

# 2R-3 音楽情景分析の処理モデル OPTIMA における 誤認識改善のための手法の考察

木下 智義      田中 英彦  
東京大学工学部

## 1 はじめに

著者らは音楽情景分析の処理モデル OPTIMA を実装した [1, 2, 3]。

OPTIMA では、複数の独立したモジュールに確率をもった仮説の組を出力させ、これを確率伝搬によって統合することによって外界の音響的事象に関する最尤推定像を求める。

現時点では実装した OPTIMA 実験システムの認識率は実用に十分であるとは言えず、その改善が課題となっている。

本稿では OPTIMA における誤認識の原因を解析し、認識率改善のための手法について考察する。

## 2 OPTIMA における低認識率の解析

処理モデル OPTIMA の大きな特徴は、確率伝搬による情報の統合することで、複数の処理モジュールが相互にその誤りを訂正することができる点にある。しかしながら、実装・評価実験を行なった結果、その特性が十分に発揮されているとは言えない。本節ではその原因を解析する。

### 2.1 評価実験結果

図 1 は、OPTIMA 実験システムにおける、和音・単音レベルでの認識率を、和音遷移情報・和音構成音情報を利用した場合と利用しなかった場合に分けて計った結果である。

	ON(○)/OFF(●)			
	○	○	●	●
和音遷移情報の利用	○	○	●	●
和音構成音情報の利用	○	●	○	●
単音認識率 (%)	87.6	87.6	87.6	87.6
和音認識率 (%)	68.8	68.8	56.3	50.0

表 1: 各モジュールが付与する情報の利用の有無による単音/和音認識率の差異

これは、和音遷移情報 (図 1 **——**) の伝搬の有無・和音構成音情報 (図 1 **——**) の伝搬の有無による和音・単音認識率の差を表にしたものである。

これを見れば、和音遷移情報や和音構成音情報によって、和音認識率が向上していることが分か

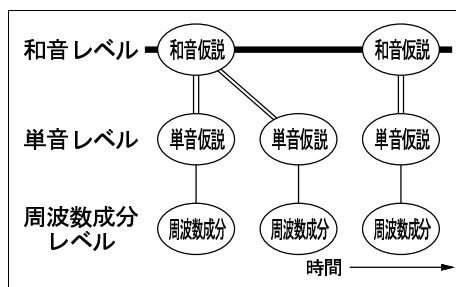


図 1: OPTIMA 仮説ネットワーク構成図

るが、一方で単音認識率の変化に寄与していないことも分かる。

この原因は、単音仮説の誤りが確率伝搬によって修正されるケースを考えると分かる。即ち、

- 単音仮説が若干の誤りを含んでおり、正しい仮説が仮説群の最上位にない。
- 誤りを含んだ単音仮説によって生成される和音仮説も同様に正解が最上位におらず、この誤りが和音遷移情報による条件つき確率の伝搬によって訂正される。
- この訂正の結果が、和音構成音情報の伝搬によって正しい単音仮説を上位へ押し上げる。

という条件が揃わなければ、和音情報が単音認識率に反映されないことになる。

また、単音誤認識の多くがオクターブのずれによるものでありながら、和音の処理ではオクターブを無視しているという点も、和音仮説から得られる情報が単音認識に活かされない原因であるといえる。

### 2.2 現在の OPTIMA が持つ構造上の限界

前項のことから分かるように、現在の仮説ネットワーク構造のみでは和音情報によって単音認識率が向上しない場合があることがわかる。そこで、他の処理モジュールを導入して処理精度の向上を考えてみる。例えば、メロディ遷移を考えて、単音レベルでも時間方向の条件つき確率を与える手法などがある。本稿ではその具体的内容については触れない。

しかしながら、処理モジュールを追加する際に、OPTIMA の致命的とも言える欠点が出てくる。OPTIMA のネットワーク構造は、単結合グラフ

として与えられるものに限られるということである。これは、グラフ内にループになっている部分があると、確率伝搬もループしてしまい、伝搬が終わらないことになるためである。(図2)

