

時間領域の畳み込みを厳密に再現する時間周波数領域の計算法*

☆ 松山直哉, 山田宏樹, 矢田部浩平 (農工大)

1 はじめに

Fourier 変換を用いると, 時間領域の畳み込みが周波数領域の積になる. 一方, Fourier 変換をする前に信号に窓をかけるため, 時間周波数領域では単純な積の形で表せない. しかし, 従来は要素積によって畳み込みが近似されることが多く, インパルス応答長と窓長の関係などによっては近似精度が悪いという課題がある. 本稿では, 時間領域の畳み込みと厳密に等価な時間周波数領域での計算式を導出することで誤差のない処理を実現した. また, 窓長などの値を変えることで近似の畳み込みの精度への影響を調べた.

2 従来の時間周波数領域の畳み込み

時間領域から時間周波数領域への変換の一例に, 離散 Gabor 変換 (DGT) がある. 時間信号 $x \in \mathbb{R}^L$ の DGT は, 分析窓 $w \in \mathbb{R}^L$ を用いて

$$X[m, n] = \sum_{l=0}^{L-1} x[l] w[l - an] e^{-j2\pi m(l-an)/M} \quad (1)$$

と定義できる. ただし, j は虚数単位であり, 定数やインデックスは後述の式のものも含め, 表-1 に示す. w のインデックスは負になる場合があるが, $w[(l - an) \bmod L]$ のように長さ L でインデックスが巡回するものとする [1]. 以降の式でも, 負のインデックスとなる場合は同様に定義する. また, 逆離散 Gabor 変換は合成窓 $\gamma \in \mathbb{R}^L$ を用いて, 式 (2) で表せる.

$$x[l] = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} X[m, n] \gamma[l - an] e^{j2\pi m(l-an)/M} \quad (2)$$

時間領域において, $x \in \mathbb{R}^L$ とインパルス応答 $h \in \mathbb{R}^L$ の巡回畳み込みは, 計算結果を $d \in \mathbb{R}^L$ として,

$$d[l] = \sum_{c=0}^{L-1} x[l - c] h[c] \quad (3)$$

と定義できる. h, d のスペクトログラムをそれぞれ $\mathbf{H}_{2D}, \mathbf{D} \in \mathbb{C}^{M \times N}$ とすると, 式 (3) は

$$D[m, n] \approx \sum_{p=0}^{N-1} X[m, n - p] H_{2D}[m, p] \quad (4)$$

のように, 時間周波数領域で各周波数ビンの巡回畳み込みで近似されることが多い. より広く用いられる

表-1 本稿で用いる定数とインデックスの記号

l, c	時間領域の時間インデックス
L	時間領域の信号長
n, p	時間周波数領域の間引き済みの時間インデックス
m, q	時間周波数領域の間引き済みの周波数インデックス
N	n, p の総数
M	m, q の総数 (=窓長)
a	時間の間引き量 (=窓のシフト幅, $N = L/a$)
b	周波数の間引き量 ($M = L/b$)

近似として, \mathbf{H}_{2D} が $N = 1$ の場合もあり,

$$D[m, n] \approx X[m, n] H_{2D}[m, 1] \quad (5)$$

と表せる. これらの近似は各周波数ビンが独立していることを仮定して計算される. しかし, DGT は Fourier 変換をする前に窓をかけるため時間周波数領域の周波数方向にスペクトルが広がる. その広がりが大きいほど, 畳み込みの近似精度に影響が及ぶ.

3 提案手法

本稿では, 周波数ビン間の依存性を考慮した, 時間周波数領域における厳密な畳み込みの表現を提案する.

3.1 時間周波数領域での畳み込みの厳密な計算法

式 (1) より, d の DGT は以下のように表せる.

$$D[m, n] = \sum_{l=0}^{L-1} \left(\sum_{c=0}^{L-1} h[c] x[l - c] \right) w[l - an] \times e^{-j2\pi m(l-an)/M} \quad (6)$$

$x[l - c]$ に式 (2) を代入する. インデックスのずれ $l - c$ を時間周波数領域で表現するため, 周波数インデックス q と時間インデックス p を用いると,

$$\begin{aligned} D[m, n] &= \sum_{p=0}^{N-1} \sum_{q=0}^{M-1} X[q, p] \sum_{c=0}^{L-1} h[c] \sum_{l=0}^{L-1} w[l] e^{-j2\pi ml/M} \\ &\quad \times \gamma[l - c + a(n - p)] e^{j2\pi q(l - c + a(n - p))/M} \\ &= \sum_{q=0}^{M-1} \sum_{p=0}^{N-1} X[q, p] H_{3D}[m, n - p, q] \end{aligned} \quad (7)$$

