

## 信号の特徴を利用した一般化調和解析\*

中 沢 誠<sup>\*1</sup> 服 部 永 雄<sup>\*1</sup> 大 内 康 裕<sup>\*2</sup> 山 崎 芳 男<sup>\*1</sup>

[要旨] 一般化調和解析は高い周波数分解能が得られ音響信号の解析に用いられた場合強力な手法であるが、計算に膨大な時間がかかるという欠点がある。分析対象となる音楽信号は基音となる音とその倍音から構成されるという特徴があり、音声信号の場合も特定の周波数にパワーが集中するという特徴がある。本論文ではこのような解析信号の特徴に着目して、音楽信号では音階に基づく音とその倍音に限って集中的に解析を行うことで従来の一般化調和解析に比べ 1/5 以下の演算量で周波数成分が抽出可能なことを示した。演奏が音階に沿って進行されているような状況においては非常に効率よく分析ができたが、基音のパワーが最大でない場合に基本周波数の推定を誤る場合があった。

キーワード 一般化調和解析, スペクトルピーク, 倍音構造, 音階

Generalized Harmonic Analysis, Spectrum peak, Spectrum structure, Scale

### 1. ま え が き

信号の周波数解析には高速フーリエ変換のように高速に調和解析を行うことが可能な手法を用いることが一般的であるが、近年は逐次減算する処理過程による分析法が様々な分野で見られる [1, 2]。

一般化調和解析もその一つであり、これは一般に調和的な有限個の周期関数の和で表すことのできる概周期関数のスペクトルについて “Generalized Harmonic Analysis” として 1933 年 N. Wiener が提案した分析方法 [3] で、この理論をもとに、観測区間内において原波形から残差が最小となる正弦波を逐次減算処理するアルゴリズムが考案された [4]。その分析過程は、観測区間内で残差エネルギーを最小とする正弦波を見つけ出し、信号から減算し、更に残差信号に対して同様に残差エネルギーを最小とする正弦波を減算する過程によって周波数を抽出するものであり、観測区間の影響を受けることなく解析を行うことが可能である。また、正弦波モデルによって信号を分析するこの手法は膨大な計算量を必要とするものの窓の影響を受けることなく任意の周波数を分析することが可能である。

楽器音あるいは音声といった音源部、共鳴部、放射部によって発せられる音は一般に基音とその倍音成分によって構成される。従って、分析する周波数の対象

を基音及び倍音成分に限定することにより計算量の削減が期待できる。

本論文では解析信号の特徴に着目して、音楽信号では音階に、音声についてはピッチ周波数に基づく限定された周波数に限って解析を行う方法を提案し、従来の一般化調和解析の解析効率に比べ 5 倍以上の効率で周波数成分が抽出可能なことを明らかにした。

### 2. 一般化調和解析

信号処理には調和解析の一つであるフーリエ解析が広く用いられている。また、近年は、局在する信号の特徴を捉える有効な分析手法として Wavelet 変換が盛んに研究されている。更には逐次減算する処理過程による分析法が様々な分野で見られるようになった。一般化調和解析は、一般に調和的な有限個の周期関数の和で表すことのできる概周期関数のスペクトルについて “Generalized Harmonic Analysis” として 1933 年 N. Wiener が論じた解析法である [3]。

この理論をもとに、観測区間内において原波形から残差が最小となる正弦波を逐次減算処理するアルゴリズムが考案され、東山、平田らによってピアノの音や人の歌声など、非定常信号の解析に導入された。また大内、及川らは雑音と信号の分離に導入し、解析手法の有用性を示してきた [5]。

この解析手法は、単純明快な処理過程でありながら窓の影響を受けることなく任意の周波数成分について分析することが可能であり、周波数に対する自由度の高さを兼ね備えているが膨大な計算量を必要とするので、ごく最近まで実際の信号処理に応用される機会は少なかった。

\* Generalized Harmonic Analysis using signal features, by Makoto Nakazawa, Hisao Hattori, Yasuhiro Ouchi and Yoshio Yamasaki.

