に近づけていく方法があげられる。その例として、画像情報から音楽情報を作る実験[1]が挙げられる。この研究では、音楽の三大要素や音の強弱、音の長さ、曲のテンポなどの音楽の要素が、画像の要素である色相・彩度・輝度などに関連付けられている。

2.2 画像情報から音楽を生成するソフトウェア

画像から音楽を生成するソフトウェアとして、i2sm が挙げられる[2]。このソフトは、画像の色情報である色相・彩度・輝度をピクセルごとに分析していき、導き出された数値を一つずつ音に変換する方法を用いたソフトウェアである。

Generation of the music using picture information
Masatoshi Suzuki, Kazuhiro Kawasumi, Mari Sako, Rei Hamakawa
Information science and technology, Chukyo University

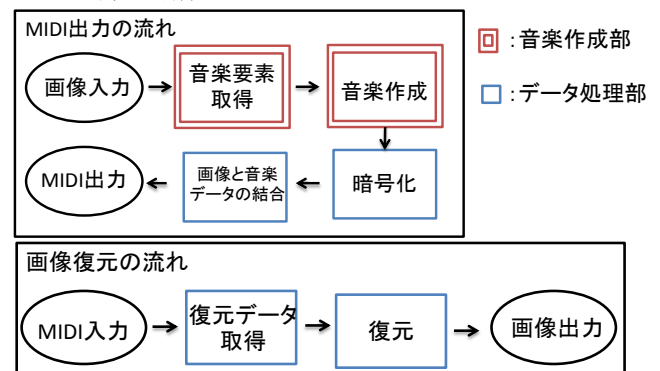
3 本システムの特徴

関連研究の[1], [2]は画像の色情報を用いた相互関係のみから音楽を生成するため、類似した色使いの場合は近似した音楽が生成される。本システムでは、ユーザが感じる画像の雰囲気を入力として音楽生成に反映させるため、画像と音楽の相互関連を高めることが可能である。それにより、画像と音楽の相互関連の意味において、関連研究よりも有効である。また、生成音楽に画像情報を保持するため、相互の関連を結ぶ。

[3]の研究より、音楽のテンポや音量と感情には関連があり、画像の雰囲気と感情が対応すると考えられるため、本研究では画像の雰囲気を音楽のテンポや音量と関連付けた。

4 システムについて

本システムは、大きく分けて音楽作成部とデータ処理部で構成されている。



4.1 音楽作成部

入力画像の幅を行、高さを列として考える。RGB 値を二進数に変換したものに表 1 のように音名が割り当てる。ノートの候補は、音名が割り当てられた行の和を取り、その和が奇数の場合は割り当てられた音名の使用が可能となる。ノートの候

桁数	音名
1	ド
2	レ
3	ミ
4	ファ
5	ソ
6	ラ
7	シ

表 1 音名割り当て

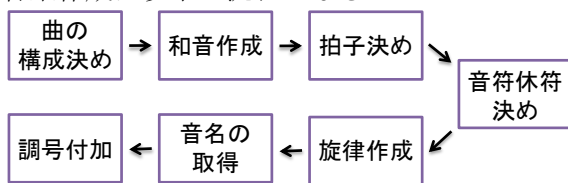
補は1行につき0~7つを列数分, R, G, B 各値についてそれぞれ求められる。

ノート候補では七音であるが, 後述する調号付加により音楽作成時には十二音となる。

ユーザの入力した雰囲気は以下の表のように音要素に変換されて反映される。

	テンポ	音の高さ	音の強さ
激しい	速い	高い	強い
賑やか	やや速い	やや高い	やや強い
穏やか	やや遅い	やや低い	やや弱い
寂しい	遅い	低い	弱い

音楽作成は以下の流れとなる。



曲の構成はあらかじめ用意しておいたパターンの中からランダムで決定する。

和音はコード進行を決めた上で作成する。コード進行は, ダイアトニックコードのみを用いたトニック・サブドミナント・ドミナントを音楽理論に基づきコードを配置することにより作成する。

拍子や音符休符の種類はランダムで生成し小節を埋める。

ノート候補の限定により使用できる音が限られるため, ラグランジェの補間多項式から生成された旋律ラインにより近い音が選択され, 並べられる。

音階は長音階・短音階を用い, 主音を決め, 調号を音楽理論に基づいて付加する。

画像の輝度が低ければ短音階を用い, それ以外なら長音階を用いる。

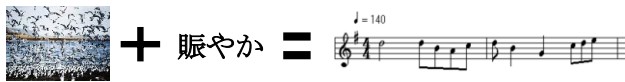


図1 生成されるメロディのイメージ

4.2 データ処理部

画像情報から音楽が生成されたが, 音楽から画像への関連がないため, 音楽生成に用いられたノートを暗号化に用いる。

元の画像は二進数化した R, G, B の各行について, 音名が使われた桁の値, 8 桁目の値, ランダムに決定される行の値を用いて XOR を取ることにより暗号化される。鍵となる値は, 暗号化の際にランダムに決定された行の値である。生成された

MIDI を特定他者に渡す際には RSA を用いて暗号化される。暗号化された RGB 値は暗号画像へと変換される。



図2 暗号化の例

暗号画像は作成された MIDI に結合される。暗号画像は, 生成された MIDI に結合されることにより閲覧が不可能となる。暗号画像が結合された MIDI は, 再生される際に画像データが結合されることによりノイズが混じることはない。

画像の復元の際には, MIDI に結合された暗号画像を分離し, MIDI のノート情報から音名の桁を求め, 使われた音名の桁の値, 8 桁目の値, 鍵の値を用いて XOR をかけることで元画像の RGB 値を求め, 画像へと復元する。

5 まとめ

本稿では, 画像情報とユーザが入力する画像の雰囲気情報を用いる方法で音楽を生成し, また, 音楽ファイルから元の画像へ復元する手法について述べた。

現在, 画像から生成された音楽には画像との関連があるか, また画像と音楽の相互関連は成り立っているかなどの評価項目で評価中である。

6 今後の課題

本システムを複数人に評価結果してもらい, 画像と音楽の相互の関連性を調査し, 調査結果からユーザが感じる画像の雰囲気へ音楽を近づけるためにテンポや音の高さ, 強さを調整していく予定である。

参考文献

- [1] 高田哲雄, 大南崇人: 画像情報から音楽情報を作る実験, 情報処理学会研究報告, Z14-1121, pp. 73-75, (2001-08)
- [2] i2sm: <http://i2sm.numberware.net/>
- [3] P. N. Juslin, J. A. Slobda: “Music and Emotion”, pp. 315 (2001)