Estimating traffic noise levels using acoustic monitoring: a preliminary study

Jean-Rémy Gloaguen[1], Arnaud Can[1], Matthieu Lagrange[2], Jean-François Petiot[2]

[1] Ifsttar-LAE, route de Bouaye-CS4, 44344 Bouguenais,FR

[2] IRCCyN, UMR CNRS 6597, École Centrale de Nantes, 1 rue de la Noë, 44321 Nantes, FR

jean-remy.gloaguen@ifsttar.fr, +33 (0)2 40 84 56 19

Introduction

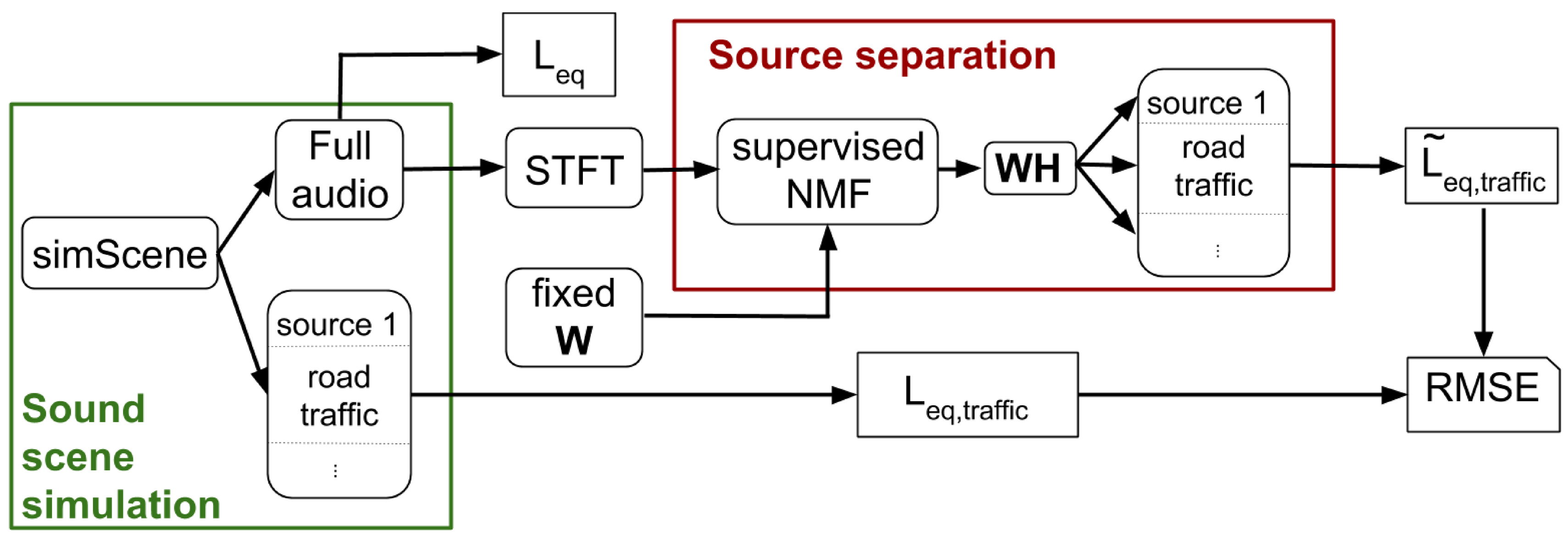
- **Context**: European Directive 2002/49/EC imposes to cities over 100 000 inhabitants to realize road traffic noise maps
- Aim to improve road traffic simulated noise maps by means of with acoustic measurements

=> Need for a method that separates road traffic from the other existing sound sources within measurements

- **Approach**: Usual methods for urban sound environments separation do not deal with overlapping sounds
- The NMF seems to suit urban sound mixture requirements but it is used yet only for audio or musical applications

=> Adapt and test the NMF on a set of simulated urban sound mixtures

Method



Sound scene simulation

Simulation of a set of sound scenes with the *simScene*[a] software:

- Realization of sound mixtures from a database of isolated sounds
- Control of the event/backgroud ratio
- Control of the time of presence of each sound source

It gives as output:

- A complete audio file with all the mixed sound sources and an audio file for each sound class,

