Contributions in the ASCETE project

D. Fourer, R. Badeau and Q. Legros

15 octobre 2022

Plan

ANR ASCETE project

SUPPORTED BY ANR

Goals : Combining deterministic and stochastic approaches to extend the proposed techniques to more complex signals

- Deterministic models combined with machine learning (e.g. deep neural networks)
- Difficult cases for mode recovery (e.g. overlapping components, noisy signals, etc.)
- Combining synchrosqueezing with Non-negative Matrix Factorization (NMF)
- Generalization to high dimension signals (images, tensors, etc.)
- New practical applications (perception, biomedicine, astronomy, etc.)

Project objectives :

- **Objective 1 :** New approaches for the study of MCSs with synchrosqueezing transforms
- **Objective 2 :** Improving signal representations using data-driven and machine learning approaches
- Objective 3 : Combining non negative matrix factorization and SST, Phase retrieval
- **Objective 4 :** Applications and Software developments

Plan

Second-order time-reassigned synchrosqueezing

- Partners : Nantes, Paris
- Involved Tasks : T1.1 : combining stochastic and deterministic approaches to improve SST

Contributions

- A new second-order group-delay estimator for horizontal synchosqueezing [Fourer, Auger 2019],
- Application to the S-transform and continuous wavelet transform [Fourer, Auger 2020]
- First analysis of the Draupner Wave signal analysis

Motivation

- Computing enhanced TFR designed for impulsive signal and strongly modulated modes
- Generalization of the time-reassigned synchosqueezed STFT first proposed by [He et al, 2019]

Main Idea

Definitions

STFT :

$$F_x^h(t,\omega) = \int_{\mathbb{R}} x(\tau) h(t-\tau)^* \, \mathbf{e}^{-j\omega\tau} \, \mathrm{d}\tau \tag{1}$$

with $j^2 = -1$

Marginalization over time of $F_x^h(t,\omega)$ leads to :

$$\begin{aligned}
\int_{\mathbb{R}} F_x^h(t,\omega)\, dt &= \iint_{\mathbb{R}^2} h(t-\tau)^* x(\tau)\, \mathbf{e}^{-j\omega\tau}\, dt d\tau & (2)\\
&= \iint_{\mathbb{R}^2} h(u)^* x(\tau)\, \mathbf{e}^{-j\omega\tau}\, du d\tau & (3)\\
&= \int_{\mathbb{R}} h(u)^* du \int_{\mathbb{R}} x(\tau)\, \mathbf{e}^{-j\omega\tau}\, d\tau & (4)\\
&= F_h(0)^* F_x(\omega) & (5)
\end{aligned}$$

with $F_x(\omega) = \int_{\mathbb{R}} x(t)\, \mathbf{e}^{-j\omega t}\, dt$ the Fourier transform of signal x.

Time-reassigned synchrosqueezing

Horizontal synchrosqueezing

$$S_{\boldsymbol{x}}^{\boldsymbol{h}}(t,\omega) = \int_{\mathbb{R}} F_{\boldsymbol{x}}^{\boldsymbol{h}}(\tau,\omega)\delta\left(t - \hat{t}_{\boldsymbol{x}}^{(2)}(\tau,\omega)\right) d\tau \quad (6)$$

where $\hat{t}_{\boldsymbol{x}}(t,\omega)$ corresponds to the time reassignment operator (group-delay).
Reconstruction :

$$x(t) = \frac{1}{2\pi F_{\boldsymbol{h}}(0)^*} \iint_{\mathbb{R}^2} S_{\boldsymbol{x}}^{\boldsymbol{h}}(\tau,\omega)\, e^{j\omega t}\, d\tau d\omega. \quad (7)$$

Second-order group-delay estimator

$$\hat{t}_{\boldsymbol{x}}^{(2)}(t,\omega) = \begin{cases} \frac{\omega - \hat{\omega}_{\boldsymbol{x}}(t,\omega) + \mathrm{Im}(\hat{q}_{\boldsymbol{x}}(t,\omega)\, \tilde{t}_{\boldsymbol{x}}(t,\omega))}{\hat{\alpha}_{\boldsymbol{x}}(t,\omega)} & \text{if } \hat{\alpha}_{\boldsymbol{x}}(t,\omega) \neq 0 \\ \hat{t}_{\boldsymbol{x}}(t,\omega) & \text{otherwise} \end{cases} \quad (8)$$

