論題(和文)	一般化調和解析における連続するフレームの関係
Title(English)	Relationship between Consecutive Frames in Generalized Harmonics Analysis
著者(和文)	野田久順, 西原明法
Authors(English)	Hisayori Noda, Akinori NISHIHARA
出典(和文)	第23回 信号処理シンポジウム, , ,
Citation(English)	, , ,
発行日 / Pub. date	2008, 11
URL	http://search.ieice.org/
権利情報 / Copyright	本著作物の著作権は電子情報通信学会に帰属します。 Copyright (c) 2008 Institute of Electronics, Information and Communication Engineers.

第23回 信号処理シンポジウム
2008年11月12日～14日（金沢）

一般化調和解析における連続するフレームの関係

Relationship between Consecutive Frames in Generalized Harmonics Analysis

野田久順 † 西原明法 ‡
† 東京工業大学大学院集積システム専攻
‡ 東京工業大学教育工学開発センター

Hisayori NODA† Akinori NISHIHARA‡
†Communications and Integrated Systems, Tokyo Institute of Technology
‡The Center for Research And DeveLopment of Educational technology, Tokyo Institute of Technology

アブストラクト 一般化調和解析では時分割された被分析信号の各フレームをパラメータを自由に決定可能な正弦波の和として表す．このとき連続する各フレームにおいて同一の正弦波を表す成分が抽出される事が予想できる．本稿では相互相関関数を重みとした二部グラフの重み付き最大マッチングを用いて同一正弦波を表す正弦波パラメータを抽出する手法を提案する．実験では上記アルゴリズムを音楽信号及び人為的に生成した音源信号に対して適用し，実際に同一正弦波を時分割フレームにまたがって抽出できるかどうかを確かめた．

Abstract In the generalized harmonics analysis, each frame of the signal is shown as sum of the sinusoids that we can arbitrarily decide the parameters. At this time, the parameters which show the same sinusoid are to be extracted in consecutive frames. We propose a method to extract the parameters which show the same sinusoid, which uses weighted bipartite graph matching algorithm whose weight is shown as cross correlation function. This algorithm was applied to several audio signals and it was confirmed that the sinusoids were able to be extracted across consecutive frames.

1 はじめに

一般化調和解析 (Generalized Harmonic Analysis)[1] は 1930 年にウィナーらによって提案された信号分析の概念である．一般化調和解析において解析対象となる信号は振幅，周波数及び位相のパラメータを任意に決定可能な正弦波の和として表現される．一般化調和解析は周波数パラメータが解析フレームサイズの逆数の整数倍に制限されることがないため，原理上短時間フーリエ変換に比べて良い周波数解像度を持つ．

一般化調和解析を音声符号化等に応用した研究として [2], [3] が挙げられる．これらの手法は各フレームについて独立に符号化を行っている．一方，近年連続したフレームの類似性を利用した予測符号化が盛んに研究されている．これらを一般化調和解析を用いた音声符号化に適用することにより，符号化時の圧縮率を向上することができると考えられる．

一般化調和解析により得られる解析結果は各フレームを構成する複数の正弦波それぞれの振幅，周波数及び位相のパラメータである．これらについて，連続するフレームにおいては同一の正弦波を表すパラメータが存在することが予想できる．これらは予測符号化のためのフレーム間の類似性として利用することが出来ると思われる．

本稿では連続する 2 つの解析対象フレームについて，一般化調和解析により得られた結果から同一の正弦波を抽出する方法を提案する．始めに連続する 2 つのフレームの正弦波の各ペアの類似度を計算し，類似度行列を作成する．このとき連続性の指標として正弦波同士の相互相関関数を使用する．続いて正弦波を頂点，相互相関関数により得られた値を重みとした二部グラフを構成し，重み付き二部グラフの最大マッチング問題に定式化して解き，それぞれの正弦波がどの正弦波と連続であるかという組み合わせを求める．最後にこのアルゴリズムを連続した複数のフレームに対して適用し，複数のフレームにまたがる正弦波の抽出を行う．

