

| | |
|--------------|---|
| Title | 振幅変調情報を用いた楽器音の音源同定に関する研究 |
| Author(s) | 川平, 勇 |
| Citation | |
| Issue Date | 2000-03 |
| Type | Thesis or Dissertation |
| Text version | author |
| URL | http://hdl.handle.net/10119/1331 |
| Rights | |
| Description | Supervisor: 赤木 正人, 情報科学研究科, 修士 |

振幅変調情報を用いた 楽器音の音源同定に関する研究

川平 勇

北陸先端科学技術大学院大学 情報科学研究科

2000年2月15日

キーワード: 振幅変調, 楽器音, 変調 - 搬送周波数マップ, マッチング, 音源同定.

1 はじめに

音源分離問題を聴覚の情景解析 (Auditory Scene Analysis : ASA) の問題としてとらえ直し, それを計算論の立場から議論する研究がある. これらの研究の多くは特徴量として振幅スペクトルを用いているので, 同じ周波数成分を含むような場合, 完全に分離できているとはいえない. そして情景解析問題の半分しか取り扱っておらず, 残り半分の音源同定の研究は数が少ない.

また, 我々が日々耳にする音のほとんどは複雑に時間変化する振幅包絡, すなわち振幅変調 (AM) を伴っている. この振幅変調は音色にも大きな影響を与え, 音源の識別や分離においても重要な情報となっていることが知られている.

本論文では, 時間的に変化する音の振幅変調周波数と搬送周波数を検出する変調 - 搬送周波数マップ検出モデルを用いる. 図 1 (左) は, その例である. 例えば図 1 (右) のように信号成分と雑音成分の搬送周波数成分 (f_c) に重なりがあっても変調周波数 (f_m) の違いを検出することによって図のように信号, 雑音の成分をグルーピングすることができる.

そして, 検出された変調 - 搬送周波数マップとあらかじめデータベースとして登録しておいたマップとのマッチングを行なうことによって楽器音の同定を試みる.

2 変調 搬送周波数マップ検出モデルとシミュレーション

図 2 に用いた変調 - 搬送周波数マップ検出モデルを示す. 変調 - 搬送周波数マップは次のように検出される

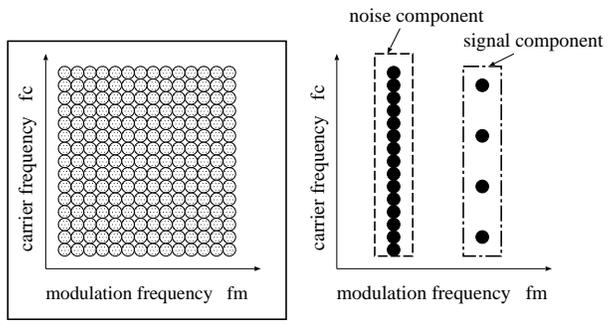
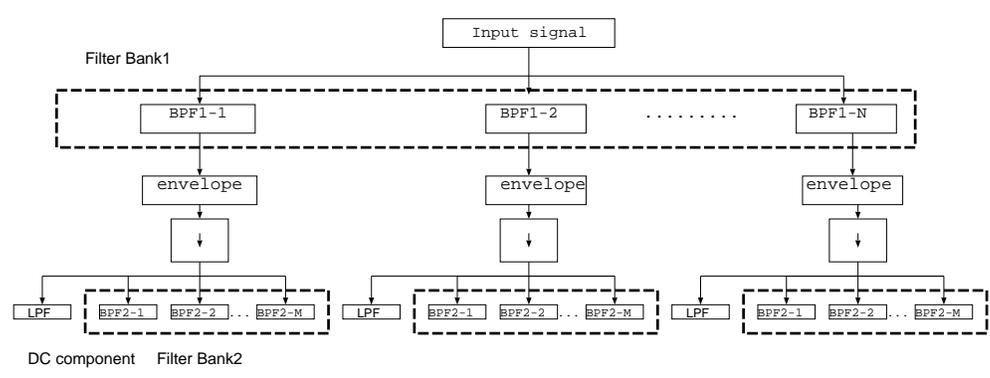


図 1: 変調 搬送周波数マップ (右) とその例 (左)



