

# ネコの心音成分を強調する畳み込み NMF の初期化戦略\*

◎ 岸本麗央, 吉野夏樹, 矢田部浩平, 岩永朋子 (農工大)

## 1 はじめに

ネコの急性心不全による突然死を予防するには、聴診による心雑音の確認が有効である。通院での聴診はネコにとってストレスとなるため、飼育者が在宅でデジタル聴診器によって記録した聴診音をもとに、獣医師が診断する方法が検討されている。しかし、在宅ではネコが過剰にリラックスすることで purring 音を鳴らすため、聴診が困難になる。そこで我々は、畳み込み NMF に基づいて聴診音に含まれる心音成分を強調する手法を提案した [1]。本稿では、[1] の手法を洗練するため、心音成分や非心音成分が有する特徴を反映した畳み込み NMF の初期化を提案する。

## 2 畳み込み NMF による心音成分の強調

提案法 [1] では、観測信号に短時間フーリエ変換 (STFT) を行って得られた振幅スペクトログラム  $\mathbf{X} \in \mathbb{R}_{\geq 0}^{F \times T}$  に対し、畳み込み NMF を適用する。畳み込み NMF は時間方向に繰り返し現れるスペクトルパターンの抽出に有効であり、基底ごとの畳み込み演算を

$$\mathbf{X} \approx \sum_{k=1}^K \mathbf{Y}_k = \sum_{k=1}^K \sum_{l=1}^L \mathbf{W}_{:,l,k} \mathcal{S}_{l-1}(\mathbf{H}_{k,:}), \quad (1)$$

のように行って得られる  $\mathbf{Y}_k \in \mathbb{R}_{\geq 0}^{F \times T}$  を全基底で総和して観測信号  $\mathbf{X}$  を近似する。ここで、 $F$  は周波数ビン数、 $T$  は時間フレーム数、 $k = 1, \dots, K$  は基底、 $l = 1, \dots, L$  はシフト量のインデックスである。また、 $\mathbf{W} \in \mathbb{R}_{\geq 0}^{F \times L \times K}$  はスペクトルパターンを表す基底行列、 $\mathbf{H} \in \mathbb{R}_{\geq 0}^{K \times T}$  は時間的な強度変化を表すアクティベーション行列であり、 $\mathbf{W}$  の  $(l, k)$  番目の列を  $\mathbf{W}_{:,l,k}$ 、 $\mathbf{H}$  の  $k$  番目の行を  $\mathbf{H}_{k,:}$  と表記する。さらに、 $\mathcal{S}_n(\cdot)$  は行列の要素を  $n$  列分シフトし、要素がなくなった箇所にゼロ埋めを行う作用素である。ただし、 $n = 0$  はシフト無し、 $n < 0$  と  $n > 0$  はそれぞれ左シフトと右シフトに対応する [1]。

各基底での  $\mathbf{Y}_k$  は、観測信号  $\mathbf{X}$  に含まれる特定の成分を構成する役割が期待される。例えば [1] では、基底ごとに心音成分や非心音成分を構成するものとし、心音成分を構成する基底の抽出によって観測信号に含まれる心音成分の強調を図った。しかし、畳み込み NMF は初期値依存性を有しているため、局所解に陥りやすい。実際に [1] においても、心音成分を構成

するための基底が非心音成分を捉えてしまい、心音成分の安定した強調が難しいという課題があった。

## 3 聴診音に現れる特徴を反映した初期化

心雑音はペアで周期性を有する第一心音 (S1) と第二心音 (S2) に付随して鳴るのに対し、非心音成分は一定時間連続して鳴るという特徴がある。そこで、提案法 [1] での心音成分の強調をより安定させるため、畳み込み NMF の基底行列に S1 と S2 のスペクトルを、アクティベーション行列に S1 が鳴る時刻を反映した初期値を与える。S1 と S2 が鳴る間隔を 3.1 節、S1 が鳴る時刻を 3.2 節の方法でそれぞれ推定し、これらの推定結果を反映させた基底行列とアクティベーション行列の初期化を 3.3 節のように行う。

### 3.1 第一心音と第二心音が鳴る間隔の推定

本節では、暫定的に観測信号から S1 と S2 を抽出した信号に対し、自己相関のピークを求めることで、S1-S2 間隔を推定する。まず、定常的に鳴る非心音成分を抽出するため、観測信号に対して平滑化処理を  $\mathbf{X}^{(\text{Extra})} = \text{medfilt}_L(\mathbf{X}) \in \mathbb{R}_{\geq 0}^{F \times T}$  のように行い、これを観測信号から  $\mathbf{X}^{(\text{Cardiac})} = \max(0, \mathbf{X} - \mathbf{X}^{(\text{Extra})}) \in \mathbb{R}_{\geq 0}^{F \times T}$  のように引き、S1 と S2 を抽出する。ただし、 $\text{medfilt}_L(\cdot)$  は窓長  $L$  の時間方向のメディアンフィルタである。次に、 $\mathbf{X}^{(\text{Cardiac})}$  の周波数方向の和をとった信号に対して自己相関を計算し、S1-S2 間隔に対応するピークを抽出する。このピークの際のラグ長を S1-S2 間隔  $\tau$  とする。

