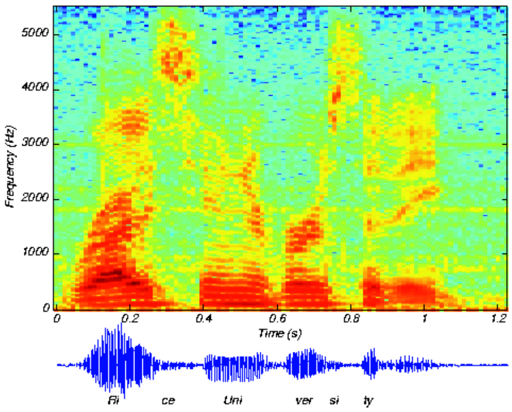
Propositions de sujets de Travail de Fin d'Etudes
Année académique 2014 - 2015

EPL-2014-044

"Débruitage de signaux vocaux à partir de la connaissance d'un bruit partiellement structuré" (en collaboration avec NXP Software, Leuven)
"Speech Signal Denoising from Partially-Structured Noise Knowledge" (collaboration with NXP Software, Leuven)

Jacques Laurent, Nilesh Madhu (NXP Software)



It is well known that speech recording can be decomposed with a few wave functions. These signals have indeed a *sparse* representation in some basis or dictionary (e.g., wavelet basis or Gabor atoms [1]). Conversely, noises (and, by extension, noisy speech) that display less structured aspects have reduced or absent sparsity in any bases.

Modern signal denoising methods leverage on this distinction in order to clean noisy signals. They obtain a clean signal estimation by promoting its sparsity in an appropriate basis while forcing it to be relatively close to the observed noisy signal (the distance between them being adjusted by the noise power estimate).

However, there exist situations where noise is not purely random. This occurs for instance when its definition includes any sources polluting the recording of the signal of interest. A typical situation is the recording of a speech signal from a mobile phone in a town environment; the noise is there constituted of structured parts (e.g., the sound of a truck engine) and of "random" parts (such as the noise produced by the wind on a microphone).

This project will investigate how such a structured noise can be removed from the recorded signal. In particular, it will be assumed that the time-frequency "energy" of the noise floor is known, such as computed by an average spectrogram over a large dataset of signals.

The model of interest will be the one sustained by the theory of *Morphological Component Analysis* (MCA) [2] or the "u+v" model of Y. Meyer. The main idea is to assume that the sensing signal model, *i.e.*, the one able to explain the noisy recording of the speech signal given the clean signal, is constituded by the following sum:

$$y = x_{\text{speech}} + x_{\text{struct.noise}} + n,$$

where x_{speech} is the (pure) speech signal, $x_{\text{struct.noise}}$ is the structured noise and n is the (remaining) unstructured noise.

The prior information about these sub-signals are that, (i) x_{speech} is *sparse* in some basis Ψ adapted to speech signal, (ii) $x_{\text{struct.noise}}$ has an estimated spectrogram (called "noise floor estimation"), and (iii) n is a remaining noise assumed Gaussian and with bounded power.

The reconstruction of x_{speech} (and also of $x_{\text{struct.noise}}$) will be intuitively achieved by promoting its sparse model under two (convex) constraints connected to the points (ii) and (iii) above.

Various algorithms from the literature will have to be considered (e.g., convex optimisation and greedy methods [1]) all these relying on a sparse description of the speech signal (e.g., in a wavelet-gabor dictionary [1,3]). This work is related to the field of sparse coding and all the codes will be developed in Matlab (inspired by the toolbox "A Numerical Tour of Signal Processing") with validation on data provided by NXP Software.

Work and Collaboration with NXP Software:

This project will be carried out in collaboration with NXP Software. The master student interested by this subject will work at NXP during several months (up to 50% of the master project duration) with daily travel and food costs covered. The whole project will be performed under the supervision of Prof. Laurent Jacques (UCL) and Nilesh Madhu (NXP Software). A NDA agreement will cover the common work.

Further Readings:

[1] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries", IEEE Transactions on Signal Processing, vol.41, n°12, December 1993, pp.3397-3415.

[2] J. Bobin, J.-L. Starck, M.J. Fadili, and Y. Moudden, "Sparsity, Morphological Diversity and Blind Source Separation", Vol 16, No 11, pp 2662-2674, 2007.

[3] Gribonval, Rémi, and Emmanuel Bacry. "Harmonic decomposition of audio signals with matching pursuit." *Signal Processing, IEEE Transactions on* 51.1 (2003): 101-111.

[4] P.-Y. Chen and I. W. Selesnick. "Translation-invariant shrinkage/thresholding of group sparse signals." Signal Processing. vol. 94, pp 476-489, January 2014.

[5] N. Parikh and S. Boyd, "Proximal Algorithms", Foundations and Trends in Optimization, 1(3):123-231, 2014.

Diplômes

ELEC, *MAP*

Nombre d'étudiants : 1

Contact persons: For more information, please contact:

Prof. **Laurent Jacques**
(office: a157, bât. Stévin, 1st floor)
homepage (laurent.jacques@uclouvain.be)

Retour page précédente - page d'introduction