

小特集 音場・音響信号のモデルとその分析

音響信号の時間周波数分析*

山崎 芳 男 (早稲田大学理工学総合研究センター)**

1. まえがき

一般に時間領域と周波数領域は直交している。これは同じ信号が時間領域では $x(t)$ として、周波数領域は $X(f)$ としてどちらの領域でも独立にかつ正確に表現することが可能であることを示している。そしてこの両者を結び付けているのがフーリエ変換対である。

音響はもちろん多くの分野にとってフーリエ解析、調和解析はなくてはならない存在となっている。特に 1965 年 J. W. Cooley と J. W. Tukey がフーリエ変換を高速化する FFT のアルゴリズムを発表⁽¹⁾したのを期にデジタル信号処理は飛躍的に発展を遂げた。

ところで人間の聴覚機能はバンドパスフィルタ群で構成されている。本来直交関係にある時間領域と周波数領域ではあるが、聴覚機構を考えると 10 ~ 100ms 程度を分析単位とする時間周波数分析がむしろ自然である。本稿では音響信号の時間周波数分析について解説する。

2. 時間周波数分析

2.1 FFT と時間窓

時間領域と周波数領域はよく知られているように
フーリエ変換

$$X(f) = \int_{-\infty}^{\infty} x(t) \cdot e^{-j2\pi ft} dt \quad (1)$$

とフーリエ逆変換

$$x(t) = \int_{-\infty}^{\infty} X(f) \cdot e^{j2\pi ft} df \quad (2)$$

とにより結びつけられている。

FFT は離散化されたフーリエ変換において分析対象を素数の積に設定することにより同じ計算

の重複が生じ高速化されるという正にコロブスの卵的発想である。城戸らは 3 を基とした FFT を提案しているが⁽²⁾ 一般には最も小さい素数 2 のべき乗が分析窓幅として広く使われている。分析窓を 2^n としたときフーリエ変換には 2^{2n} の乗算回数が必要であるが、FFT では n 回に減る。例えば 1,024 点のフーリエ変換では乗算回数は従来の方法では約 104 万回行わなければならないが FFT では約 2 万回に、4,096 点の場合には実に約 1,677 万回の乗算が約 4 万 9 千回に減る。FFT は電子計算機のソフトウェアとしてばかりではなく専用ハードウェアとしても広く使われている。

しかし連続信号の短時間フーリエ変換 STFT はあくまでも分析窓幅を周期とする信号を解析しているということを常に念頭に置かなければならない。すなわち連続信号を適当に切り出して FFT 分析すると、起点と終点を結ぶ周期関数と見なされるので、切り出し窓幅が周期の整数倍でない限り不連続を生じ原信号にないスペクトルが観測されてしまう。

2.2 ウェーブレット変換とスカログラム

ウェーブレット (Wavelet) 変換

$$\sqrt{\frac{f}{f_1}} \int_{f_1}^f x(t') w \cdot \left\{ \frac{f}{f_1} (t-t') \right\} dt' \quad (3)$$

の 2 乗絶対値

$$\left| \sqrt{\frac{f}{f_1}} \int_{f_1}^f x(t') w \cdot \left\{ \frac{f}{f_1} (t-t') \right\} dt' \right|^2 \quad (4)$$

をスカログラム (Scalogram) と呼ぶ。

短時間フーリエ変換 (STFT: Short Time Fourier Transform) から求められるスペクトログラム (Spectrogram) が定周波数バンド幅分析であるのに対し、ウェーブレット変換によるスカログラムはオクターブ分析のような定対数バンド幅分析である。

2.3 ウィグナー分布

ウィグナー分布は 1932 年、量子力学の分野で

* Time frequency analysis of acoustical signals.

