

# 解説

## 音声認識技術を用いた音楽情報処理\*

嵯峨山茂樹, 武田晴登, 亀岡弘和,  
西本卓也 (東京大学・情報理工学系研究科)\*\*

43.75.-z; 43.72.-p; 43.60.-c

### 1. はじめに

音楽は、多量の情報を音響信号に載せて人間に情報を伝えているが、その点では言語音声も同じであり、これらは双壁をなすと言えよう。音声分野では、様々な手法が音声信号処理及び音声認識技術に開発・投入され、高度な発達を遂げた。その考え方や手法 (例えば [1]) は、音楽情報処理においても共通に利用できるのではないだろうか。実は、これが音声認識や音声信号処理を専門としていた筆者らが、音楽情報処理を始める動機であった。

コンピュータを用いた音楽の研究や技術全般に関する解説には良書 [2, 3] があるので、本稿では、著者らの近年の研究を紹介しつつ、音声認識技術を用いた音楽信号処理・音楽情報処理の可能性をアピールしたい。そのため、関連研究の紹介 [4, 5] は完全に省略するが、ご容赦いただきたい。以下、音楽信号解析、音楽演奏解析、自動作曲の順に議論しよう。

### 2. Cepstrum から Specmurt へ

音楽 (特に西洋音楽) では、かなりの部分を多重音 (和音, コード) が占めており、自動採譜や音楽検索などの問題では、構成音が何であるかを知る多重ピッチ解析は不可欠である。音声認識のための音響分析で多用されている cepstrum 分析は、ピッチ抽出手法としても標準手法の一つだが、スペクトルを対数変換するので、多重ピッチ解析には原理的に適していない。

Cepstrum を手本に何かできないだろうか。Cepstrum は、パワースペクトルの対数を取って

逆フーリエ変換したもので、そのため spectrum の最初 4 文字を反転して cepstrum と呼ばれる。では、パワースペクトルは線形のままで周波数軸を対数にして逆フーリエ変換したらどうなるだろうか<sup>1</sup>。とにかく、名前としては spectrum の後の 4 文字を反転してそれを specmurt と呼ぼう [6]。

周波数  $\omega$  の対数  $x = \log \omega$  の上では、 $n$  次高調波周波数は基本周波数の値にかかわらず  $\log(\text{基本周波数}) + \log n$  である (「調和拘束」と呼ぼう)。ここで、周波数成分間のパワー比が基本周波数にかかわらず共通である調波構造を想定し、基本周波数に相当する位置を原点とし基本波パワーを 1 とした共通調波構造パターン  $h(x)$  を定義する。基本周波数がどの値でどれだけの成分を持つかを表した基本周波数分布  $u(x)$  を定義し、パワースペクトルの加法性を仮定すると、多重音スペクトル  $v(x)$  は  $u$  と  $h$  の畳み込み  $v(x) = h(x) * u(x)$  になる。Wavelet 変換などにより観測された  $v(x)$  に対して  $h(x)$  の逆畳み込みを、Fourier 変換領域 (specmurt 領域) の除算により行えば、基本周波数分布  $u(x)$  が得られる。

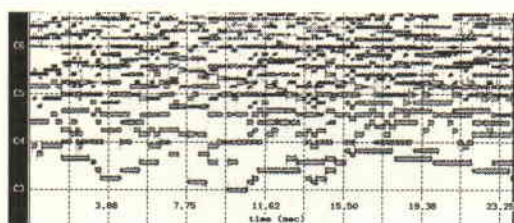
実際には調波構造は共通ではないが、平均的な特性を  $h(x)$  に仮定することにより、高調波を抑圧して基本周波数を強調する効果はある。例えば、 $h(x)$  を  $1/f$  特性とした場合の実演奏信号に対する音高可視化結果の例を図-1 に示す。パワースペクトルでは楽器音の倍音が複雑に重なり合うため、目視でも構成音の同定が困難なのに対し、specmurt 分析結果では、正解である参照用の MIDI データとかなり近くなっていることが分かる。

