

パート間の調和を利用した合唱の基本周波数推定の検討

A Study of Pitch Estimation for Chorus Using Harmony between Voice Parts

河合 彬弘[†]
Akihiro Kawai

酒向 慎司[†]
Shinji Sako

北村 正[†]
Tadashi Kitamura

1. はじめに

自動採譜に関する要素技術として、音響信号に対する基本周波数推定が盛んに研究されている。特に多重音に対する基本周波数推定は、音源数や調波構造など、収録状況や演奏位置によって変動し得る未知パラメータの存在から困難な問題となっている。これまでの研究では、主に楽器音を対象とするものが多かったが、合唱における多重音の基本周波数推定でも同様の難しさがあり、また歌詞によって音響信号が大きく変化する歌声では、別種の難しさがあると考えられる。

合唱の基本周波数を取得することができれば、プロの合唱団の技術向上や小中学校の音楽教育の支援など、様々な応用が期待できる。合唱は、複数の旋律がハーモニーを成しながら進行していくという特徴がある。そこで本研究では、合唱における各パートの基本周波数は調和するような関係にある傾向が強いと仮定し、その傾向を利用した推定精度の改善を試みる。

2. 多重音の基本周波数推定

本研究では音響信号をフレーム分割し、各々のフレームに対してクラスタリングに基づく多重音の基本周波数推定 [1] を行う。この手法ではスペクトルの加法性を仮定し、各々の平均が倍音構造の制約を持つような混合正規分布を用いて多重音のスペクトルから単音のスペクトルを分離する(図1)。基本周波数推定は分離された単音のスペクトルに対して行う。しかし、単音のスペクトルを適切に分離できるような混合正規分布を構成するためには、そのスペクトルの基本周波数がわかっている必要がある。そこで、以下のような反復により推定を行う。

1. 各基本周波数を初期化する。
2. 現時点における各基本周波数の推定値を基に混合正規分布を構成する。
3. 構成した混合正規分布を用いて観測スペクトルを分離し、分離されたスペクトルの基本周波数を推定する。
4. 終了判定を行う。終了しない場合は推定値を更新して2に戻る。

この手法には構成音の倍音構成比に依存しないという利点があり、発声内容によって調波構造が変化する合唱においても有用と考えられるが、基本周波数の反復推定における初期値の設定問題が存在する。

3. 合唱におけるパート間の調和

合唱の趣きの1つとして、歌声の調和を楽しむことが挙げられる。一般に、2つ音が調和するかどうかは倍音成分の重なり具合によって決定されることが知られている。すなわち、人間の耳では基本周波数の比が単純な整

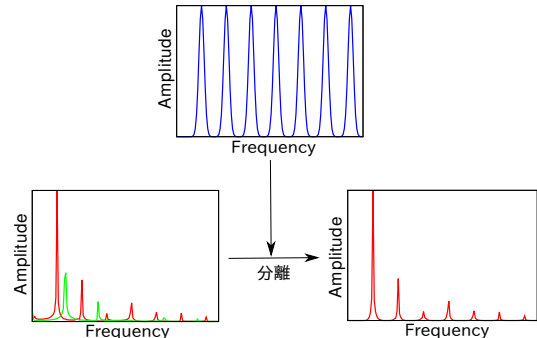


図1: 平均が倍音構造の制約を持った混合正規分布を用いたスペクトルの分離

表1: 協和音程における基本周波数の比率

基本周波数の比	音程
1:1	完全1度
1:2	完全8度
2:3	完全5度
3:4	完全4度
4:5	長3度
5:6	短3度
3:5	長6度
5:8	短6度

数となる場合に調和して聞こえる [2]。このような場合の音程(協和音程)を表1に示す。合唱曲は歌声が美しく調和するように意識して作曲されているため、ある音に注目したとき、他の音は協和音程の関係にある傾向が強いと予想される。したがって、前節の手法で反復推定の初期値も主旋律と協和音程のような関係にある基本周波数を与えることで、推定精度の改善が期待できる。

4. 実験

4.1 実験内容

前節で述べた初期値の与え方の効果を確かめるための実験を行った。今回はソプラノを主旋律と見做し、主旋律の基本周波数は既知という条件で、主旋律以外のパートの基本周波数を推定した。初期値は既知である主旋律の基本周波数に対して $1, \frac{4}{5}, \frac{2}{3}, \frac{1}{2}, \frac{2}{5}, \frac{1}{3}$ となる周波数を与えた。また、比較のため初期値を同じ数だけランダムに与える場合の実験も行った。

この段階では、基本周波数を推定したいパート数に関係なく推定結果は主旋律を除いて5個得られる、したがって、不要な推定結果を削減する必要がある。そこで、推定結果から得られる分離された5つのスペクトルから必要なパート数を選ぶ際に考えられるすべての組み合わせに対してスペクトルの和を計算し、分離前のスペクトルとの相関が最も高い組み合わせを選び出すことで、推定結果の削減を行った。

[†]名古屋工業大学, Nagoya Institute of Technology

表 2: 実験条件

サンプリング周波数	16 kHz
フレーム長	128 ms (2048 点)
フレーム周期	64 ms (1024 点)
窓関数	blackman 窓

4.2 実験条件

実験データとしては混声 3 部 (ソプラノ, アルト, テナー) の合唱曲 4 曲 (believe, 贈る言葉, 巣立ちの歌, 翼をください) を用いた。さらに, 音響信号の複雑さに対する推定精度の変化を確かめるため, 各合唱曲に対し以下のような 4 種類の実験データを用意した。