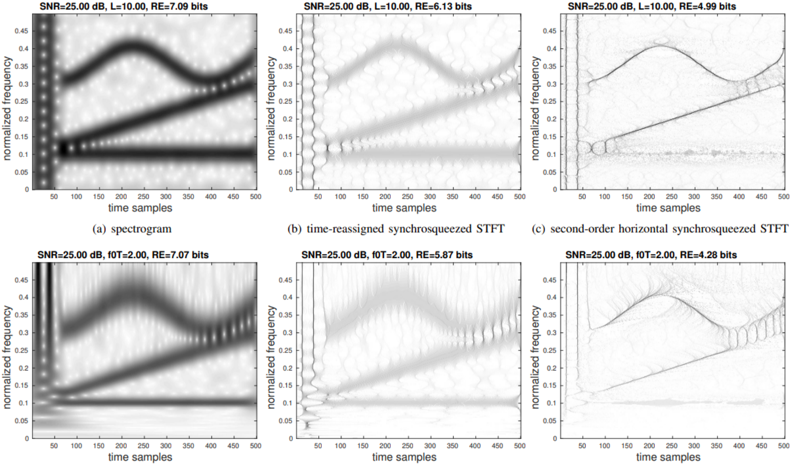
with :

$$\hat{t}_{\boldsymbol{x}}(t,\omega) = \mathrm{Re}\left(\tilde{t}_{\boldsymbol{x}}(t,\omega)\right), \text{with} \quad \tilde{t}_{\boldsymbol{x}}(t,\omega) = t - \frac{F_{\boldsymbol{x}}^{\mathcal{T}\boldsymbol{h}}(t,\omega)}{F_{\boldsymbol{x}}^{\boldsymbol{h}}(t,\omega)} \quad (9)$$

$$\hat{\omega}_{\boldsymbol{x}}(t,\omega) = \mathrm{Im}\left(\tilde{\omega}_{\boldsymbol{x}}(t,\omega)\right), \text{with} \quad \tilde{\omega}_{\boldsymbol{x}}(t,\omega) = j\omega + \frac{F_{\boldsymbol{x}}^{\mathcal{D}\boldsymbol{h}}(t,\omega)}{F_{\boldsymbol{x}}^{\boldsymbol{h}}(t,\omega)} \quad (10)$$

$$\hat{\alpha}_{\boldsymbol{x}}(t,\omega) = \mathrm{Im}(\hat{q}_{\boldsymbol{x}}), \text{with} \quad \hat{q}_{\boldsymbol{x}}(t,\omega) = \frac{F_{\boldsymbol{x}}^{\mathcal{D}^{n}\boldsymbol{h}} F_{\boldsymbol{x}}^{\boldsymbol{h}} - F_{\boldsymbol{x}}^{\mathcal{D}^{n-1}\boldsymbol{h}} F_{\boldsymbol{x}}^{\mathcal{D}\boldsymbol{h}}}{F_{\boldsymbol{x}}^{\mathcal{T}\boldsymbol{h}} F_{\boldsymbol{x}}^{\mathcal{D}^{n-1}\boldsymbol{h}} - F_{\boldsymbol{x}}^{\mathcal{T}\mathcal{D}^{n-1}\boldsymbol{h}} F_{\boldsymbol{x}}^{\boldsymbol{h}}} \quad (11)$$

Numerical results (comparison)

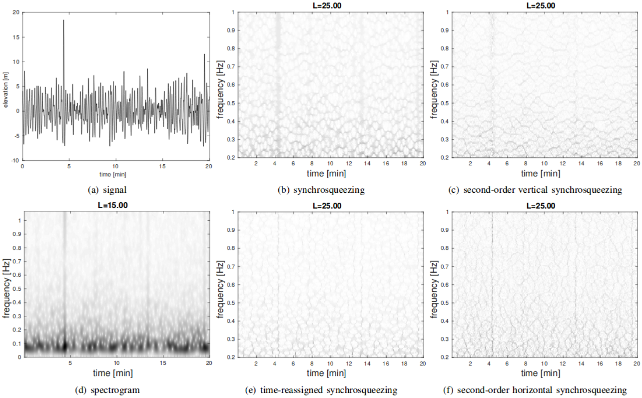


(a) spectrogram (b) time-reassigned synchrosqueezed STFT (c) second-order horizontal synchrosqueezed STFT

(d) stockwellogram (e) time-reassigned synchrosqueezed S-transform (f) second-order horizontal synchrosqueezed S-transform

Fig. 1. Comparison of the resulting TFRs with Rényi Entroy (at order $\alpha = 3$) of a synthetic multicomponent signal. The TFRs of the synchrosqueezing methods (b),(c), (e) and (f) correpond to their squared modulus.