<sup>\*1</sup> 早稲田大学大学院国際情報通信研究科

<sup>\*2</sup> 千葉工業大学

(問合先: 中沢 誠 〒367-0035 本庄市西富田大久保山 1011 早稲田大学国際情報通信研究科 山崎研究室 e-mail: nakazawa@acoust.rise.waseda.ac.jp)

(2004年2月18日受付, 2004年10月27日採録決定)

2.1 分析過程

一般化調和解析は次に示す過程によって信号を逐次分析する。

信号  $x_0(t)$  から (1) 式に示す周波数  $f$ [Hz] の正弦波  $W(t, f)$  を減算したときのエネルギー減少量  $e$  は (2) 式によって表される。このとき  $L$ [s] は観測区間長とする。

$$W(t, f) = S_f \sin 2\pi ft + C_f \cos 2\pi ft \quad (1)$$

$$e = \frac{2z}{L\{z^2 - 2(1 - \cos z)\}} \{(z + \sin z)s^2 - 2(1 - \cos z)cs + (z - \sin z)c^2\} \quad (2)$$

ただし,  $z, s, c$  は (3) 式で示される。

$$\begin{aligned} z &= 4\pi ft \\ s &= \int_0^L x_0(t) \sin 2\pi ftdt \\ c &= \int_0^L x_0(t) \cos 2\pi ftdt \end{aligned} \quad (3)$$

(2) 式の  $e$  を最大とする周波数  $f$ [Hz] を求め (4) 式により  $\sin$  振幅値  $S_f$ ,  $\cos$  振幅値  $C_f$  を得ることができる。

$$\begin{aligned} S_f &= \frac{2z}{L} \cdot \frac{s(z + \sin z) - c(1 - \cos z)}{z^2 - 2(1 - \cos z)} \\ C_f &= \frac{2z}{L} \cdot \frac{s(z - \sin z) - c(1 - \cos z)}{z^2 - 2(1 - \cos z)} \end{aligned} \quad (4)$$

得られた  $\sin$  振幅値  $S_f$ ,  $\cos$  振幅値  $C_f$  より導出した正弦波  $W(t, f)$  を  $x_0(t)$  から減算することで (5) 式により残差信号を得る。

$$x_1(t) = x_0(t) - W(t, f) \quad (5)$$

以上の処理の繰り返しによって解析を行う一般化調和解析は、周波数分解能の自由度が高い一方で解析周波数分解能によっては膨大な演算が必要となる。

3. 信号の特徴に着目した処理

一般に、楽曲は調律された楽器によって音階に沿って演奏されることから、音楽信号は調律とその音階の特徴をもっている。音楽の調律周波数や音声の発声者のピッチ周波数をあらかじめ知ることにより、分析対象の周波数の限定が可能となり、一般化調和解析の計算量の大幅な削減が期待できる。本手法における解析の流れを図-1 に示す。すなわち、本手法は、まず分析対象とする信号に頻繁に現れる周波数成分を特定し、楽曲や話者及びピッチ周波数が変わらない限りその周

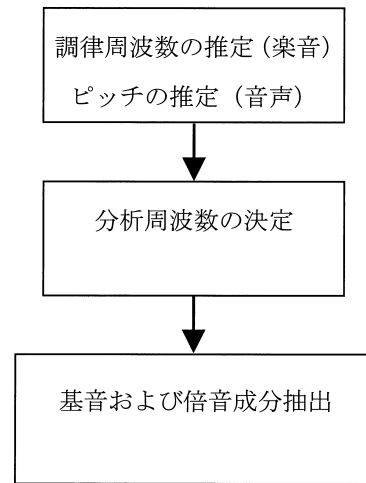


図-1 本手法による解析の流れ

波数成分が引き続き頻繁に現れるという特徴を利用した分析、解析手法である。

ギターのベンディングや、スライド奏法、音声のビブラートに対しては残差が多くなり収束が遅くなるが、最悪の場合でも従来の一般化調和解析に必要な計算量となるだけである。これらの奏法が曲全体に及ぶわけではないので、部分的にこれらの奏法が使われていても計算量の削減が期待できる。

3.1 調律周波数の推定

音階の基準になる  $A_4$  は必ずしも 440 Hz になるとは限らない。そこで、 $A_4$  に相当する音階の基本周波数すなわち調律周波数を以下の手法により求める。

まず、解析区間長  $L$ [s] の信号  $x_0(t)$  から (6) 式により求められる各音階の周波数  $f_n(x)$  に相当する成分の  $\sin$  振幅値  $S(f_n(x))$ ,  $\cos$  振幅値  $C(f_n(x))$  を一般化調和解析により抽出し (7) 式により残差エネルギー  $E_1(f_0(x))$  が求められる。

$$\begin{aligned} f_n(x) &= x \times 2^{\frac{n}{12}} \\ (-\infty < n < \infty, x : \text{調律周波数 } (A_4)) \end{aligned} \quad (6)$$

$$\begin{aligned} E_1(f_0(x)) &= \int_0^L \left\{ x_0(t) \right. \\ &\quad \left. - \sum_m (S(f_m(x)) \sin 2\pi f_m(x)t \right. \\ &\quad \left. + C(f_m(x)) \cos 2\pi f_m(x)t) \right\}^2 dt \end{aligned} \quad (7)$$

