Infinite Positive Semidefinite Tensor Factorization for Source Separation of Mixture Signals

Kazuyoshi Yoshii[1)] Ryota Tomioka[2)]
Daichi Mochihashi[3)] Masataka Goto[1)]

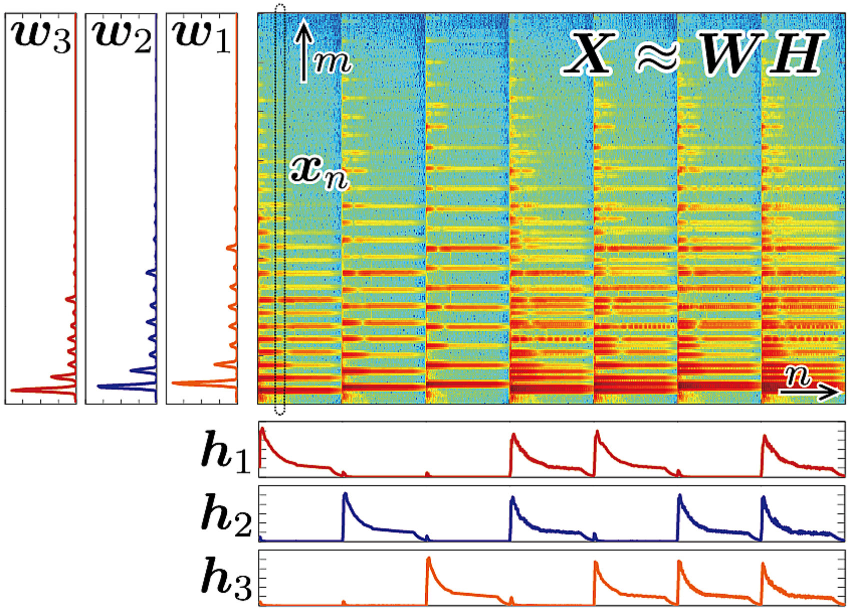
1) National Institute of Advanced Industrial Science and Technology (AIST)
2) The University of Tokyo 3)The Institute of Statistical Mathematics (ISM)

Take-Home Messages

- We proposed positive semidefinite tensor factorization (PSDTF)
  - Tensor extension of nonnegative matrix factorization (NMF)
    - Nonnegative tensor factorization (NTF) is a naive extension of NMF
  - Bayesian nonparametrics
    - The gamma process is used for Bayesian PSDTF that can deal with an infinite number of bases
- A special case: log-determinant PSDTF (LD-PSDTF)
  - Elegant variational inference
    - Closed-form MU and VB updates were derived
  - Various applications
    - Single-channel audio source separation
    - Multi-channel EEG signal analysis

Nonnegative Matrix Factorization

- Each <u>nonnegative vector</u> is approximated by a convex combination of <u>nonnegative vectors</u> (bases)



Observed matrix

$$\boldsymbol{X} = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N] \in \mathbb{R}^{M \times N}$$

Basis matrix

$$\boldsymbol{W} = [\boldsymbol{w}_1, \cdots, \boldsymbol{w}_K] \in \mathbb{R}^{M \times K}$$

Activation matrix

$$\boldsymbol{H} = [\boldsymbol{h}_1, \cdots, \boldsymbol{h}_K]^T \in \mathbb{R}^{K \times N}$$

Vector-wise factorization

$$\boldsymbol{x}_n \approx \sum_{k=1}^{K} \boldsymbol{w}_k h_{kn} \stackrel{\text{def}}{=} \boldsymbol{y}_n$$