また、共通調波構造パターンを自動的に決定する方法 [7] は効果大きい。これを HMM による note-on/off 検出処理をすることで、MIDI データに変換でき、その結果、音色や速度や音高を操作

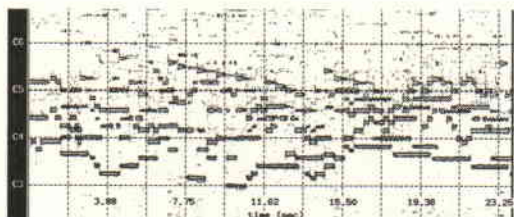
<sup>1</sup> 数式の上では、虚数軸へ写像する複素 Mellin 変換に等価だが、そう呼ぶと実際の処理が分りにくい。

\* Music information processing using speech recognition techniques.

\*\* Shigeki Sagayama, Haruto Takeda, Hirokazu Kameoka and Takuya Nishimoto (Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, 113-8656)



(a) Wavelet 解析スペクトログラム



(b) Specmurt 法による基本周波数分布の濃淡表示



(c) 参照用標準 MIDI 信号のピアノロールウィンドウ表示

図-1 Specmurt 法による J.S. バッハ「音楽の捧げ物」(BWV 1079) から六声のリチェルカーレ (RWC 研究用音楽データベース収録; 管弦楽演奏) の 5 声部部分の音高可視化結果例

することができる。RWC 音楽データベース [8] 中の単一楽器 (ギター, ピアノ) による数曲を処理して MIDI データに変換した結果の評価では, 多重ピッチの同定率 (音符正解率) はフレーム単位で 75~80%であった。

### 3. 混合ガウス分布の推定学習の利用

#### 3.1 調波拘束クラスタリング多重ピッチ推定

音声認識の音響モデルでは混合ガウス分布モデル (GMM) を多用する。これは, 言わば確率分布の濃淡を複数のガウス分布の和で表現するものであり, EM (Expectation Maximization) アルゴリズム [1] によって学習推定できる。一方, 音楽信号スペクトルの倍音構造は, 言わば周波数軸上の音響エネルギーの濃淡である。これにも GMM を使って, 調和拘束下で学習ができるだろう。更に, そのような倍音構造の複数個の和をモデルとして, EM アルゴリズムによってその構成要素を推定することができるのではないだろうか。

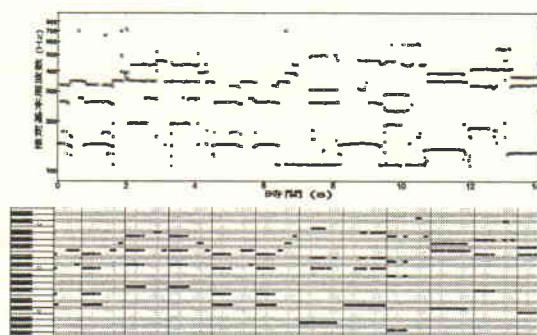


図-2 Harmonic Clustering によるジャズ曲 “Crescent Serenade (Guitar Solo)” (RWC 研究用音楽データベース収録) の基本周波数推定結果 (上段) と付属の参照用標準 MIDI 信号のピアノロール表示 (下段)

これは, 周波数軸上に散らばった音響エネルギーを調波関係拘束条件下でクラスタリングすることになるので, この手法を Harmonic Clustering と呼ぶ。前章同様に  $x$  を対数周波数とすると, wavelet 変換によって求めたスペクトル密度  $v(x)$  は, 対数周波数  $x$  の微小 wavelet の個数ヒストグラム値と理解できるので, 第  $k$  結びクラスタ (倍音構造) の  $n$  倍音成分に帰属する確率を  $p_{k,n}(x)$  とし,  $k$  番目の結びクラスタのうちの基本周波数推定値に相当するクラスタ重心を  $\mu_k$  とすると, クラスタリングの目的関数は

$$\int \sum_{k=1}^K \sum_{n=1}^{N_k} \varphi(x, \mu_k + \log n) p_{k,n}(x) v(x) dx \quad (1)$$