と表せる. ただし, $\mathbf{H}_{3D} \in \mathbb{C}^{M \times N \times M}$ は

$$\begin{aligned} H_{3D}[m, p, q] &= \sum_{c=0}^{L-1} h[c] \sum_{l=0}^{L-1} w[l] e^{-j2\pi ml/M} \\ &\quad \times \gamma[l - c + ap] e^{j2\pi q(l - c + ap)/M} \end{aligned} \quad (8)$$

*Rigorous computation of time-domain convolution in the time-frequency domain. By Naoya MATSUYAMA, Koki YAMADA and Kohei YATABE (Tokyo University of Agriculture and Technology)

Algorithm 1 \mathbf{H}_{3D} を求めるアルゴリズム

Input: $w, \gamma, \mathbf{h} \in \mathbb{R}^L$
Output: $\mathbf{H}_{3D} \in \mathbb{C}^{M \times N \times M}$

- 1: $\mathbf{F}_h = \text{fft}(\mathbf{h})$
- 2: $\mathbf{F}_w = \text{fft}(\text{circshift}(w, -M/2))$
- 3: $\mathbf{F}_\gamma = \text{fft}(\text{circshift}(\gamma, -M/2))$
- 4: $\mathbf{W} = \mathcal{S}(\mathbf{F}_w) \in \mathbb{C}^{M \times L}$
- 5: $\mathbf{\Gamma} = \mathcal{S}(\mathbf{F}_\gamma) \in \mathbb{C}^{M \times L}$
- 6: **for** $q = 0, \dots, M-1$ **do**
- 7: $\mathbf{H}_{3D}[:, :, q] = \text{ifft}(\mathbf{F}_h \odot \mathbf{W} \odot \mathbf{\Gamma}[q, :]) \in \mathbb{C}^{M \times L}$
- 8: $\mathbf{H}_{3D}[:, :, q] = \text{decimate}(\mathbf{H}_{3D}[:, :, q], a) \in \mathbb{C}^{M \times N}$
- 9: **end for**

である。畳み込みの性質より、 \mathbf{X} , \mathbf{H}_{3D} の時間インデックス p , $n-p$ は入れ替えられるため、式 (8) では $H_{3D}[m, p, q]$ と示した。この式をそのまま計算すると、指数関数の計算を含むこともあって、計算時間と数値精度の面で不利である。よって、 \mathbf{H}_{3D} が効率良く計算できるアルゴリズムについて考える。

3.2 FFT を用いた \mathbf{H}_{3D} の実装方法

式 (8) について、 c に対し $h[c]$ と $\gamma[l-c+ap]$ が、 l に対し $w[l]e^{-j2\pi ml/M}$ と $\gamma[l-c+ap]e^{j2\pi q(l-c+ap)/M}$ が畳み込みの関係である。そこで、高速 Fourier 変換 (FFT) を用いる。時間領域での巡回畳み込みより、FFT による周波数領域での要素積の方が、計算時間と数値精度の両方で効果的だからである [2]。FFT による \mathbf{H}_{3D} の実装方法を Algorithm 1 に示す。ただし、 $[\cdot]$ は配列のその次元における要素のベクトル、 \odot は要素積、 $\text{decimate}(\cdot, a)$ は配列の 2 次元目の要素を a で間引く操作を表す。 \mathcal{S} は窓の周波数スペクトル \mathbf{F}_w , $\mathbf{F}_\gamma \in \mathbb{C}^L$ を周波数シフトさせる作用素で、

$$\begin{aligned} \mathbf{W} &= \mathcal{S}(\mathbf{F}_w) \in \mathbb{C}^{M \times L} \\ &= [\mathcal{S}_b^0(\mathbf{F}_w), \mathcal{S}_b^1(\mathbf{F}_w), \dots, \mathcal{S}_b^{M-1}(\mathbf{F}_w)] \end{aligned} \quad (9)$$