=> Creation of specific urban sound environments (animated or quiet streets, parks…)

Source separation

Supervised Non-negative Matrix Factorization:

$$\mathbf{V} \approx \mathbf{WH}$$

- $\mathbf{V_{F\times N}} \in \mathbf{R}^+$, the audio spectrogram of the sound mixture,
- $\mathbf{W_{F\times K}} \in \mathbf{R}^+$, fixed dictionary composed of a set of $\mathbf{K}$ audio spectra of urban sound sources
- $\mathbf{H_{K\times N}} \in \mathbf{R}^+$, the feature matrix standing for the temporal variation of the spectrum found iteratively by the minimisation of the cost function

$$\min \mathrm{D}(\mathbf{V}||\mathbf{WH}) \quad \mathrm{w.r.t.}\, \mathbf{W} > 0, \mathbf{H} > 0$$

3 tested cost functions :

- The Euclidian distance (Euc. Dist)
- The Kullback-Leibler divergence (K-L div)
- The Itakura –Saïto divergence (I-S div)

- $\mathbf{W}$ is composed of 3 classes (*car*, *bird*, *horn*)
- $\mathbf{H}$ is updated with the maximisation-minimisation algorithm[b]

[a]M. Rossignol & al., simScene: a web-based acoustic scenes simulator, *1st Web Audio Conference*, 2015

[b]C. Févotte & J. Idier, Algorithms for nonnegative matrix factorization with the β–divergence, *Neural Computation*, vol 23, no 9, 2011

Experiment and first results

- 20 simple scenes composed of 3 sound classes (car, horn, bird)
- $\mathbf{W}$ is composed of theses 3 classes
- NMF is performed with 100 iterations

=> $\mathrm{L}_{\mathrm{eq,traffic}}$ and $\tilde{\mathrm{L}}_{\mathrm{eq,traffic}}$ estimated for each scene

