

## 時間周波数領域における近似畳み込みの性質に関する調査\*

☆ 村田佳斗, 矢田部浩平 (農工大), 泉悠斗, 高橋祐, 近藤多伸 (ヤマハ株式会社)

### 1 はじめに

ハウリング抑圧やエコーキャンセラなどの信号処理において、空間伝達系の推定は重要である。残響時間の長い伝達系の推定では、計算量の削減や推定問題の簡単化のために、時間領域の畳み込みを時間周波数領域の畳み込みで近似したモデルを用いることがある。しかし、この近似畳み込みで生じる誤差が空間伝達系の推定に与える影響は明らかでない。そこで本稿では、この近似誤差が与える影響を調べる準備として、畳み込みを用いる時間信号及びインパルス応答の標準基底及び係数を用いた畳み込み表現を導出した。

### 2 時間周波数領域での近似畳み込み

時間信号  $\mathbf{x} \in \mathbb{R}^L$  の時間周波数表現は、次式に示す離散 Gabor 変換 (DGT) によって得られる。

$$X[m, n] = \sum_{l=0}^{L-1} x[l] w[l - an] e^{-i2\pi m(l-an)/M} \quad (1)$$

ここで、 $n = 0, 1, \dots, N-1$ ,  $m = 0, 1, \dots, M-1$  はそれぞれ時間フレームと周波数ビンのインデックス、 $N, M \in \mathbb{N}$  はそれぞれ時間フレーム数と周波数ビン数、 $a \in \mathbb{N}$  は窓関数のシフト幅、 $i = \sqrt{-1}$  である。窓関数  $\mathbf{w} \in \mathbb{R}^L$  は信号長  $L$  と等しくなるようにゼロ埋めし、インデックスは  $w[(l - an) \bmod L]$  のように長さ  $L$  で巡回するものとする [1]。信号  $\mathbf{x}$  とインパルス応答  $\mathbf{h} \in \mathbb{R}^L$  の巡回畳み込み結果  $\mathbf{d} \in \mathbb{R}^L$  は

$$d[l] = \sum_{c=0}^{L-1} x[c] h[l - c] \quad (2)$$

と定義される。この畳み込みは、 $\mathbf{d}, \mathbf{x}, \mathbf{h}$  のスペクトログラムを  $\mathbf{D}, \mathbf{X}, \mathbf{H} \in \mathbb{C}^{M \times N}$  とすると

$$D[m, n] \approx \sum_{r=0}^{N-1} X[m, n] H[m, n - r] \quad (3)$$

のように、各周波数ビンの巡回畳み込みで近似される。しかし、DGT では窓関数の影響でスペクトルが周波数方向に広がりを持つため、近似誤差が生じる。この誤差はインパルス応答長に対する窓長に依存するため、インパルス応答長が長くなるほど近似誤差は大きくなる [2]。この近似誤差が後段の処理にどのような影響を与えるかは明らかでない。

### 3 標準基底と係数を用いた畳み込み表現

本稿では、近似誤差が後段の処理に与える影響を調べる準備として、畳み込みを用いる信号を標準基底とそれに対応する係数に分離した計算を行い、時間領域及び時間周波数領域における畳み込み表現を導出する。この表現によって、時間周波数領域での近似畳み込みで生じる誤差に関する厳密な議論や評価が可能となることが期待される。

#### 3.1 時間領域での畳み込み表現

時間信号  $\mathbf{x}$  及びインパルス応答  $\mathbf{h}$  は、標準基底とそれに対応する係数  $\alpha, \beta \in \mathbb{R}^L$  を用いると、

$$\mathbf{x} = \sum_{l=0}^{L-1} \alpha[l] \mathbf{e}_l, \quad \mathbf{h} = \sum_{l=0}^{L-1} \beta[l] \mathbf{e}_l \quad (4)$$

と表せる。ここで標準基底は  $\{\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_{L-1}\}$  で表され、 $\mathbf{e}_l[l] = 1$  の単位ベクトルである。また、式 (4) を用いて式 (2) を表現すると

$$\begin{aligned} \mathbf{d} &= \sum_{l=0}^{L-1} \left\{ \sum_{c=0}^{L-1} (\alpha[l] \mathbf{e}_l) (\beta[l - c] \mathbf{e}_{l-c}) \right\} \\ &= \sum_{l=0}^{L-1} \left( \mathbf{e}_l \sum_{c=0}^{L-1} \alpha[c] \beta[l - c] \right) \end{aligned} \quad (5)$$