実験においては複数の解析対象に対して上記のアルゴリズムを適用し，抽出された正弦波及び同一正弦波と識別された正弦波ペアの可視化を行う．

2 一般化調和解析

一般化調和解析においては解析対象信号はフレームに時分割された後，各フレームについて以下に示す式によっ

て表現される．

$$x_0(n) \cong \sum_{k=1}^{K} A_k \sin(\omega_k n + \phi_k) \tag{1}$$

ここで A_k，ω_k 及び ϕ_k はそれぞれ k 番目の正弦波の振幅，周波数及び位相を表わし，K は抽出する正弦波の数を表す．信号の再合成は上記のパラメータを用いて容易に行うことができる．解析時には解析対象信号と再合成した信号の誤差が最小となるようにパラメータを決定していく．

一般化調和解析で使用されるアルゴリズムのほとんどは解析対象信号から正弦波を一本ずつ引き抜くというものである．この際，引き抜いた後の残差信号のエネルギーが最小となるように引き抜く正弦波のパラメータを推定していく．一般化調和解析において知られているアルゴリズムは [3]–[8] 等が挙げられる．以上に挙げたアルゴリズムは正弦波パラメータ推定の方法が異なっているが，一本ずつ引き抜くという点は共通である．

著者は過去の研究において，以上に挙げた手法に比べ高速かつ再合成した信号の GDL(Generalized Distortion Level) が良い手法を提案した．以下では著者の提案した正弦波パラメータ推定の手法を述べる．

2.1 周波数推定

正弦波パラメータの推定を行うにあたり，振幅，周波数及び位相の 3 つのパラメータを同時に求めることは難しい．そこで，初めに周波数パラメータを推定し，その後位相及び振幅を推定する．

周波数の推定では

$$X_k(\omega) = \sum_{n=0}^{N-1} x_k e^{-j\omega n} \tag{2}$$

で表される離散時間フーリエ変換を使用する．ここでは離散時間フーリエスペクトルの最大値 $|X_k(\omega)|$ を与える ω の値を求め，この値を正弦波の周波数とする．$|X_k(\omega)|$ は

$$|X_k(\omega)| = \sqrt{X_{k_r}^2(\omega) + X_{k_i}^2(\omega)} \tag{3}$$

$$X_{k_r}(\omega) = \sum_{n=0}^{N-1} x_k(n) \cos(\omega n) \tag{4}$$

$$X_{k_i}(\omega) = \sum_{n=0}^{N-1} x_k(n) \sin(\omega n) \tag{5}$$

で与えられる．ここで $X_{k_r}(\omega)$ と $X_{k_i}(\omega)$ はそれぞれ (2) 式の実部と虚部を表す．

実際の周波数の推定では周波数軸上での探索を行う．解析対象信号に FFT を適用し，FFT スペクトルの最大値を与える周波数を $\omega(0)$ とする．この $\omega(0)$ を解析開始点とし，$\omega(0) - 2\pi/N$ から $\omega(0) + 2\pi/N$ の範囲内で最適な周波数 ω を探索する．ここで N はフレームサイズを表す．

探索範囲内では $|X_k(\omega)|$ が上に凸の関数となっていると仮定する．この仮定より極大となる点が一つのみ存在すると考える事ができる．よって

$$\frac{\partial |X_k(\omega)|}{\partial \omega} = 0 \tag{6}$$

となる ω を求めれば良い事になる．この式は

$$\begin{aligned} \frac{\partial}{\partial \omega}|X_k(\omega)| &= \frac{\partial}{\partial \omega}\sqrt{X_{k_r}^2(\omega) + X_{k_i}^2(\omega)} \\ &= \frac{X_{k_r}(\omega)\frac{\partial}{\partial \omega}X_{k_r}(\omega) + X_{k_i}(\omega)\frac{\partial}{\partial \omega}X_{k_i}(\omega)}{\sqrt{|X_k(\omega)|}} \\ &= 0 \end{aligned} \tag{7}$$

と表すことができる．ここで $\frac{\partial}{\partial \omega}X_{k_r}(\omega)$ と $\frac{\partial}{\partial \omega}X_{k_i}(\omega)$ は

$$\frac{\partial}{\partial \omega}X_{k_r}(\omega) = -\sum_{n=0}^{N-1} n x(n) \sin(\omega n) \tag{8}$$

$$\frac{\partial}{\partial \omega}X_{k_i}(\omega) = \sum_{n=0}^{N-1} n x(n) \cos(\omega n) \tag{9}$$