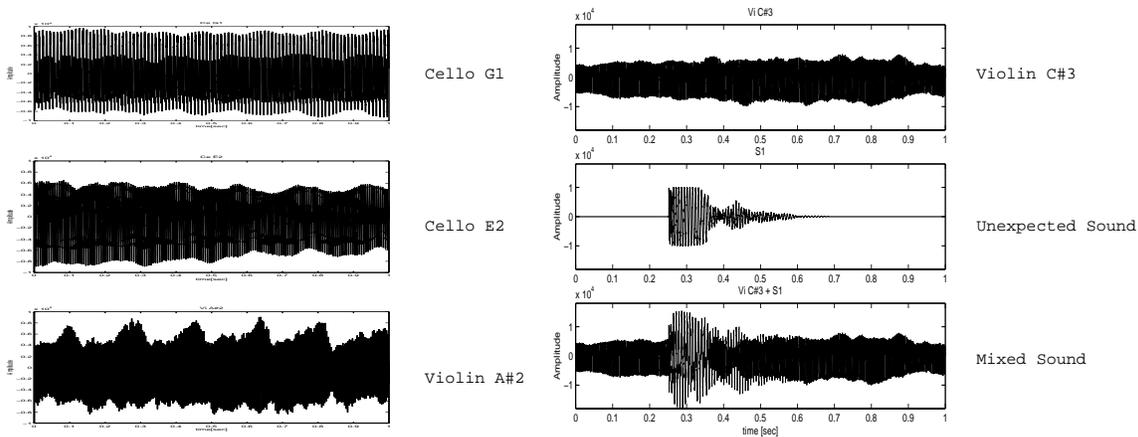


図 3: 解析信号 (左上: チェロ G1 音, 左中: チェロ E2 音, 左下: ヴァイオリン A#2 音, 右段: 混合信号)

分析合成系で構築されている。

次に楽器音の解析を行ない, 各音の特徴を調べた。解析に用いた信号の一例を図 3 に示す。図 3 左側は上から順にチェロ G1 音, チェロ E2 音, ヴァイオリン A#2 音である。図 3 右側はヴァイオリン C#3 音, 突発音, およびこの 2 つを混合した信号である。

図 4 に検出した変調 - 搬送周波数マップの検出結果を示す。変調周波数 f_c 軸の高い位置にそれぞれの音階の基本周波数が表れていて, この場所は変動が激しい。また搬送周波数 f_m 軸の全域および変調周波数 f_c 軸の番号 1 から 40 にわたる領域 (64×40 の領域) は基本周波数が表れている位置の変動と比較すると時間的变化はほとんどない。

したがって, 時間的变化がほとんどない領域を時間平均したものをデータベースとして登録しておき, この部分でマッチングを行なうことにする。

3 音源同定のためのマッチング

図 4 の検出結果を基に, マッチング手法を用いて音源同定を行なった。マッチングは, 全体マッチングと部分マッチングの 2 種類を行なった。全体マッチングは, 各音階ごとの類似度がどのくらいなのかを確認し, 同定のための評価尺度を確認するために行なった。部分マッチングは, 突発音が付加されるとマップ上にどのような影響が表れているのかを調べるために行なった。

全体マッチングは, 前節で述べた 64×40 の領域を一つのベクトルとして取り扱い, 各時間ごとの領域のベクトルと, データベースを同じようにベクトルとしたものとの類似度を計算した。部分マッチングは, 全体マッチングで用いた領域を, 搬送周波数 f_m 軸で 64 分割し, 分割した領域どうしで類似度を計算した。