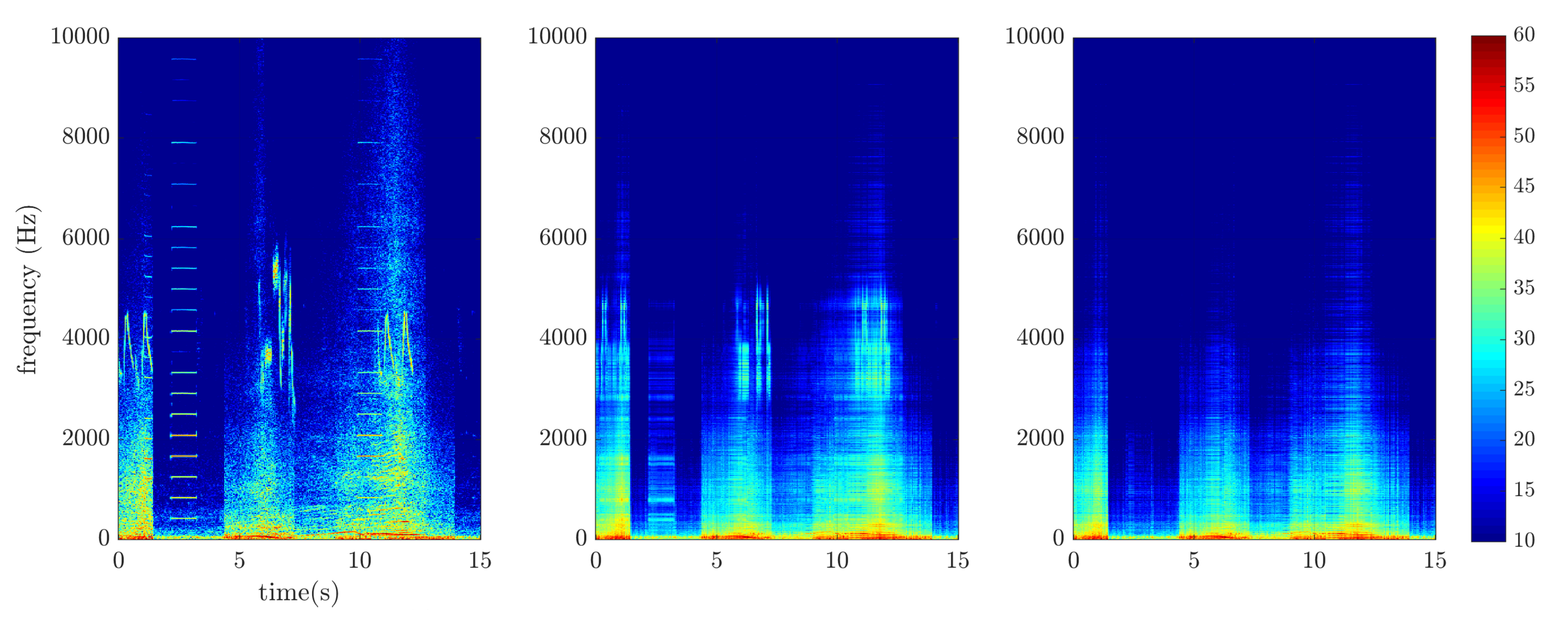


Fig 1: on the left the spectrogram of a scene created with simScene, in the middle the estimated spectrogram after 100 iterations, on the right the road traffic spectrogram estimated after source separation

- Estimation of the performance of the implemented NMF by the RMSE depending on number of iterations and the cost function:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{20}\left(\mathrm{L}^{i}_{\mathrm{eq,traffic}} - \tilde{\mathrm{L}}^{i}_{\mathrm{eq,traffic}}\right)^2}$$

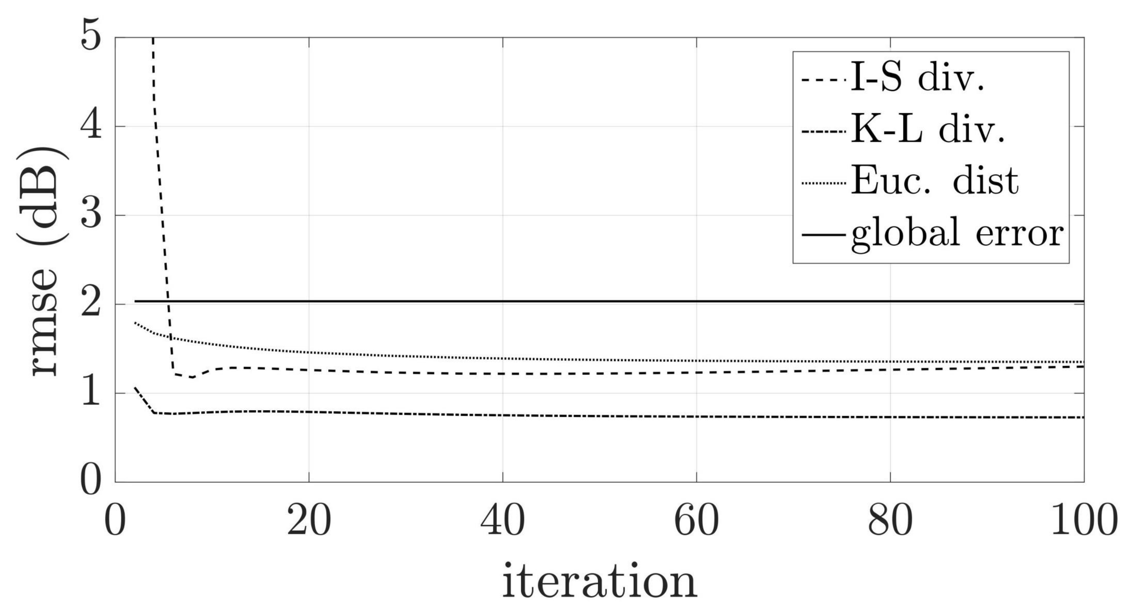


Fig 2: comparison of the RMSE for the 3 cost functions. We add the RMSE between L_{eq} and $\mathrm{L}_{\mathrm{eq,traffic}}$ without source separation

Conclusion and investigations

- First results for the road traffic source separation in an sound urban context within simple simulated sound scene
- Low RMSE values and improvement of the estimated road traffic sound levels show the interest of the NMF in that context
- The K-L divergence seems the most promising approach

- **Future works**:
  - Building a set of more complete and realistic sound scenes
  - Design an experimental plan in order to optimize the modeling parameters
  - Add temporal contraints on $\mathbf{H}$ to modelize more realistic comportements (smooth NMF)