となる。ただし、標準基底のインデックスは  $\bmod L$  で巡回している。 $\mathbf{x}, \mathbf{h}$  の標準基底は、畳み込み後の  $\mathbf{d}$  の標準基底となるので、 $\alpha, \beta$  の畳み込み結果との積として表せる。また、 $\mathbf{d}$  のスペクトログラムは

$$\text{DGT}(\mathbf{d}) = \sum_{l=0}^{L-1} \left\{ \text{DGT}(\mathbf{e}_l) \sum_{c=0}^{L-1} \alpha[c] \beta[l - c] \right\} \quad (6)$$

となり、各時間フレーム  $l$  における標準基底の DGT 及び畳み込み係数の積の総和で表すことができる。ただし、DGT を行う関数を  $\text{DGT}(\cdot)$  とした。

#### 3.2 時間周波数領域での畳み込み表現

時間信号とインパルス応答のスペクトログラム  $\mathbf{X}, \mathbf{H}$  は、先に標準基底のみの DGT を行うことで

$$\mathbf{X} = \text{DGT}(\mathbf{x}) = \sum_{l=0}^{L-1} \alpha[l] \text{DGT}(\mathbf{e}_l) \quad (7)$$

$$\mathbf{H} = \text{DGT}(\mathbf{h}) = \sum_{l=0}^{L-1} \beta[l] \text{DGT}(\mathbf{e}_l) \quad (8)$$

\*Investigating characteristics of approximate convolution in the time-frequency domain. By Keito MURATA, Kohei YATABE (Tokyo University of Agriculture and Technology), Yuto IZUMI, Yu TAKAHASHI and Kazunobu KONDO (Yamaha Corporation).