となり, EM アルゴリズムによって効率的に解が得られる。このような手法によって高精度な多重ピッチ推定手法 [9, 10] が導かれた。

このように多重ピッチ推定に確率推定の手法を導入することで, 条件つき確率を導入して滑らかなピッチ推定を行ったり, AIC (Akaike's Information Criterion) [11] などの情報量規準を用いて, 同時発音数の推定や倍ピッチ/半ピッチ誤りを回避するオクターブ位置推定などができるようになった [10]。この推定結果を前章同様に MIDI データに変換して精度評価を行ったところ, 更に高い精度が得られた。結果例を図-2 に示す。また, 調和拘束を緩めて, 非調和音の解析もできるようになった。

#### 3.2 調波時間拘束クラスタリング

前章の考え方を更に進めれば, 音楽信号のスペ

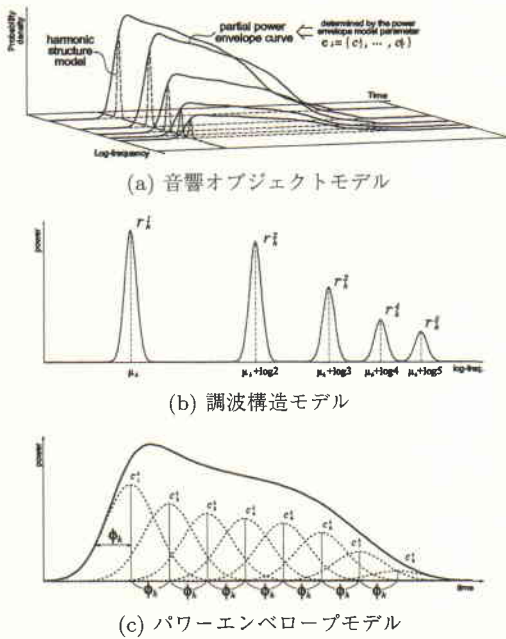


図-3 ガウス基底による単音の調波時間拘束モデル

クトログラムは周波数と時間の2次元上に音響エネルギーが分散している濃淡図形であることに思い至る。ならば、これも拘束条件付きの混合ガウス分布で表現できないだろうか。

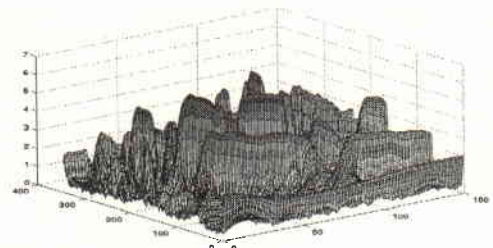
音楽信号を構成する各単音を図-3(a)のような周波数・時間構造を持つとすれば、周波数方向には前章同様に調波拘束(図-3(b))を使おう。時間方向には、ガウス分布を接続して蛇腹のように伸び縮みするパワーエンベロープのモデル(図-3(c))を導入しよう。こうすれば、このような拘束つき混合ガウス分布によって表現される単音モデルが、更に複数個の和となって、統計的変動を伴って観測されるときにそのパラメータを推定する問題として、各単音の基本周波数と開始・終了時刻が推定できるだろう。

図-4に、観測したスペクトルと、この手法(Harmonic-Temporal-structured Clustering)によって推定したモデルの例を示す。音符の推定精度は、前節より更に向上し、同条件では90%前後の精度に達し、自動採譜の実現可能性が高まって来ている。