ここでは、 $430 < x < 450$  において  $x$  を 1 Hz 刻みで変化させ  $E_1(f_0(x))$  を最小とする  $f_0(x)$  すなわち  $x$ [Hz] を真の調律周波数とする。

この処理過程をある時間間隔、例えば 1s で行うこ

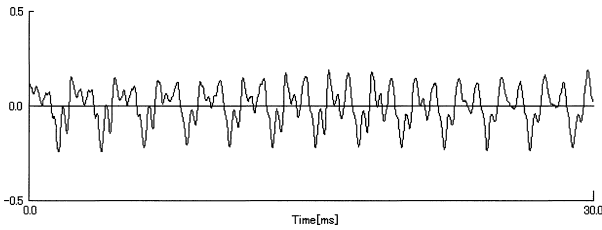


図-2 バイオリン音の時間波形

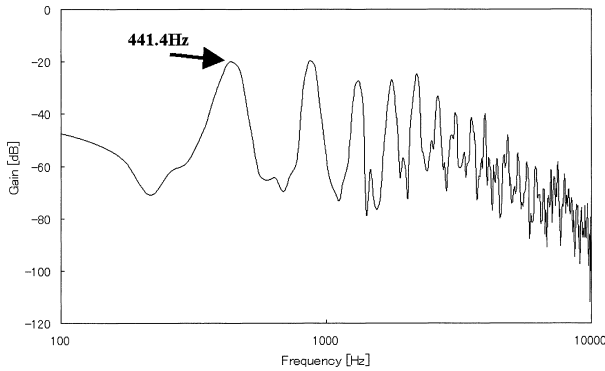


図-3 バイオリン音のスペクトル

とにより基本周波数の揺らぎに追従した分析も可能となる。

### 3.2 分析に用いる周波数の決定

3.1 節により推定された調律周波数  $x$  を用い、分析に用いる周波数を決定する。まず、基本周波数  $F$  [Hz] を基に基音から  $k$  倍音  $kF$  [Hz] までの  $k$  個の  $\sin$  振幅値  $S_{kF}$ 、 $\cos$  振幅値  $C_{kF}$  を抽出する。そのときの信号  $x_0(t)$  から  $k$  個の倍音成分を減算した信号のエネルギー  $E_2(F)$  を (8) 式により求める。

$$E_2(F) = \int_0^L \left\{ x_0(t) - \sum_{m=1}^k (S_{mF} \sin 2\pi mFt + C_{mF} \cos 2\pi mFt) \right\}^2 dt \quad (8)$$

この  $F$  を (6) 式のような音階に合わせて変化させ (8) 式の  $E_2(F)$  を最小とする  $F$  を真の基本周波数と推定する。本手法においては  $k = 3$  として解析を行う。

図-2 に解析に用いたバイオリン音の時間波形を示す。図-3 にバイオリン音のスペクトルを示す。解析区間長は 1,024 点 (23.2 ms) でハニング窓を適用した信号に後半に 0 を付加した 65,536 点に対してフーリエ変換を行った。図-4 に本手法によって求められた音階と  $E_2(F)$  の関係を示す。 $E_2(F)$  の最小値となる音階は  $A_4$  であることが確認でき、そのときの基本周波数は 440 Hz となった。図-5 に本手法により抽出された基音及び倍音成分を示す。