- data 1 : MIDI で単純な波形の音源から作成 (歌声部分のみ)
- data 2 : MIDI で歌らしい音源から作成 (歌声部分のみ)
- data 3 : MIDI で歌らしい音源から作成 (歌声部分 + ピアノ伴奏)
- data 4 : 文献 [3] に収録されている実際の合唱 (歌声部分 + ピアノ伴奏)

その他の実験条件を表 2 に示す。なお, 実験データには音がない区間が含まれるため, 音の有無を区別できるように短時間パワーの閾値を設定し, 各フレームにおける短時間パワーが閾値を上回った場合にのみ基本周波数推定を行った。

正解率の計算は推定結果の削減前と削減後の両方で行った。推定された基本周波数を最も近い音高に割り当て, 1 フレームに 6 つ (または 3 つ) 得られる推定結果の中からパート毎の正解音高系列 $b = \{b_1, b_2, \dots, b_M\}$ に最も近い音高を選び出した音高系列 $a = \{a_1, a_2, \dots, a_N\}$ を作成し, a と b に対して式 (1) と式 (2) を用いた DP マッチングを行い, 式 (3) で正解率を計算した。

$$d(i, j) = \min \left\{ \begin{array}{l} d(i-1, j) + e(i, j) \\ d(i-1, j-1) + e(i, j) \\ d(i, j-1) + 1 \end{array} \right\} \quad (1)$$

$$e(i, j) = \begin{cases} 1 & (a_i \neq b_j) \\ 0 & (a_i = b_j) \end{cases} \quad (2)$$

$$accuracy(\%) = \frac{N - d(N, M)}{N} \times 100 \quad (3)$$

4.3 実験結果

2 つの実験における正解率を図 2 に, 推定結果の一例 (data1 に当たる実験データ) を図 3 に示す。

図 2 を見ると, 音響信号の複雑さの程度に関係なく調和を考慮して初期値を与えた場合の正解率が, ランダムに初期値を与える場合の正解率を上回っていることがわかる。また, 正解率が低くなっている実験データに関しては, 正解音高との音高差が 1 で不正解と判定される推定結果が多いことも確認された。

推定結果の削減に関しては, 実際の合唱を用いた場合と, それ以外の場合では正解率が大きくことになっており, この手法は音響信号の複雑さに結果が大きく左右されることがわかる。また, 分離に用いる混合正規分布の分散の値は予備実験により 20 を用いたが, 分散の変更が削減に与える影響は大きく, 正解率が安定しない傾向が見られた。

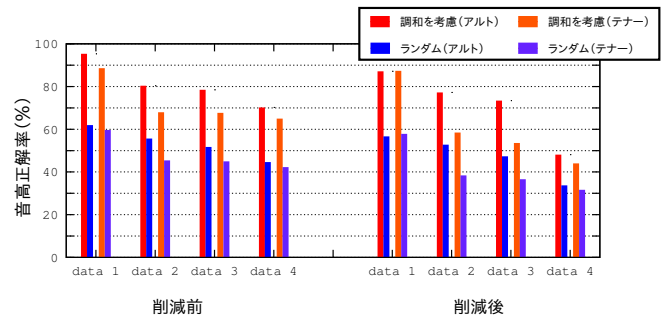


図 2: 正解率 (4 つの楽曲の平均)

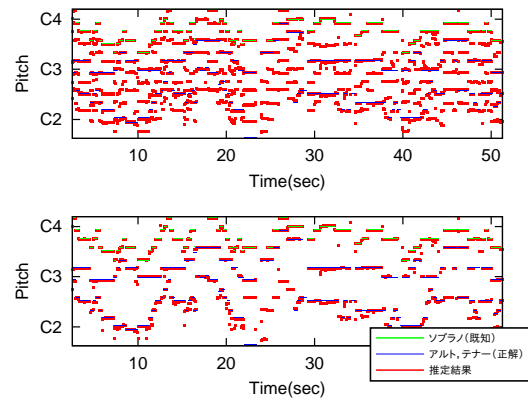


図 3: 巣立ちの歌の推定結果 (上: 削減前, 下: 削減後)

5. むすび

本研究では合唱における基本周波数の反復推定の際に, 主旋律に対して調和する初期値を与える実験を行い, 初期値をランダムに与える場合と比べて推定精度が向上することを確認した。しかし, 今回の実験は主旋律の基本周波数を既知として与えているため, 実用的であるとは言い難い。今後の課題として, 上記の問題を解決するため, 主旋律に対する他のパートの基本周波数の出現頻度を考慮しながら反復推定を行う手法を検討することが挙げられる。

謝辞

本研究の一部は, 文部科学省科学研究費補助金 (課題番号: 21700191) の支援を受けて行われたものである。

参考文献

- [1] 亀岡 弘和 他, “ハーモニッククラスタリングによる多重音の基本周波数推定アルゴリズム”, 情報処理学会研究報告, 2003-MUS-50, pp.33-38, 2003.
- [2] 石桁真礼生, 末吉保雄, 丸田昭三, 飯田隆, 金光威和雄, 飯沼信義, “楽典”, 音楽之友社, 2008.
- [3] 若松 正司, “クラス合唱 旅立ちの日に”, ヤマハミュージックメディア, 2009.