

Still Music: Kinect を用いた 空間認識による音楽の自動生成

宮下恵太^{†1} 串山久美子^{†1}

概要：私達の生活する空間に音は常に存在している。本来空間と音とは密接に関わり合っているものだが、それが音ではなく音楽となった場合、その音楽と空間との結びつきについて考えられることは少ない。そこで、空間の情報にもとづき音楽を生成することによって音楽と空間との関係を深めることができるのではないかと考えた。本稿は、Kinect を用いた空間認識情報にもとづき自動で音楽を生成する手法について述べるものである。

Still Music: Automatic Music Generation by Spatial Recognition Using Kinect

KEITA MIYASHITA^{†1} KUMIKO KUSHIYAMA^{†1}

Abstract: Sound always exists in the space we live in. Originally space and sound are closely related. However, when the sound becomes music, it is rare to think about the connection between music and space. Therefore, I thought that the connection between music and space is strengthened by generating music based on space information. In this paper, we propose a method of automatically generating music based on spatial recognition information using Kinect.

1. はじめに

本稿では、Kinect による空間認識技術を用いて、空間の状態から音楽を自動で生成するシステムについて述べる。

私達の生活に音はとても身近なものである。生活音や歩くときの足音など、私達がいる空間には常に音が存在している。川の近くでは川のせせらぎが、街の中では人々の声や車の音が聞こえるなど、その空間と音とは密接に関わり合っている。しかし、店内で流される BGM など、音楽によって空間をデザインする場合、そこで流される音楽とその音楽の流れる空間との結びつきについて考えられることは少ない。そこで、空間をセンシングし、その情報から音楽を生成することによって、その音楽が鳴っている空間と、流れている音楽との結びつきを深めるサウンドデザインの自動生成手法を提案する。

2. 関連研究

あるシステムやルールにしたがって自動的に作曲された音楽を生成音楽(generative music)と呼ぶ。画像情報から音楽を生成する研究として、鈴木らの楽師 μ [1]は画像の RGB 情報を暗号化したデータなどをもとにメロディ、和音を生成し、Midi データとして出力する。また、リアルタイムに音楽を生成するサービスに、〈SUUMO SOUND VIEW[2]〉がある。これは、ユーザのスマートフォンから位置情報を取得し、ユーザの移動速度、進行方向、その地点の混雑度、周辺環境、最寄りの施設、天候などをリアルタイムに音楽に変換

するもので、ユーザは今歩いている街の風景を「聴く」事ができる。

Microsoft の Kinect を音表現に利用した例に、〈The V Motion Project[3]〉、Chris Vik による 〈Dance Controlled Kinect Music (Part1)[4]〉などが挙げられる。これらはいずれも人間の動きから音楽を生成し、人間の動きと音楽を視覚的に結びつけたものである。特に Kinect を用いた空間認識情報から音楽を生成する研究に高橋らの実空間の情報を用いた背景音楽の自動生成[5]がある。これは人間の動きだけではなく空間全体の情報を音楽に変換するものである。

3. システムの特徴

本システムでは通常の楽曲にみられるようなメロディや和音を用いての作曲ではなく、単純なフレーズを何度も反復し、幾重にも重ねていくことによって音楽を組み立てていく、ミニマルミュージックの技法で曲を作る。Kinect の RGB カメラの画像から検出した物体 1 つに対して 1 つのフレーズを割り当て、物体の数に応じて演奏されるフレーズの数が増える。これにより空間の状態と音楽の状態の関係をよりわかりやすく示すことができる。

本システムには、Kinect との情報の受け渡し及び情報処理を openFrameworks で行い、画像処理と物体の検出には openFrameworks 上で動作するオープンソースの画像処理ライブラリである OpenCV を用いた。また演奏全体の制御、フレーズの生成、Midi への出力には Midi データの取扱いが比較的容易な、ビジュアルプログラミング言語 PureData を利用した。

^{†1} 首都大学東京
Tokyo Metropolitan University

4. システム概要

本システムは大きく分けて、全体の演奏をコントロールする制御部、空間上の物体検出、諸データを処理する情報処理部、フレーズの生成、Midi への出力を行う音楽生成部に分けられる。Midi を用いることで、フレーズを演奏する音色などがある程度任意に変更することができる。