** Yoshio Ymasaki (Advanced Research Center for Science and Engineering, Waseda University, Tokyo, 169)

E. Wigner により紹介され³⁾、近年、電子計算機の性能向上と共に過渡現象の解析にしばしば使われるようになった。

時間信号 $f(t)$ の自己ウィグナー分布 $W(t, f)$ は、

$$W(t, f) = \int_{-\infty}^{\infty} f\left(t + \frac{\tau}{2}\right) \cdot f^*\left(t - \frac{\tau}{2}\right) \cdot \exp(-j2\pi f\tau) d\tau \quad (5)$$

となる。ここで $*$ は複素共役を示す。 $f(t)$ が $0 < t \leq a$ 以外で 0 の時間信号であるとすれば、その積分区間は $-a \sim +a$ となる。ウィグナー分布は時間と周波数の 2 次元関数であり、時間関数やそのフーリエ変換との関係が数学的にも明確であるのでたいへん扱い易い関数である。ウィグナー分布を周波数軸方向に積分すると時間信号のインパルス応答に、時間軸方向に積分するとパワースペクトルとなる。ウィグナー分布は周波数成分の時間変化という人間の聴覚機構とよくなじむ分析手

法であるが、ウィグナー分布は周波数と時間という本来直交関係にある 2 変数を使って一つの関数で表しているのでクロス項が生じる。

音場のように複雑は伝送系のウィグナー分布には多くのクロス項が現れる。平均化操作によってクロス項を軽減することもできるが、クロス項を音場の特徴をきわだたせる手段として積極的に利用することも可能である。すなわち、音場がデッドで単純な場合にはクロス項が少なく、逆に複雑な場合にはクロス項も多くなり音場の特徴を助長するので、可視化には有効な手段となるのではないかと考えたわけである⁴⁾。