Numerical results on Draupner signal and codes



(a) signal (b) synchrosqueezing (c) second-order vertical synchrosqueezing (d) spectrogram (e) time-reassigned synchrosqueezing (f) second-order horizontal synchrosqueezing

Fig. 2. Waveform (a) and TFRs of the Draupner wave signal. spectrogram (d), synchrosqueezing (b), second-order vertical synchrosqueezing (c), time-reassigned synchrosqueezing (e) and second-order horizontal synchrosqueezing (f).

Matlab Codes

```
https://fourer.fr/hsst
https://fourer.fr/sthsst
```

Possible extensions

- Robust to noise estimators using regularization
- High-order group-delay estimators (Pham, Meignen et al.)
- Theoretical analysis of time-reassigned synchrosqueezing wavelet transform [Li, Zhang, Auger et al. 2022]
- Self-matched extracting wavelet transform and signal reconstruction [Li, Auger, et al. 2022]
- etc.

Pseudo-Bayesian approach for mode extraction

- Partners : Paris
- Involved Tasks : T1.1 + T1.2 : New ridge extraction technique

Contributions

- A new noise-robust mode extraction method based on the STFT [Legros, Fourer, 2021],
- Extension using alpha-beta divergence and a new signal detector (submitted [Legros, Fourer, 2022])
- Application on both synthesized and real-world signals

Motivation

- Robust mode extraction method usable to any time-frequency representation
- Robust Instantaneous Frequency estimator and Ridge detection method

Main Idea

- Given a discrete-time spectrogram $s_{n,m}$ (n the time index and m the frequency bin), model each spectrogram slice as :

$$s_{n,m}|\bar{m}_n \sim g\left(m - \bar{m}_n\right), \quad \text{with} \quad g(m) = \frac{2\sqrt{\pi}L}{M}\,\mathbf{e}^{-\left(\frac{2\pi mL}{M}\right)^2} \tag{12}$$

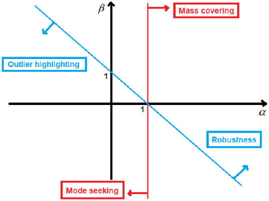
- Compute the joint likelihood (assuming independence between successive frames)

$$p(\mathbf{s}_n|\bar{m}_n) = \prod_{m=0}^{M-1} p(s_{n,m}|\bar{m}_n). \tag{13}$$

- Instead of maximizing posterior given by $p(\bar{m}_n|\mathbf{s}_n) = \frac{p(\mathbf{s}_n|\bar{m}_n)p(\bar{m}_n)}{p(\mathbf{s}_n)}$ which is equivalent to minimize the KL divergence between $p(\mathbf{s}_n)$ and $p(\mathbf{s}_n|\bar{m}_n)$, we replace the KL-divergence by a robust $\alpha - \beta$-divergence leading to the following pseudo-posterior :
- Use the pseudo-posterior expressed in terms of alpha-beta cross-entropy :

$$p(\bar{m}_n|\mathbf{s}_n) \propto e^{-M\mathrm{CE}^{\alpha,\beta}_{\mathbf{AB}}(\bar{m}_n)}p(\bar{m}_n). \tag{14}$$

Main Idea



Algorithm 1

1: Input: TFR S_0, GRW mean m_0 and variance σ_0, K, g.
2: **for** $k = 1, \ldots, K$ **do**
3: **for** $n = 0, \ldots, N-1$ **do**
4: Compute $p(m_n)$ by matching moments from Eq. (18).
5: Compute the pseudo-posterior $p(m_n|s_n)$ from Eq. (16).
6: Perform MMSE estimation of $\hat{m}_n$.
7: **end for**
8: Repeat steps 4 to 6 iterating from $n = N-1, \ldots, 0$
9: Update the TFR by subtracting the kth ridge (TFR support set to 0).
10: **end for**

