

ベイジアンネットワークを用いた弦楽器の音の群化

1140307 岡崎 健士 【 福本研究室 】

1 はじめに

複数の音が混ざりあった音響信号の波形をコンピュータが元の波形に復元することは非常に困難である。一方で人間は複数の音が混ざりあった音響信号から、それぞれの音を見分けることができる。これは人間が一つ一つの音をまとまりとしてとらえることができ、音同士の境界を把握することができるためである。つまり混ざりあった音響信号の分離を行う過程の中で音同士を一つ一つのまとまりとしてとらえることができれば、音響信号の分離を正確に行うことができると考えられる。この音同士の境界を把握することを音の群化という [1]。

本研究では 2 種類の弦楽器の混ざりあった音響信号をグラフ構造を持つ確率モデル (グラフィカルモデル) の一つであるベイジアンネットワークを用いて群化を行う。

2 ベイジアンネットワーク

ベイジアンネットワークとはグラフィカルモデルの一つで確率変数 (ノード) 間の条件付き独立性を有向非循環グラフを用いて表現する。一般的に音声認識のための音響モデルとしては隠れマルコフモデルが広く用いられているが、近年、隠れマルコフモデルよりも柔軟なフレームワークであるグラフィカルモデルの研究が行われつつある。

N 個の確率変数 $X_1, X_2, \dots, X_N$ に対してベイジアンネットワークのグラフが定義されていて、そのグラフに置けるノード X_i の親ノードの集合を $Pa(X_i)$ とする。 $Pa(X_i)$ に変数が与えられた時の X_i の条件付確率を $P(X_i|Pa(X_i))$ と表すと、 $X_1, X_2, \dots, X_N$ の同時確率分布は

$$P(X_1, X_2, \dots, X_N) = \prod_{i=1}^N P(X_i|Pa(X_i)) \quad (1)$$

となる。

このベイジアンネットワークを時間方向に拡張したものをダイナミックベイジアンネットワークと呼び、音声認識の分野で用いられることが多い。ダイナミックベイジアンネットワークは時間軸方向にエッジが存在するということを除けばベイジアンネットワークと同じものであり、また隠れマルコフモデルはダイナミックベイジアンネットワークの一種と考えることができる。このことよりダイナミックベイジアンネットワークは少なくとも隠れマルコフモデル以上の表現力を持っていると考えられる [2]。つまりベイジアンネットワークを用いることにより一度混ざりあった音響信号同士の境界がより正確に分かることが予想される。

3 実験

実験にはエレクトリックギター (ギター) とエレクトリックベースギター (ベース) の弦楽器の音響信号を用いる。この 2 種類の楽器の音が混ざりあった音響信号に対して信号処理を施し周波数解析を行う。周波数解析には一般化調和解析 (GHA) を用いる。解析を行った信号にベイジアンネットワークの処理を行うことで、その出力されたデータが混ざり合う前の音響信号を示しているかどうかを確認する。

また図 1 に実験を行う際の大まかな流れを示す。



図 1 出力するまでの流れ

4 まとめ

本報告ではベイジアンネットワークを用いたギターとベースという 2 種類の弦楽器での音の群化の方法を提案した。今後行うべきこととして、まず現実のデータを用いて構築し、提案が妥当であるかということを検証する必要がある。またベイジアンネットワークが抱える問題点の一つとしてノードの数が膨大になるとノード間の関連性を判断するのが困難になり、また確率テーブルをすべて埋めることも困難になる。推論精度の高いベイジアンネットワークを人の手で作成するのは非現実的であるといえる。

今回は似た 2 種類の弦楽器を想定したが、異なる複数の楽器で行えるようにし、またベイジアンネットワークが抱える問題点を解決することが今後の発展として考えられる。

参考文献

- [1] 亀岡弘和, “電子情報通信学会「知識ベース」”, 2 群 9 編 2 章, 2012
- [2] 小林 隆二, 篠田 浩一, 古井 貞熙, “基本周波数情報を用いたダイナミックベイジアンネットワークによる音声認識”, “電子情報通信学会技術研究報告 105 巻 第 686 号, pp.77-82, 2006