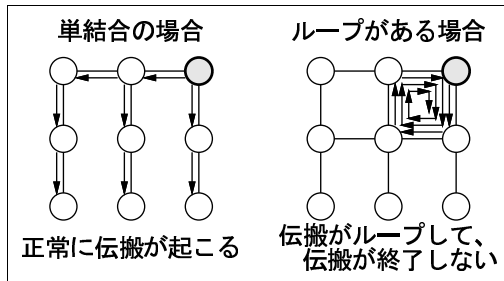


図2: 現在のOPTIMAの仮説ネットワークは単結合でなければならない

そこで、本稿では、まずこの欠点を改善するためにいくつかの手法を提案してみる。

### 3 OPTIMAの改良

#### 3.1 確率伝搬の収束を待つ方法

一番簡単に、見ために全く手を加えない方法を考える。これは、図2のように伝搬がループした際に、伝搬を繰り返して行くうちに、確信度の変化が少なくなると考えられることから、変化量が一定値以下になったら伝搬を終了するという方法である。

この方法の長所は、処理が簡単で、現システムに変更をほとんど加えずにすむ点である。

逆に短所は、まずそもそも確信度の変化量が減衰して行くことが示されていない点<sup>1</sup>。また、収束までに必要な計算量の予想がつかない点も挙げられる。

実際のデータを用いて実験を行ない、収束の速度と計算量を確認する必要がある。

#### 3.2 伝搬処理を複数の段階に分割する方法

##### 3.2.1 仮説レベルでの分割

評価実験での結果から、和音情報が単音情報へ反映されることがほとんどないということが分かるので、和音↔単音の確率伝搬を行なわない方法を考える(図3)。

つまり、最初に周波数成分レベルと、単音レベルだけで処理を行ない、その後で単音レベルと和音レベルで処理を行なうようにする。

この場合、長所として現在の処理系をそのまま流用できる点に加え、前項の手法のように、計算量の予想がつかないこともない。

しかし、一度全体の処理を行なった上で2段階の処理を行なうので、リアルタイム的な処理がで

<sup>1</sup>もともと、OPTIMAの評価実験を行なうために実験システムを走らせた時には、減衰して行くという印象を受けた

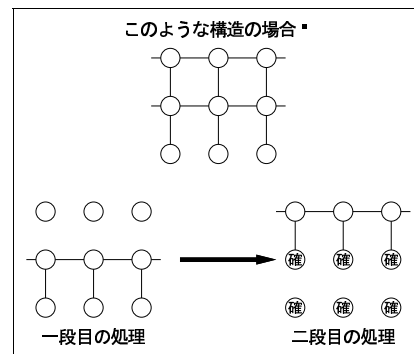


図3: OPTIMA 仮説レベル分割方式

きなくなるという欠点がある。

##### 3.2.2 時間方向の分割

前項と同様に、構造を時間方向に分割して同様の処理を行なうという方法もあるが、この場合は時間遷移情報が活かされにくい(図4)。

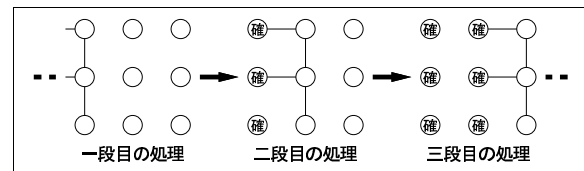


図4: OPTIMA 時間分割方式

#### 3.3 完全に新しい枠組を用いる方法

以上2つの手法はいずれも短所を含んでおり、そのままOPTIMAに応用することは難しい。従って、従来の枠組とは異なる全く新しい手法を考える余地もある。

### 4 おわりに

OPTIMA 実験システムの評価実験の結果から、低認識率の原因を解析し、OPTIMAの欠点を克服するための新しい処理方式について考察を行なった。今回考察した方法はいずれも問題点を含んでおり、そのまま実装を行なう訳には行かない。それぞれの問題点を解決して実装へ持ち込むことが今後の課題となる。

### 参考文献

- [1] 柏野 邦夫, 中臺 一博, 木下 智義, 田中 英彦: “音楽情景分析の処理モデル OPTIMA の実装”, 情処 '95 春全大, 6D-2 (1995).
- [2] 木下 智義, 柏野 邦夫, 中臺 一博, 田中 英彦: “音楽情景分析の処理モデル OPTIMA における音楽シーン情報の抽出と利用”, 情処 '95 春全大, 6D-3 (1995).
- [3] 中臺 一博, 柏野 邦夫, 木下 智義, 田中 英彦: “音楽情景分析の処理モデル OPTIMA における統計的単音仮説生成処理”, 情処 '95 春全大, 6D-4 (1995).