

ウィグナー分布とフーリエ変換のハイブリッドスペクトル解析法 による音声信号解析

鈴木 宏 *

A Voice Signal Analysis with A Method of Spectral Analysis using A Hybrid of Wigner Distribution and Fourier Transform

SUZUKI Hiroshi

Wigner Distribution (WD) has recently attracted attention from many researchers as a tool for time-frequency signal analysis. WD offers higher frequency resolution but with interference terms for multiple frequency components. A method of spectral analysis as a hybrid of WD and FT is proposed for time-varying signals. The method offers higher frequency resolution with neither interference terms nor aliasing. This method is applied to measurement of a voice signal.

キーワード : Wigner Distribution, Fourier Transform, Spectrum Analysis, Voice Signal

1. ま え が き

時変信号の解析法の一つとしてウィグナー (Wigner) 分布^{(1),(2)}があり、いくつかの分野に応用されている^{(3),(4)}。このウィグナー分布は、特殊な前処理を行った信号のフーリエスペクトルを求めるものである。この方法は、周波数分解能がフーリエ変換のそれに比べ優れており、分解能が高いため、短いウインドウ長でも時変信号の局所的な変化を調べられる。しかしこの前処理により、複数の周波数成分を含む信号で、その成分間に干渉項が生じる。また、高周波数成分によるエイリアシングも発生する。

これらを防ぐために、ウィグナー分布は一般に、原信号の実部と虚部が互いにヒルベルト変換 ($\cos(t)$ に対して $\sin(t)$ を与えるような、位相を 90 度シフトさせる効果をもつ変換) の関係になっている解析信号に変換したのちに計算を行う。この解析信号を用いれば、負周波数成分との干渉項およびエイリアシングの影響を取除くことはできるが、正周波数成分同士の干渉項は依然残り、スペクトルを汚す原因となる。

ウィグナー分布がフーリエ変換の実部と虚部のそれぞれの畳込みの和で表せることに注目し、畳込みを狭い範囲で行うことにより、この 2 つの分析法を融合

させた解析法であるハイブリッドスペクトル解析法を提案した⁽⁵⁾。この手法は、ウィグナー分布と同様の非常に高い周波数分解能が得られる上に、範囲を変えてフーリエスペクトルを畳込んだ中の最小値を求めることで、ウィグナー分布でスペクトルを汚す正周波数成分同士の干渉項が軽減されることが特徴である。この解析法を用いることで、小さなウインドウでの低周波数域のスペクトル解析や、時変信号の局所的な周波数変化の解析ができる。

本論文では、この手法を音声信号解析に適用するための、解析に必要な処理手順を示し、実際の音声データに適用し、この解析法をフーリエ変換やウィグナー分布と比較して本手法の有効性を示す。また、サンプリングタイムを変えることで、時変信号解析に及ぼす影響や解析時間に対する考えを考察する。

2. 各種解析法の概要

2-1 ウィグナー分布

サンプル間隔 t_0 の 1 次元時間離散信号 $f(nt_0)$ の離散ウィグナー分布 (以下 DWD) : $W(nt_0, k\omega_0)$ は

$$W(nt_0, k\omega_0) = \sum_{m=-\frac{M}{2}}^{\frac{M}{2}} f(t_0(n+m)) f^*(t_0(n+m)) e^{-jmt_0k\omega_0}$$

* 工学科 情報エレクトロニクス系 教授

原稿受付 2022 年 5 月 20 日

$$= \sum_{m=-\frac{M}{2}}^{\frac{M}{2}} R(t_0, n, m) e^{-jmt_0 k \omega_0} \quad (1)$$

ここで、 M : ウィンドウ点数,

* : 複素共役,

$$\omega_0 : \text{基本周波数} (= \frac{2\pi}{Mt_0})$$

で定義されている⁽⁵⁾.

2-2 ウィグナー分布と

フーリエ変換のハイブリッド法⁽⁵⁾

パワースペクトルと DWD の間に、連続的に存在する変換

(a) $(2l)$ 位置 [偶数項]

$$CS_{(2l, C_N)} = \frac{2}{C_N + 1} \sum_{r=-\frac{C_N-1}{2}}^{\frac{C_N-1}{2}} (A_{l-r} A_{l+r} + B_{l-r} B_{l+r}) \quad (2)$$

ただし、 $C_N = 1, 3, 5, \dots, \frac{M}{2} - 1$ で畳み込む離散

フーリエ変換 (以下 DFT) の成分数を表す。

(b) $(2l+1)$ 位置 [奇数項]

$$CS_{(2l+1, C_N)} = \frac{2}{C_N} \sum_{r=0}^{\frac{C_N-2}{2}} 2(A_{l-r} A_{l+1+r} + B_{l-r} B_{l+1+r}) \quad (3)$$

