Improving biodiversity monitoring through soundscape information retrieval

Yu Tsao, Tzu-Hao Lin

Research Center for Information Technology Innovation, Academia Sinica

Mao-Ning Tuanmu, Joe Chun-Chia Huang, Chia-Yun Lee

Biodiversity Research Center, Academia Sinica

Chiou-Ju Yao

National Museum of Natural Science

Outline

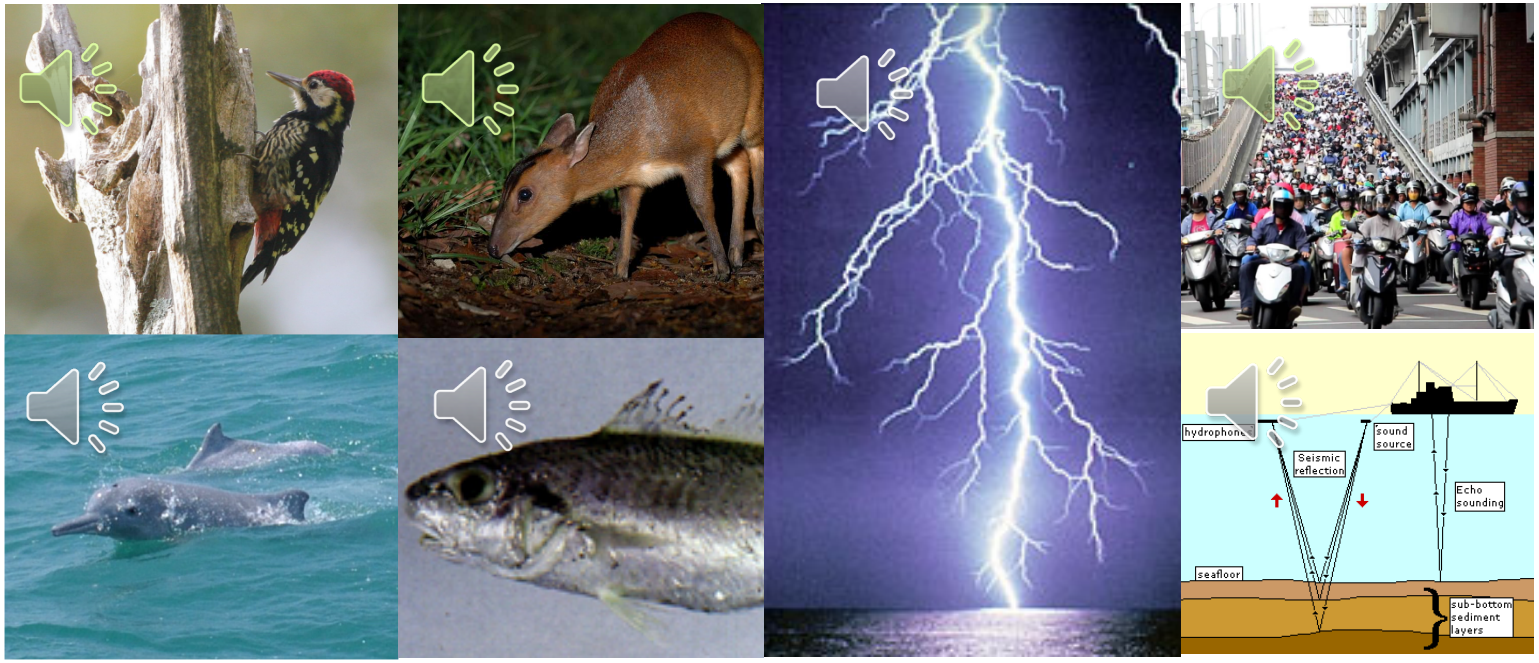
- Introduction
  - Soundscape information retrieval (SIR)
  - A critical challenge in SIR
- The multi-layered non-negative matrix factorization (MLNMF) blind source separation (BSS)
- Three applications of using MLNMF BSS
  - Separating soundscape components from long-duration recordings
  - Separating different species of animal vocalizations
  - Searching target signals from noisy recordings
- Conclusion

Outline

- Introduction
  - Soundscape information retrieval (SIR)
  - A critical challenge in SIR
- The multi-layered non-negative matrix factorization (MLNMF) blind source separation (BSS)
- Three applications of using MLNMF BSS
  - Separating soundscape components from long-duration recordings
  - Separating different species of animal vocalizations
  - Searching target signals from noisy recordings
- Conclusion

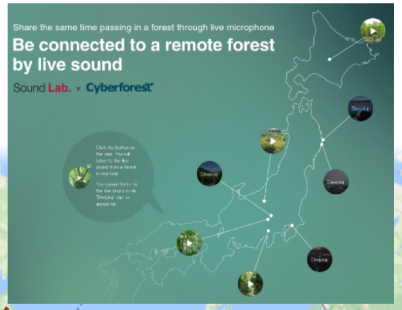
Soundscape: data science of ecosystem

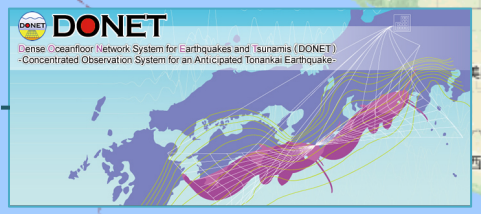
- **Biophony:** biodiversity
- **Geophony:** geophysical events and weather conditions
- **Anthrophony:** impacts of human activities on ecosystem



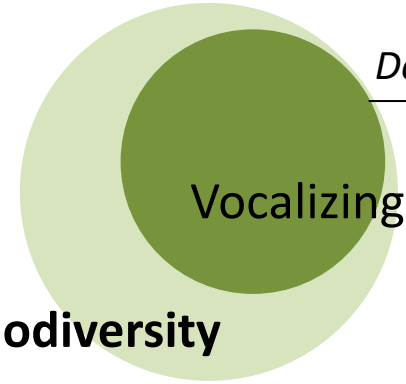
Large-scale soundscape projects

- **Big data:** large-scale, long-term monitoring
- **Interdisciplinary integration:** sensor network, signal processing, ecological research





One important application of SIR: biodiversity monitoring



Detection and classification

Composition of calling species

Statistical modeling

Dynamics of biodiversity

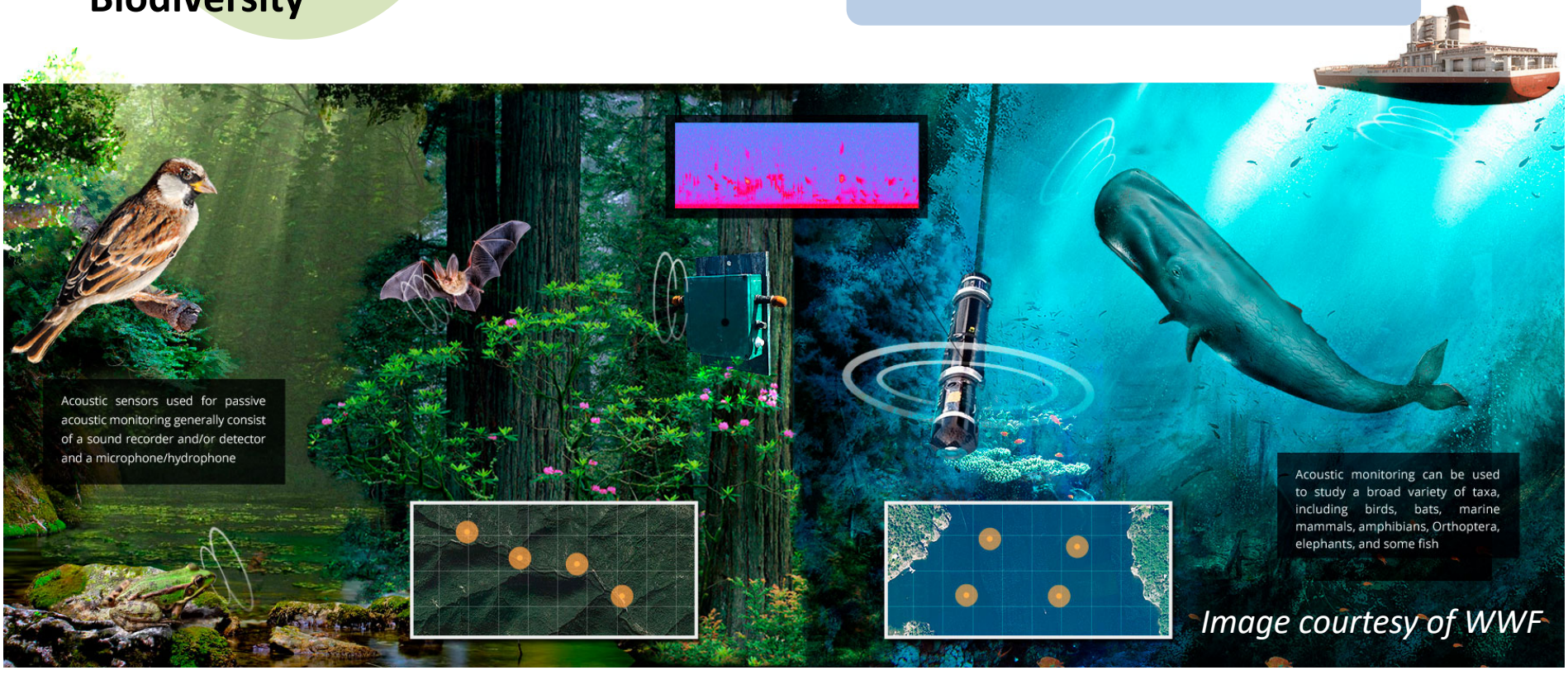
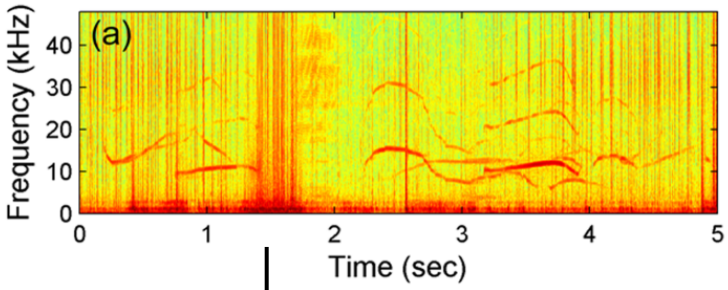


