

低周波数分解能スペクトル群からの高周波数分解能スペクトルの合成による 複数分解能 Gabor 変換の計算量削減

Complexity Reduction of Multi Resolution Gabor Transform by Synthesis of High Frequency Resolution Spectrum from Low Frequency Resolution Spectra

高山亮輔[†]
Ryousuke Takayama

荒井秀一[†]
Shuichi Arai

1. まえがき

楽音や音声などの非定常信号の時間-周波数解析には、短時間フーリエ変換 (STFT) [1] がよく利用される。信号の解析には高い時間分解能、周波数分解能が望ましいが、STFT では、フーリエ変換の不確定性原理に縛られ、時間分解能と周波数分解能がトレードオフの関係にあるため、高い時間分解能、周波数分解能を同時に得ることはできない。

そこで、Lukin らは複数の分解能を適応的に選択する手法を提案し、ノイズリダクションなどの応用事例でその効果を確認している [2]。しかし、この方法は、ある解析したい時刻、周波数において複数の分解能でスペクトルを求める必要がある。そのため、フレームシフト幅は最も時間分解能が高いものに、周波数 bin は最も周波数分解能が高いものにあわせている。この様に、細かいフレームシフト幅、周波数 bin での STFT を複数回行う必要があり、計算量が膨大となる問題が存在する。

そこで、本稿では複数分解能による Gabor 変換 [3] の計算量を削減するため、Gaussian の数学的性質に着目して、低周波数分解能によるスペクトルから、高周波数分解能によるスペクトルを算出する方法を提案する。

2. Gabor 変換の計算量削減

Gabor 変換は、STFT の窓関数に Gaussian を用いる解析法である。そのため、Gabor 変換の時間-周波数分解能は Gaussian の標準偏差 σ によって決まる。複数分解能による Gabor 変換では、複数の σ を用いて解析を行う。本章では、低周波数分解能を用いて算出したスペクトルから、高周波数分解能を用いて算出するスペクトルを合成し、計算量を削減する手法を提案する。

2.1. スペクトル合成法

Gabor 変換によって求まる時刻 τ のスペクトルは連続域では、式 1 のように表される。

$$X_{\sigma}(\tau, \omega) = \int_{-\infty}^{\infty} x(t) g_{\sigma}(t - \tau) e^{-i\omega(t - \tau)} dt \quad (1)$$

ここで、 g_{σ} は標準偏差 σ の Gaussian、 ω は中心角周波数である。

Gabor 変換の σ を操作するために、我々は Gaussian 同士を畳み込むと任意の倍率 α ($\alpha > 1$) の標準偏差の Gaussian が得られる数学的性質に着目した。この性質は式 2 で表される。

$$g_{\alpha\sigma}(t) = \int_{-\infty}^{\infty} g_{\sigma}(t - \xi) g_{\sqrt{\alpha^2 - 1}\sigma}(\xi) d\xi \quad (2)$$

[†] 東京都市大学大学院 工学研究科
Graduate School of Engineering, Tokyo City University

ここで、 $\alpha\sigma$ の時のスペクトルは式 1, 2 より、式 3 のように表される。

$$X_{\alpha\sigma}(\tau, \omega) = \int_{-\infty}^{\infty} x(t) \int_{-\infty}^{\infty} g_{\sigma}(t - \tau - \xi) g_{\sqrt{\alpha^2 - 1}\sigma}(\xi) d\xi \cdot e^{-i\omega(t - \tau + \xi - \xi)} dt \quad (3)$$

この式を整理すると、

$$X_{\alpha\sigma}(\tau, \omega) = \frac{\int_{-\infty}^{\infty} x(t) g_{\sigma}(t - \tau - \xi) \cdot e^{-i\omega(t - \tau - \xi)} dt}{\int_{-\infty}^{\infty} g_{\sqrt{\alpha^2 - 1}\sigma}(\xi) \cdot e^{-i\omega\xi} d\xi} \quad (4)$$