と表すことができる．

ここで仮定より $\frac{\partial}{\partial \omega}|X_k(\omega)|$ は単調減少関数であるので，ω を実数空間上での二分探索により求めることができる．二分探索では一回の計算で探索空間を 1/2 に狭めることができる．よって ω の計算精度は 2^{-M} となる．ここで M は探索回数を表す．言い換えれば ω を時間計算量 $O(-N \log \varepsilon)$ で計算することが可能である．ここで ε は推定される周波数の計算精度を表す．

実際の二分探索では (7) の符号のみを見て探索を行うため，(7) の分子の符号のみを考慮すればよい．

二分探索により推定された周波数を k 番目の正弦波の周波数とする．

2.2 位相推定

位相は

$$\phi_k = \arctan\left(\frac{X_{k_i}(\omega_k)}{X_{k_r}(\omega_k)}\right) \tag{10}$$

で表わされる．

2.3 振幅推定

振幅は残差の二乗和が最小となるように決定する．残差の二乗和 E_k は

$$E_k = \sum_{n=0}^{N-1} \{e_k(n)\}^2 \tag{11}$$

$$e_k(n) = x_{k-1}(n) - A_k \sin(\omega_k n + \phi_k) \tag{12}$$

で表わされる．

この値を最小化するために $\frac{\partial E_k}{\partial A_k}=0$ となる A_k を計算する．この式は

$$\begin{aligned} \frac{\partial E_k}{\partial A_k} &= \frac{\partial}{\partial A_k}\sum_{n=0}^{N-1}\{x_{k-1}(n)-A_k\sin(\omega_k n+\phi_k)\}^2 \\ &= \sum_{n=0}^{N-1}\{2(x_{k-1}(n)-A_k\sin(\omega_k n+\phi_k)) \\ &\quad \sin(\omega_k n+\phi_k)\} \\ &= 0 \qquad (13) \\ A_k &= \frac{\sum_{n=0}^{N-1}x_{k-1}(n)\sin(\omega_k n+\phi_k)}{\sum_{n=0}^{N-1}\sin^2(\omega_k n+\phi_k)} \qquad (14) \end{aligned}$$

で表わされる．これにより振幅の値が求まる．

2.4 パラメータの再計算

以上で述べた手法は周波数の推定に窓付き離散時間フーリエ変換を使用している．この窓付き離散時間フーリエ変換は矩形窓の影響を受け，スペクトル中でサイドローブを発生させる．一般化調和解析では対象信号を複数の正弦波の和として表現しているが，ある正弦波の周波数パラメータを推定する際，残りの正弦波のサイドローブの影響を受け，本来推定されるべき周波数から若干ずれた値が推定されるという問題が考えられる．

この問題を防ぐため，一度抽出された正弦波のパラメータを別の正弦波が抽出された後に再度計算する．実際には抽出した正弦波を他の正弦波を抽出した後に対象信号に足し合わせ，その信号から再度正弦波を抽出し，始めのパラメータに上書きする．これにより周波数パラメータを推定を抽出した他の正弦波のスペクトルのサイドローブの影響が少ない状態で行うことができる．

正弦波の抽出と再抽出を行う順番には様々なものが考えられる．以下では二つ挙げる．

2.4.1 単独再計算

単独再計算は正弦波パラメータを一本ずつ再計算するアルゴリズムである．新しく正弦波を抽出する前に，それまでに抽出した正弦波のパラメータを再計算していく．具体的には k 本目の正弦波を抽出する前に，既に抽出した 1 本目から $k-1$ 本目の正弦波をそれぞれを順番に，対象信号に再合成した後，1 本の正弦波を抽出し，抽出した正弦波のパラメータを再合成した正弦波パラメータに上書きするということを繰り返す．$i(1 \le i \le k-1)$ 本目の正弦波のパラメータを再計算する際，1 本目から $k-1$ 本目の正弦波が対象信号からほぼ取り除かれた状態で周波数推定を行うことができるため，1 本目から $k-1$ 本目の正弦波のスペクトルのサイドローブの影響が小さくなる．これにより i 本目の周波数推定の結果がより本来推定されるべき値に近づくことになる．

2.4.2 二重再計算

これまでに述べた問題の他にも，周波数及び位相が同一又は π ずれた 2 つの正弦波が抽出されるという問題が考えられる．これを避けるため，以下の二重再計算を用いて 2 つの正弦波を統一することを考える．