Image courtesy of WWF

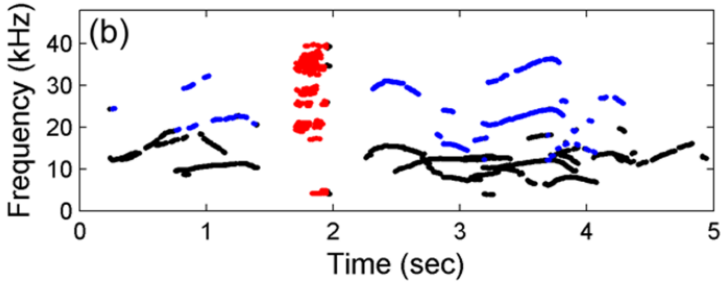
Searching based on existing information

Choose targets → Rule-based detectors → Long-term monitoring

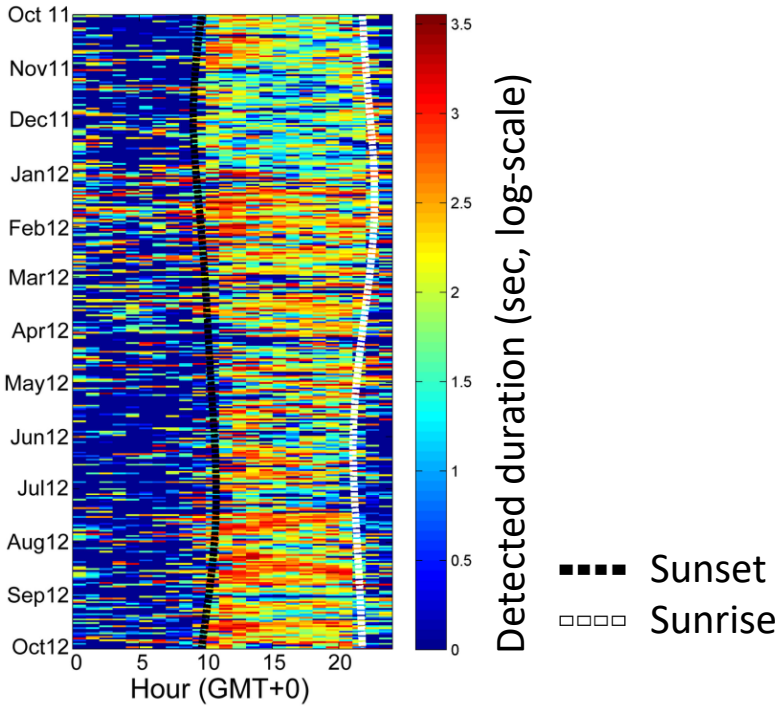
Spectrogram of a recording clip



Detection result of tonal sounds

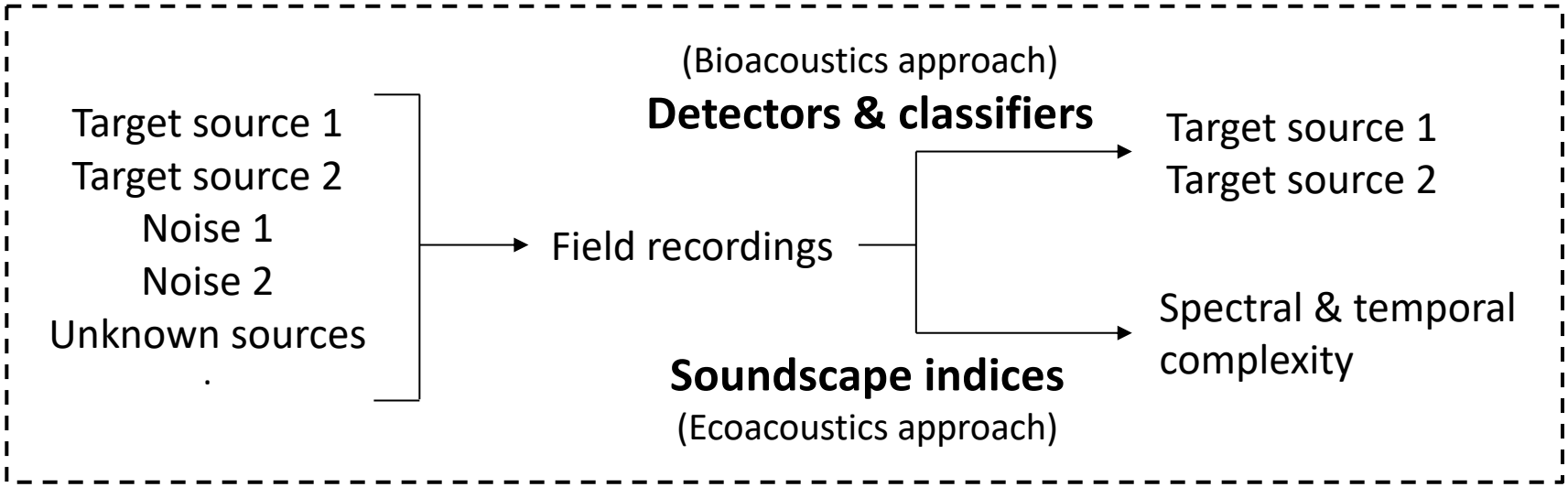


Year-round presence of toothed whales



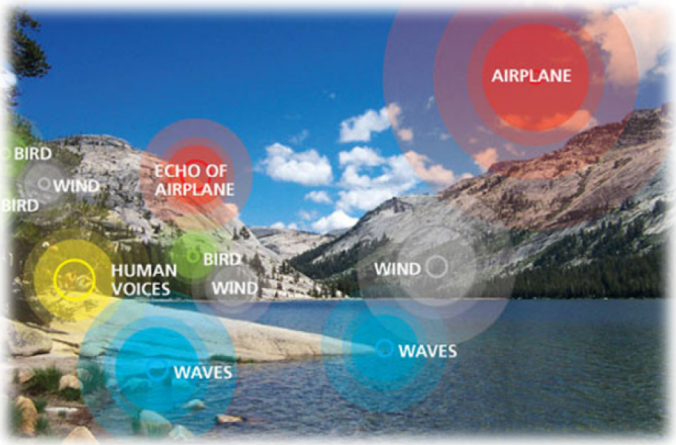
Lin et al. (2015) PLOS ONE

A critical challenge in soundscape analysis



- **Precision of acoustical analysis**
  - Noise interference
  - Simultaneous sound sources
- **Low precision may lead to a biased ecological interpretation**

Source separation can be applied to address the issue



https://www.nps.gov/yose/learn/nature/soundscape.htm

Outline

- Introduction
  - Soundscape information retrieval (SIR)
  - Challenge in soundscape
- The multi-layered non-negative matrix factorization (MLNMF) blind source separation (BSS)
- Three applications of using MLNMF BSS
  - Separating soundscape components from long-duration recordings
  - Separating different species of animal vocalizations
  - Searching target signals from noisy recordings
- Conclusion

Source separation may improve the analysis of soundscape

- **Supervised approach**
  - Model building based on training data
  - Powerful but require labeled and clean data

- **Unsupervised approach (blind source separation)**
  - Dictionary learning
  - Matrix factorization