となる。この式の下線部は、 $X_{\sigma}(\tau + \xi, \omega)$ である。以上より、Gabor 変換のスペクトル $X_{\sigma}(\tau + \xi, \omega)$ に異なる σ の Gaussian と複素正弦波との積である $g_{\sqrt{\alpha^2 - 1}\sigma}(\xi) \cdot e^{-i\omega\xi}$ を畳み込むことで、 $\alpha\sigma$ のガボール変換による解析結果を算出できる。

しかし、式 4 の畳み込み積分は計算量が多い。そこで、合成元の解析結果を間隔 T ごとに算出し、間引きながら畳み込むことで計算量を削減する。式 4 の畳み込みを離散化するため $\xi = mT$ と置くと、

$$X_{\alpha\sigma}(\tau, \omega) \simeq \sum_{m=-\infty}^{\infty} \{X_{\sigma}(\tau + mT, \omega) \cdot g_{\sqrt{\alpha^2 - 1}\sigma}(mT)\} \cdot T \quad (5)$$

と書ける。ここで m は、スペクトルを合成する際に用いる合成項数である。このように、異なる分解能によるガボール変換の結果を求める際の計算量を削減でき、間引き間隔 T が広いほど計算量は削減できる。しかし、畳み込みを間引くと誤差が生じる。この誤差の主な要因は、周波数域で周期化するために生じるサイドローブである。本稿では、サイドローブのピークとメインローブとの差であるダイナミックレンジが 11bit 量子化ノイズレベル $\epsilon = \frac{1}{2^{10}}$ を下回るよう設計した式 6 から T を決定した。

$$T \leq \frac{\sqrt{2\pi}}{\sqrt{-\ln \epsilon}} \cdot \frac{\sqrt{\alpha^2 - 1}}{\alpha} \cdot \sigma \quad (6)$$

2.2. 合成項数の算出

合成項数である m は本来無限大の長さをとるべきであるが、畳み込む Gaussian は $t = 0$ から単調減少する関数であるため、量子化幅を下回る減衰率の範囲まで畳み込めば、十分であると考えられる。畳み込む Gaussian は、

$$g_{\sqrt{\alpha^2 - 1}\sigma}(m) = e^{-\frac{(mT)^2}{2(\sqrt{\alpha^2 - 1}\sigma)^2}} \quad (7)$$

と表せる. 式 7 と量子化幅 ϵ を等しいと置いて m について解くと,

$$m = \lceil \alpha / \pi \cdot \ln \epsilon \rceil \quad (8)$$

と解ける.

3. 計算量の評価

本章では, 提案手法の計算量を評価するため, 提案手法による複数分解能解析の積和演算量と通常の複数分解能による Gabor 変換の積和演算量の比較を行う.

まず通常の Gabor 変換の計算量を考える. 通常の Gabor 変換は, 高速フーリエ変換 (FFT) を用いて行われる. また, 1 章で論じたとおり, 複数分解能によるガボール変換ではフレームシフト幅を最も時間分解能が高いものに, 周波数 bin 数は最も周波数分解能が高いものに合わせる必要がある. よって, DFT の点数を N とし, 用いる分解能数を n とすると窓掛けを含めたガボール変換の計算量は,

$$n \cdot 2^{n-1} N \log_2 2^{n-1} N + N \sum_{i=0}^{n-1} 2^i \quad (9)$$

と表せる. 次に, 提案手法の計算量を考える. 提案手法では, $2m+1$ 個のスペクトルから 1 つのスペクトルを合成する. よって, 計算量は合成項数 m によって決まる. また, 提案手法においても, 最も時間分解能がよいものは通常どおり計算を行うため, 計算量は

$$2^{n-1} N \log_2 2^{n-1} N + (n-1) 2^{n-1} N (2m+1) + N \quad (10)$$

と表せる. よって, 提案手法と通常の Gabor 変換の計算量比は式 11 のように表せる.

$$\frac{2^{n-1} N \log_2 2^{n-1} N + (n-1) 2^{n-1} N (2m+1) + N}{n \cdot 2^{n-1} N \log_2 2^{n-1} N + N \sum_{i=0}^{n-1} 2^i} \quad (11)$$

ここで, f_s をサンプリング周波数とすると N は周波数分解能 $\Delta f = f_s / N = 5.38 [\text{Hz}]$ を満たすものを用いる.

