

# 音響データの分析による音の可視化について

内田 早紀<sup>†</sup>, 川島 正行<sup>‡</sup>,

岡山理科大学大学院 総合情報研究科<sup>†</sup>, 岡山理科大学 総合情報学部<sup>‡</sup>

## 1 はじめに

近年科学技術の発展に伴い、様々なものが自動化されてきている。その一方で芸術の分野、特に楽曲の演奏・公演は人間が行うことで、演者の感情や聴く人の心理状態、さらに可聴域と呼ばれる、人間が把握できる音の範囲にも個人差があるため、同じ楽曲でも捉え方が様々である。それをより効果的に表現するとき用いるのが演出である。近年ではプロジェクションマッピングなどの高度な技術を用いたものや大規模公演における聴衆参加型など音楽公演における演出は多岐にわたる。その多くが光と音をリンクさせて効果を生み出している。さらに [1] や [2] にあるようにグローブと呼ばれる目には見えないが人間が感じる音楽特徴の研究も行われている。

本研究では楽曲の音量と各周波数帯の音量を組み合わせ、楽曲の自動演出を行うための基本的なシステムの構築を行う。

## 2 準備

### 2.1 音と周波数

時間  $t \in I \subset \mathbb{R}_{\geq 0}$  を変数とする実数値関数  $x: I \rightarrow \mathbb{R}$  のことを信号という。  $x(t)$  が連続関数であるときアナログ信号と呼ばれレコードの音量などが例に上げられる。音は空気が振動することによって生まれる波が生み出すものである。本研究では音を音響信号として捉える。

音の高低は振動の速さによって決定される。振動の速さは1秒間の振動回数を表した値、周波数によって表され、単位は Hz で表す。

通常、コンピュータで音を扱う場合には、「標本化」と「量子化」という処理によってアナログ信号をデジタル信号に変換する A/D 変換を行う。ここで標本化とは信号  $x(t)$  の定義域  $I$  をある間隔  $\Delta t > 0$  ごとに分割しその値を求めることである。即ち、信号  $x(t)$  の定義域を  $I' := I \cap \{k\Delta t \mid k \in \mathbb{N}\}$  に制限することに対応する。この間隔  $\Delta t$  のことを標本化間隔という。また量子化とは、

$q_1, q_2, \dots, q_k \in \mathbb{R}$  として、もし  $a_i \leq x(k\Delta t) \leq b_i$  ならば  $x(k\Delta t) = q_i$  のように値域  $x(I')$  の値を離散化することである。この離散化の個数  $k$  は量子化ビット数と呼ばれる。例として音楽 CD の標本化間隔は  $\frac{1}{44100}$  秒であり量子ビット数は  $2^{16} = 65536$  である。

一般に人に聞こえる周波数の範囲のことを可聴域という。可聴域は、低い音で約 20Hz、高い音で約 20kHz までとされている。また加齢にともない可聴域の幅は狭まることが知られている。こうした人に聞こえる音を可聴音と呼び、人の耳に聞こえないほど高い音を超音波、人の耳に聞こえないほど低い音を超低周波音という。

### 2.2 音楽ファイル形式の種類と特長

人の可聴域は約 20Hz から約 20kHz であるが、生演奏を録音したレコードの最大周波数は約 40kHz であり可聴域以外の範囲の音も含まれている。アナログ音源を A/D 変換によってデジタル信号に変換した音楽ファイルの形式は様々であり、それぞれに違った特長を持つ。以下、代表的な音楽ファイル形式の種類とその特長を示す。

ファイル形式	wave	mp3 (128kbps)	FLAC
標本化間隔	$\frac{1}{44100}$	$\frac{1}{44100}$	$\frac{1}{48000}$
量子ビット数	$2^{16}$	$2^{16}$	$2^{24}$
周波数 (kHz)	~ 44.1	~ 15.2	~ 48

表 1: 音楽ファイル形式

### 2.3 周期信号とフーリエ級数

$x(t)$  を周期が  $T$  の周期信号とする。このとき

$$A_0 = \int_0^T x(t) dt,$$

$$A_k = \int_0^T x(t) \cos\left(2\pi \frac{k}{T} t\right) dt,$$

$$B_k = \int_0^T x(t) \sin\left(2\pi \frac{k}{T} t\right) dt$$

を  $x(t)$  のフーリエ係数といい

$$x_K(t) := \frac{A_0}{2} + \sum_{k=1}^K \left( A_k \cos\left(2\pi \frac{k}{T} t\right) + B_k \sin\left(2\pi \frac{k}{T} t\right) \right)$$

On sound visualization by analysis of musical data

<sup>†</sup>Saki Uchida · Graduate School of Informatics, Okayama University of Science

<sup>‡</sup>Masayuki Kawashima · Faculty of Informatics, Okayama University of Science