### 3.2 第一心音が鳴る時刻の推定

S1-S2 間隔を  $\tau$  と推定したため、S1 が鳴る時刻の  $\tau$  だけ後に S2 が鳴ると仮定できる。一方、S2-S1 間隔 (心室拡張期) は S1-S2 間隔 (心室収縮期) よりも長いいため、S2 が鳴った時刻の  $\tau$  だけ後は心音と比べると微小な心雑音やノイズが鳴る。この性質から、S1 のピークが現れる時刻を推定する。まず、S1 と S2 の両方のピークが含まれる  $\mathbf{p} = \sum_{f=1}^F \mathbf{X}_{f,:}^{(\text{Cardiac})} \in \mathbb{R}_{\geq 0}^T$  を得る。次に、S1 のピークを保持しつつ、S2 のピークを小さくするため、 $\mathbf{p}$  と  $\mathbf{p}$  を左に  $\tau$  シフトした  $\mathcal{S}_{-\tau}(\mathbf{p})$  に関して、 $\mathbf{q} = \mathbf{p} \odot \mathcal{S}_{-\tau}(\mathbf{p}) \in \mathbb{R}_{\geq 0}^T$  を求める。ただし、 $\odot$  は要素ごとの積である。さらに、信号全体における振幅の変動に対して頑健な S1 のピーク検出を行うため、 $\mathbf{q}$  に対して局所正規化を適用する。具体的には、 $\mathbf{q}$  の  $t$  番目の要素

\*Initialization strategy for convolutive NMF that enhances feline cardiac sound components. By Reo KISHIMOTO, Natsuki YOSHINO, Kohei YATABE and Tomoko IWANAGA (Tokyo University of Agriculture and Technology).

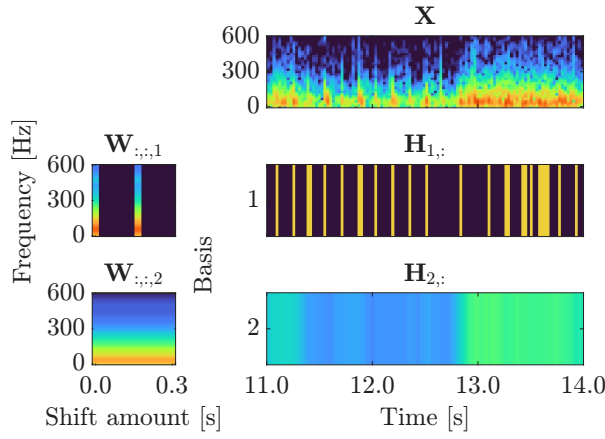


図-1 観測信号（右列，1 段目）をもとに初期化した基底行列（左列）とアクティベーション行列（右列，2，3 段目）。2，3 段目の基底行列の左右のパルスはそれぞれ S1 と S2 に対応し，アクティベーション行列でのパルスは，S1 が鳴る推定時刻に立っている。右列は信号全体のうち，11.0–14.0 s を拡大している。

$q_t$  に対して，局所最大値  $q_{\max}^{t,L} = \max(q_t, \dots, q_{t+L-1})$  を求め， $\tilde{q}_t = q_t / (q_{\max}^{t,L} + \varepsilon)$  とした  $\tilde{\mathbf{q}} \in \mathbb{R}_{\geq 0}^T$  を得る。ただし， $\varepsilon$  はゼロ除算を防ぐための微小な定数である。最後に，閾値  $\eta$  より大きい  $\tilde{\mathbf{q}}$  の要素を 1，それ以外を 0 とした  $\mathbf{r} \in \mathbb{R}_{\geq 0}^T$  を得る。得られた  $\mathbf{r}$  の要素が 1 の時刻を S1 が鳴る時刻と考える。

### 3.3 基底行列とアクティベーション行列の初期化

推定した S1–S2 間隔  $\tau$  と  $\mathbf{r}$  をもとに，基底行列  $\mathbf{W}$  とアクティベーション行列  $\mathbf{H}$  の初期値を

$$\mathbf{W}_{:,l,k} = \begin{cases} \frac{\sum_{t=1}^T \mathbf{x}_{:,t}^{(\text{Cardiac})}}{[T/L]} & (l = 1, 1 + \tau, k < K) \\ 0 & (l \neq 1, 1 + \tau, k < K), \\ \sigma^{(\text{Extra})} \mathbf{u}^{(\text{Extra})} & (k = K) \end{cases} \quad (2)$$

$$\mathbf{H}_{k,:} = \begin{cases} \mathbf{r}^\top & (k < K) \\ (\mathbf{v}^{(\text{Extra})})^\top & (k = K) \end{cases} \quad (3)$$