図 - 1 に時間幅 1.5ms、周波数幅 1 ~ 1/240oct. で平均化を行い振幅の絶対値をデシベル表示したコンサートホールのウィグナー分布を示す。

2.4 累積 (cumulative) スペクトル
イギリス KEF 社の Fincham らがスピーカシステムの過渡特性を評価する時間周波数分析手法

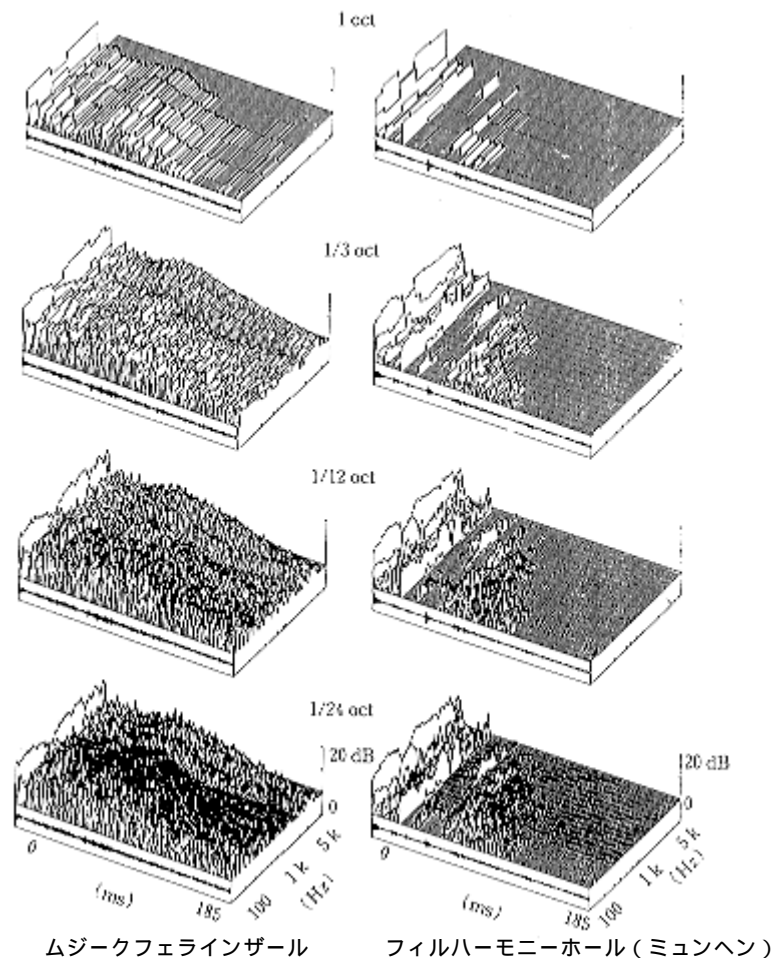


図 - 1 ウィグナー分布の平均化

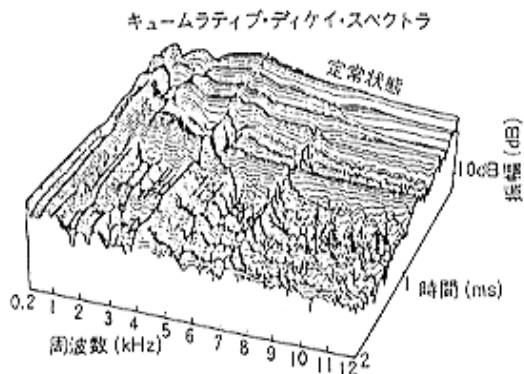


図 - 2 KEF の累積スペクトル処理例

として提案した方法で⁵⁾ ,インパルス応答 $h(t)$ から立ち上がり累積スペクトル,

$$\int_0^t h(t) e^{-j2\pi f t} dt \quad (6)$$

立ち下がり累積スペクトル

$$\int_t^\infty h(t) e^{-j2\pi f t} dt \quad (7)$$

として計算される。

図 - 2にスピーカの立ち下がり累積スペクトル処理例を示す。

2.5 一般化調和解析 (GHA)

前述のようにフーリエ解析は 周期的で調和的な信号に対する基礎理論でFFTなど広く利用されているが 信号は観測時間外では観測時間内の波形が周期的に繰り返されると仮定しているので 観測区間によって多数の異なる周波数分析が抽出されるという欠点がある。

一方、1958年にWienerの提案した一般化調和解析は観測区間内で原波形から残差エネルギーが最小となる正弦波を抽出し、残差成分に同様の処理を繰り返すという単純明快な解析方法である⁶⁾。従って定常的でない僅かな周波数変動に対しても正確な周波数成分の抽出が可能であり、観測区間を越えて信号の予測が可能であるなどの特徴がある。

このような多くの特徴を持つ一般化調和解析ではあるが、膨大な演算時間を必要とするので実用の信号の分析に使われることはほとんどなかった。平田らは⁷⁾効率的な算出手法を提案し楽器音の分析などを試みている。

一般化調和解析は信号 $x(t)$ を $0 \sim L$ 間で観察したとき:

