

## 音色空間を用いた音源分離システム

日比 啓文 半田 伊吹 武藤 誠 坂井 修一 田中 英彦  
 東京大学工学系研究科

### 1. はじめに

近年、コンピュータで音楽から楽音を認識し、楽譜をつくる採譜システムの研究がなされている[柏野 96]。しかし、それらのシステムでは楽音同士の周波数成分が重なったときを検出することは難しく、適応処理を施しても誤認識が発生する機会が多い[木下 00]。一方そのような誤認識を減らすため楽器が奏でる楽音を、遷移確率などを利用して一連の流れとして扱うことも提案されている。これを解決する手法として楽器の音色を用いるアプローチは今まででなされていない。本稿では音色空間をもちいて楽器ごとの奏でる楽音を追跡し、楽音の重なりを検出し、それを採譜に適応する手法についての提案を行う。

### 2. 調波構造ストリーム[中谷 95]

ストリームとはある楽器が旋律を奏するとき、その旋律の奏でる音をつなげたもののことをいう。このストリームに着目した音源分離を行う手法として、調波構造ストリームエージェントを用いた手法があげられる。調波構造ストリームエージェント[中谷 95]とは、楽器の音が調波構造からなることを利用して、1) 新しい調波構造ストリームの発見、2) 調波構造の追跡、3) ストリーム間の干渉の低減、4) 調波構造の終了の検知、の過程を行う独立のマルチエージェントである。

ここで注目すべきは3)のストリーム間の干渉の低減方法なのだが、

- 1) 自分が追跡している音成分に対応したフィードフォワード信号をつくり、他のエージェントの入力から引く
- 2) 自分が追跡している音成分に対応した周波数領域の閾値を上昇させる

という処理を行って、精度を向上させている。これが調波構造ストリームエージェントである。

### 3. 楽器の音色空間[McAdams 95]

かねてから、心理学の分野では楽器の楽音は3次元空間にマッピングできると言われてきた。Grey は楽器の音色距離をアンケート等から測定し、それらの距離を元に3次元空間上にマッピングを行った[Grey 77]。しかし、この研究では各軸には傾向は現れたが数値として表現するに至らなかった。そこで、McAdams や Winsberg 等により18種類の楽器を対象にして、88種類の楽音を、CLASCAL アルゴリズムにより3次元空間にマッピングする研究が行われた[McAdams 95]。この研究では、尤度関数からパラメータとし、潜在クラスごとに値を求め、その値を元にして3次元空間を形成している。以下の図1が音色空間である。

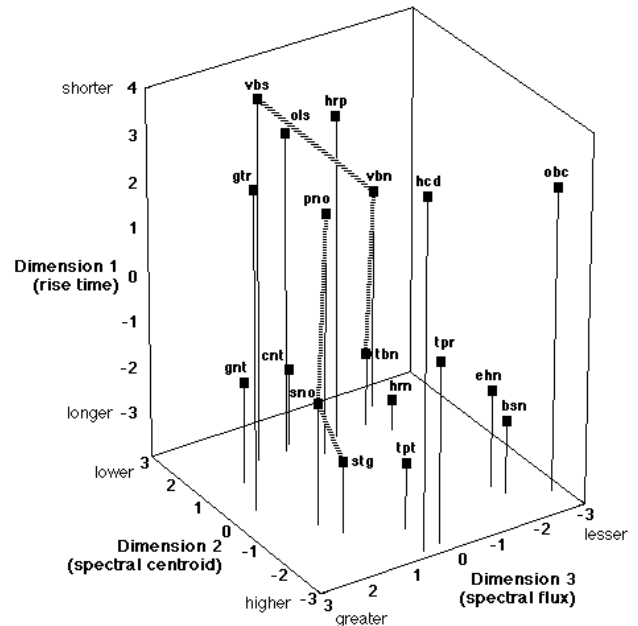


図1:3次元音色空間

### 4. 音色空間を用いた音源分離システム

本研究は複数パートの楽器でそれぞれが単旋律を奏でた音楽を想定している。システムは1)楽器特定部と2)楽音特定部に別れ、1)楽器特定部では音楽を最初から最後まで一通り奏で音色空間の出現位置を記憶し、その出現位置と出現確率から出現する楽器のパート数と楽

器を推定する。2)楽音特定部では、1)楽器特定部で求められた音色出現位置が正しいと仮定し、どの楽器がどの時間にどの楽音を奏でているかを認識する。図2にシステムの概要を示す。