Numerical results

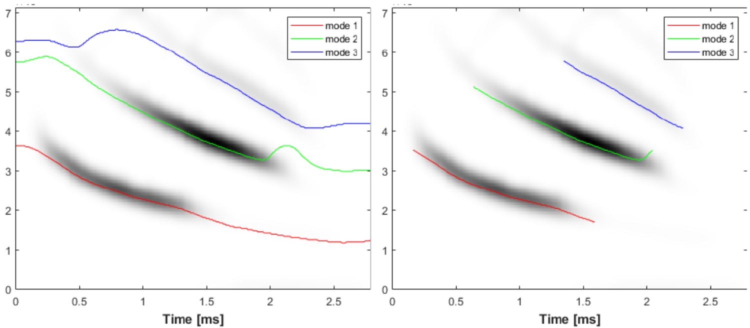


Fig. 13. Estimation of $K = 3$ signal components of the real bat record signal using the proposed ABD method with $\alpha = 0.4, \beta = 0.7$ with (right) and without (left) performing detection.

Numerical results and codes

Table 1. RQF of each reconstructed components (averaged over 100 realizations of noise) for the different competing approaches for a SNR = 10dB. In second rows are displayed, for each case, the std of the estimators.

	C1	C2	C3	Average
Brevdo [5]	15.86 ±0.84	16.60 ±0.84	6.52 ±2.12	12.70 ±1.4
ABD, $\alpha = 0.4, \beta = 0.4$	**17.12** ±1.81	16.63 ±0.81	11.04 ±0.81	**14.31** ±0.63
ABD, $\alpha = 0.2, \beta = 0.4$	17.07 ±0.79	16.64 ±0.78	10.89 ±0.62	14.26 ±0.74
ABD, $\alpha = 0.4, \beta = 0.2$	16.93 ±0.89	**16.52** ±0.76	**10.90** ±0.57	14.28 ±0.75
ABD, $\alpha = 0.2, \beta = 1.2$	14.96 ±6.74	16.78 ±0.80	9.33 ±0.64	13.35 ±3.93
ABD, $\alpha = 0.7, \beta = 1.2$	11.62 ±4.84	16.37 ±0.72	9.51 ±0.33	12.20 ±8.03
ABD, $\alpha = 0.2, \beta = 1.5$	5.52 ±8.90	16.28 ±2.43	8.64 ±0.24	9.85 ±2.83
RD [24]	16.22 ±7.26	12.28 ±7.14	5.05 ±7.92	11.18 ±7.45

Matlab Codes

https://fourer.fr/eusipco21
https://codeocean.com/capsule/8693890/tree/v1

EM approach for mode extraction and IF estimation

- Partners : Paris, Grenoble
- Involved Tasks : T1.1 + T1.2 : New ridge extraction technique

Contributions

- A new noise-robust mode extraction method based on the STFT [Legros, Fourer, 2021]
- Application on synthesized signals

Motivation

- Noise-robust mode extraction method
- Deal with overlapping components

Main Idea

- Given a spectrogram $s_{n,m}$, a ssume a mixture model :

$$p(s_{n,m}|\boldsymbol{w}_n, \hat{\boldsymbol{m}}_n) = \sum_{k=1}^{K} w_n^k g(m - \hat{m}_n^k) + \frac{1}{M}\left(1 - \sum_{k=1}^{K} w_n^k\right), \tag{15}$$

- Compute the joint-likelihood :

$$p(\boldsymbol{S}|\boldsymbol{W}, \hat{\boldsymbol{\mathcal{M}}}) = \prod_n \prod_m p(s_{n,m}|\boldsymbol{w}_n, \hat{\boldsymbol{m}}_n). \tag{16}$$

- Consider two distinct prior :
  - Total Variation :

$$p(\hat{\boldsymbol{\mathcal{M}}}|\epsilon) \propto \exp\left[-\epsilon \sum_{k=1}^{K} \|\hat{\boldsymbol{m}}_{k,:}^{\top}\|_{TV}\right], \tag{17}$$