$$\varepsilon(t, f) = x_0(t) - A(f) \cdot \sin(2\pi f t + \Phi) \quad (8)$$

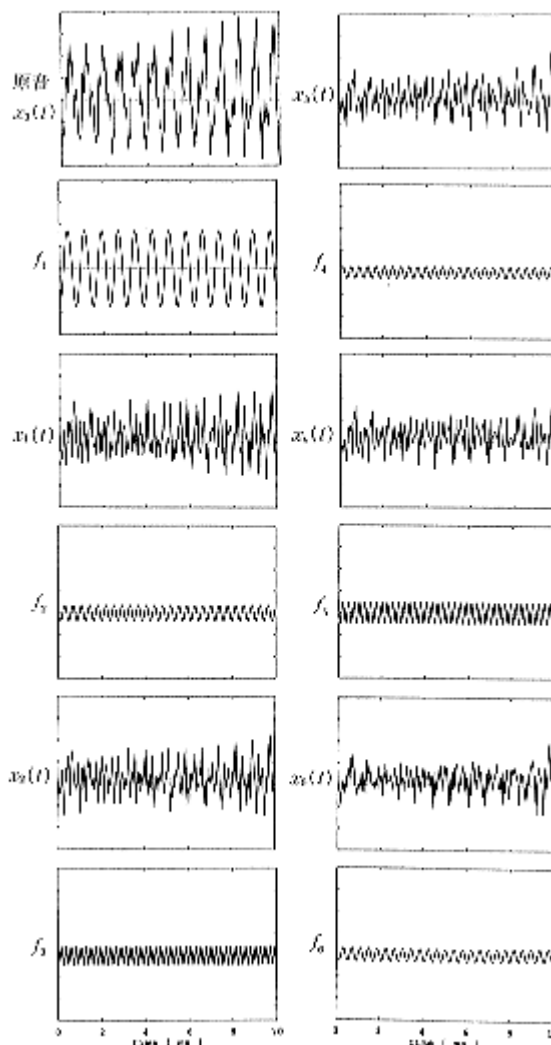


図 - 3 一般化調和解析の手順

のエネルギー

$$E(f) = \int_0^T \varepsilon(t, f)^2 dt \quad (9)$$

を最小とする純音の周波数 f_1, A_1, Φ_1 を決定する。

次に残差

$$x_1(t) = x_0(t) - A_1(f_1) \cdot \sin(2\pi f_1 t + \Phi_1) \quad (10)$$

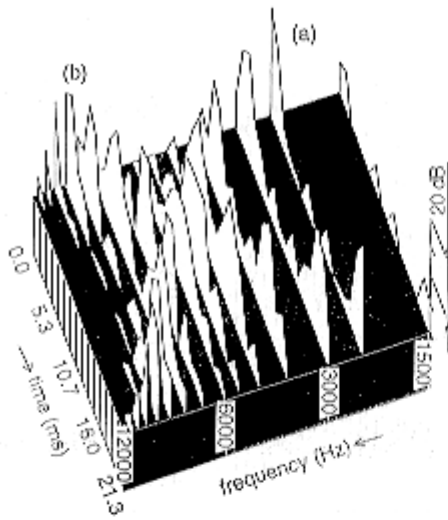
について同じを操作を繰り返し、順次周波数成分を決定していく。図 - 3にバイオリン独奏を例に一般化調和解析の手順を示す⁸⁾。