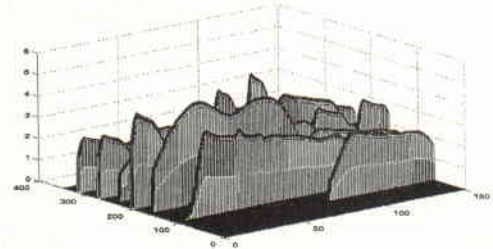
#### 4. 連続音声認識手法による演奏の解析

##### 4.1 リズム認識

連続音声認識は、音響モデルと言語モデルを統合して尤度最大の経路探索により行われている。



(a) 音楽信号の観測スペクトル分布



(b) ガウス基底による混合音の推定

図-4 調波時間拘束クラスタリングによる推定結果

この枠組は、音声認識に限らず極めて汎用のモデルであって、音楽情報処理の多方向で活用できそうである。まず、リズム認識への応用から議論しよう。

例として、MIDI鍵盤上で演奏入力した楽曲のMIDIデータから楽譜を復元する(採譜する)問題を考えよう。各音符の音高情報は正確に得られる。しかし音長は、演奏に表情を付けるためのテンポやリズムの変動などを反映して連続的に伸縮するので、音符の音価(記述された音符の長さ)を復元することは容易でない。市販ソフトによって量子化した例を図-5(a)に示す。観測された個々の音長を、単純に量子化処理しただけでは、演奏に極めて忠実な楽譜は得られても、演奏者が意図し聴き手に伝わる元の楽譜(同図(c))は復元できない。

では、人間はどのようにリズムを認知しているのか。図-5(a)のような複雑なリズムとは考えず、もっと簡単でありふれたリズムに基づいて音長が変動していると理解するのではないか。これは、音素のスペクトルパターンが毎度変動とする音素モデルと、それを理解する語彙や文法(言語モデル)を持つ連続音声認識[1]と同型の問題ではないか。

このように考えて、我々は連続音声認識における音響モデルに対応して、隠れマルコフモデル





(a) 市販ソフトによる楽譜化結果



(b) HMM によるリズム認識結果



(c) 元の楽譜

図-5 R. シューマン「トロイメライ」(Op.15, No.7) の MIDI ピアノ演奏データからの楽譜復元例

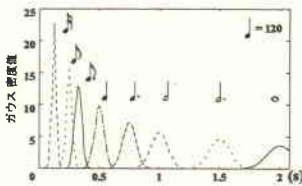


図-6 各音符音長の変動モデル：音響モデルに対応

(HMM) の隠れ状態を音価, 状態出力を音長 (図-6), 言語モデルに対応して音価の連鎖のリズム語彙 (図-7) あるいは音価の n-gram 連鎖モデルを適用して, 新しいリズム認識の手法 [12-15] を試みた。

HMM の状態出力分布は図-8 に示すように, 観測された音長 (実際は IOI (inter-onset interval) を用いる) の分布モデル (音響モデルに対応) とし, 状態間の遷移は, あり得る音価列 (リズム単語と呼ぶ) の集合 (リズム語彙) を並列に持つ言語モデル (図-8), あるいは音価の n-gram 確率 (図-9) により与える (言語モデルに相当)。こうして, Bayes の定理により事後確率最大の経路を Viterbi アルゴリズムにより求める (経路探索に対応) ことにより, 事後確率最大の音価列が得ら

頻度順	1 小節単位	2 拍単位
1 位	♪ ♪ ♪ ♪	♪ ♪
⋮		
10 位	♪ ♪ ♪ ♪ ♪ ♪	♪ ♪ ♪
⋮		

図-7 音価列の出現頻度例 (4/4 拍子) : 単語に対応

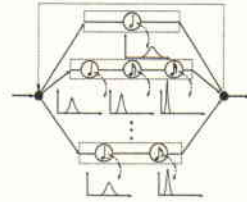


図-8 2 拍単位リズムモデル例 : 言語モデルに対応

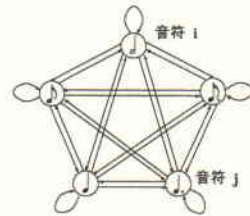


図-9 音符接続のリズムモデル例 : 任意音素列に対応

れる。

演奏者のスキルあるいは演奏のスタイルなどにより音長の確率変動が大きい場合は, 音長分布の分散を大きくする。音楽のジャンルやスタイルによりリズムに「くせ」がある場合には, リズム語彙の構成や確率に反映する。

