Naïve Unsupervised Bioacoustics Signal Processing with Nonnegative Matrix Factorization

Axel Marmoret[1], Nicolas Farrugia[1], Dorian Cazau[2]

[1] IMT Atlantique, Lab-STICC, Brest, France, [2] CNRS Lab-STICC, ENSTA Bretagne, Brest, France

axel.marmoret@imt-atlantique.fr

Take home message

NMF is an unsupervised learning technique that can extract repetitive patterns in acoustic data.

NMF - Nonnegative Matrix Factorization

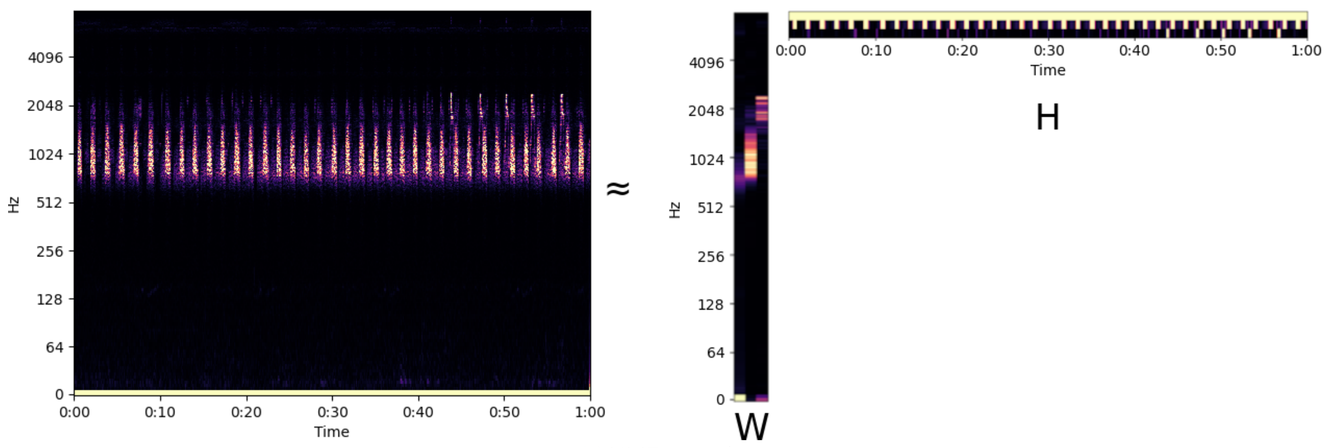
NMF is an unsupervised learning technique that factorizes a matrix V into two nonnegative matrices W and H, such that [1]:

$$V \approx WH \tag{1}$$

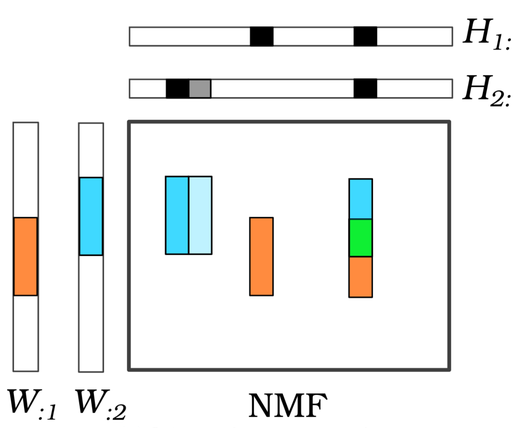
This is obtained by minimizing some distance function $d()$ between V and WH:

$$\arg\min_{W,H \geq 0} d(V, WH) \tag{2}$$

For instance, using the Euclidean distance, the optimization problem becomes: $\arg\min_{W,H \geq 0} \sum_{i,j} (V_{ij} - (WH)_{ij})^2$



(a) NMF example.



(b) Schematic representation.

Source Separation and Sound Event Detection

Source Separation can be obtained by applying NMF on the spectrogram of an audio signal [2]. It has already been studied for bioacoustics signals [3, 4], notably for underwater [5, 6] signals.

When sources are accurately separated, one can detect them by thresholding the activation matrix H, leading to Sound Event Detection [7]. Still, this needs appropriate post-processing to be efficient.

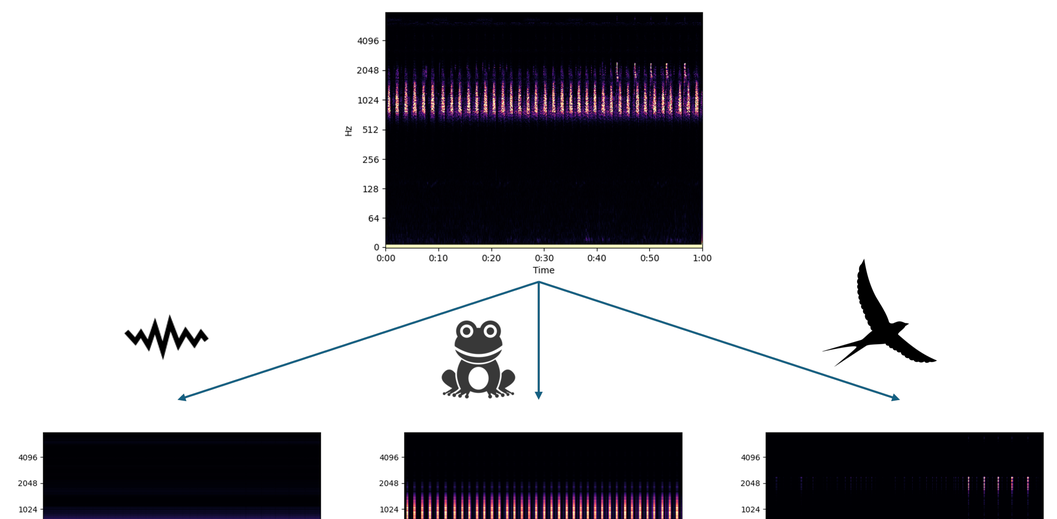
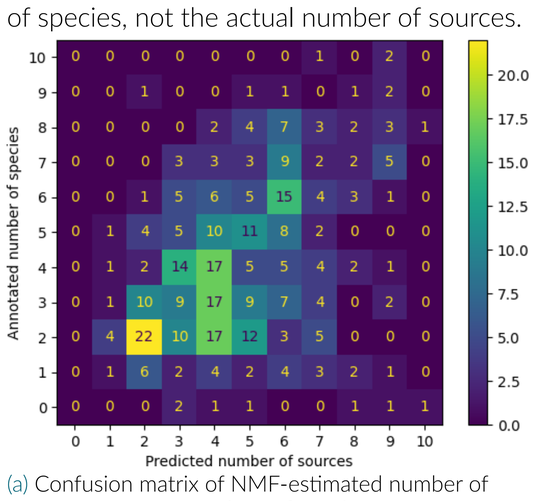