  - Laplacian :

$$p(\hat{\boldsymbol{\mathcal{M}}}|\lambda) \propto \exp\left[-\frac{\lambda}{2} \sum_{k=1}^{K} \|\boldsymbol{L}_a \hat{\boldsymbol{m}}_{k,:}^{\top}\|_2^2\right], \tag{18}$$

Main Idea

- Approximate joint posterior as :

$$p(\boldsymbol{W}, \hat{\boldsymbol{\mathcal{M}}}|\boldsymbol{S}) \propto p(\boldsymbol{S}|\boldsymbol{W}, \hat{\boldsymbol{\mathcal{M}}})p(\hat{\boldsymbol{\mathcal{M}}})p(\boldsymbol{W}). \tag{19}$$

- (modified) Expectation-Maximization at iteration (i) :

$$\begin{aligned} \widehat{Q}(\boldsymbol{W}|\boldsymbol{W}^{(i)}) &= \log p(\boldsymbol{S}|\boldsymbol{W}, \widetilde{\boldsymbol{\mathcal{M}}}) + \log \left[p(\hat{\boldsymbol{\mathcal{M}}})p(\boldsymbol{W})\right], \\ \boldsymbol{W}^{(i+1)} &= \underset{\boldsymbol{W}}{\operatorname{argmax}} \;\; \widehat{Q}(\boldsymbol{W}|\boldsymbol{W}^{(i)}). \end{aligned} \tag{20}$$

Numerical results

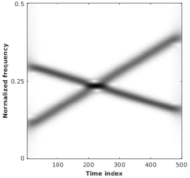
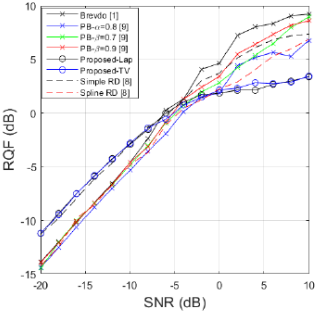
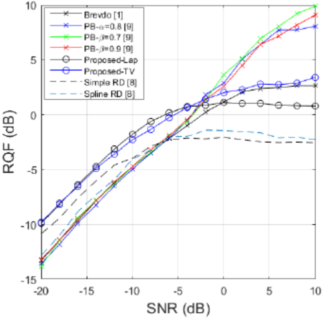


Figure 1: Spectrogram of the analyzed multicomponent signals with overlapping components.





Finite Rate of Innovation (FRI) approach for mode extraction

- Partners : Paris
- Involved Tasks : T1.2 : New ridge extraction technique

Contributions

- FRI-based approach for mode extraction and IF estimation (submitted to SPL [Legros, Fourer, 2022])
- Possibly combined with synchrosqueezing
- Application on both synthesized and real-world signals

Motivation

- Robust IF estimation and Ridge detection method

Main Idea

- Observation model :

$$s_{n,m} \approx \sum_{k=0}^{K-1} a_k(n) g(m - \phi'_k(n)) \tag{21}$$

where $a_k(n) = \alpha_k^2(n)$ and $g(m) = \mathbf{e}^{-\left(\frac{2\pi mL}{M}\right)^2}$

- Assumes the signal to retrieve as a stream of Dirac Pulses :

$$f_n(m) = \sum_{k=0}^{K-1} a_k(n) \delta(m - \phi'_k(n)) \quad . \tag{22}$$

- Estimation of $\hat{f}_n = \boldsymbol{D}_g^{-1} \boldsymbol{V}^{-1} \boldsymbol{s}_n$ where $[\boldsymbol{V}]_{m,\lambda} = e^{j2\pi(\frac{m\lambda}{M})}$ is a $(M \times 2M_0 + 1)$ matrix and $\boldsymbol{D}_g$ is a diagonal matrix gathering the discrete time Fourier series coefficients of g.