以上は単旋律の場合であったが, 多旋律あるいは和声の場合も同様な扱いができる。まず, MIDI 信号のすべての声部にわたって “note-on” の間隔 (IOI) の時系列を得る。音長やリズム単語のモデルもそれに対応して学習しておく。前節と同様の原理により, すべての声部にわたるリズム譜が推定できるから, 各音の継続時間を考慮して同時発音を推定し, 多声部の楽譜を得ることができる。同時発音は自己遷移としてモデル化できる。このようにして得られた楽譜の例を図-5(b) に示す (ただし, 両手への分割と調号は人手で行った)。

連続認識手法に基づく種々のモデルについて, 人間の演奏を MIDI 信号で記録したデータを収集し, モデルの性能評価を行った一例 [14, 15] を挙げると, MIDI キーボードによる単旋律の演奏では, 市販のシーケンサソフトの量子化機能 40.7~

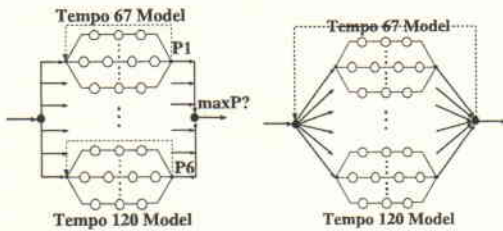


図-10

一定テンポモデル

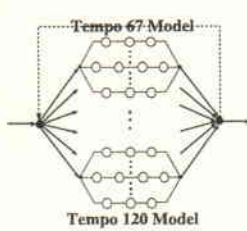


図-11

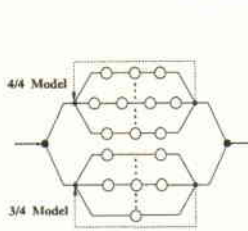


図-12

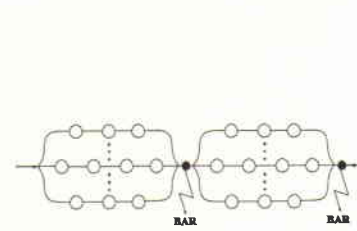


図-13

59.7%を上回る正解率 85.9~97.3%を得た。また、多声音楽の電子ピアノ演奏に対しても市販ソフトの 14.4%~45.4%を上回る 36.8%~92.2%の正解率を得た。連続音声認識の枠組は有効である。

#### 4.2 テンポ推定, 拍子推定, 拍節推定

テンポが未知の場合は, 以上のテンポ既知のモデルを複数のテンポ候補について並列に持つ (図-10) ことで, 事後確率最大のテンポと音価列を得ることができる [12-14]。テンポが変動する場合は, 異テンポ間の遷移確率を与えて Viterbi 探索することにより, 音価列とともに部分的なテンポが推定できる (図-11)。あるいは, 音符の  $n$  連鎖の音長比 (リズムベクトル) を観測量とすることにより, テンポに依存しない扱いも可能である [15]。

2 拍子系と 3 拍子系のリズム語彙を並列に持ち (図-12), Viterbi 経路を探索することで拍子を推定することもできる。また, 揚げ拍リズムのモデルを先頭に加える (図-13) ことにより, 小節線位置を推定することも可能である。

また, リズム認識とテンポ推定を交互に繰り返すことで不可分な両者を分離推定することができる。テンポの変化部分のセグメンテーションも, セグメンタル  $k$ -means 法や one-pass DP 法などの音声認識の手法 [1] に帰着することができる。

### 5. 連続音声認識手法による自動編曲

#### 5.1 自動和声づけの問題

音楽の愛好者は多く, 旋律の作曲には多くの人が興味を持つが, それに適切な和声付けをして編曲するにはやや専門的な音楽の知識や経験が必要である。そのような知識や経験を持たない初心者でも編曲ができる自動ツールや支援ツールには大きな用途が想定できる。ここでも連続音声認識手法を有効に使えないだろうか。