とおき  $\lim_{K \rightarrow \infty} x_K(t) = x(t)$  が成立するとき,  $\lim_{K \rightarrow \infty} x_K(t)$  を  $x(t)$  のフーリエ級数という. フーリエは周期  $T$  の  $C^\infty$  周期アナログ信号  $x(t)$  はフーリエ級数展開可能であることを示した ([3] [4]). ここで  $C_k = \frac{A_k - iB_k}{2}$  と置くことにより, 上記フーリエ級数は以下のように書くことができる.

$$x(t) = \sum_{k=-\infty}^{\infty} C_k e^{(2\pi i \frac{k}{T} t)}, \quad C_k = \frac{1}{T} \int_0^T x(t) e^{-(2\pi i \frac{k}{T} t)} dt.$$

## 2.4 離散フーリエ変換 (DFT) と窓関数

周期  $T$  の周期信号  $x(t)$  を標本化間隔  $\Delta t = \frac{T}{N}$  で標本化した  $x\left(n \frac{T}{N}\right)$  のフーリエ変換を離散フーリエ変換という:

$$X_k = \frac{1}{N} \sum_{n=0}^{N-1} x\left(n \frac{T}{N}\right) e^{-2\pi i \frac{k}{N} n}.$$

デジタル信号を扱う場合には DFT を用いて周波数分析を行う.

実際に取り扱う信号が仮に周期関数だとしてもその周期が与えられているということなどは考えにくく, 処理対象が非周期信号の場合もある. そのため実数  $T_0$  と  $T_1$  を任意に固定し, 処理対象が区間  $[T_0, T_0 + T_1]$  で周期的であると仮定して処理を進める.

このとき出来上がった信号は一般には連続ではないので, 窓関数と呼ばれる関数を掛けることによって処理対象を DFT が行える形にする. 窓関数には様々なものがあり, ここでは主にハニング窓を使用する.

## 3 研究の方針

本研究では公演の自動演出を目標として mp3 音源から左右の音量  $L$ ,  $R$  と各周波数の音量  $F_i$  ( $i = 0, \dots, 512$ ). 以下の情報を取り出す. 実装には processing3 ver.3.5.3 を用いた ([5]).

人間の可聴域は約 20Hz ~ 約 20kHz だがこれには個人差があり, 高音部の 15kHz を越えるとうまく聞き取れないことも多い. ここでは可聴域のなかで聞き取りにくい音を超高音ということにする.

1 フレームあたり 513 個の周波数帯音量を低音, 中音, 高音, 超高音の 4 つにわけ, ここから 126 個のデータを取り出す. これに左右の音量  $L$  と  $R$  を加えた  $2^7 = 128$  個の数値を基に可視化を行う. その手法として 128 個のデータを組み合わせ人間の心臓を模したものの作成し,  $L$ ,  $R$  と低音, 中音, 高音, 超高音にそれぞれ役割を与える. (図 1) それぞれの役割は以下の通りである.

- $L$  と  $R$  は右心室と左心室の色の変化
- 低音は右心室に入る青い雫の落下速度と大きさ
- 中音は左心室に入る赤い雫の落下速度と大きさ
- 高音は右下の赤い雫の落下速度と大きさ
- 超高音はハートの鼓動

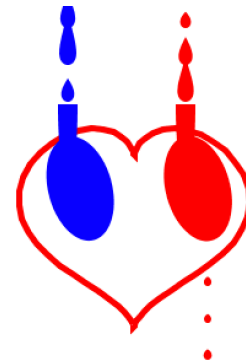


図 1: 心臓

この心臓の動きから, 日常的に聞いている楽曲の直感的なイメージと実際に流れている音量と周波数の関係性を考察する.

## 4 今後の課題

周波数分析にウェーブレット変換を用いることで, フーリエ変換の場合との違いを考察する. また今回は mp3 に音楽ファイルを固定したがハイレゾ音源や実際の演奏でこのシステムがどのようなふるまいをするかを考察する. さらにその評価実験, ユーザインタフェースの構築などを今後の課題とする.

## 参考文献

- [1] 芳賀 直樹, 中山 雅紀, 藤代 一成, “様々な音楽ジャンルにおけるリズムパターンのグローブ可視化”, Vol.2013-MUS-101 No.19, 2013-12-24
- [2] 芳賀 直樹, 中山 雅紀, 藤代 一成, “音楽演奏におけるグローブの可視化”, 全国大会講演論文集 2013(1), 265-267, 2013-03-06
- [3] 高橋 信, 信号処理のための数学, オーム社, 2007
- [4] 城戸 健一, 音響入門シリーズ デジタルフーリエ解析 1 基礎編, コロナ社, 2007
- [5] Daniel Shiffman(訳: 尼岡 利崇), 初めての Processing 第 2 版, オライリー・ジャパン, 2018