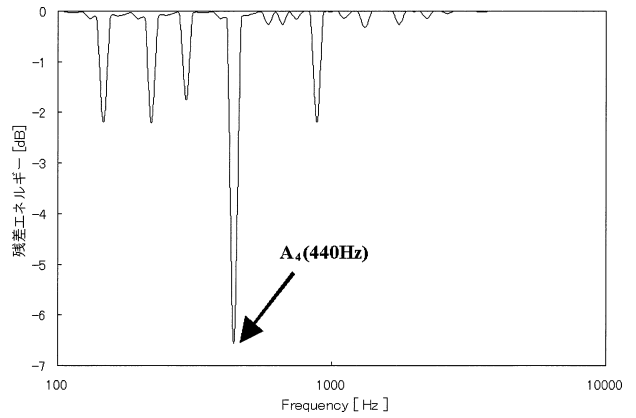


図-4 本手法によって求められた音階と  $E_2(F)$  の関係

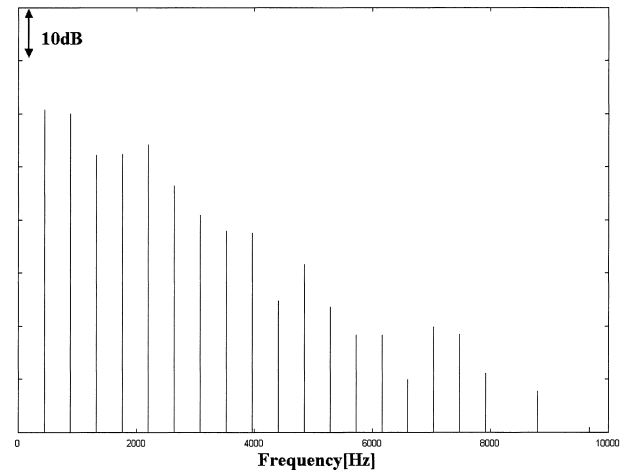


図-5 本手法により抽出された基音及び倍音成分 (解析区間長 30 ms, 抽出周波数成分数 20)

### 3.3 基音及び倍音成分抽出

3.2 節により推定された基本周波数を基に倍音成分を抽出する。また、ピアノを代表とする弦楽器のスペクトルの特徴として倍音の非整数倍性がある。弦の振動数とこれによって出る音を初等力学で扱う場合は弦の太さあるいは剛性を無視しているので倍音は基音の整数倍として解析される。しかし、実際にはこのような弦は存在せず正確には整数倍にはならない。H. Fletcher によると第  $n$  倍音の周波数  $f_n$  は (9) 式のように表される [6]。

$$f_n = nf_0(1 + Bn^2)^{\frac{1}{2}} \quad (9)$$

ここで  $f_0$  は剛性のない理想弦としたときの基本周波数、 $B$  は各弦に固有の定数であり低音純銅線を巻いた巻銅線では裸線に比べて大きい。そこで、本手法において倍音成分を抽出する際に基本周波数の整数倍のみではなく、整数倍音及びその周辺の複数の周波数に対しても一般化調和解析による抽出を行い、残差エネルギーが最小になる周波数を真の倍音成分として抽出する。

表-1 解析に用いたバイオリン音の音階

音階	周波数 [Hz]	音階	周波数 [Hz]
G <sub>3</sub>	196.0	D <sub>5</sub>	587.3
A <sub>3</sub>	220.0	E <sub>5</sub>	659.3
B <sub>3</sub>	246.9	F <sub>5</sub>	698.5
C <sub>4</sub>	261.6	G <sub>5</sub>	784.0
D <sub>4</sub>	293.7	A <sub>5</sub>	880.0
E <sub>4</sub>	329.6	B <sub>5</sub>	987.8
F <sub>4</sub>	349.2	C <sub>6</sub>	1,046
G <sub>4</sub>	392.0	D <sub>6</sub>	1,174
A <sub>4</sub>	440.0	E <sub>6</sub>	1,318
B <sub>4</sub>	493.9	F <sub>6</sub>	1,397
C <sub>5</sub>	523.3	G <sub>6</sub>	1,568

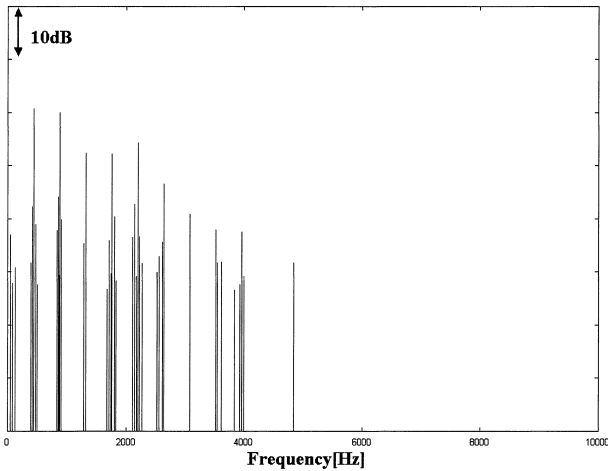


図-6 従来の一般化調和解析による周波数成分 (解析区間長 30 ms, 抽出周波数成分数 40)

本手法を用いて楽音の解析を行った。バイオリン音を G<sub>3</sub> (196 Hz) から G<sub>6</sub> (1,568 Hz) までを表-1 に示すように順に演奏した信号を用いて、標本化周波数 44.1 kHz, 量子化ビット数 16 bit で収録された信号を 1 解析区間長 30 ms で解析を行った。