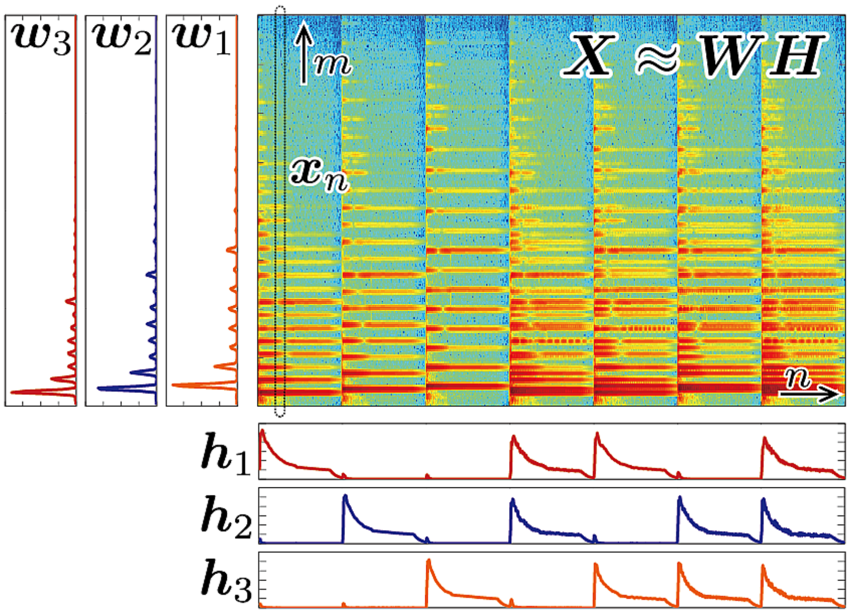
Bregman divergence: $\mathcal{D}_\phi(\boldsymbol{x}_n|\boldsymbol{y}_n) = \phi(\boldsymbol{x}_n) - \phi(\boldsymbol{y}_n) - \phi'(\boldsymbol{y}_n)^T(\boldsymbol{x}_n - \boldsymbol{y}_n)$ → Minimize

Kullback-Leibler divergence: $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_n|\boldsymbol{y}_n) = \sum_m \left(x_{mn} \log x_{mn} y_{mn}^{-1} - x_{mn} + y_{mn}\right)$

Itakura-Saito divergence: $\mathcal{D}_{\mathrm{IS}}(\boldsymbol{x}_n|\boldsymbol{y}_n) = \sum_m \left(-\log x_{mn} y_{mn}^{-1} + x_{mn} y_{mn}^{-1} - 1\right)$

A Major Limitation of NMF

- The elements of each basis vector are assumed to be independent
  - The correlations between those elements are ignored



Element-wise representation

$$X_{mn} \approx \sum_{k=1}^{K} W_{mk} H_{kn} \overset{\text{def}}{=} Y_{mn}$$

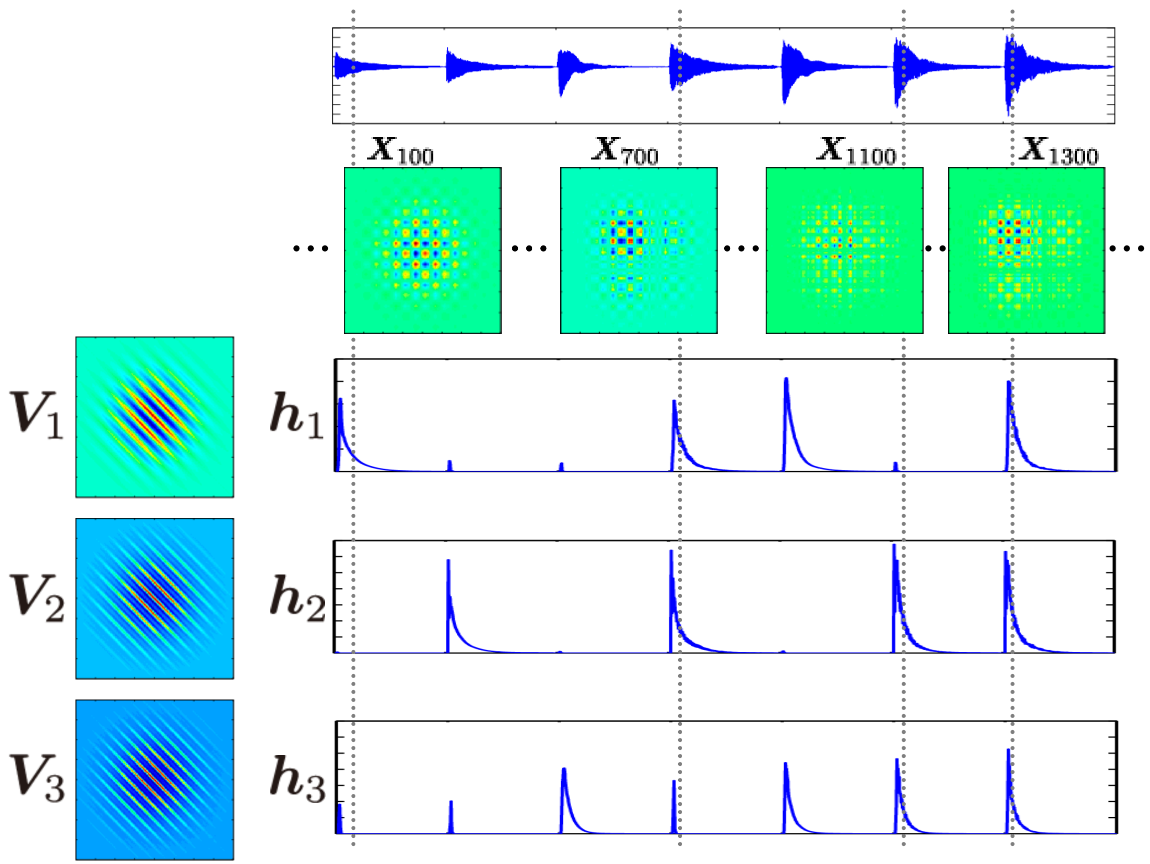
The cost function is defined in an element-wise manner