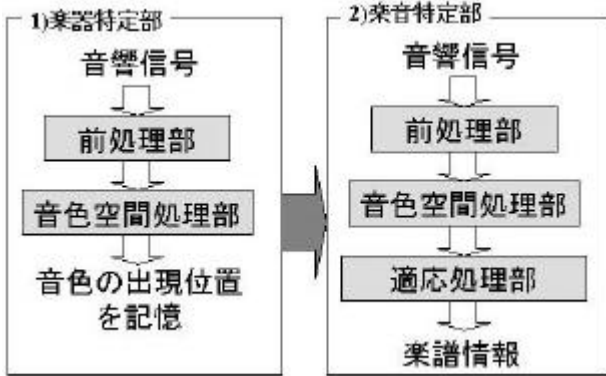


図2:音色空間を用いた音源分離システム

#### 4.1. 楽器特定部

音楽の中で互いの周波数成分が重ならない楽音の出現頻度は、周波数成分が重なる楽音の出現頻度より多いと思われる。このことを利用して楽器推定を行う。まず、a)前処理部で時間周波数解析を行い、周波数成分ごとにトレースを行う。そして、その中で現れる調波構造に着目し、その調波構造を一つの音色情報としてb)音色空間処理部に送りその音色をマッピングする。

b)音色空間処理部において、ある楽器で奏でられかつお互いに周波数成分が重ならない調波構造の音色は、音色空間上に同じもしくは近接した位置にマッピングされる。他方、周波数成分が重なる調波構造の音色は音色空間上に違う位置にマッピングされる。また、周波数成分のどの成分が重なるかで音色空間にマッピングされる位置が変わってくるため、ある位置での出現頻度は非常に小さくなる。周波数成分の重ならない調波構造と重なる調波構造の音色空間における位置を図3に示す。図3の黒点は調波構造の音色をマッピングした点、白点は正しい楽器の音色をマッピングした点とする。よって、出現頻度を求めることにより、楽器の数と種類が決定される。

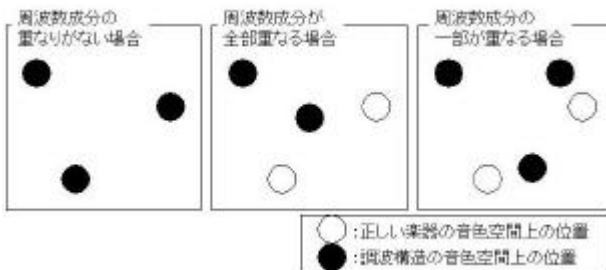


図3:周波数成分の重なり方と音色空間上の位置

#### 4.2. 楽音特定部

まず楽器特定部で、楽器の数と音色の位置が特定できたとする。そしてこの部分では、2で紹介した調波構造ストリームを生成した後、これらを音色空間にマッピングすることにより、楽器の音色ごとに新たなストリームを形成する処理を行う。この部分の処理は、楽器特定部と同じくa)前処理部で時間周波数解析を行い、周波数成分ごとにトレースを行う。そして、その中で現れる調波構造に着目し、その調波構造を一つの音色情報として、b)音色空間処理部におくり、その音色をマッピングする。そして、音色の位置と楽器音色の位置の整合性を求め、その情報をc)適応処理部に送り、それぞれの事象に対応した適応処理を行う。ここで、適応処理は音色空間処理部から得られる情報により、どの楽器とどの楽器の音色が重なっているかという情報に基づき行われる。

#### 5. まとめ

本稿では音色空間を用いて楽器の重なりを認識する手法の提案を行った。この手法によって、どの楽器がどの楽音を鳴らしているかという採譜にとって大きな情報を得られる可能性がある。しかし、現段階では実装及び、実験が行っていないため、このシステムの安定性の議論は行うことができない。今後の課題として1)音楽の中で周波数成分が重なる楽音の頻度と重ならない楽音の頻度の統計を出すこと、2)システムの実装及び実験により、システムの安定性を導き出すこと、3)周波数成分が重なったときの音色空間における位置のずれ具合を導くこと、4)音色空間から得られた情報による効果的な適応処理を考案すること、などが挙げられる。

#### 参考文献

- [柏野 96] 柏野, 中臺, 木下, 田中: 音楽情景分析の処理モデル OPTIMA における単音の認識, 電子情報通信学会論文誌, Vol. J79-D (11), No.11, pp.1751-1761, 1996
- [木下 00] 木下 智義, 坂井 修一, 田中 英彦: 周波数成分の重なり適応処理を用いた複数楽器の音源同定処理, 信学論(D- ), Vol. J83-D- , No.4, pp.1073-1081, 2000
- [McAdams 95] S. McAdams, S. Winsberg, S. Donnadieu, G. De Soete, J. Krimphoff: Perceptual scaling of synthesized musical timbres: common dimensions, specificities, and latent subject classes, Psychological Research, Vol.58, pp.177-192, 1995
- [Grey 77] J. M. Grey: Multidimensional perception scaling of musical timbres, J. Acoust. Soc. Am., Vol.61, No.5, pp.1270-1277, 1977
- [中谷 95] 中谷 智広, 奥乃 博, 川端 豪: 音環境理解のためのマルチエージェントによる調波構造ストリーム分離, 知能誌, Vol.10, No.2, pp.68-77, 1995