- Estimation of the mode location by estimating the annihilating filter h such as :

$$
\begin{aligned}
(\hat{f}_n * \boldsymbol{h})(l) &= \sum_{i \in \mathbb{Z}} h(i) \hat{f}_n(l-i) = 0 \\
&= \sum_{k=0}^{K-1} a_k(n)\, \mathbf{e}^{\frac{-j2\pi l \phi'_k(n)}{M}} \underbrace{\sum_{i \in \mathbb{Z}} h(i)\, \mathbf{e}^{\frac{j2\pi i \phi'_k(n)}{M}}}_{H\left(e^{\frac{-j2\pi \phi'_k(n)}{M}}\right)} = 0
\end{aligned}
\tag{23}
$$

  with $H(z)$ the Z-transform of $\boldsymbol{h}$, whose roots are $e^{\frac{-j2\pi\phi'_k(n)}{M}}$.
- In presence of noise, h minimizes $||\hat{f}_n * \boldsymbol{h}||^2$ (Total Least Squares method)
- The FRI-TLS method can be combined with vertical synchrosqueezing (FRI-SST)

Numerical results

(b) SNR=0dB

	C1	C2	C3	Average
Brevdo [10]	4.06 ±0.82	4.34 ±0.94	−0.22 ±0.51	2.71 ±0.76
PB, $\beta = 0.7$ [9]	3.45 ±5.54	3.60 ±3.77	1.17 ±1.84	2.74 ±3.72
PB, $\alpha = 0.5$ [9]	3.66 ±3.78	4.18 ±1.36	0.59 ±1.87	2.81 ±2.33
RD [8]	9.00 ±1.72	8.21 ±1.04	6.56 ±1.41	7.92 ±1.39
FRI	3.71 ±1.47	1.43 ±1.09	0.13 ±0.76	1.76 ±1.11
FRI TLS (proposed)	8.81 ±1.08	**8.28** ±2.37	**7.37** ±0.98	**8.15** ±1.48
FRI SST (proposed)	**9.16** ±2.65	7.53 ±2.09	6.03 ±0.87	7.57 ±1.87

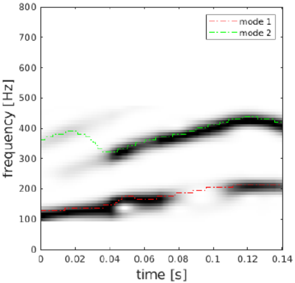


Fig. 2. Estimation of the first $K = 2$ signal components of the speech signal using the proposed TLS method.

Malab Codes

https://codeocean.com/capsule/7022037/tree/v1

Harmonic/Percussive Source Separation

- Partners : Paris
- Involved Tasks : T1.2 : New ridge extraction technique, T4.2 : application of SST to audio signals

Contributions

- Computes AM-FM estimator in the TF plane for separating harmonic / percussive sources [Fourer, 22]
- Combines Machine-learning with time-frequency analysis
- Application to music audio signals

Motivation

- Disentangling the harmonic deterministic part from the percussive and noisy part
- Enhancing audio signal for MIR applications

Main Idea

- Assumes an instantaneous mixture

$$x(t) = s_h(t) + s_p(t) \tag{24}$$

with

$$x(t) = \mathbf{e}^{\lambda_x(t)+j\phi_x(t)}, \quad \text{with } j^2{=}{-1}, \tag{25}$$

- Uses previously proposed estimators in [Fourer, Auger et al. 2018]

$$\hat{q}_x^{(tn)}(t,\omega) = \frac{F_x^{\mathcal{D}^n h} F_x^h - F_x^{\mathcal{D}^{n-1} h} F_x^{\mathcal{D} h}}{F_x^{\mathcal{T} h} F_x^{\mathcal{D}^{n-1} h} - F_x^{\mathcal{T}\mathcal{D}^{n-1} h} F_x^h} \tag{26}$$

$$\hat{\Psi}_x^{(tn)}(t,\omega) = \frac{F_x^{\mathcal{D} h} F_x^{\mathcal{T}\mathcal{D}^{n-1} h} - F_x^{\mathcal{T} h} F_x^{\mathcal{D}^n h}}{F_x^{\mathcal{T}\mathcal{D}^{n-1} h} F_x^{h}(t,\omega) - F_x^{\mathcal{T} h} F_x^{\mathcal{D}^{n-1} h}} + j\omega \tag{27}$$

-

$$\hat{\nu}_x(t,\omega) = \mathrm{Re}\Big(\hat{q}_x(t,\omega)\Big), \qquad \hat{\alpha}_x(t,\omega){=}\mathrm{Im}\Big(\hat{q}_x(t,\omega)\Big) \tag{28}$$

$$\hat{\dot{\lambda}}_x(t,\omega) = \mathrm{Re}\left(\hat{\Psi}_x(t,\omega)\right), \quad \hat{\dot{\phi}}_x(t,\omega) = \mathrm{Im}\left(\hat{\Psi}_x(t,\omega)\right) \tag{29}$$