具体的には k 本目の正弦波を抽出する前に，すでに抽出した 1 本目から $k-1$ 本目の正弦波を周波数パラメータの昇順で並び変えた後，それぞれの正弦波パラメータを順番に再計算していく．初めに初期処理として 1 本目の正弦波を対象信号に再合成する．続いて 2 本目から $k-1$ 本目の正弦波について，i 本目の正弦波を対象信号に再合成し，$i-1$ 本目の正弦波パラメータを再計算するという手順を繰り返す．そして最後に後処理として $k-1$ 本目のパラメータを再計算する．これにより近い周波数をもった二つの正弦波を統一し，無駄な正弦波の抽出を避けることができる．

2.5 計算量

FFT，周波数推定，位相推定及び振幅推定にかかる時間計算量はそれぞれ $O(N\log N)$，$O(-N\log\varepsilon)$，$O(1)$，$O(N)$ 及び $O(N)$ である．正弦波を K 本抽出した場合，全体の時間計算量は

$$\begin{aligned} & O(K(N\log N - N\log\varepsilon + 1 + N + N)) \\ = \quad & O(NK\log\tfrac{N}{\varepsilon}) \qquad (15) \end{aligned}$$

となる．ここで ε は周波数推定における許容誤差を表す．

3 同一正弦波の抽出

一般化調和解析では対象信号をフレームサイズ N で時分割した後，各フレームについてフレームを構成する正弦波を K 本抽出する．この際，微小時間内で定常と考えられる解析対象信号においては，隣接するフレームにおいて周波数及び振幅がほぼ等しく，かつ波形がほぼ連続な同一と考えられる正弦波が抽出されると予想される．

以下では隣接するフレームにおける同一と考えられる正弦波の抽出アルゴリズムと，それを利用した複数のフレームにまたがる正弦波の抽出アルゴリズムについて述べる．

3.1 正弦波の連続性

始めに隣接するフレームに含まれる二本の正弦波の類似度を定義する．ここでは周波数及び振幅がほぼ等しく，かつ波形がほぼ連続している状態を類似している状態 (図 1) であると定義する．

類似の指標として相互相関関数によって定義される以下の類似度関数 $S(s_{i,k}(n), s_{i+1,l}(n))$ を導入する．

$$s_{i,k} = A_{i,k}\sin(\omega_{i,k}n+\phi_{i,k}) \qquad (16)$$

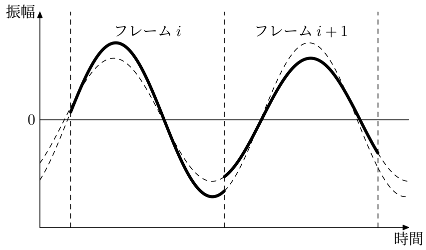

図 1: 正弦波の類似性

$$\begin{aligned} &S(s_{i,k}, s_{i+1,l}) \\ &= \sum_{n=0}^{2N-1} \left(s_{(i,k)}(n) s_{(i+1,l)}(n-N)\right) \end{aligned} \tag{17}$$

ここで $s_{i,k}$ は i フレーム目から抽出された k 本目の正弦波，$A_{i,k}$，$\omega_{i,k}$ 及び $\phi_{i,k}$ はそれぞれ $s_{i,k}$ の振幅，角周波数及び位相，N はフレームサイズを表す．この式は連続する 2 つのフレームに含まれる独立した二本の正弦波を互いのフレームまで延長し，これらの相互相関を取ったものを表している．この式の値は二つの正弦波の周波数が十分近くかつ連続性が大きい場合は大きくなり，逆の場合は低くなる性質を持っている．これをもって二つの正弦波の類似度とする．

3.2　正弦波ペアの組み合わせ

上に述べた方法を連続した二つのフレーム内に含まれる正弦波のペアに適用すると，正弦波同士の類似度を要素とした行列を作成することができる．この行列の中で各行各列に重複無く，かつ選んだ類似度の和が最大になるように類似度を選び，正弦波のペアの集合を選び出す．

この問題は重み付き二部グラフにおける最大マッチング問題として定式化することができる．この問題は Hopcroft-Karp 法 [9] 等により解くことができる．

これによって得られるマッチングは考えうる全ての正弦波のペアを列挙したものであり，周波数が大きく違う等，論理的に非同一であることが明らかな正弦波のペアも含んでいる．得たい結果は論理的に同一と考えられる正弦波のペアの集合であるので，これらのペアを相対的な周波数の変化比のハードスレッショルディングにより取り除く．残ったマッチングを結果として得る．