ただし、 $C_N = 2, 4, 6, \dots, \frac{M}{2}$ で畳み込む DFT

の成分数を表す。

ここで、 $l = 1, 2, 3, \dots, \frac{M}{2} - 1$, A_k, B_k は DFT の値

$$(1 \leq k \leq \frac{M}{2} : \text{整数値, これ以外の } k \text{ では } 0)$$

がある。

C_N の変化に対する畳み込みスペクトルの最小値を求めるスペクトル解析法を、

$$CSM(nt_0, k\omega_0) = \min[CS_{(k, C_N)}] \quad (4)$$

ここで、 $\min[]$ は、 $[]$ 内の最小値を表す。

と定義する。これをハイブリッド法と呼ぶ。

この解析法は、

1. ウィグナー分布と同様の特性を持つために、周波数分解能が非常に高い。

2. 畳み込みスペクトルの最小値を求めているため、正周波数同士の干渉項が軽減される。

の特徴がある。

この解析法を用いることで、小さなウィンドウでも低周波数のスペクトル解析や時変信号の局所的な周波数変化の解析ができる。

3. 音声信号解析への適用

今回は、GRAPHTEC 製のロガーで測定を行い、Matlab により解析を行った。信号のスペクトル解析において、対象とする信号の最大周波数がいくつであるかが重要となる。音声を解析する場合、① 声の最大周波数を 5kHz と考える場合と、② 人間の可聴範囲を考慮して、最大周波数を 20kHz と考える場合がある。今回はこの①と②について、以下の手順で実際の測定および解析を行った。

1. サンプリングタイムの設定

① 声の最大周波数を 5kHz と考え、サンプリング定理より 10kHz (100 μ s) とした。

② 人間の可聴範囲を 20Hz~20kHz と考え、サンプリング定理より 40kHz (25 μ s) としたいが、ロガーの仕様より 50kHz (20 μ s) と変更した。

2. フィルタの値

エイリアシング (高周波数成分の低周波数成分への折り返し) を防ぐために、サンプリング周波数の半分以上の周波数をカットする必要がある。

① サンプリング周波数が 10kHz のため、5kHz 以上の周波数をローパスフィルタで取り除く必要があり、5kHz のローパスフィルタを用いた。

② サンプリング周波数が 50kHz のため、25kHz のローパスフィルタを用いたが、この場合も装置の関係で 20kHz のローパスフィルタを用いた。

3. 解析時間

何点の解析を行うかで、基本周波数 (最低周波数) と周波数分解能が決まる。ピアノの音を解析する場合、中央のラの音が 440Hz で半音上がった音は 466.16Hz である。この 2 音の差 26.16Hz を分けて周波数上に表したい場合、周波数分解能としての周波数の刻み幅は 13Hz は欲しい。

① 声の場合、高い分解能は必要ではなく基本周波数を 20Hz 程度考え、解析点数をサンプリングタイム 100 μ s で 2 の N 乗の 512 点 (51.2ms) の基本周波数 19.53Hz とした。

② サンプリングタイムが 20 μ s のため、基本周

波数は、1024 点 (20.48ms) の場合 48.83Hz, 2048 点 (40.96ms) の場合 24.41Hz, 4048 点 (81.92ms) の場合 12.21Hz となる. 今回は①と同じ 20Hz 程度を基本周波数としたいため, 点数を 2048 点 (40.96ms) として解析を行った. 基本周波数を同じくらいにしたため, 解析時間は同じくらいになったが, サンプルングタイムが違うため, 解析する点数は 4 倍になっている. この違いが結果にどう反映するかは 4.3 節で述べる.

4. ウィンドウの設定

DFT を計算する場合, 有限長の窓関数を掛け合わせてフーリエ変換を行う. 窓関数は, 方形窓, ハニング窓, ハニング窓などがあるが, 今回は代表的に使用されているハニング窓

$$w_N = 0.5 - 0.5 \cos\left(\frac{2\pi n}{N}\right) \quad 0 \leq n \leq N - 1$$

(5)

ここで, N は窓の点数 (解析点数) を使用した.

4. 解析結果

4-1 サンプルングタイム 100 μ s の場合 (①)

図 1 に示すような「あ・い・う」と順番に続けて発音をした信号を 100 μ s でサンプルングし, 窓関数 (512 点, 51.2ms) により切り出し, DFT を行った後に窓関数を 3ms ずらしながら 0.5s まで 150 回の DFT を行った結果をショートタイムフーリエ変換 (以下 STFT) として図 2 に示す. 同様に, 512 点の解析点数で 3ms ずつずらして解析した DWD とハイブリッド法の結果をそれぞれ図 3 と図 4 に示す.