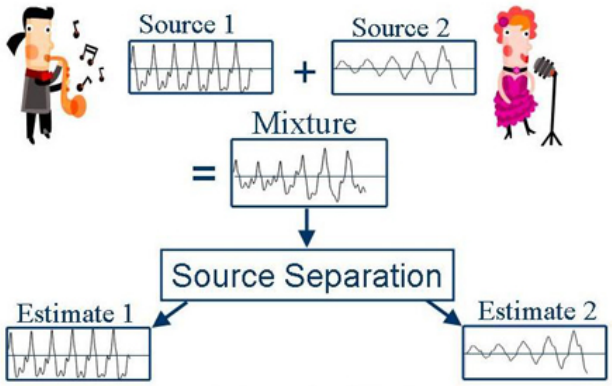
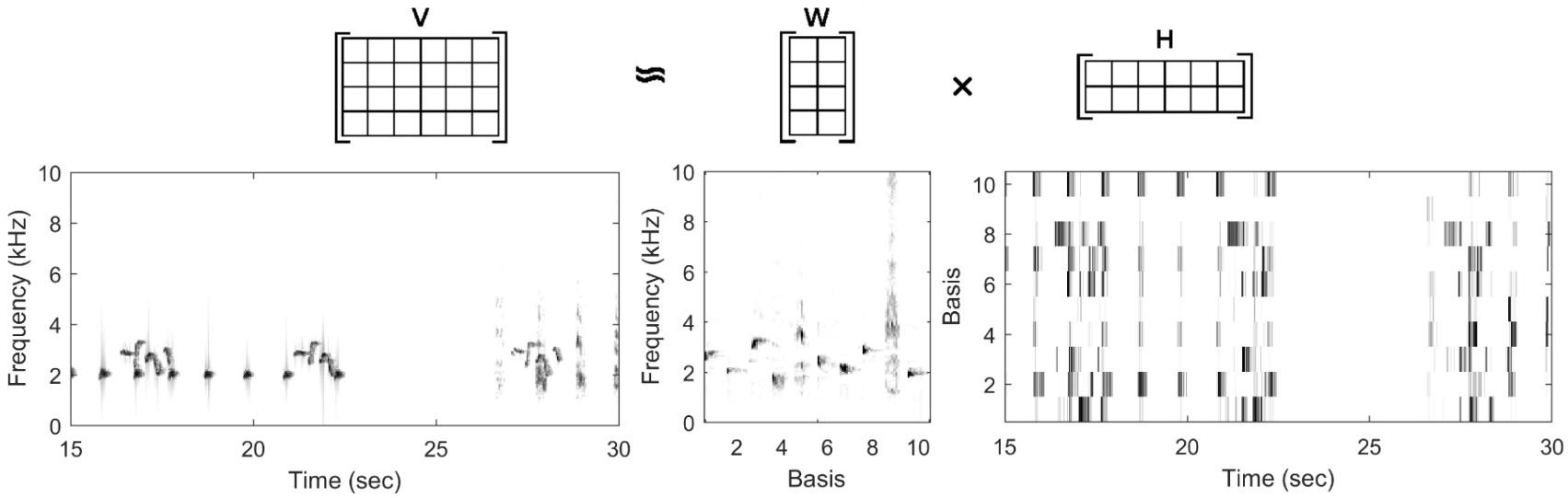


Image courtesy of music.cs.northwestern.edu

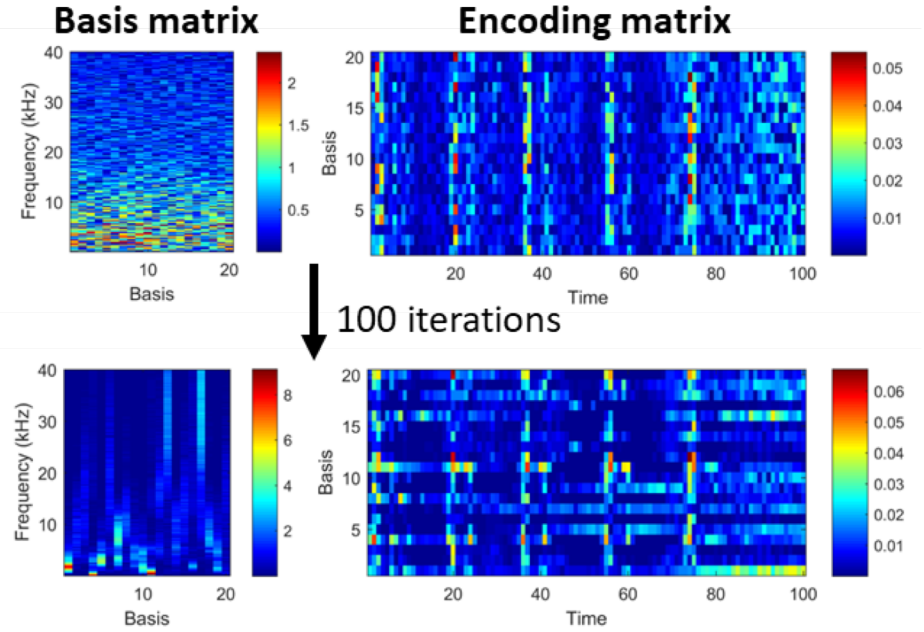
Non-negative matrix factorization (NMF)

- Decomposition of a non-negative matrix (spectrogram)
  - **Basis matrix (W)** : spectral feature (bricks)
  - **Encoding matrix (H)** : temporal information (number of bricks)



Self-learning of NMF

- Initialize by random values or prior knowledge
- Update the basis matrix (*W*) and encoding matrix (*H*) through iterations



Update procedure:

$$W_{ia} \leftarrow W_{ia} \sum_{\mu} \frac{V_{i\mu}}{(WH)_{i\mu}} H_{a\mu}$$

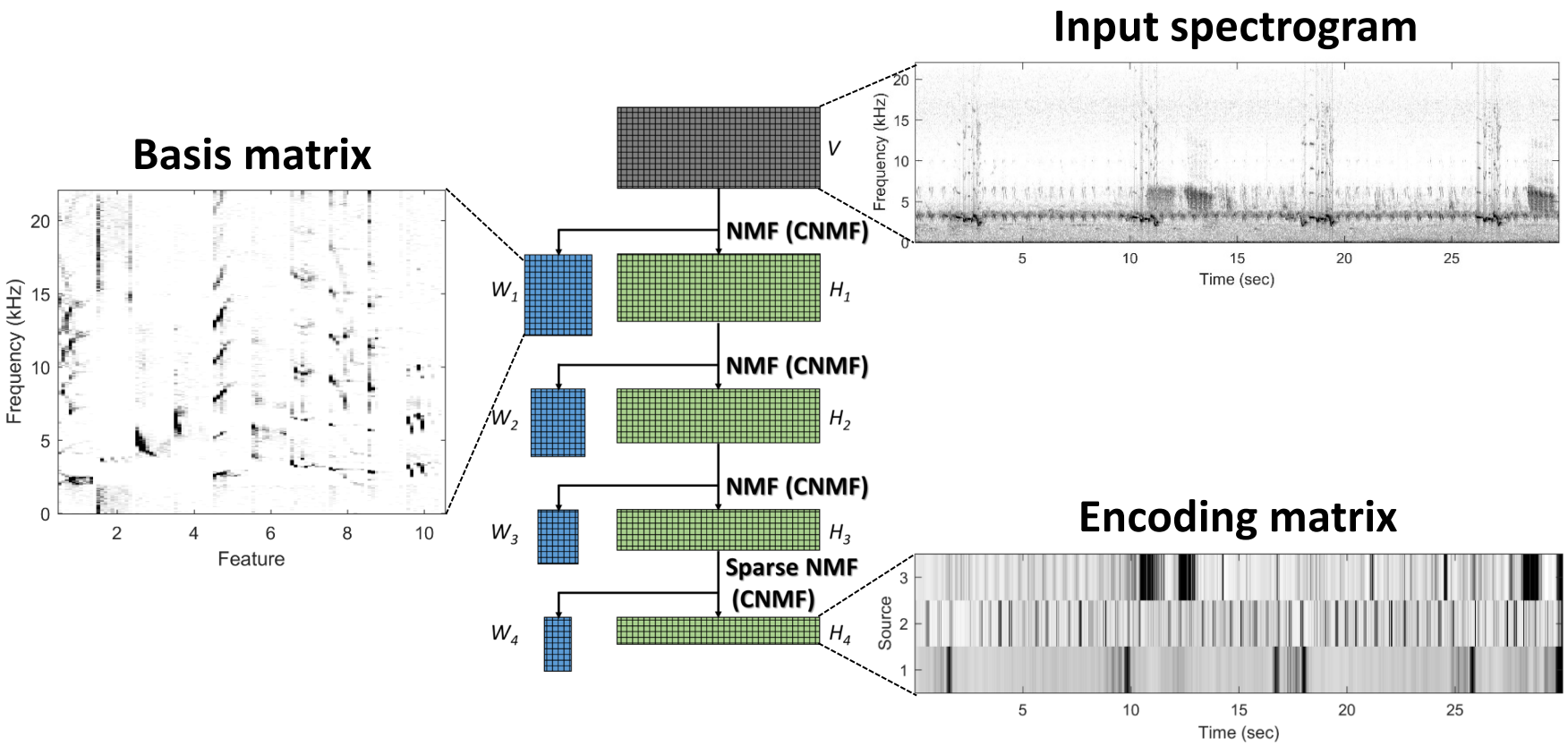
$$H_{a\mu} \leftarrow H_{a\mu} \sum_{i} W_{ia} \frac{V_{i\mu}}{(WH)_{i\mu}}$$

$$H = \underset{H>0}{\operatorname{argmin}}\, d(V, WH) + \lambda \|H\|_1$$

Sparsity constraint

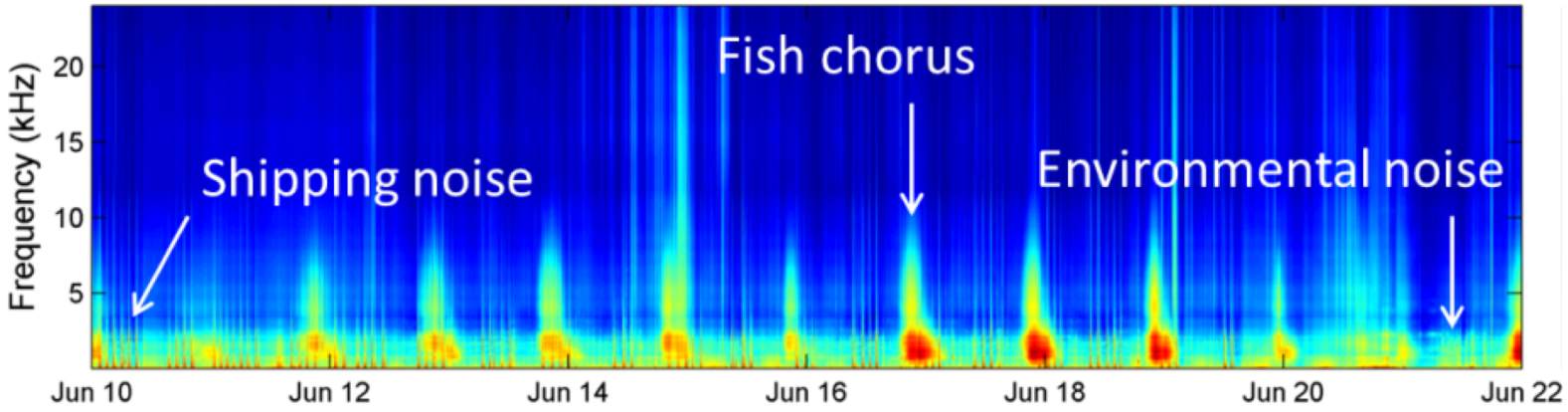
Multiple layers of NMF (MLNMF)