与えられた旋律への自動和声づけ問題に関して,

連続音声認識手法 [1] に基づく我々の手法 [16, 17] の原理を一言で言えば, 旋律は背後の和声進行を想定して生み出された, と見なす仮説に基づく。これは, 西洋古典音楽についてはかなり妥当と考えられるが, 仮に和声を想定せずに旋律を作曲した場合でも, あたかも旋律がそのようにして作られたかのように, 背後の和声を推定することに相当する。これは, 無意味な言語音連鎖を入力してもそれを音声認識すると文法にかなった文章が出力されることと同様であり, 音楽の和声づけにおいては意味がある。

具体的な手法は前章とほぼ同じである。HMM における隠れ状態を和声, 状態遷移を和声進行, 状態出力をその区間に対応する旋律断片としてモデル化する。従って, 和声から生まれる旋律断片に含まれる経過音・刺繍音・倚音・掛留音などの使用の確率により出力確率を評価する。状態間の遷移は, 和声進行 (和声常套句と呼ぶ) の集合 (和声語彙: 図-7 と同形) を並列に持つ言語モデル (図-8 と同形), あるいは和声の  $n$ -gram 確率 (図-9 と同形) により与える (言語モデルに相当)。こうして, Bayes の定理により事後確率最大の経路を Viterbi アルゴリズムにより求める (経路探索に対応) ことにより, 事後確率最大の音価列が得られる。このようにして得られる和声づけの単純な例を図-14 に示す。

この枠組では, 作曲者や時代のスタイルなどにより異なる旋律と和声の関係は, 状態出力確率に組み入れられる。同様にジャンルやスタイルによって和声進行に「くせ」がある場合には, 和声常套句の構成や確率に反映する。これらは, 多量の学習データから学習することも可能であるし, 音声認識で大いに議論された少量のデータからモデルパラメータを推定する手法 (話者適応なども含む) を利用することもできる。

## Zeuch ein zu deinen Thoren



図-14 自動和声付け結果例：旋律はコラール“Zeuch ein zu deinen Thoren”（声部進行は考慮せず和音のみ示している）

しかし、HMMや和声常套句の学習のためのデータは、現実には不足しがちである。音声認識における言語モデルの学習データ不足への対処と同様に、品詞クラス  $n$ -gram などに対応して和声進行常套句や非和声音クラスの  $n$ -gramなどを考慮し、各種の仮定に基づく近似を導入すれば、比較的良い結果が得られる。

## 5.2 調認識と転調認識

調性が未知の場合は、以上の調性既知のモデルをすべての調性候補について並列に持つ（図-10と同形）ことで、事後確率最大の調性と和声づけを得ることができる。転調する場合は、転調確率を調性モデル間の遷移確率として与えて Viterbi 探索することにより、和声づけ結果とともに各部分の調性が推定できる。音声認識では言語識別に相当する。

このように、和声づけを連続音声認識と同型の問題と捉えることにより、和声進行、スタイル学習、転調認識などを統合的な枠組で扱うことができる。

## 6. 自動対位法

多声音楽の作曲技法として対位法があるが、職人的な技量が必要とされる。自動対位法が実現できれば、効果的な作編曲や、自動和声づけの結果得られる和声進行を具体的な声部の旋律、特に低旋律を生成し、和音の転回形を決定するなどにも有用である。我々は、自動対位法を確率モデルと DP 経路探索の問題 [1] として定式化 [18] した。大



図-15 自動対位法の例：上段は与えられた旋律（cantus firmus）、下段は自動生成された対旋律

量の譜例によってモデルを学習すれば、熟達した作曲者が決して行わない動き（禁則など）には低い確率、大に行いそうな動きは高い確率を学習できると仮定する。

17世紀の対位法の規範に則った一対一対位法では、与えられた定旋律（cantus firmus）と対旋律の間で成す一対一の音程、両旋律進行の禁則遵守などを確率評価し、動的計画法の問題として  $N$ -best の音高を決定する。二声一対多、あるいは三声一対一対位法も同様な解法が得られた。