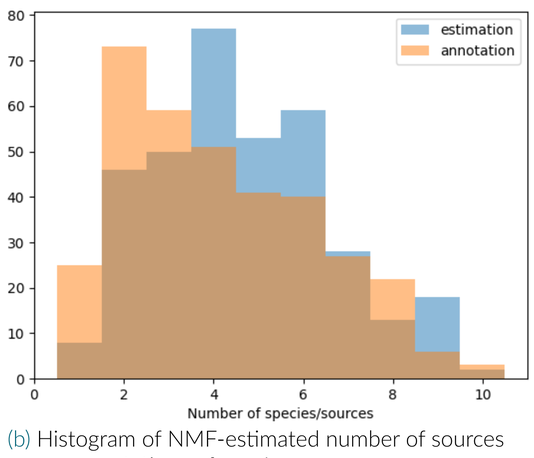
Figure 2. Source separation with NMF.

Estimating the number of sources

Once NMF is computed, one can estimate the number of sources by clustering the NMF outputs. Hereafter are presented some results obtained on the AnuraSet dataset [8] (INCT17 subset). Results are obtained by clustering the H rows to find simultaneous events, using the DBSCAN algorithm [9]. These results are intended as a **Proof Of Concept**. Indeed, annotations indicate the presence/absence of species, not the actual number of sources.



(a) Confusion matrix of NMF-estimated number of sources versus annotations of species.



(b) Histogram of NMF-estimated number of sources versus annotations of species.

Listenable examples

Follow this link to listen to source separation examples on the AnuraSet dataset and on marine mammal recordings:

ax-le.github.io/resources/examples/JJBA_source_separation

Open-source toolbox

gitlab.imt-atlantique.fr/a23marmo/nmf_bioacoustic

First Conclusions, from the data viewpoint

- ✓ NMF can extract interpretable patterns from bioacoustic data.
- ✓ NMF can help investigate in an unsupervised way the content of recordings.
- ✓ NMF is supported by numerous previous research and can be further improved.
- × When annotated data are available, other techniques (neural networks) will outperform NMF.
- × The post-processing of NMF results is crucial and may be complex.
- × NMF is not suited when events are sparse, as it appears to be for marine mammal recordings.

This was a first approach, and a lot of work remains to be done!

Future work - as a bioacoustician (you?)

- Have a better look at data pre-processing,
- Find the best way to post-process NMF results,
- Leveraging other data and other tasks.

Future work - as a computer scientist (me!)

- Improve the NMF algorithm (constraints or dictionary learning),
- Extend NMF with the most recent algorithms (Deep NMF in particular),
- Extend to tensor decomposition (multi-way arrays, for instance 3D-arrays).

Contact me! :)

You can help me extend this work beyond a naïve computer scientist study, and make it more useful for bioacoustics!

References

[1] D. Lee and H. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[2] P. Smaragdis and J. C. Brown, "Non-negative matrix factorization for polyphonic music transcription," in *2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (IEEE Cat. No. 03TH8684)*, pp. 177–180, IEEE, 2003.

[3] J. Ludena-Choez, R. Quispe-Soncco, and A. Gallardo-Antolin, "Bird sound spectrogram decomposition through non-negative matrix factorization for the acoustic classification of bird species," *PloS one*, vol. 12, no. 6, p. e0179403, 2017.

[4] N. Hassan and D. A. Ramli, "Underdetermined blind source separation of bioacoustic signals.," *Pertanika Journal of Science & Technology*, vol. 31, no. 5, 2023.

[5] T.-H. Lin, S.-H. Fang, and Y. Tsao, "Improving biodiversity assessment via unsupervised separation of biological sounds from long-duration recordings," *Scientific reports*, vol. 7, no. 1, p. 4547, 2017.

[6] D. Li, M. Wu, L. Yu, J. Han, and H. Zhang, "Single-channel blind source separation of underwater acoustic signals using improved nmf and fastica," *Frontiers in Marine Science*, vol. 9, p. 1097003, 2023.

[7] S. Mohmmad and S. K. Sanampudi, "Exploring current research trends in sound event detection: a systematic literature review," *Multimedia Tools and Applications*, pp. 1–43, 2024.

[8] J. S. Cañas, M. P. Toro-Gómez, L. S. M. Sugai, H. D. Benítez Restrepo, J. Rudas, B. Posso Bautista, L. F. Toledo, S. Dena, A. H. R. Domingos, F. L. de Souza, *et al.*, "A dataset for benchmarking neotropical anuran calls identification in passive acoustic monitoring," *Scientific Data*, vol. 10, no. 1, p. 771, 2023.

[9] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *kdd*, vol. 96, pp. 226–231, 1996.