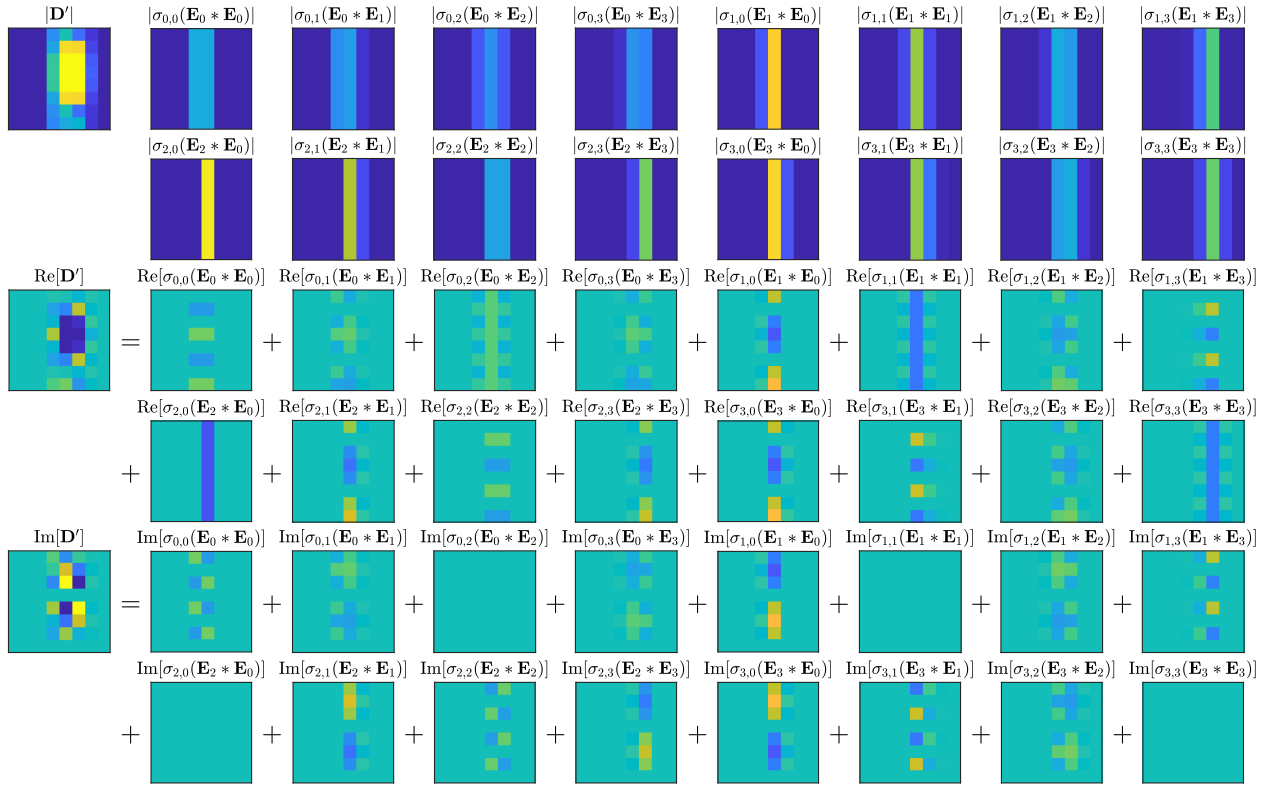


図-1 標準基底とそれに対応する係数を用いた時間周波数領域での畳み込み表現。標準基底  $e_p, e_q$  の DGT を  $E_p, E_q$  とし、それらの巡回畳み込みを  $E_p * E_q$  で表す。 $\sigma_{p,q} (= \alpha[p] \beta[q])$  は、 $E_p * E_q$  に付随する係数である。スペクトログラムの計算には、窓長が  $2^3$  のハン窓を用い、シフト幅  $a = 4$  とした。上二段は、時間周波数領域での畳み込み結果の DGT (左端) 及び  $\sigma_{p,q}(E_p * E_q)$  の絶対値を示す。中二段は、 $\sigma_{p,q}(E_p * E_q)$  の実部及びそれらの和 (左端) を示す。下二段は、 $\sigma_{p,q}(E_p * E_q)$  の虚部及びそれらの和 (左端) を示す。

のように表せる。式 (7) と式 (8) を用いると、式 (3) の時間周波数領域での近似畳み込みは

$$\mathbf{D} \approx \sum_{p=0}^{L-1} \sum_{q=0}^{L-1} \left\{ \alpha[p] \beta[q] \sum_{r=0}^{N-1} \text{DGT}(e_p)[m, n] \text{DGT}(e_q)[m, n-r] \right\} \quad (9)$$

と表せる。式 (7) と式 (8) から分かるように、式 (9) の標準基底  $e_p, e_q$  と係数  $\alpha[p], \beta[q]$  はそれぞれ対応している。つまり、標準基底の DGT の巡回畳み込み結果に付随する係数は、 $\alpha[p], \beta[q]$  の単純な積の形となる。よって、時間周波数領域での近似畳み込みは、標準基底の DGT の畳み込みとそれに付随する係数との積の総和で表すことができる。

図-1 に、単純な信号を用いた概説図を示す。ただし、図-1 の計算に用いたスペクトログラムは図-2 に示す  $\mathbf{X}', \mathbf{H}'$  である。ここで、 $\mathbf{X}', \mathbf{H}'$  は  $\mathbf{x}', \mathbf{h}' \in \mathbb{R}^{32}$  の DGT である。また、 $\mathbf{x}', \mathbf{h}'$  は、連続した 4 サンプルでのみランダムな値を持つ信号である。 $\mathbf{D}'$  は、 $\mathbf{X}', \mathbf{H}'$  の近似畳み込み結果である。図-1 から、標準基底とそれに対応する係数を用いた表現により、時間周波数領域での畳み込み信号のスペクトログラム  $\mathbf{D}'$  が、単純な構造を持つスペクトログラム  $\sigma_{p,q}(E_p * E_q)$  に分解されていることが確認できる。

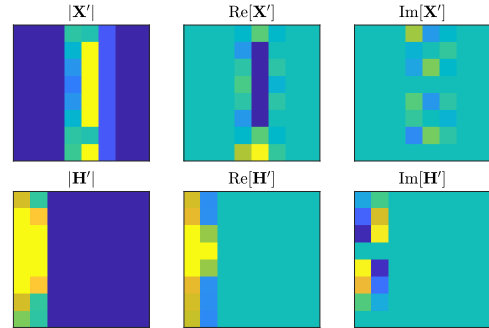


図-2 図-1 に用いた時間信号及びインパルス応答のスペクトログラムの絶対値 (左)、実部 (中)、及び虚部 (右)。上段が時間信号、下段がインパルス応答である。

## 4 むすび

本稿では、時間領域での畳み込みの時間周波数表現及び時間周波数領域での畳み込みを標準基底とそれに対応する係数の畳み込みの積の形で表せることを確認した。今後は、時間周波数領域での近似畳み込みで生じる誤差の定式化を検討し、時間周波数領域におけるインパルス応答の推定への応用について考える。

### 参考文献

- [1] 矢田部浩平, “第三回: 短時間フーリエ変換入門,” 音響学会誌, **77**, 396–403 (2021).
- [2] 松山直哉, 山田宏樹, 矢田部浩平, “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,” 音講論集, pp. 203–204 (2024.3).