- Estimate basis weights by multiple layers of NMF or CNMF
- Learn the encoding information of *k* sources by the sparse layer



Outline

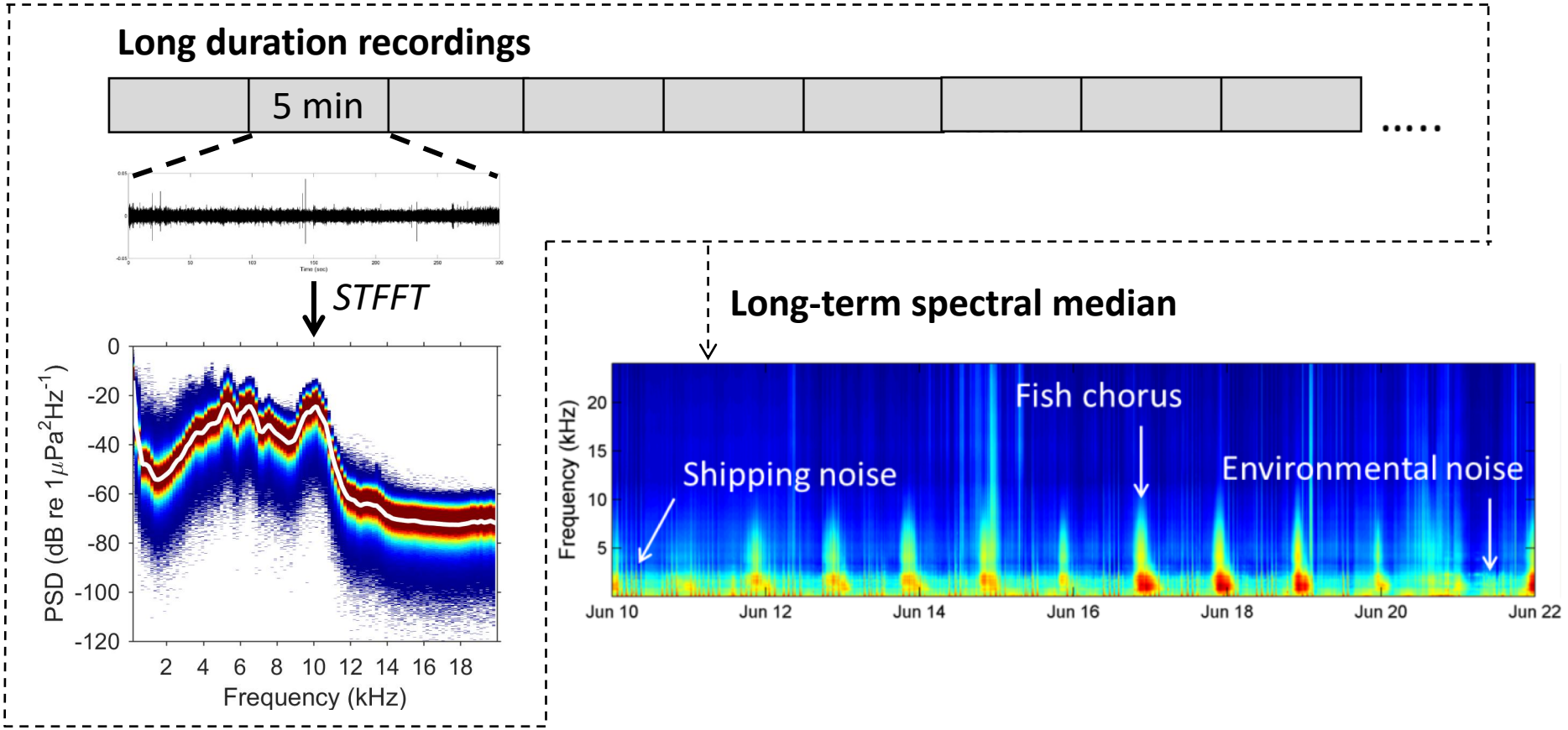
- Introduction
  - Soundscape information retrieval (SIR)
  - Challenge in soundscape
- The multi-layered non-negative matrix factorization (MLNMF) blind source separation (BSS)
- Three applications of using MLNMF BSS
  - Separating soundscape components from long-duration recordings
  - Separating different species of animal vocalizations
  - Searching target signals from noisy recordings
- Conclusion



SEPARATION OF SOUNDSCAPE COMPONENTS

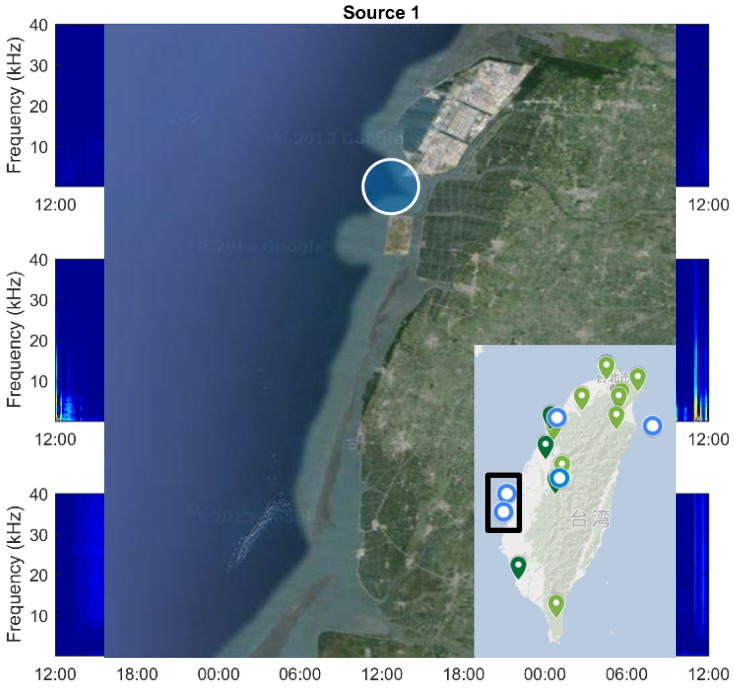
Visualization of long-duration recordings

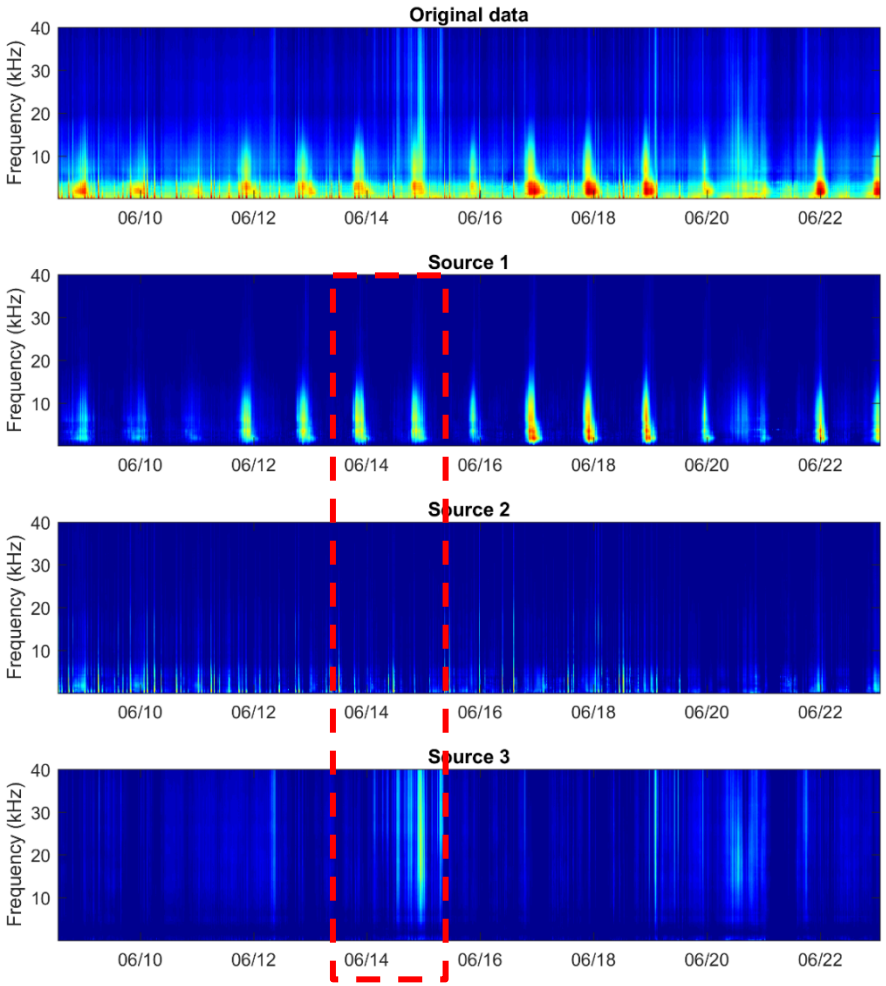
- Visualization of marine soundscape using a long-term spectrogram