3.3　複数のフレームにまたがる連続性

上に述べたアルゴリズムを複数のフレームに連続して適用することにより，複数のフレームにまたがり論理的に同一と考えられる正弦波を抽出することができる．

4　実験

4.1　方法

実験では音声信号に対しこれまでに述べたアルゴリズムを適用し，複数のフレームにまたがった正弦波を検出できるかどうかを調べた．

一般化調和解析には第二章で述べた手法のうち再計算を行わない手法を NVIDIA 社より提供されている GPGPU 向け開発言語 CUDA を用いて実装したものを使用した．またフレームサイズは 512，正弦波の抽出本数は 16 本とした．重み付き二部マッチングには実装の都合上，最短経路探索に Bellman-Ford 法を使用した Primal-Dual 法を応用したアルゴリズムを使用した．これによる計算結果は Hopcroft-Karp 法による物と同等である．

実験に使用した音声信号は 3 つで，いずれも 44100Hz, 16bit，2ch の PCM 形式である．実験では 2 つの音声チャンネルのうち左側のチャンネルについてのみ解析を行っている．

音源 1 は実際に演奏したクラリネットの単音である．これはクラリネットの倍音の出現パターンが非常に特徴的であり，周波数解析を行うにあたり非常に特徴的な結果が得られることが予想されたために使用した．

音源 2 は周波数が線形変化及び対数変化する chirp 信号 2 本を合成したものである．提案手法が周波数に変化のある音声信号に対しても有効であることを確かめるために使用した．

音源 3 は JPOP の CD 音源から 1 秒間抜き出したものである．抜き出した部分には MML 表現で「O4L8D♯FD♯」と表されるシンセサイザーの演奏と，その他の楽器の音声が含まれている．

4.2　結果

実験の結果を図 2-7 に示す．周波数軸は計算上使用した各周波数単位である．図 2 及び図 3 は音源 1，図 4 及び図 5 は音源 2，図 6 及び図 7 は音源 3 の解析結果を表している．

図 3，図 5 及び図 7 中で緑で示した破線は一般化調和解析の結果であり，それぞれが各フレームにおいて抽出された正弦波を表す．各正弦波からは振幅，周波数及び位相パラメータが得られる．図 2-7 中で赤で示した実線は提案手法により検出されたフレーム間で連続する正弦波を結んだものである．提案手法により得られたマッチングを実線で結んでいる．

音源 1 では周波数 0.06 付近に基本音となる正弦波が含まれ，その整数倍の周波数を持つ正弦波が倍音として含まれている様子が示されている．また基本音や第二，第三倍音の周辺に細かな正弦波が現れている．これは実際