一般化調和解析は非調和周波数分析の分析が可能であるという特長を持っている。

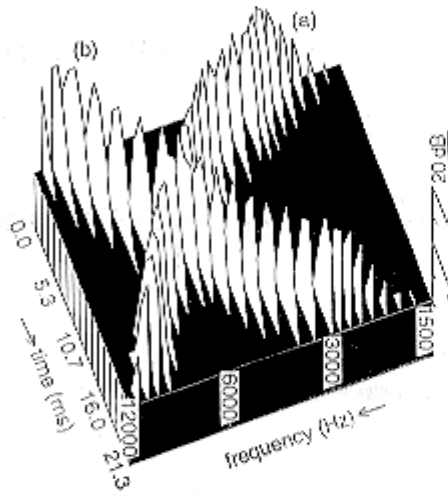
パワースペクトル $P(f_k)$ は

$$P(f_k) = A_k^2 \quad (11)$$

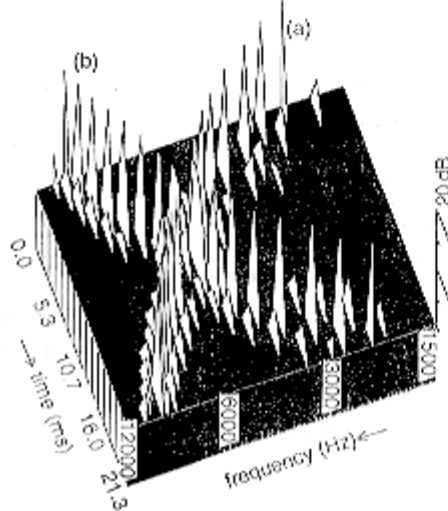
によって推定される。



(a) STFT

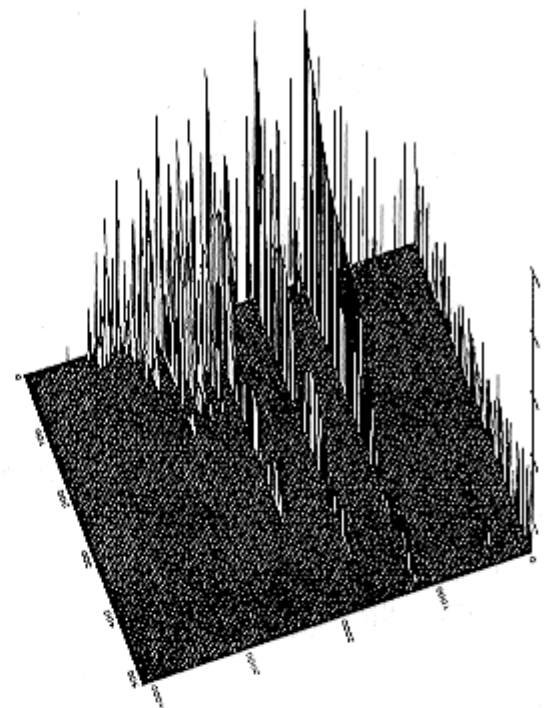


(b) Wavelet

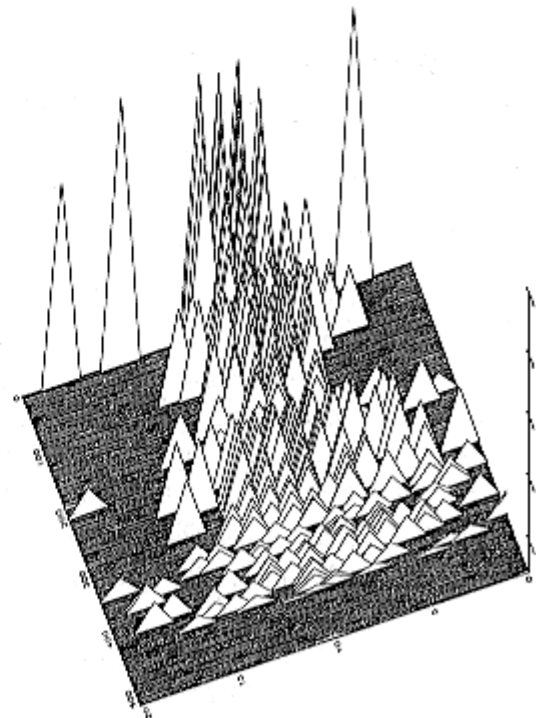


(c) GHA

図 - 4 時間周波数分析の比較 (東山らによる)



(a) 帯域 : 47 (Hz) - 12 (kHz), $N/oct : 32$



(b) 帯域 : 47 (Hz) - 12 (kHz), $N/oct : 32$

図 - 5 ピアノ単音の一般化調和解析 (平田, 東山)

図 - 4 に東山らによる二つのチャープ (Chirp) 信号を (a) 短時間フーリエ変換して得たスペクトログラム, (b) ウェーブレット分析したスペクト

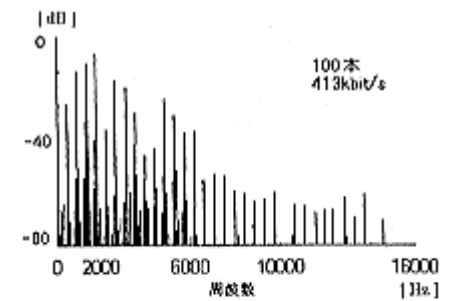
ログラム (c) 一般化調和解析した結果を示す。
 東山, 平田らの行った一般化調和解析によるピアノ音の分析結果を図 - 5 に示す。