BSS of biophony and anthrophony

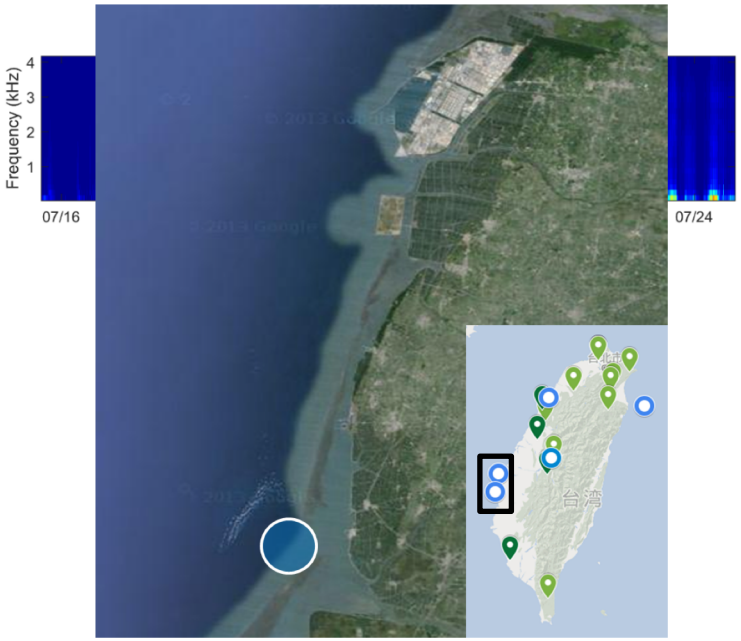
- Recordings collected near by the Mailiao industrial harbor, Yunlin County
  - Fish chorus
  - Shipping noise
  - Snapping shrimp sounds

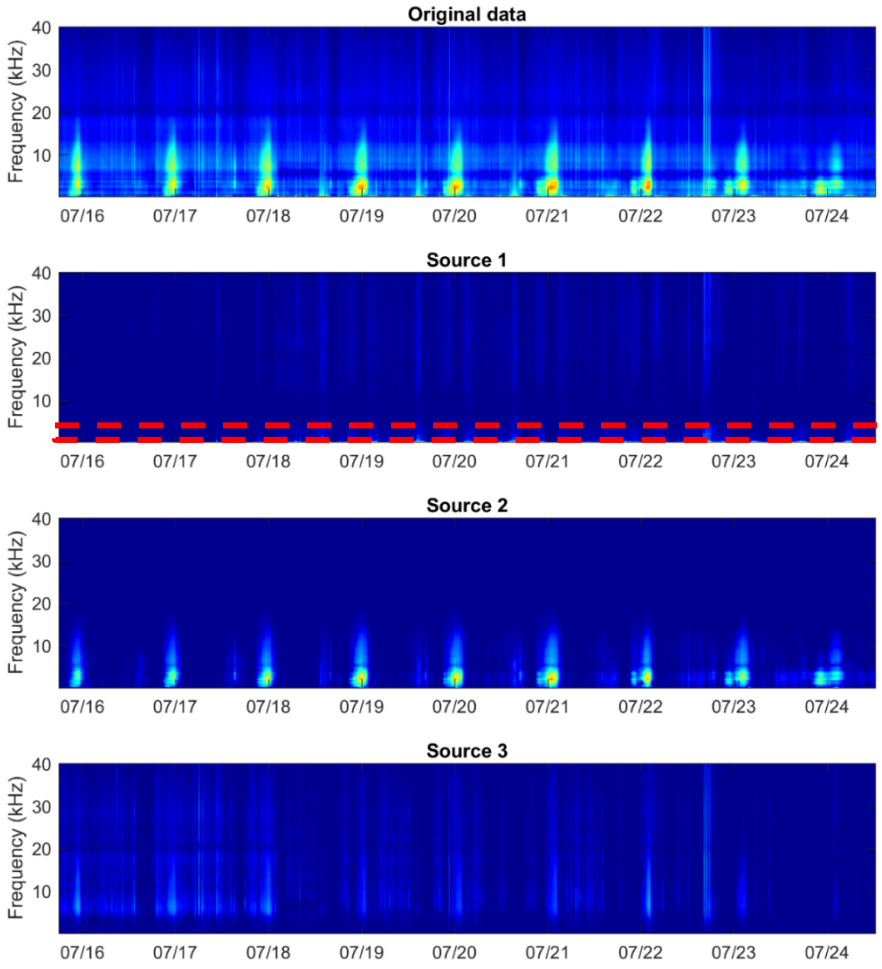


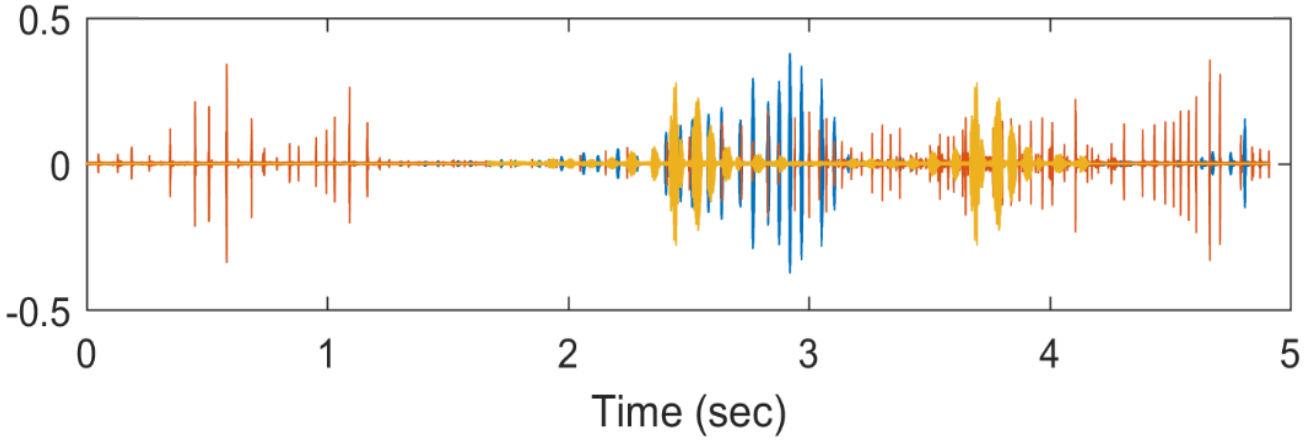


BSS of biophony and geophony

- Recordings collected near by Waisanding sandbar, Yunlin County
  - Current noise
  - Fish chorus
  - Snapping shrimp sounds



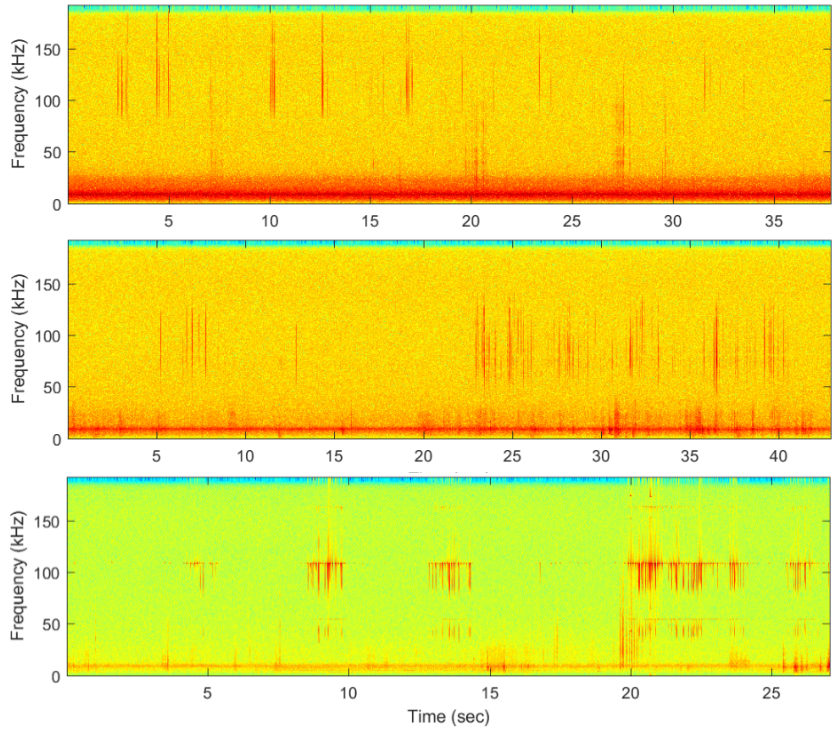




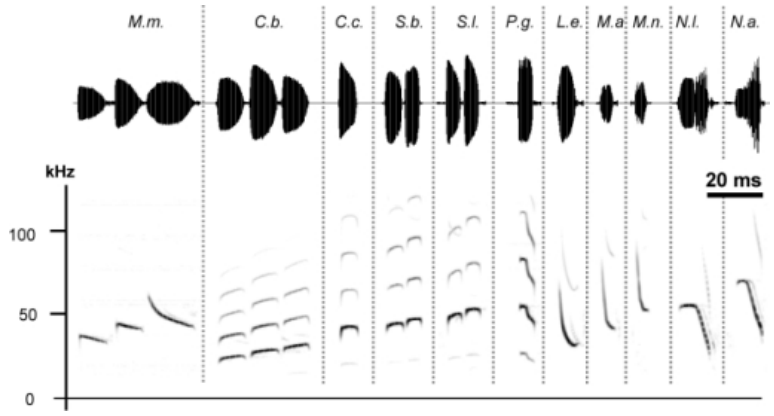
SEPARATION OF SPECIES-SPECIFIC CALLS

Separation of different bat echolocation calls

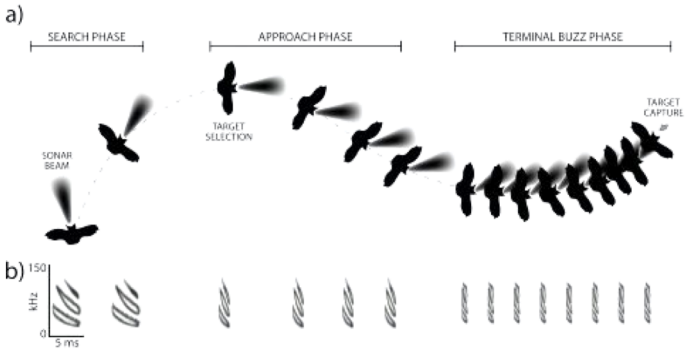
- Reference calls of 13 bat species, that differed in spectral and temporal features