図 3 より DWD 正周波数成分同士の干渉項があり主成分がどこであるかが他の手法に比べ分かりづらくなっている. 図 2 と図 4 を比較すると同じ周波数の位置でピークがでていますが, ハイブリッド法の方はピークが鋭く主成分位置がはっきりしていることが分かる.

4-2 サンプルングタイム 20 μ s の場合 (②)

①の場合と同様に, 図 5 に示すような「あ・い・う」と順番に続けて発音をした信号を 20 μ s でサンプルングし, 窓関数 (2048 点, 40.96ms) により切り出し, DFT を行った後に窓関数を 3ms ずらしながら 0.5s まで 150 回の DFT を行った結果を STFT として図 6 に示す. 同様に, 2048 点の解析点数で 3ms ずつずらして解析した DWD とハイブリッド法の結果をそれぞれ図 7 と図 8 に示す.

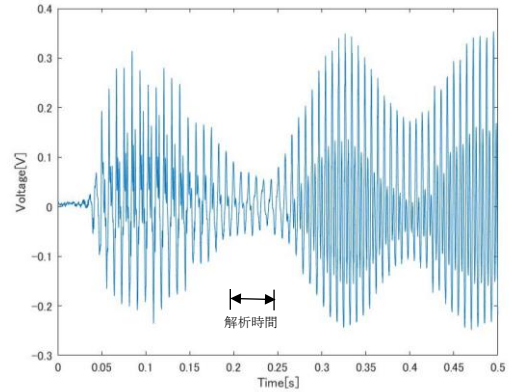


図 1 「あ・い・う」の音声信号 (Sampling Time 100 μ s)

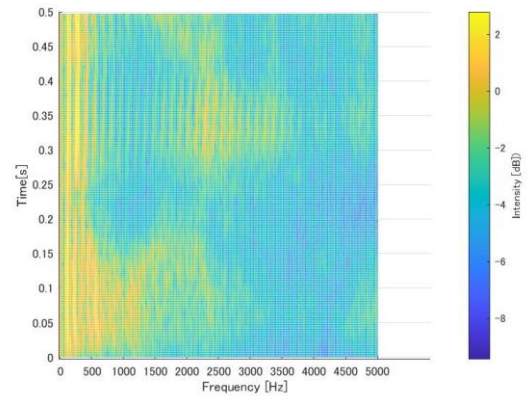


図 2 図 1 のショートタイムフーリエ変換

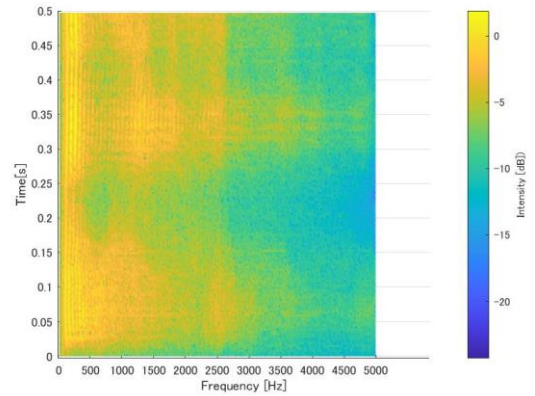


図 3 図 1 のウィグナー分布による解析結果

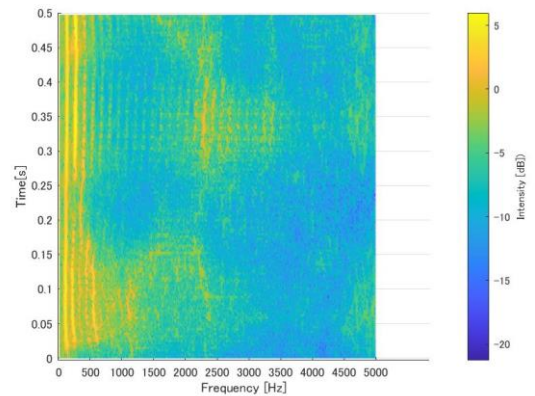


図 4 図 1 のハイブリッド法による解析結果

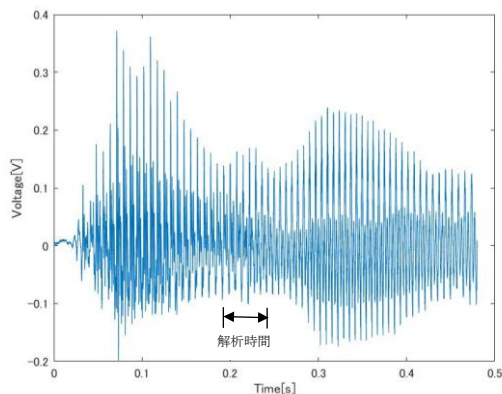


図5 「あ・い・う」の音声信号 (Sampling Time 20 μ s)