本稿では, (1) $f_s = 44.1 [\text{kHz}]$, 分解能数 4, (2) $f_s = 96.0 [\text{kHz}]$, 分解能数 4, (3) $f_s = 44.1 [\text{kHz}]$, 分解能数 5, の 3 パターンで計算量を求めた. 各パターンでの計算量比を表 1 に示す. 表 1 より, 提案手法により計算量の

表 1: 各パターンでの計算量比

パターン	f_s [kHz]	分解能数	計算量比 [%]
(1)	44.1	4	88.5
(2)	96.0	4	83.9
(3)	44.1	5	87.7

削減が行えている. また, サンプリング周波数, 分解能数を増やすことで削減率が向上している.

サンプリング周波数を上げると, 通常の Gabor 変換ではサンプリング点が増加するため計算量が増加するが, 提案手法では間引き間隔 T ごとに畳み込みを行うためサンプリング周波数の影響を受けない. そのため, 計算量削減率が向上している.

分解能数を増やした場合, 通常の Gabor 変換では $2^{n-1} N \log_2 2^{n-1} N$ のオーダーで計算量が増えるのに対し, 提案手法では $2m+1$ のオーダーとなり, 通常の Gabor 変換に比べ計算量が増えないため, 計算量削減率が向上した.

4. 実験

本章では, 実際に複数分解能による解析を行った結果を示す. 図 1 は, ピアノ音で構成される楽曲に対して, 複数分解能解析を行った結果のスペクトログラムである.

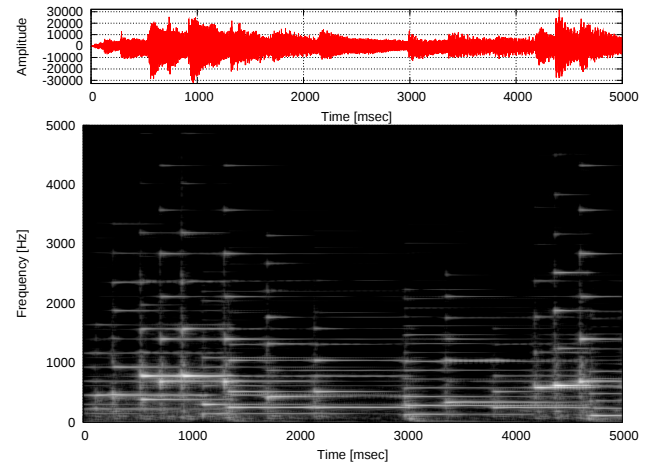


図 1: 複数分解能解析を行った際のスペクトログラム

図 1 のように, 複数分解能による解析を行うことで, より鮮明なスペクトログラムを得ることができる.

5. 結論

本稿では, 複数分解能による Gabor 変換の計算冗長性に着目し, 計算量削減のためのスペクトル合成法を提案した. 実際に FFT によるガボール変換の計算量と比較し, 複数分解能による解析の計算量が削減可能であることを示した.

本稿での計算量比は Lukin らの手法と同条件で最大 83.9% となったが, さらに高サンプリングレート化したり, 多数の分解能を用いればさらに削減率は向上する.

参考文献

- [1] J.B. Allen, L.R. Rabiner, "A unified approach to short-time Fourier analysis and synthesis", Proc. IEEE, vol.65, pp.1558-1564, Nov. 1977.
- [2] Alexey Lukin, Jeremy Todd, "Adaptive Time-Frequency Resolution for Analysis and Processing of Audio", AES Convention, May.2006.
- [3] D. Gabor, "Theory of communication", J. Inst. Elect. Eng., vol.93, no.111, pp.429-457, London, 1946.