Data collected by Biodiversity Research Center, Academia Sinica, Taiwan



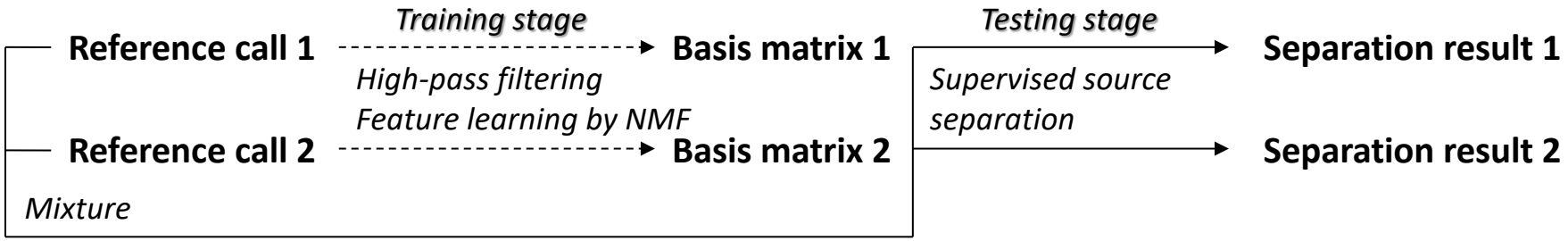
Surlykke and Kloko (2008) PLOS ONE



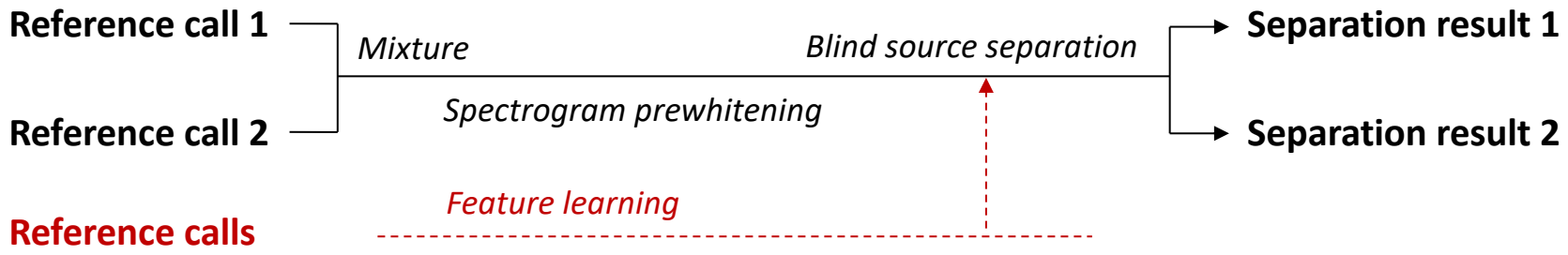
http://tcabasa.org/?page_id=2718

Comparing Supervised NMF SS and MLNMF BSS

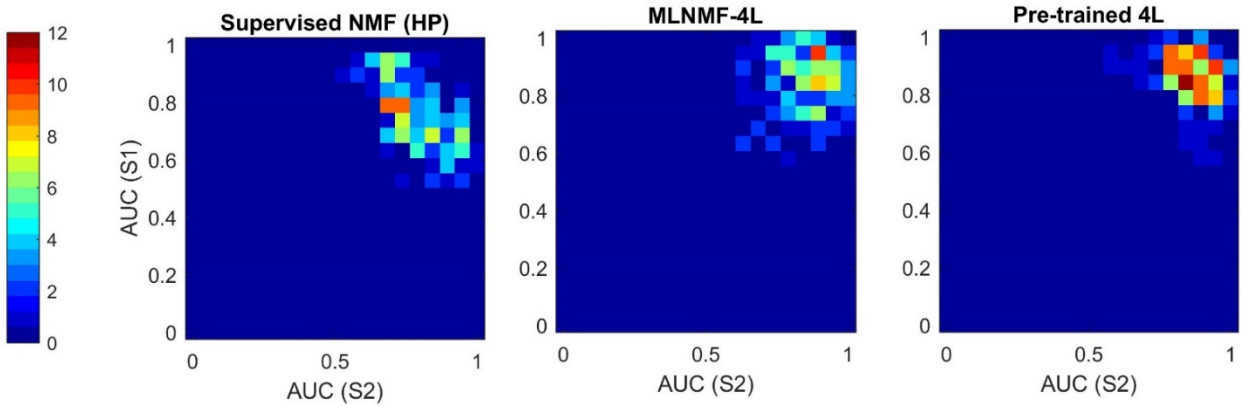
- **Supervised NMF**



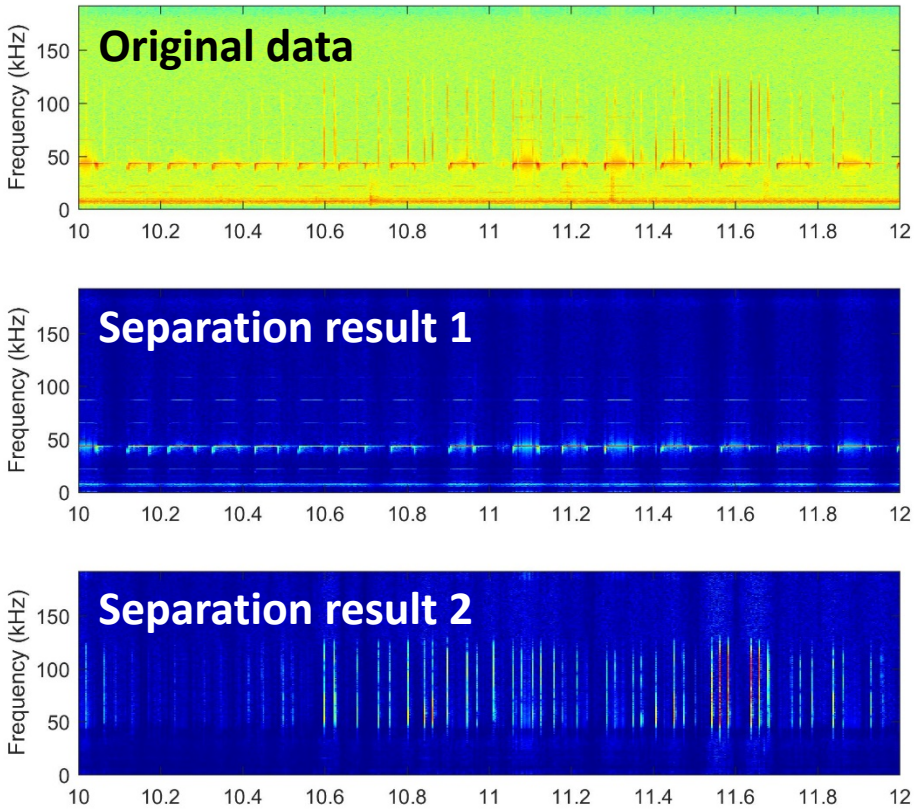
- **MLNMF-4L (pre-trained)**

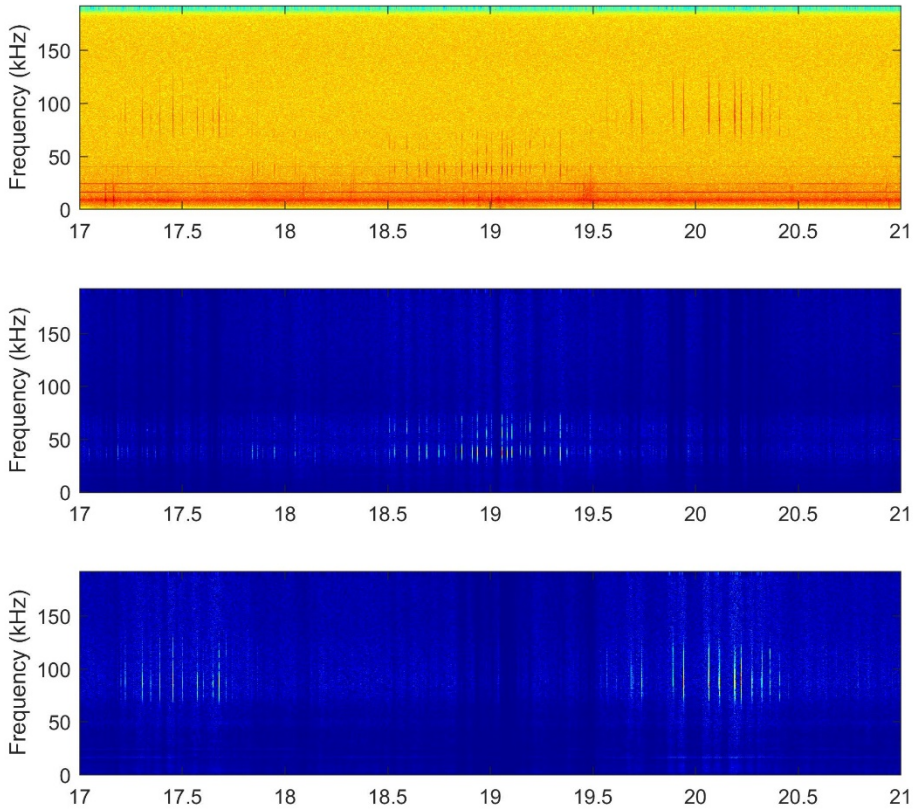


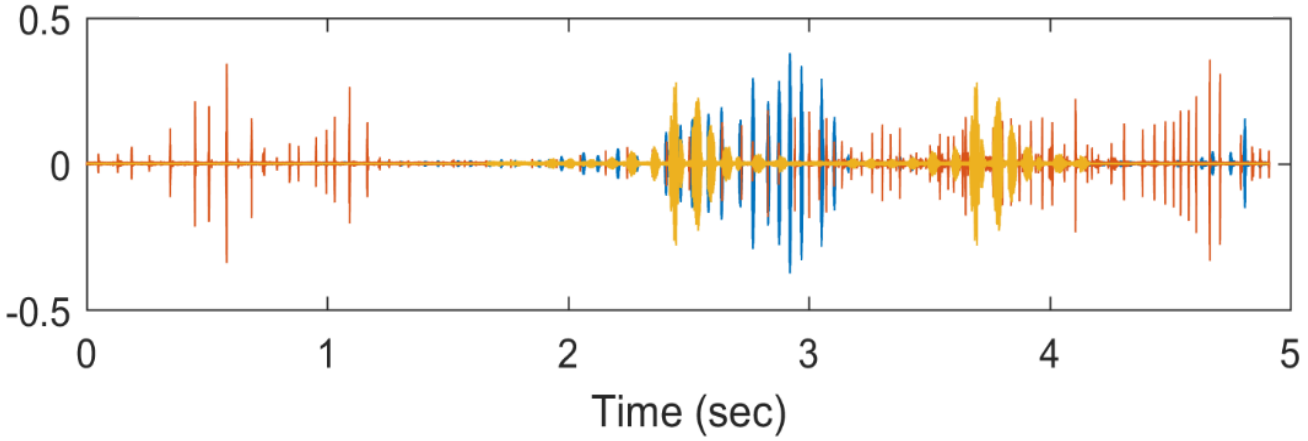
Distribution of AUC scores



Examples of separation result





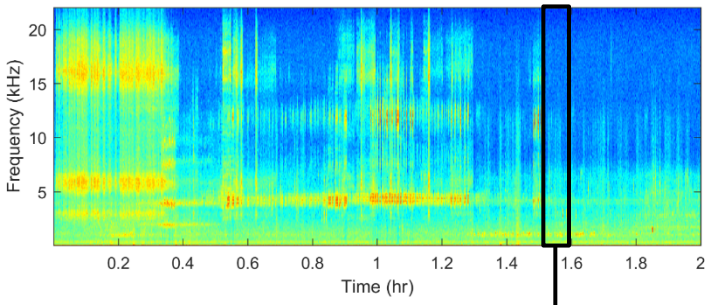


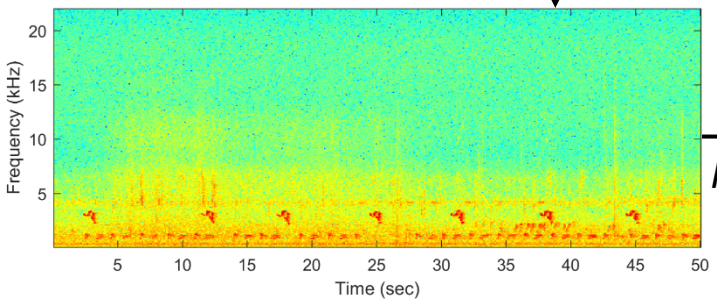
SEARCHING TARGET SIGNAL IN A NOISY SOUNDSCAPE

Bird calls in a subtropical forest

- **Dusky fulvetta (*Schoeniparus brunnea*)**
  - Common species < 2000 m elevations