解析の対象とされる周波数成分数は約 6,000 成分であり、従来の一般化調和解析で一つの周波数を求める処理過程と同様な処理過程によって基音と倍音成分を一括して分析できるので大幅な計算量の削減が可能となる。

#### 4. む す び

本論文では、楽音や音声の周波数成分に偏りがあることを利用して、一般化調和解析の計算量を大幅に削減可能であることを示した。信号の特徴を得る一つの指標として音階を導入して基本周波数を推定し、それを基に基音及び倍音成分を抽出する手法について論じた。

図-5 に本手法により抽出された基音及び倍音成分を示す。図-5 から 20 倍音まで抽出されていることが確認できる。

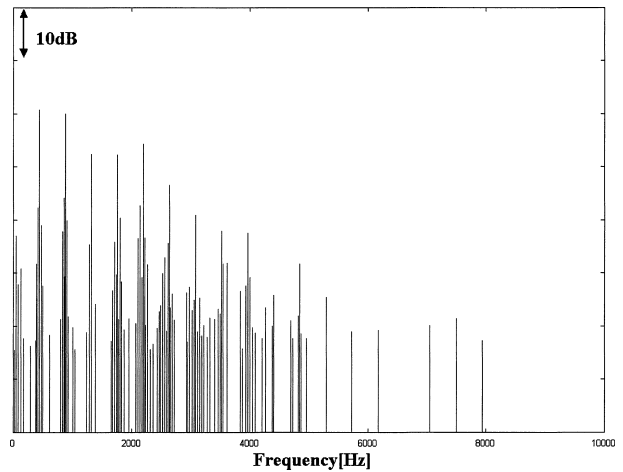


図-7 従来の一般化調和解析による周波数成分 (解析区間長 30 ms, 抽出周波数成分数 100)

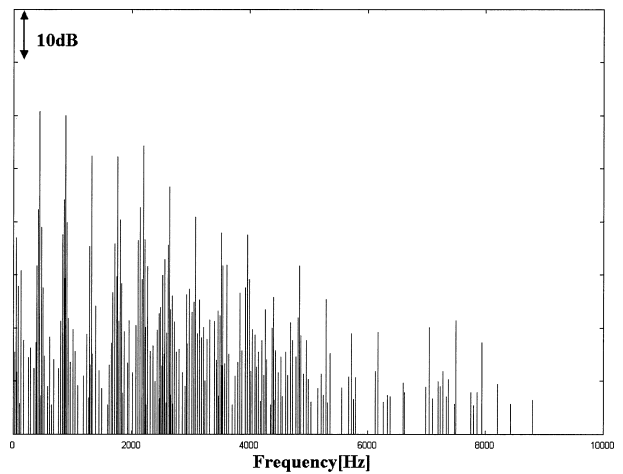


図-8 従来の一般化調和解析による周波数成分 (解析区間長 30 ms, 抽出周波数成分数 200)

一方、図-6、図-7、図-8 は従来の一般化調和解析により周波数成分を抽出しており、抽出した成分の数はそれぞれ 40、100、200 である。図-6 では図中に確認できる倍音数は 11 個程度であるのに対し、図-7、図-8 では 20 個程度の倍音が図中に確認できる。

本手法による解析結果と従来の一般化調和解析による解析結果を比較すると本手法により抽出された倍音成分を従来の一般化調和解析により得るには本手法で抽出する周波数成分数の 5 倍以上の周波数成分数を抽出することになる。この結果から、楽音や音声の大きな特徴といえる調和構造の効率的な抽出には本手法が適していると考えられる。

#### 文 献

[1] S.G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal. Process.*, 3397-3415 (1993).  
 [2] 及川靖広, 山崎芳男, "Matching Pursuit を用いた 2 マイクロホン収音での多音源方向推定," 音講論集, pp.501-502 (2003.9).  
 [3] N. Wiener, *The Fourier Integral and Certain of*

- Its Applications* (Cambridge University Press, London, 1933), pp. 150–199.
- [4] 牛山 聡, 東山三樹夫, 飯塚昌弘, 平田能睦, “一般調和解析による波形分析,” 信学技報, EA93-103, pp. 39–44 (1994).
- [5] 大内康裕, 及川靖広, 山崎芳男, “一般化調和解析を用いた基本周波数の推定および雑音と信号の分離,” *JAS journal*, Vol. 2, pp. 30–36 (2000).
- [6] H. Fretcher, “Normal vibration frequencies of a stiff piano string,” *J. Acoust. Soc. Am.*, 36, 203–209 (1964).
-