のように設定する。ただし， $\sigma^{(\text{Extra})}$  は  $\mathbf{X}^{(\text{Extra})}$  に特異値分解を適用して得られた最大特異値， $\mathbf{u}^{(\text{Extra})} \in \mathbb{R}_{\geq 0}^F$  と  $\mathbf{v}^{(\text{Extra})} \in \mathbb{R}_{\geq 0}^T$  はそれぞれ  $\sigma^{(\text{Extra})}$  に対応する非負左特異ベクトルと非負右特異ベクトルである。また， $[\cdot]$  は最近傍丸めであり， $(\cdot)^\top$  は転置である。さらに，心音に付随して鳴る心雑音を捉えるため，式 (2) で初期化した  $\mathbf{W}$  に微小な定数  $\varepsilon$  を加える。図-1 に基底行列  $\mathbf{W}$  とアクティベーション行列  $\mathbf{H}$  の初期値を可視化した例を示す。

## 4 実験

リットマン社の電子聴診器である Littmann CORE Stethoscope を用いて，サンプリング周波数 48 kHz でネコの聴診音を収録し，1.2 kHz にリサンプリングして提案法を適用した。STFT には窓長 64 サンプル

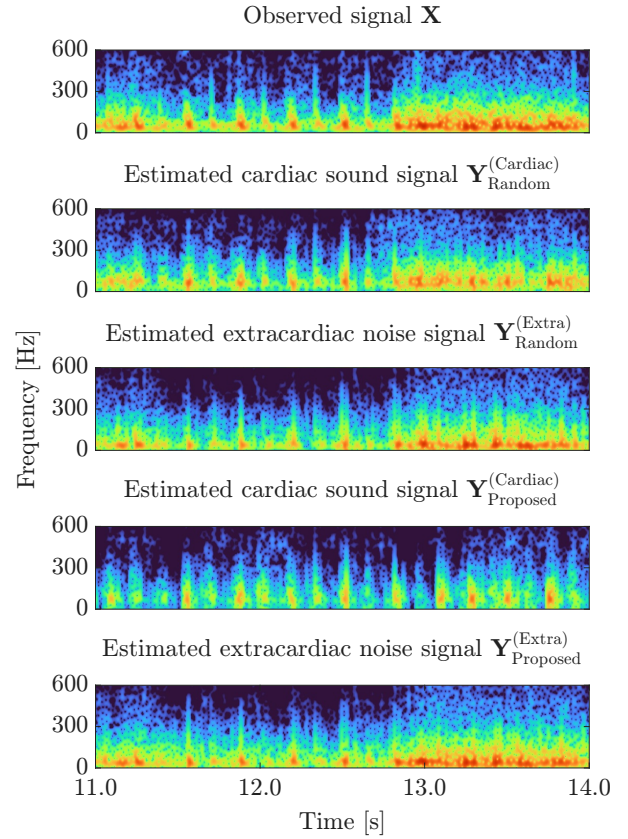


図-2 観測信号（1 段目）に畳み込み NMF を適用して得られた心音成分と非心音成分の推定信号。2，3 段目が 0 から 1 の乱数，4，5 段目が提案法で基底行列とアクティベーション行列を初期化した際の，それぞれ心音成分の推定信号と非心音成分の推定信号である。

(53 ms) の Hann 窓を用い，シフト長は 32 サンプル (27 ms) とした。畳み込み NMF の基底数は  $K = 2$ ，最大シフト量は  $L = 12$  とし，反復回数は 1000 回とした。閾値は  $\eta = 0.3$  とした。提案法の有効性を確認するため，基底行列とアクティベーション行列を 0 から 1 の一様分布乱数で初期化した場合と比較した。

畳み込み NMF の初期化方法ごとの結果を図-2 に示す。2 段目より，基底行列とアクティベーション行列を 0 から 1 の乱数で初期化した場合には，心音成分の推定信号  $\mathbf{Y}_{\text{Random}}^{(\text{Cardiac})}$  において，非心音成分の混在が目立った。一方，4 段目より，提案法で初期化した場合には，心音成分の推定信号  $\mathbf{Y}_{\text{Proposed}}^{(\text{Cardiac})}$  にほとんど非心音成分は見られず，観測信号  $\mathbf{X}$  の 13–14 s で purring 音に埋もれた心音成分を抽出できた。

## 5 むすび

本稿では，聴診音に現れる特徴を反映して畳み込み NMF の初期化を行うことで，心音成分の安定的な強調を実現した。今後は，様々な個体や症状のデータに対する手法の有効性を検証し，頑健性を向上させる。

### 参考文献

- [1] 岸本麗央，矢田部浩平，岩永朋子，“ネコの聴診音信号に含まれる心音成分の強調，”音講論集，pp. 159–160 (2026.3).