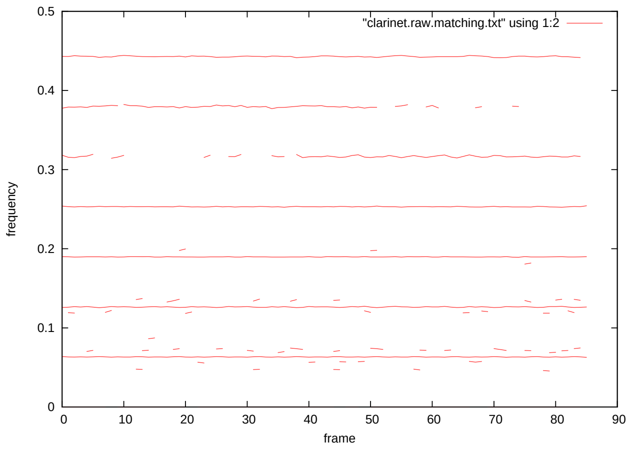

図 2: 音源 1

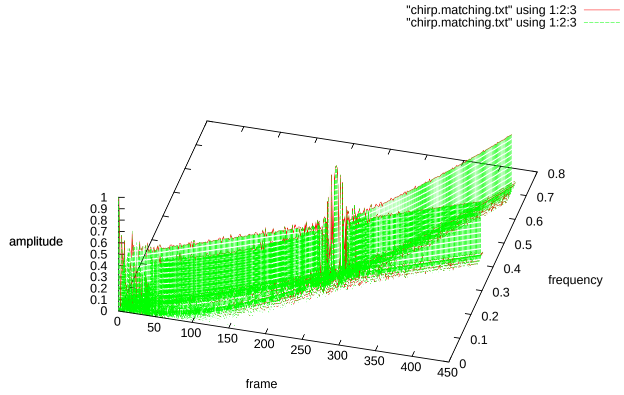

図 5: 音源 2

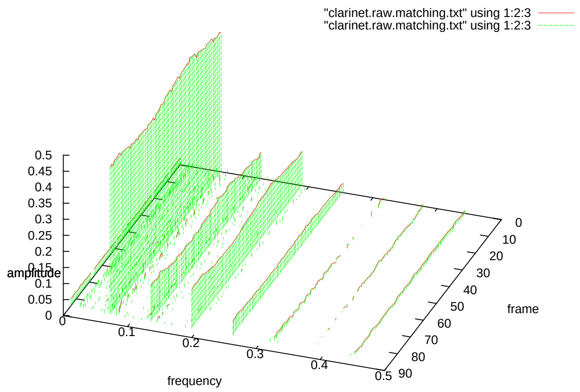

図 3: 音源 1

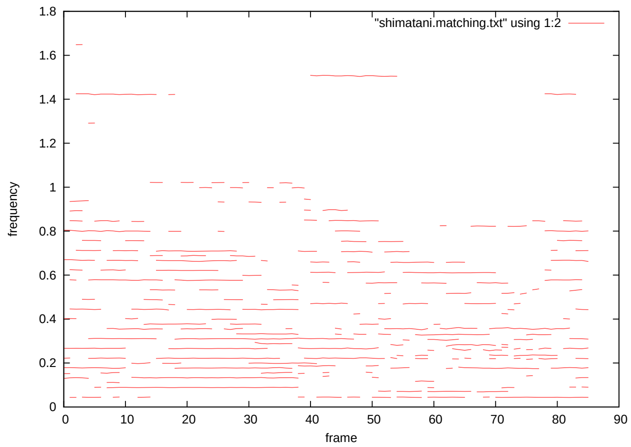

図 6: 音源 3

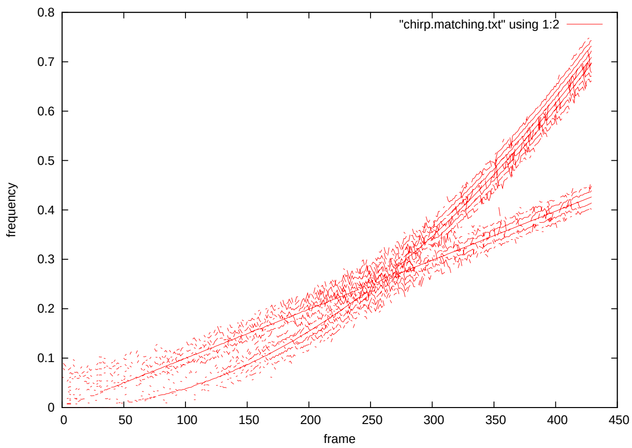

図 4: 音源 2

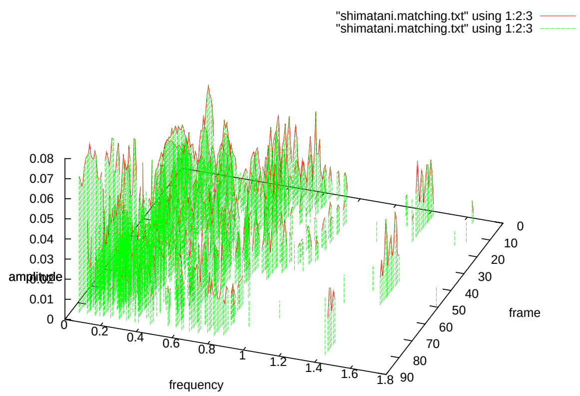

図 7: 音源 3

に演奏した際に生じたブレスや様々な揺らぎの成分だと思われる．特徴的な点は基本及び倍音の殆どがフレーム間をまたいで一本の正弦波として抽出されている点である．抽出された正弦波の周波数と振幅は揺らぎが多く，実際に演奏した音であることの特徴が出ている．