と定義する。ただし、 \mathcal{S}_b^m は bm だけインデックスを巡回シフトさせる作用素である。 $q = 0, \dots, M-1$ の各値について、 $\mathbf{\Gamma}[q, :] \in \mathbb{C}^L$ が $\mathbf{\Gamma}$ の q 行目のベクトルを、 $\mathbf{H}_{3D}[:, :, q] \in \mathbb{C}^{M \times L}$ が時間方向を間引く前の \mathbf{H}_{3D} の q における平面を表すものとし、7 行目のように要素積をとる。FFT の逆変換で周波数領域を時間領域に戻して間引くことで \mathbf{H}_{3D} を得る。

2, 3 行目では窓の振幅スペクトル \mathbf{F}_w , \mathbf{F}_γ に線形位相が乗らないよう、FFT をする前に窓のピークがデータの初めに来るように巡回シフトする [3]。よって、 \mathbf{W} , $\mathbf{\Gamma}$ は対角付近にメインローブの成分が連なり、要素積をとった結果、メインローブのピークは式 (8) における $q = m$ に対応し、 \mathbf{H}_{2D} はそのときに得られる配列にあたる。メインローブには幅があるため、近似で使われなかった $q = m$ 以外の要素に含まれる情報を使うことで、周波数ビン間の依存性を考慮した、時間周波数領域での厳密な畳み込みができる。

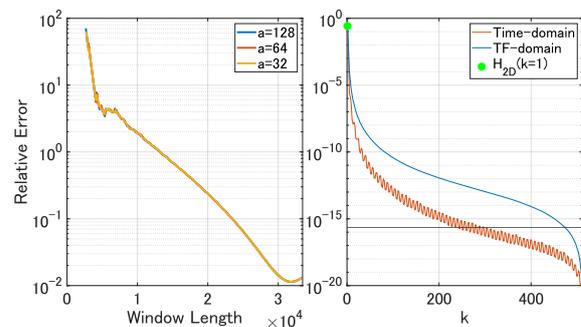


図-1 式 (5) の計算における窓長を変えたときの相対誤差の変化 (左) と、畳み込みに用いる、窓の振幅スペクトルの使う情報を増やしたときの相対誤差の変化 (右)。時間領域、時間周波数領域で相対誤差をとった結果をそれぞれ Time-domain, TF-domain で表した。また、式 (4) と等価な条件である $k = 1$ で畳み込んだ場合を緑点で、計算機イプシロンを黒の横線で示した。

4 実験

\mathbf{x} には VCTK データセットから約 10 秒で切り取った音源を用いた。 w, γ は Hann 窓とその双対窓とし、 \mathbf{h} はインパルス応答のシミュレーション関数で生成した。 \mathbf{x} と \mathbf{h} の時間周波数領域での近似の畳み込みの精度を、巡回畳み込みに対する相対誤差で評価した。

1 つ目の実験では、式 (5) で計算した際に、窓長を変えたことによる相対誤差の変化を調べた。結果を図-1 の左図に示す。窓長を大きくすることで誤差は下がるが、 10^{-2} 程度までしか下がらなかった。

2 つ目の実験では、 $k = 1, 3, \dots, M$ として、

$$H_{3D}[m, p, q] = \begin{cases} 0 & \text{if } |m - q| > k, \\ H_{3D}[m, p, q] & \text{if } |m - q| \leq k. \end{cases} \quad (10)$$

とし、式 (7) の計算をしたときの相対誤差を調べた。窓長 $M = 512$ 、窓のシフト幅 $a = 128$ とした。結果を図-1 の右図に示す。 $k = 1$ のときが式 (4) による近似計算、 $k = M (= 512)$ のときが式 (7) の提案手法での計算に対応する。 k が大きいほど誤差が下がり、最終的に計算機イプシロン以下にまで下がった。

5 むすび

本稿では、時間周波数領域での畳み込みの厳密な表現を導出し、近似の畳み込みの精度を調べた。時間領域で誤差をとった際、上下に値が振れる原因を調査するとともに、 \mathbf{H}_{3D} の応用を今後考える。

参考文献

- [1] 矢田部浩平, “第三回: 短時間フーリエ変換,” 音響学会誌, **77**, 396–403 (2001).
- [2] P. Duhamel and M. Vetterli, “Fast Fourier transforms: A tutorial review and a state of the art,” *Signal Process.*, **19**, 259–299 (1990).
- [3] 矢田部浩平, “第五回: 実装における諸注意,” 音響学会誌, **77**, 537–544 (2021).