3. 一般化調和解析による時間周波数分析の応用

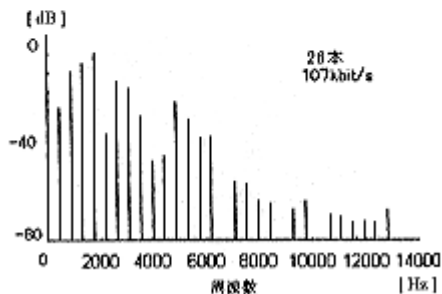
3.1 聴覚特性を考慮した高速1bit信号の高エネルギー符号化

著者らは高速1bit音響信号処理を提案し, 最も高エネルギー符号化, 時間軸圧伸・ピッチ変換等の信号処理についても検討を加えてきた^{8),9)}。高速1bit信号は量子化雑音が高域に集中するので一般化調和解析では量子化雑音の成分が抽出されてしまう。そこで, 周波数成分を可聴領域に限定する。オーボエ演奏の信号を窓幅11msで一般化調和解析を試みた。図 - 6(a) に示すように100成分抽出し, それを合成すると聴感上, 原音と違いのない音が合成可能であった。

聴覚特性にはマスキングや臨界帯域がある。次に聴覚に着目し一般化調和解析により高速1bit信号区間ごとにデータレートを変化させる不等長符号化を検討した。分析区間により違いはあるものの図 - 6(b) に示す平均で約26本程度, 伝送容



(a) 一般化調和解析による周波数成分と波形



(b) 聴覚特性を考慮した周波数成分と波形

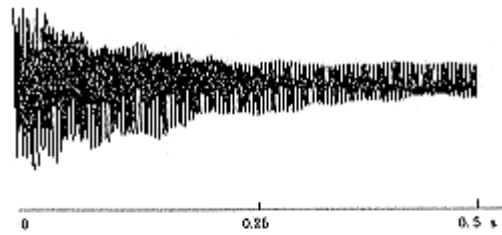
図 - 6 一般化調和解析による高エネルギー符号

量的107kbit/sで聴感上原音と変わらない高エネルギー符号化が可能であることが確認された。

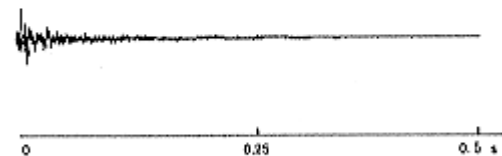
3.2 時間軸圧伸・ピッチ変換

一般化調和解析の主周波数成分の時間変化が正確に抽出できる特長を生かして, 時間軸圧伸・ピッチ変換を紹介する¹⁰⁾。

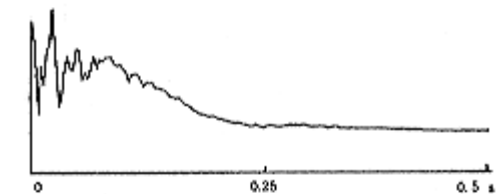
時間軸圧伸・ピッチ変換は抽出された主周波数成分の時間変化に基づいて, ピッチを変換するには振幅変化を保存したまま周波数のみ変換比率に応じて変更する。



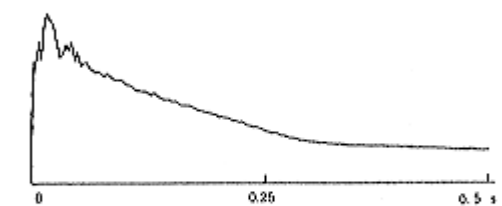
(a) ピアノの原波形



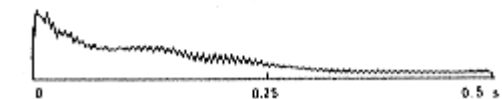
(b) 原波形から主周波数成分30本を抜いた残差



(c) 439 Hz の振幅の時間変化



(d) 880 Hz の振幅の時間変化



(e) 2,659 Hz の振幅の時間変化

図 - 7 周波数成分の時間変化

一方時間軸圧伸には周波数成分をそのまま振幅の時間変化を変換比率に応じて圧縮・伸張する。

ピアノの音(基音 880Hz)を処理した結果を図 - 7 に示す。(a) は原音の時間波形であり,(b) はこの方法で周波数成分を 30 本抽出した算差である。(c) は、抽出した主周波数成分(439, 880, 2,659Hz)である。図 - 8 の(a) は原波形,(b) は主周波数成分 30 本を抜いた残差,(c) は主周波数成分は抽出した時間変化に基づいて合成した波形,(d) はピッチを 1.5 倍に上昇させた,(e) は