Gamma priors are placed in an element-wise manner

Problem: STFT cannot completely decorrelate frequency bins because finite windows are used for analyzing non-stationary signals

Positive Semidefinite Tensor Factorization

- Each positive semidefinite matrix is approximated by a convex combination of positive semidefinite matrices (bases)



Observed tensor

$$\boldsymbol{X} = [\boldsymbol{X}_1, \cdots, \boldsymbol{X}_N] \in \mathbb{R}^{M \times M \times N}$$

Basis tensor

$$\boldsymbol{V} = [\boldsymbol{V}_1, \cdots, \boldsymbol{V}_K] \in \mathbb{R}^{M \times M \times K}$$

Activation matrix

$$\boldsymbol{H} = [\boldsymbol{h}_1, \cdots, \boldsymbol{h}_K]^T \in \mathbb{R}^{K \times N}$$

Matrix-wise factorization

$$\boldsymbol{X}_n \approx \sum_{k=1}^{K} \boldsymbol{V}_k h_{kn} \stackrel{\text{def}}{=} \boldsymbol{Y}_n$$

Bregman matrix divergence can be used as a cost function

PSDTF: A Natural Extension of NMF

- Nonnegative Matrix Factorization (NMF)
  - Vector-wise factorization
  - Bregman divergence

$$\boldsymbol{x}_n \approx \sum_{k=1}^{K} \boldsymbol{w}_k h_{kn} \stackrel{\text{def}}{=} \boldsymbol{y}_n$$

    - Kullback-Leibler (KL) divergence
      $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_n|\boldsymbol{y}_n) = \sum_m \left(x_{mn} \log x_{mn} y_{mn}^{-1} - x_{mn} + y_{mn}\right)$
    - Itakura-Saito (IS) divergence
      $\mathcal{D}_{\mathrm{IS}}(\boldsymbol{x}_n|\boldsymbol{y}_n) = \sum_m \left(-\log x_{mn} y_{mn}^{-1} + x_{mn} y_{mn}^{-1} - 1\right)$
- Positive Semidefinite Tensor Factorization (PSDTF)
  - Matrix-wise factorization
  - Bregman matrix divergence