しかし、二声の自由対位法（華麗対位法）[19]では、あり得るすべてのリズムの対旋律を解探索の対象とするのでは組合せが爆発する。そこで、対旋律のリズム生成と音高決定の二つの問題に分解する。まず定旋律からリズムを抽出し、小節を単位として適合するリズム列を動的計画法により求める。言わば、二つの旋律の独立感が出るように相補的な音の刻みを与えるわけである。そしてそのリズムの中身を音高で埋める。この際に、定旋律との垂直音程適合確率、旋律の適正さの確率、音程遷移確率（進行禁則遵守など）を統計学習において、確率計算に用いる。これも動的計画法で求める。このようにして生成した自由対位法の例を図-15に示す。

## 7. 結 語

現代の連続音声認識の技術は確率的な逆問題の解法と捉えることができ、その主な構成要素は、音声分析、音響モデル、言語モデル、最適経路探索の4要素であり、その原理は音声認識に留まらず汎用的な解法である。音楽情報処理の各種の問題も確率的な逆問題として捉えれば、上記の全部あるいは一部の構成要素を利用して解ける可能性がある。また、音楽特有の問題も、音声分析やモデル学習の手法をヒントにして新手法を開発することができる。

もっとも、音声と音楽では問題が時に同型であっても同一ではなく、対象とする情報が対照的あるい



は双対的である場合も多い。音声スペクトル（特に有声音）は音韻性を反映した包絡形状と、基本周波数 ( $F_0$ ) を反映した微細構造からなるが、音声認識ではおもに前者に関心があるのに対し、音楽信号では後者やリズムなどの時間要素が重要である。また、音声認識では多重音声は扱われることが少ないが、音楽信号処理では和音や多声部音楽は不可避の問題である。このような音楽情報処理の研究を通して、音声認識における韻律や時間要素や重畳音声などの情報の利用につながる研究としてフィードバックできることも期待できる。

今回扱わなかったが、音声認識技術を活用できそうな問題として、楽器・音色認識、旋律認識、楽譜追跡、歌詞認識、ドラム譜作成、口ドラム認識、指使い自動決定なども考えられる。今後の展開が楽しみな分野である。

#### 謝 辞

当該分野の研究を進展させた我々の研究グループの現・旧メンバの川上隆、齋藤直樹、吉川響、大槻知史、高橋佳吾、菅原啓太、中潟昌平、齋藤翔一郎の各氏に感謝する。

#### 文 献

- [1] X.-D. Huang, A. Acero and H.-W. Hon, *Spoken Language Processing: A Guide to Theory, Algorithm, and System Development* (Prentice-Hall, Upper Saddle River, N.J., 2001).
- [2] M. Mathews and J. Pierce, *Current Directions of Computer Music Research* (MIT Press, Cambridge, Mass., 1985).
- [3] 長嶋洋一, 橋本周司, 平賀 謙, 平田圭二編, “コンピュータと音楽の世界,” bit 別冊, 共立出版 (1999).
- [4] A. Klapuri, “Multiple fundamental frequency estimation based on harmonicity and spectral smoothness,” *IEEE Trans. Speech Audio Process.*, 11, 806–816 (2003).
- [5] 嵯峨山茂樹, 武田晴登, 亀岡弘和, 西本卓也, “音楽情報処理と音声認識,” 音講論集, 2-6-9, pp. 785–788 (2004.9).
- [6] 高橋佳吾, 西本卓也, 嵯峨山茂樹, “対数周波数逆畳み込みによる多重音の基本周波数解析,” 情報処理学会研究報告, 2003-MUS-53-13, pp. 61–66 (2003).
- [7] 亀岡弘和, 齋藤翔一郎, 西本卓也, 嵯峨山茂樹, “Specmurt 法による音楽信号の音高可視化における共通調波構造パターンの自動決定,” 音講論集, 2-6-15, pp. 803–804 (2004.9).
- [8] M. Goto, H. Hashiguchi, T. Nishimura and R. Oka, “RWC music database: Popular, classical, and jazz music database,” *Proc. Int. Conf. Music Information Retrieval (ISMIR2002)*, pp. 287–288 (2002).
- [9] 亀岡弘和, 西本卓也, 篠田浩一, 嵯峨山茂樹, “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,” 音講論集, 3-7-3, pp. 837–838 (2003.3).
- [10] 亀岡弘和, 西本卓也, 嵯峨山茂樹, “ハーモニック・クラスタリングによる多重音信号音高抽出における音源数とオクターブ位置推定,” 情報処理学会研究報告, 2003-

MUS-51-5, pp. 29–34 (2003).