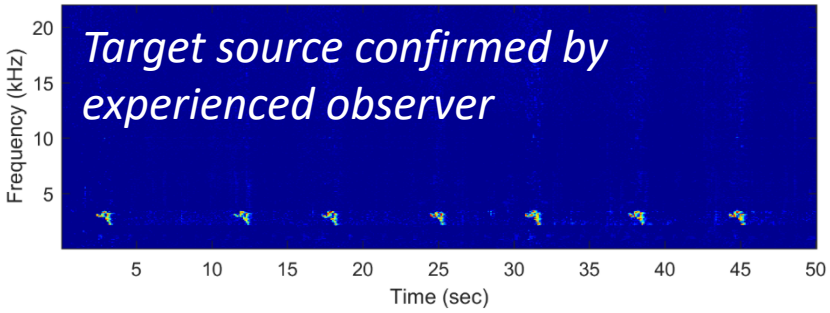
Field recordings from 5 AM – 7 AM

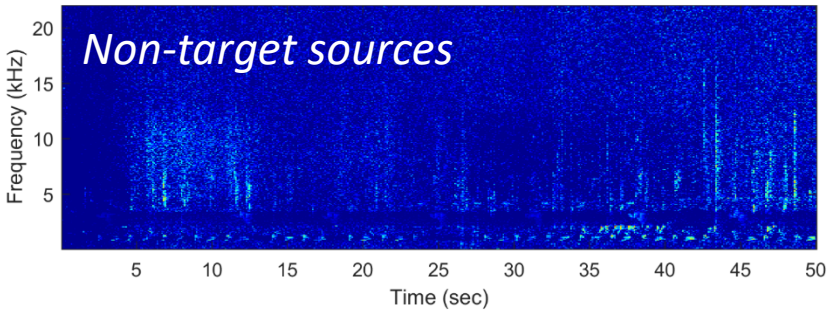




Prewhitening

MLNMF

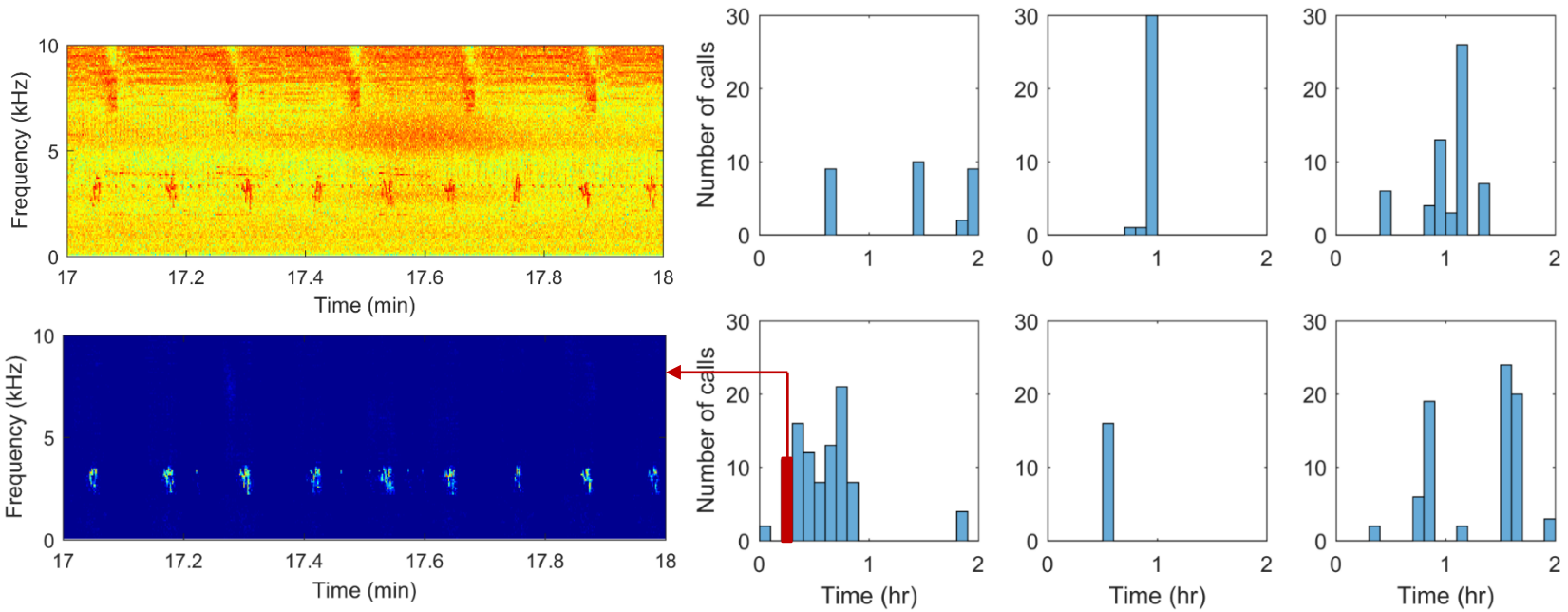




Data collected by National Museum of Natural Science, Taiwan

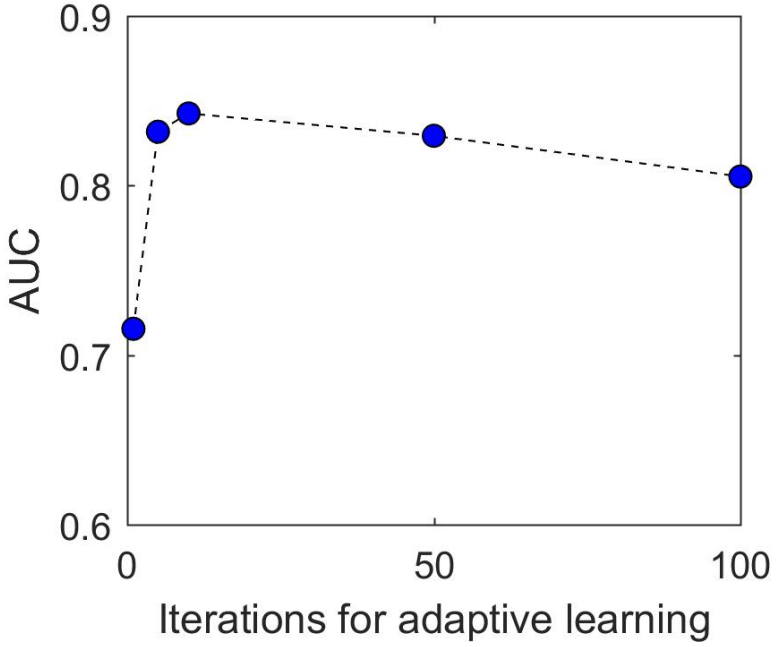
Spatial-temporal distribution of bird calls

- MLNMF model perform well in searching the target bird call by using a small training data (unlabeled!)
- Noise types not encountered in the training data can also be separated effectively

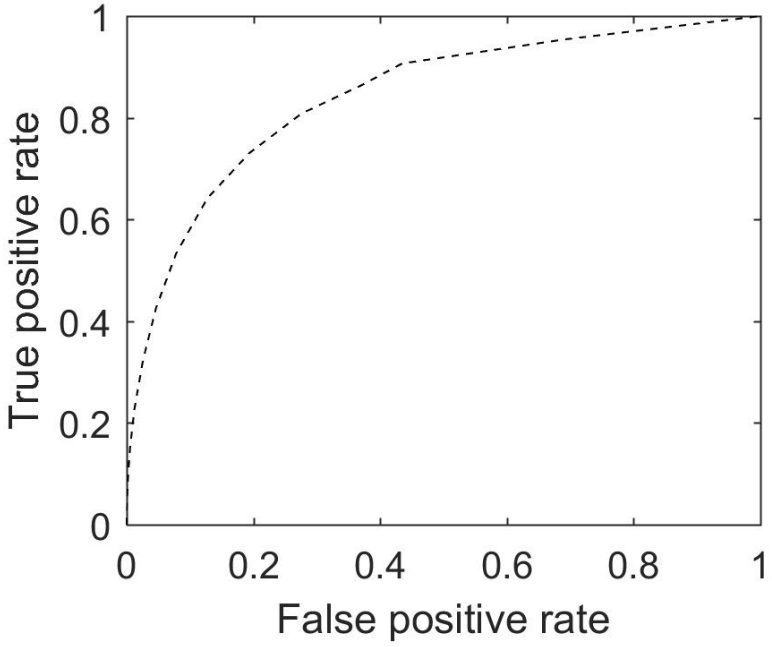


Performance of “audio search”

Separation performance of using different iterations



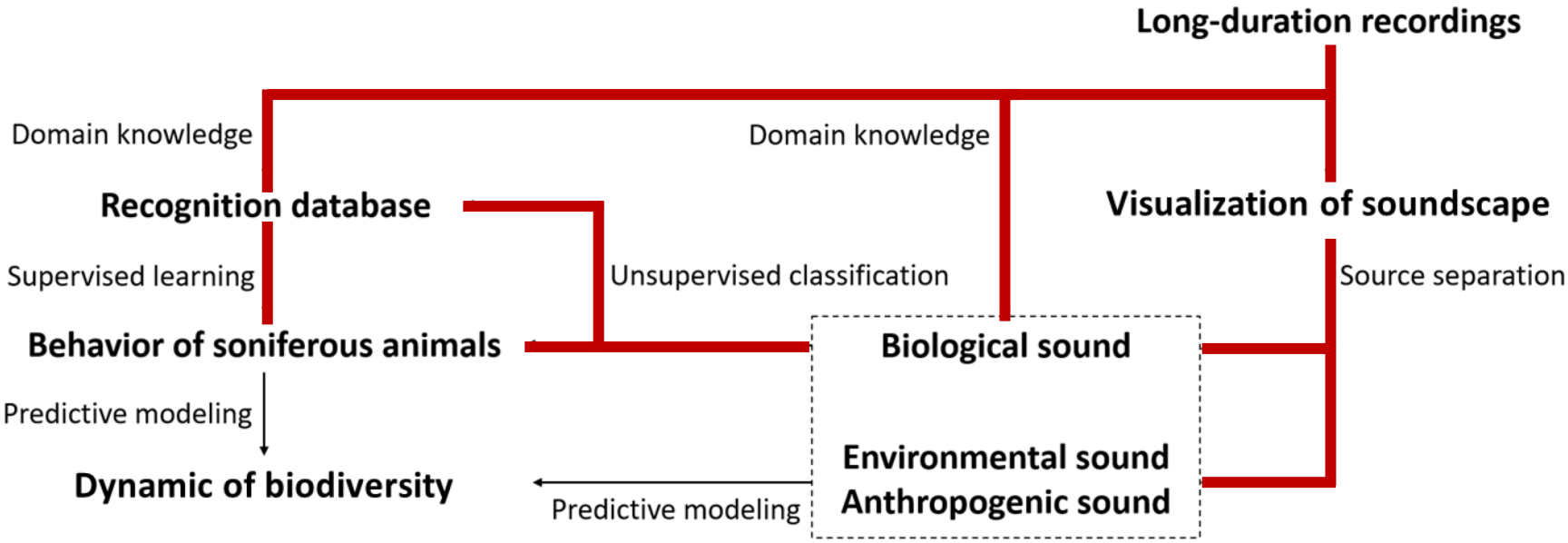
Detection performance of 10 iterations of adaptive learning



Outline

- Introduction
  - Soundscape information retrieval (SIR)
  - Challenge in soundscape
- The multi-layered non-negative matrix factorization (MLNMF) blind source separation (BSS)
- Three applications of using MLNMF BSS
  - Separating soundscape components from long-duration recordings
  - Separating different species of animal vocalizations
  - Searching target signals from noisy recordings