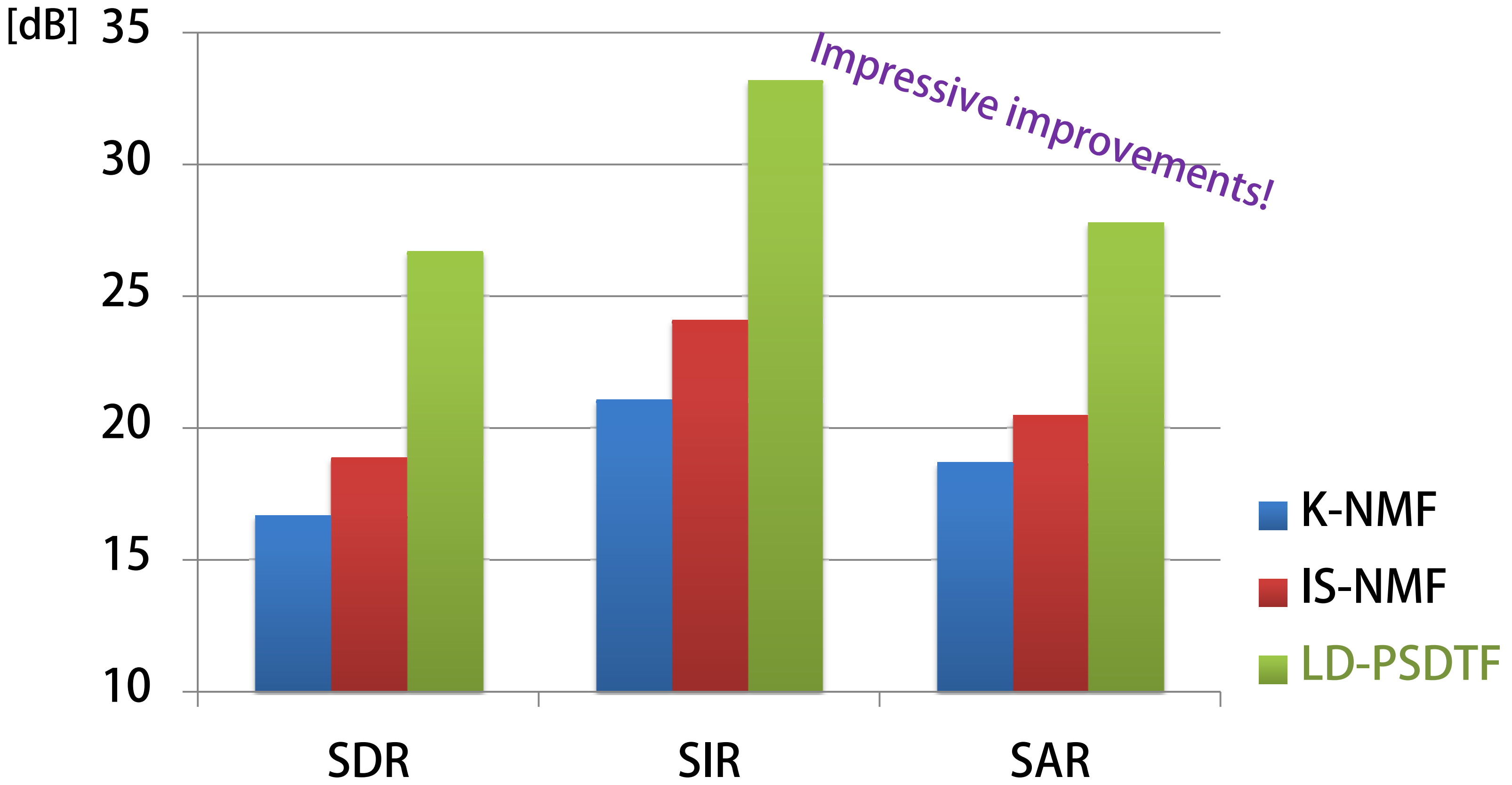
$$\boldsymbol{X}_n \approx \sum_{k=1}^{K} \boldsymbol{V}_k h_{kn} \stackrel{\text{def}}{=} \boldsymbol{Y}_n$$

    - von Neumann (vN) divergence
      $\mathcal{D}_{\mathrm{vN}}(\boldsymbol{X}_n|\boldsymbol{Y}_n) = \mathrm{tr}\left(\boldsymbol{X}_n \log \boldsymbol{X}_n \boldsymbol{Y}_n^{-1} - \boldsymbol{X}_n + \boldsymbol{Y}_n\right)$
    - Log-determinant (LD) divergence
      $\mathcal{D}_{\mathrm{LD}}(\boldsymbol{X}_n|\boldsymbol{Y}_n) = -\log\left|\boldsymbol{X}_n \boldsymbol{Y}_n^{-1}\right| + \mathrm{tr}\left(\boldsymbol{X}_n \boldsymbol{Y}_n^{-1}\right) - M$

$K \rightarrow \infty$
Nonparametric Bayesian infinite extension feasible

Single-Channel Audio Source Separation

- LD-PSDTF outperformed KL-NMF and IS-NMF
  - Tested on a toy mixture signal consisting of three piano sounds



Take-Home Messages

- We proposed positive semidefinite tensor factorization (PSDTF)
  - Tensor extension of nonnegative matrix factorization (NMF)
    - Nonnegative tensor factorization (NTF) is a naive extension of NMF
  - Bayesian nonparametrics
    - The gamma process is used for Bayesian PSDTF that can deal with an infinite number of bases
- A special case: log-determinant PSDTF (LD-PSDTF)
  - Elegant variational inference
    - Closed-form MU and VB updates were derived
- Another special case: von-Neumann PSDTF (vN-PSDTF)
  - This is worth investigating (closed-form solution exists?)