類似度は, 以下に示す式で計算している。

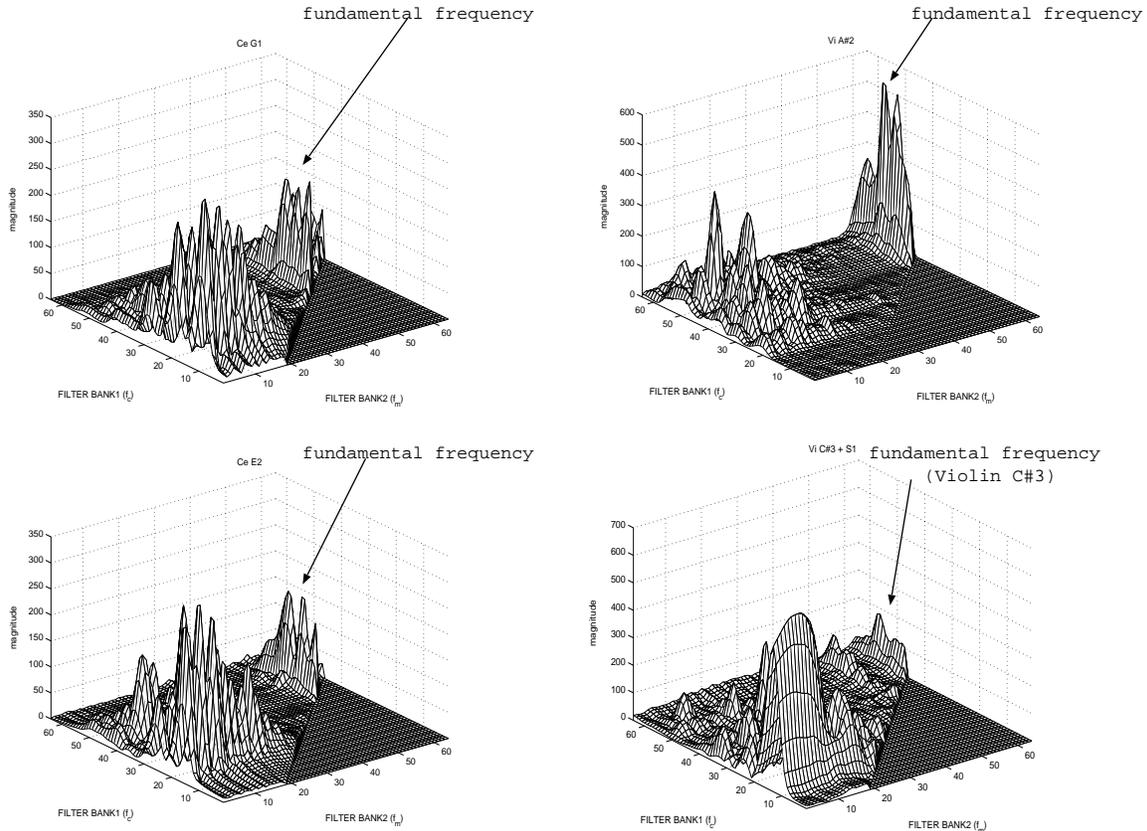


図 4: 変調 – 搬送周波数マップの検出結果 [時間平均] (左上: チェロ G1 音, 左下: チェロ E2 音, 右上: ヴァイオリン A#2 音, 右下: 混合信号)

$$s^{(l)} = \cos \theta^{(l)} = \frac{(\mathbf{x}, \mathbf{r}^{(l)})}{\|\mathbf{x}\| \cdot \|\mathbf{r}^{(l)}\|} \quad (0 \leq s^{(l)} \leq 1)$$

\mathbf{x} : 入力マップ

$\mathbf{r}^{(l)}$: データベースにあるマップ

$s^{(l)}$: 類似度 ($1 \leq l \leq 12$)

l は 1 ~ 3 がチェロの各音階, 4 ~ 12 がヴァイオリンの各音階を表している.

全体マッチングによる各音階とのマッチングは, 正しい楽器音の音階との類似度はすべて 0.9 以上であった. 突発音を付加した場合には, 類似度は 0.9 以上のものはなかった. この結果から, 全体マッチングで類似度が 0.9 以上であればその楽器の音であると同定することができる.

部分マッチングによる各音階とのマッチングは, 正しい楽器音の音階であると, 64 個中ほとんどが 0.9 以上の類似度であった. 突発音を付加した場合には, 各領域における類似

度は、突発音が始まると 0.9 以下になってしまう。この結果から、同じ楽器・音階でのマッチング結果と、突発音が付加された場合のものと各領域ごとで比較すれば、どの成分のどの時間に突発音の影響が表れているのかという情報がおおまかではあるが得られる。また、搬送周波数 (f_c) 軸側のチャンネル番号の 33 から 64 の中で、類似度が全時間にわたってほぼ 0.9 以上の値を示していれば、楽器音の成分がマップ上に存在し、その領域は突発音の影響を受けていないといえる。

図 5 に 部分マッチングの結果の一例を示す。

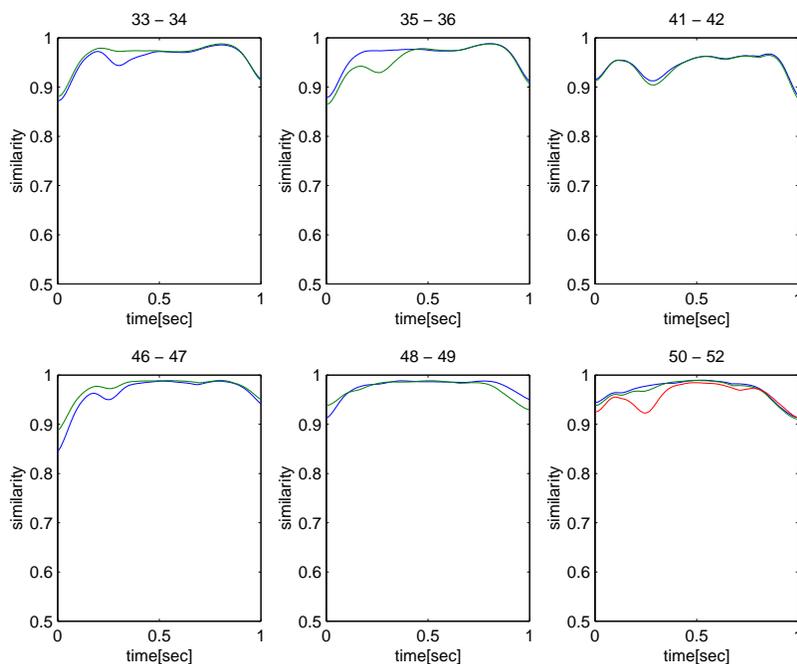


図 5: 部分マッチングの結果 (Cello G2)

4 まとめ

本論文では、楽器音および突発音が付加された楽器音を、変調 - 搬送周波数マップ検出モデルを用いて周波数分解後得られた振幅情報に着目し、またデータベースとのマッチングを行なうことで、楽器音を同定する方法を提案した。本方法では、楽器音の音階をマップの f_m 軸側において探索することで同定できた。また、部分マッチングの結果から突発音の影響がどの成分のどの時間に表れているのかが確認された。