- Conclusion

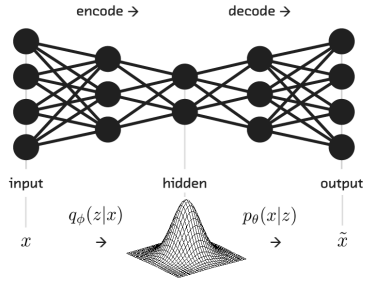
Improving biodiversity monitoring using SIR



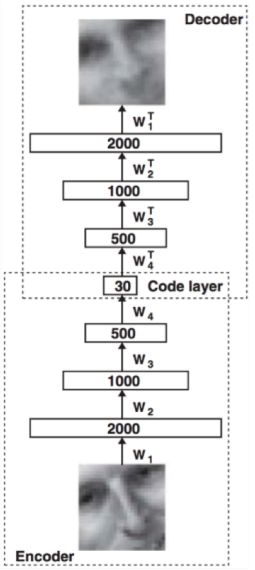
- Reduce the difficulty of acoustical analysis for ecologists
  - Effective blind source separation
    - Soundscape components on a long-term spectrogram
    - Animal vocalizations with different spectral and temporal characteristics
  - Efficient supervised separation by using a small training data

Future integration of deep learning and SIR

- **Unsupervised learning (Deep AE, Variational AE)**
  - Improve NMF-based blind source separation
  - Facilitate the collection of labeled data with minimum manpower?

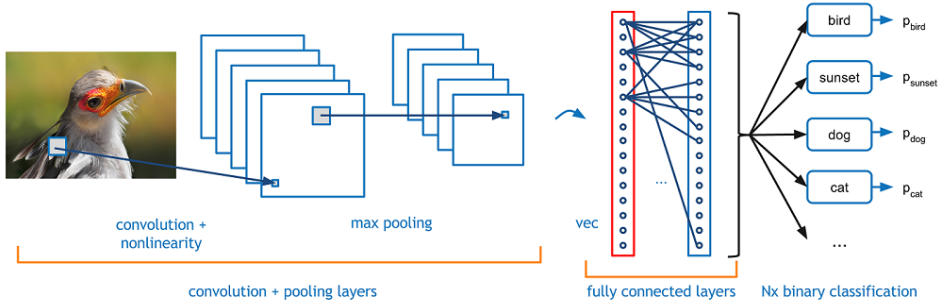


blog.fastforwardlabs.com



Hinton & Salakhutdinov (2006) Science

- **Supervised learning (CNN, RNN…)**
  - Identification of species by learning from a large amount of training data
  - Identification of behavior or individual?



adeshpande3.github.io

Acknowledgement

- Ministry of Science and Technology
  - Independent postdoctoral research fellowship
- Academia Sinica Data Science Seed Grant Program

Welcome for collaboration!

Thank you very much!