- [11] H. Akaike, “Information theory and an extension of the maximum likelihood principle,” *2nd Int. Symp. Information Theory*, Akademia Kiado, Budapest, pp. 267–281 (1973).
- [12] 齋藤直樹, 中井 満, 下平 博, 嵯峨山茂樹, “隠れマルコフモデルによる音楽演奏情報からの音符列推定,” 平成 11 年度電気関係学会北陸支部連合大会講論集, F-62, p. 362 (1999).
- [13] 齋藤直樹, 中井 満, 下平 博, 嵯峨山茂樹, “隠れマルコフモデルによる音楽演奏からの音符列の推定,” 情報処理学会音楽情報科学研究会, 99-MUS-33, pp. 27–32 (1999).
- [14] 大槻知史, 齋藤直樹, 中井 満, 下平 博, 嵯峨山茂樹, “隠れマルコフモデルによる音楽リズムの認識,” 情報論, 43, 245–255 (2002).
- [15] 武田晴登, 西本卓也, 嵯峨山茂樹, “確率モデルによる多声音楽演奏の MIDI 信号のリズム認識,” 情報論, 45, 670–679 (2004).
- [16] 川上 隆, 中井 満, 下平 博, 嵯峨山茂樹, “隠れマルコフモデルを用いた旋律への自動和声付け,” 情報処理学会音楽情報科学研究会, 99-MUS-34, pp. 59–66 (2000).
- [17] 菅原啓太, 西本卓也, 嵯峨山茂樹, “HMM と音符連鎖確率を用いた旋律への自動和声付け,” 情報処理学会研究報告, 2003-MUS-53 (2003).
- [18] 中潟昌平, 西本卓也, 嵯峨山茂樹, “動的計画法と音列出現確率を用いた対位法の対旋律の自動生成,” 情報処理学会研究報告, 2004-MUS-56, pp. 65–70 (2004).
- [19] 中潟昌平, 西本卓也, 嵯峨山茂樹, “動的計画法に基づく自動対位法,” 音講論集, 3-7-12, pp. 587–588 (2005.3).

#### 嵯峨山茂樹

昭 49 東大院・工・計数 修士課程了。同年日本電信電話公社 (現 NTT) 入社。平 2-5 ATR 自動翻訳電話研究所音声情報処理研究室長。平 10 北陸先端大 教授。平 13 東大院・情報理工 教授。音声・音楽・文字・エージェントの信号処理・情報処理の教育研究に従事。博士 (工学)。発明協会発明賞, 科学技術庁長官賞などを受賞。電子情報通信学会, 情報処理学会, 日本音響学会, 人工知能学会, SICE, IEEE, ISCA 各会員。

#### 武田 晴登

平 13 年慶應大学理工学部卒業, 東京大学大学院情報理工学系研究科在学中。日本音響学会, 情報処理学会各会員。

#### 亀岡 弘和

平 14 年東京大学工学部計数工学科卒業, 東京大学大学院情報理工学系研究科博士課程在学中。音楽情報処理, 音声言語情報処理, デジタル信号処理などに興味を持つ。第 20 回電気通信普及財団テレコムシステム技術学生賞受賞。日本音響学会, 電子情報通信学会, 情報処理学会, IEEE, ISCA 各学生会員。

#### 西本 卓也

平 5 早大・理工卒。平 7 早大院・理工 修士了。平 8 京都工繊大・工芸 助手。平 14 東大・院・情報理工 助手。パターン情報処理, 音声対話システム, 福祉情報工学, 擬人化音声対話エージェントの研究に従事。人工知能学会 1998 年度研究会奨励賞。日本音響学会 第 20 回粟屋潔学術奨励賞。情報処理学会, 電子情報通信学会, 人工知能学会, ヒューマンインタフェース学会会員。