音源 2 では 2 本の chirp 信号がそれぞれ独立した正弦波として抽出されている様子が示されている．抽出された正弦波の周波数の時間変化は信号生成時のパラメータと一致している．250 フレーム付近で一度 2 本の正弦波が途切れているが，これは一般調和解析の際に 2 本の正弦波が合成され 1 本の正弦波として抽出されてしまったためである．また抽出された正弦波の周囲に振幅が小さく，かつ周波数変化の大きい正弦波が抽出されているが，これは一般調和解析において chirp 信号として抽出されなかった残差信号であると考えられる．これらの残差信号は短時間の chirp 信号と，これを正弦波で近似したもの差であるため，周波数変化の多い信号となっている．

音源 3 では細かい正弦波が多数抽出されているが，これらの多くはフレーム方向で補完する事により一本の正弦波として繋ぎ合わせることが出来る事が予想される．これは一般化調和解析による抽出本数が少ないため，本来抽出されるべき一本の正弦波の途中の部分が抽出されなかったために起こったものと思われる．この部分についても改善する必要があると思われる．

音源 3 の結果で注目すべき点は周波数 1.4 から 1.6 における三本の正弦波である．この三本は伴奏中の「O4L8D♯FD♯」という譜面に完全に一致している．これにより本手法が自動採譜等，基本周波数を重要視する解析手法に対して有効であると考えられる．

音源 1 から音源 3 までに共通して，フレーム間をまたいで単一の正弦波として抽出されている正弦波が多いという点が上げられる．これはあるフレームで抽出された正弦波が次のフレームでも抽出される可能性が高いと言い換えることが出来る．この特性により一般化調和解析アルゴリズム中の FFT の結果の代替として前フレームの解析結果を使用して高速化を行うことが出来るものと考えられる．また連続したフレームで同一の正弦波が得られるという点は予測符号化にも応用可能だと考えられる．

5 まとめ

一般化調和解析の結果を用いて対象信号から複数のフレームにまたがる正弦波を抽出する手法を提案した．提案手法を 3 種類の音声信号に対して適用した結果，楽器の単音からは基本音及び倍音に当たる正弦波を連続する複数のフレームにまたがって抽出することができ，2 本の chirp 信号を合成した音声信号からは独立した正弦波を複数のフレームにまたがって抽出することができた．また JPOP の前奏部分からは，演奏されている譜面と一致する正弦波を抽出することができた．

提案手法により得られた結果は，一般化調和解析によりあるフレームで抽出された正弦波の一部は次のフレームでも抽出されやすいということを示している．これらの特性は一般化調和解析の高速化や予測符号化に応用可能であると考えられる．

参考文献

[1] N. Wiener, “Generalized harmonic analysis,” *Acta Mathematica*, Vol. 55, pp.117-285, 1930

[2] 中沢誠, 山崎芳男, “1/12N オクターブ分析を用いた音の符号化,” *GITS/GITI research bulletin*, pp.81-85, July 2003

[3] 平田能睦, 小池恒彦, “一般調和解析を用いた音声帯域圧縮法,” **電子情報通信学会技術研究報告**. *EA,* **応用音響**, Vol. 98, No. 277, pp.17-24, Sept. 1998

[4] E. B. George, “Analysis-by-synthesis/overlap-add sinusoidal modeling applied to the analysis and synthesis of musical tones,” *J. Audio Eng. Soc.*, Vol. 40, No. 6, pp.497-515, 1992

[5] T. Terada, “Nonstationary waveform analysis and synthesis using generalized harmonic analysis,” *IEEE TF/TS Symp.*, pp.429-432, 1994

[6] 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,” **電子情報通信学会技術研究報告**. *EA,* **応用音響**, Vol. 93, No. 527, pp.39-44, Mar. 1994

[7] 村岡輝雄, 桐生晋也, “一般化調和解析 (GHA) の高速化に関する検討 (信号処理, LSI, 及び一般),” **電子情報通信学会技術研究報告**. *CAS,* **回路とシステム**, Vol. 103, No. 142, pp.1-6, June 2003

[8] 東山三樹夫, 小池恒彦, “高い周波数分析精度の信号分析手法,” **日本音響学会誌**, Vol. 54, No. 8, pp.568-574, Aug. 1998

[9] J. E. HOPCROFT, “An $n^{5/2}$ algorithm for maximum matchings in bipartie graphs,” *SIAM J. Comput.*, Vol. 2, pp.225-231, 1973