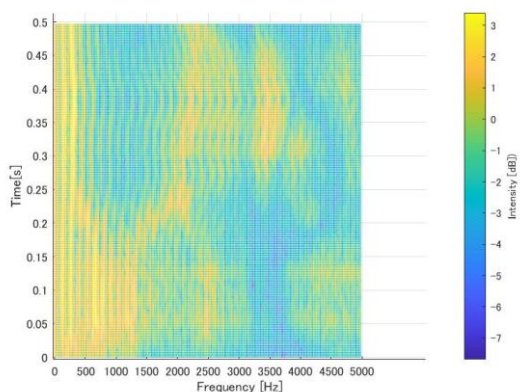


図6 図5のショートタイムフーリエ変換

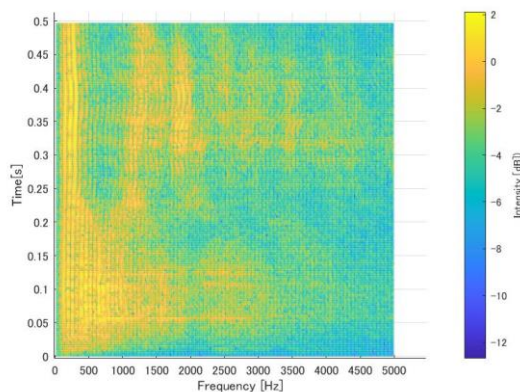


図7 図5のウィグナー分布による解析結果

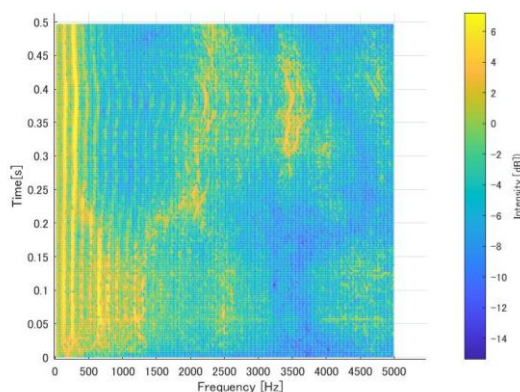


図8 図5のハイブリッド法による解析結果

3つの手法の解析結果の違いは①の場合と同じで、ハイブリッド法が一番ピークが鋭く主成分位置がはっきりしていることが分かる。

4-3 サンプリングタイムが及ぼす影響

①の100 μ sのサンプリングタイムと②の20 μ sのサンプリングタイムの違いにより、解析結果がどのように異なるかを考察する。

サンプリングタイムが同じ場合、解析点数(解析時間)が多いほど基本周波数が小さくなり周波数分解能(周波数軸の刻み幅)が良くなるが、窓関数長が長くなり信号が大きく切り取られるため、時間軸の局所の変化を解析することができない。反対に解析点数が少ないと時間軸の局所変化を見て取れるが、基本周波数が高くなり周波数分解能が悪くなる。今回は基本周波数をほぼ同じ値とし解析点数を変えた。そのため周波数分解能がほぼ同値で図2~図4と図6~図8を比較しても見た目はあまり変わらない。しかし、測定する点数は、②が25000点で①が5000点と5倍の差があり、測定時間も解析時間も②が多くかかる。

解析結果を見るとハイブリッド法が主成分が明確になっており特徴抽出ができていていると考えられるため、音声信号解析には100 μ sのサンプリングタイムでハイブリッド法での解析が最適と考える。

5. あとがき

音声信号解析に必要な処理手順を示し、実際の音声データに適用し、各種解析方法での結果を示した。また、サンプリングタイムを変えても基本周波数が同値の場合、解析結果にはあまり影響がないことが分かった。さらに、音声信号解析においてもハイブリッド法が有効であることが確認できた。ハイブリッド法を他の信号へ適用することや解析時間の短縮が今後の課題である。

参考文献

- 1) T.A.C.M. Claasen and W.F.G. Mecklenbauer : The Wigner Distribution -Tool for Time-Frequency Signal Analysis-PART 1.2.3, Philips J.Res.,35, (1980), pp.217-250,pp.276-300, pp.372-389
- 2) B.Boashash : Note on the Use of the Wigner Distribution for Time-Frequency Signal Analysis, IEEE Trans. Acoust., Speech & Signal Process., ASSP-36,9,(1988),pp.1518-1521, pp.1006-1013
- 3) LEON COHEN : Time-Frequency Distribution -A Review, Proceedings of the IEEE, Vol.77, No.7,(1989),pp.941-981
- 4) Boualem Boashash : Time-Frequency Signal Analysis, Longman Cheshire (1992)
- 5) 鈴木宏, 小林史典, 戸谷順信: ウィグナー分布とフーリエ変換のハイブリッドスペクトル解析法による表面波動現象の解析, 日本機械学会論文集 (C編), 68 巻 671 号 (2002.7), 2037-2044