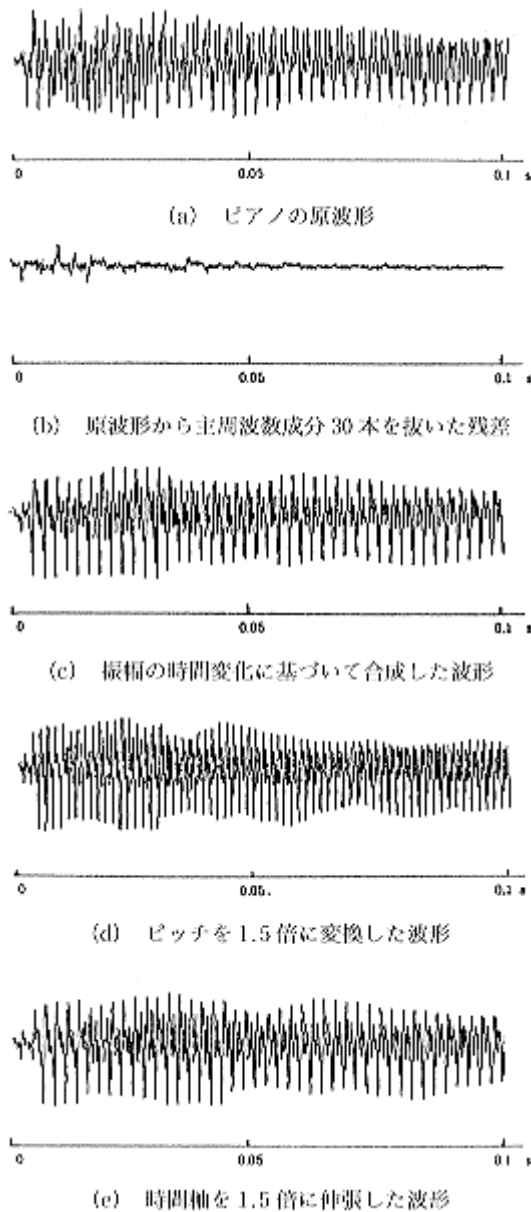


図 - 8 ピアノ単音の時間軸・ピッチ変換処理

時間軸を 1.5 倍に伸張した波形である。

3.3 室内音場解析への一般化調和解析の導入
音場のインパルス応答を短い区間に分割し、区間ごとに一般化調和解析することにより時間と共に変動する周波数成分の分析を試みた¹¹⁾。

図 - 9 は愛知県芸術劇場において測定したインパルス応答である。分析はデータ点数 1,024 点(標本化周波数 48kHz) 約 21ms の短区間に矩形窓によってインパルス応答をそのまま切り出して一般化調和解析を行った。それぞれの区間にお



