

E-030

Doubly Nested Circle of Fifths を用いた和音認識の特性改善 Improvement of Chord Recognition Using Doubly Nested Circle of Fifths

植村 あい子[†] 甲藤 二郎[†]
Aiko Uemura Jiro katto

1. まえがき

和音は自動採譜やカバーソングの検索の手がかりになり、和音は音楽の構成を決める重要な要素であることから、和音認識の研究は多く行われている。そこで本研究は、自動採譜や音楽情報検索への応用を目的として音響信号からの和音認識を試みる。

和音認識の従来研究には、特徴量にクロマベクトル、和声モデルに HMM を用いたもの[1]や、チューニングを行ったクロマベクトルに対しテンプレートマッチングにより和音名を求めるもの[2]などがある。

一方、音楽知識に特徴ベクトルを写像する試みとして、筆者らは Circle of Fifths の調性の類似に着目し、写像によって得られた特徴量を用いて調性の推定を行った[3]。

本研究では[3]の手法を踏襲し、和音の類似性を表す Doubly Nested Circle of Fifths (以下 DNCOF) を用いた和音認識を試みる。本稿では、ベクトルの写像時に調性を反映して改善を行い、得られた特徴ベクトルを用いて和音認識を行い、その有効性を確認した。

2. 提案手法

実音源 DNCOF ベクトルの生成と和音認識までの手順を提案する。この流れを図 1 に示す。

処理は 11250Hz にダウンサンプリングした wav 信号を、複数のフレームに切り出して行う。各フレームは 8192 個のサンプルから構成され、1つのフレームに対して1つの DNCOF ベクトルが求められる。

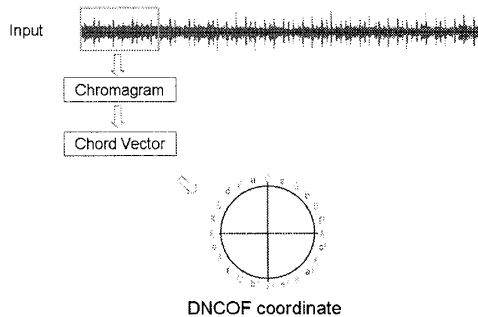


図 1 DNCOF ベクトルと和音認識の流れ

2.1 クロマベクトル

クロマベクトルは、周波数パワースペクトルを特定のピッチクラスに振り分けたものである。

はじめに、各フレームに切り出された信号に対して定 Q 変換[4]を行う。音響信号は 11025Hz にダウンサンプリングをし、ウィンドウサイズは 8192 サンプル(0.74s)、ホップサイズはウィンドウサイズの 1/4 とした。定 Q 変換で得られたスペクトル $X_{cg}(k)$ のパワーのオクターブを吸収して、36bin クロマベクトルを求める。

[†] 早稲田大学大学院基幹理工学研究科

$$CH(b) = \sum_{m=0}^M |X_{cg}(b+12m)| \quad 1 \leq b \leq 36 \quad (1)$$

M は定 Q スペクトルの総オクターブ数、 b はクロマベクトルの bin インデックスである。次にスペクトルのパワーが他の bin にも分配されてしまうことを避けるため、Harte ら[2]の手法を用いてチューニングを行い、フレーム t における 12bin クロマベクトル $Chroma(t)$ を得る。

2.2 コードベクトル

コードベクトルは major と minor の 24 種の和音がどのような尤度を持つかを表す 24 次元のベクトルである。コードベクトルの各要素は、12bin クロマベクトル $Chroma(t)$ を入力とし、その重み付け和とする。和音の構成音のうち、各音の重要度は異なるため、この重要度 w_i を反映させる。この w_i は[5]の和音テンプレートを参考にした。

$$C(t) = \begin{bmatrix} C_C(t) \\ \vdots \\ C_{B_{\min}}(t) \end{bmatrix} \quad (2)$$

$$C_{P_n}(t) = \sum_{i=0}^{11} w_i Chroma_{P_{(i+n) \bmod 12}}(t) \quad (3)$$

$$P_0 = C, P_1 = C\#, \dots, P_{11} = B \text{ または } P_0 = C_{\min}, P_1 = C\#_{\min}, \dots, P_{11} = B_{\min}$$

2.3 DNCOF 平面への射影

Doubly Nested Circle of Fifths[1]は三和音の関係を示している。これは、隣り合う三和音は似ており、対角上の三和音は似ていないという特徴がある。

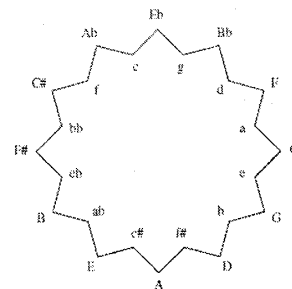


図 2 Doubly Nested Circle of Fifths[1]

ここで、2.2 節で得られたコードベクトルを図 2 の平面に射影する。このとき、調性を反映させるため、ある調で用いられる和音(調固有和音)の成分のみの合成を行う。この手順は次の通りである。

(1) DNCOF を円に見立て、調固有和音の方向を向いたベクトルを基底とする行列 u を構成する。これは角度の異なる 24 個の単位ベクトルを並べた行列 U を調固有和音に基づいてビットマスク w_{key_i} を掛けることで作成される。 w_{key_i} はある調 key における和音 i のマスクである。