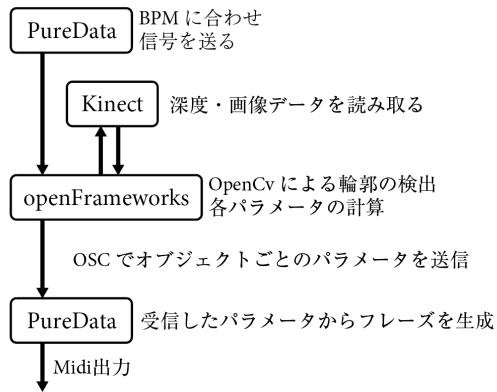


図 1 システム概要図
Figure 1 System Image

4.1 制御部

全体のコントロールは PureData 側の制御部で行う。制御部では演奏のテンポを表す BPM、フレーズの生成に用いるスケールを変更することができる。設定された BPM に合わせ各部に信号を送る。

4.2 情報処理部

空間を認識するためのデバイスとして、前述の Kinect を用いる。空間全体を認識するために Kinect は空間の天井部に設置する。Kinect の RGB カメラから取得した画像を処理し、物体の輪郭を検出する。それと同時に検出した物体の数だけ、フレーズが割り当てられたオブジェクトを作成する。

4.3 オブジェクト

オブジェクトには個別にフレーズ、音色をコントロールするためのパラメータを持たせた。各パラメータはオブジェクトの位置や深度から計算される。

チャンネル: 出力する際の Midi チャンネルを変更する。オブジェクトの中心の RGB 値によって決定する。

ベロシティ: フレーズの音量を決定する。オブジェクトの高さが高いほど音量が大きくなる。

レンジ: フレーズの音の長さを決定する。オブジェクトの Y 座標によって決まる。

パン: 左右のチャンネルごとの音量を決定する。オブジェクトの X 座標によって決まる。

パルス: フレーズが繰り返されるまでの拍数を決定する。オブジェクトの空間の中心からの距離で決まる。中心に近いほど短く、遠いほど長く繰り返される。

オクターブ: フレーズが演奏される音域を決定する。オブジェクトの大きさによって決まり、大きいほど低く、小さい

ほど高い音になる。

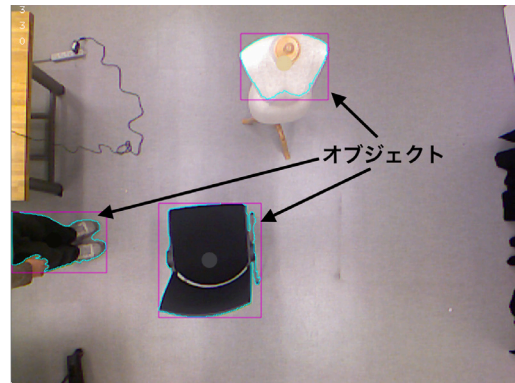


図 2 画像認識画面

Figure 2 Image recognition

4.4 音楽生成部

オブジェクトの各パラメータは、openFrameworks から OSC (Open Sound Control) というプロトコルによって PureData の音楽生成部に送信される。物体を検出すると信号が送られ、あらかじめ設定したスケールとオブジェクトのパラメータにもとづきフレーズがランダムに生成され、設定した BPM で演奏される。オブジェクトが複数作成されることによって演奏されるフレーズが増えていき音楽の厚みを増していく。演奏データは PureData の Midi 出力からリアルタイムで出力される。

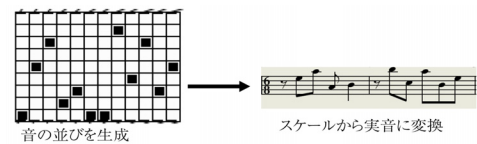


図 3 フレーズの生成

Figure 3 Generation of phrase.

5. おわりに

本稿では、Kinect を用いて空間を認識し、その情報に基づいて音楽を生成するシステムについて述べた。今後は、フレーズだけでなく BPM や調性などにも変化をつけ演奏に展開を持たせることによって、より多様で音楽的な演奏を可能にすることを目指す。

参考文献

- [1] 鈴木 正敏, 川角 和弘, 酒光 真理, 濱川 礼. 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~. 情報処理学会第 74 回全国大会論文講演集. 2012, p.327-328.
- [2] "SUUMO SOUND VIEW". <https://soundview.suumo.jp/>, (参照 2016-12-24)
- [3] "The V Motion Project". <http://www.custom-logic.com/blog/v-motion-project-the-instrument/>, (参照 2016-12-24).
- [4] "Dance Controlled Kinect Music". <https://chrisvik.wordpress.com/2012/03/06/dance-controlled-kinect-music-part-1/>, (参照 2016-12-24).
- [5] 高橋 弦太, 笹岡 久之. 実空間の情報を用いた背景音楽の自動生成. 情報処理学会第 74 回全国大会論文講演集. 2012, p.377-378.