図 - 9 愛知芸術劇場のインパルス応答

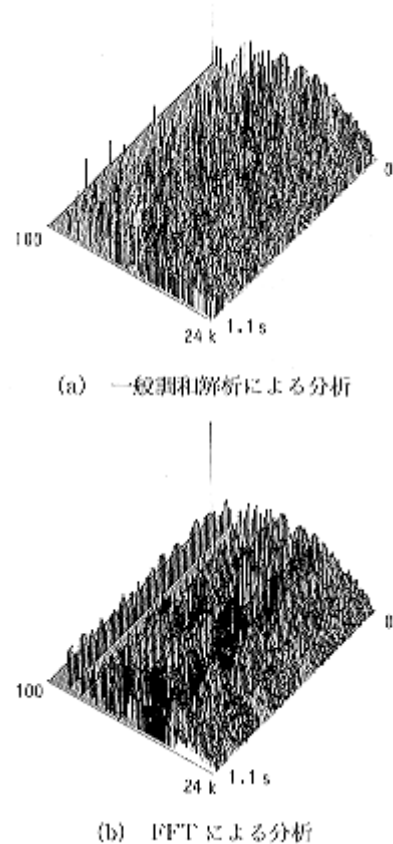


図 - 10 愛知芸術劇場の時間周波数分析

いて200本の周波数成分を抽出した結果を図 - 10 (a)に示す。

同様にデータ点数1,024点の矩形窓によってインパルス応答を切り出し,FFTによって解析を行った結果を図 - 10(b)に示す。周波数は対数軸で表している。

周波数分布の時間変化については,FFT,一般化調和解析,いずれの場合も高域の周波数成分は低域の周波数より早く減衰しているのが分かる。FFTによる解析では解析する周波数は分析窓幅により決定される。従って,解析結果は分析点数で決まる周波数に現れ,周波数成分の変動等の細かな分析を行うことが不可能である。一方,一般化調和解析は短い観測区間を用いて微細な周波数変動を検出可能であることから,モードの変化の様子,音場の拡散性の様子を観測することに適していると考えられる。

4.むすび

以上音響信号の時間周波数分析について解説した。本来直交している時間,周波数領域での分析であるから,それぞれの方法の前提条件や制約を十分理解して使う必要がある。

コンピュータの処理能力の飛躍的な向上によりWienerが1950年代に提唱した調和解析の原点とも言える一般化調和解析ですら,時間はかかるもののパーソナルコンピュータで実行することが可能となった。

一般化調和解析は定常的でないわずかな周波数

変動に対しても正確な周波数成分の抽出が可能であり,観測区間を越えて信号の予測が可能であるなどの特徴を持ち,人間の聴覚特性によくなじむ時間周波数分析手法であり,音響信号処理への応用が期待される。

文 献

- 1) J.W.Cooley and J.W.Tukey, "An algorithm for the machine computation of complex Fourier series," Math. Comput. 19, 297 - 301 (1965).
- 2) 城戸健一, デジタル信号処理入門(丸善,東京,1985).
- 3) E.Wigner, "On the quantum correction for thermodynamics equilibrium," Phys. Rev. 40, 749-759 (1932).
- 4) 伊勢史郎, 清山信正, 岡田利一, 山崎芳男, 伊藤 毅, "近接4点法により得られた仮想音源の展開一昔場の過渡状態の把握 -," 音講論集, 57ト572 (1987).
- 5) J.M.Berman and L.R.Fincham, "Loudspeaker evaluation and simulation system using digital techniques," AES Convention Pre Print (1975.3).
- 6) N.Wiener, The Fourier Integral and Certain of Its Application (Dover, New York, 1958).
- 7) 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, 39 - 44 (1994).
- 8) 唐川周三, 遠藤一夫, 山崎芳男, 西川明成, 太田弘毅, "帯域分割一般調和解析による高速1bit 高能率符号化," 音講論集, 585 - 586 (1996.3).
- 9) 天田 武, 大内康裕, 及川靖広, 工藤修二, 山崎芳男, "聴覚特性を考慮した一般調和解析による高速1bit 信号の高能率符号化," 音講論集, 599 - 600 (1996.9).
- 10) 林 雅尚, 山崎芳男, 及川靖広, "一般調和解析を用いた高速1bit 音響信号の時間軸圧伸・ピッチ変換," 音講論集, 603 - 604 (1996.9).
- 11) 及川靖広, 遠藤一夫, 工藤修二, 山崎芳男, "室内音場解析への一般調和解析の導入一大空間音場の解析 -," 音講論集, 757 - 758 (1996.9).