- AM-FM Parameters (after discretization) :
  - AM : $\hat{\dot{\lambda}}_x[k,m]$
  - FM : $\hat{\alpha}_x[k,m]$
  - AM-FM : $C_x[k,m] = \sqrt{\hat{\dot{\lambda}}_x[k,m]^2 + \hat{\alpha}_x[k,m]^2}$

Main Idea

Training based on Linear Discriminant Analysis

- Reference Separation Mask using ground truth :

$$M_{\boldsymbol{h}}^{(\boldsymbol{true})}[k,m] = \begin{cases} 1 & \text{if } |F_{\boldsymbol{s_h}}^{\boldsymbol{h}}[k,m]|^2 > |F_{\boldsymbol{s_p}}^{\boldsymbol{h}}[k,m]|^2 \\ 0 & \text{otherwise} \end{cases}, \tag{30}$$

$$M_{\boldsymbol{p}}^{(\boldsymbol{true})}[k,m] = 1 - M_{\boldsymbol{h}}^{(\boldsymbol{true})}[k,m] \tag{31}$$

- Compute the centroid of each source (*i.e.* $\mu_{\boldsymbol{h}}$ ou $\mu_{\boldsymbol{p}}$) in the discriminant space.

$\Rightarrow$ The trained model corresponds to eigenvectors and $\mu_{\boldsymbol{h}}$ et $\mu_{\boldsymbol{p}}$.

Separation

- For each TF point, Compute the descriptors $\mathcal{Q}_{\boldsymbol{x}}[k,m]$
- Compute the linear projection $P_{\mathcal{Q}}$
- Separation masks :

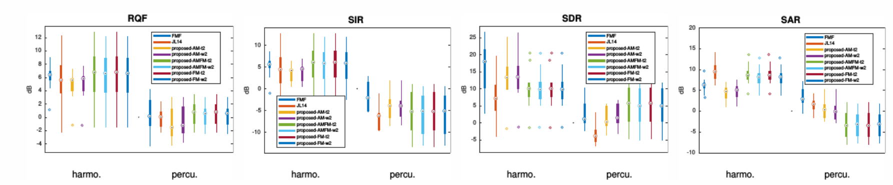
$$M_{\boldsymbol{h}}[k,m] = \begin{cases} 1 & \text{if } ||P_{\mathcal{Q}}[k,m] - \mu_{\boldsymbol{h}}|| < ||P_{\mathcal{Q}}[k,m] - \mu_{\boldsymbol{p}}|| \\ 0 & \text{otherwise} \end{cases}, \quad M_{\boldsymbol{p}}[k,m] = 1 - M_{\boldsymbol{h}}[k,m]. \tag{32}$$

- Reconstruction :

$$\hat{s}_{\boldsymbol{h}} = \text{TFCT}^{-1}(F_{\boldsymbol{x}}^{\boldsymbol{h}}[k,m]M_{\boldsymbol{h}}[k,m]) \tag{33}$$

$$\hat{s}_{\boldsymbol{p}} = \text{TFCT}^{-1}(F_{\boldsymbol{x}}^{\boldsymbol{h}}[k,m]M_{\boldsymbol{p}}[k,m]) \tag{34}$$

Numerical results



(a) (b) (c) (d)

Audio results

`https://fourer.fr/publi/gretsi22/`

Biblio : Objective 1 : New approaches for the study of MCSs

- Fourer, D., & Auger, F. (2019, September). Second-order time-reassigned synchrosqueezing transform : Application to Draupner wave analysis. In 2019 27th European Signal Processing Conference (EUSIPCO) (pp. 1-5). IEEE.
- Fourer, D., & Auger, F. (2019, August). Synchrosqueezing temporel d'ordre 2 de la transformée de Fourier à court terme. In XXVIIeme Colloque francophone de Traitement du Signal et des Images (GRETSI 2019).
- Fourer, D., & Auger, F. (2021, January). Second-Order Horizontal Synchrosqueezing of the S-transform : a Specific Wavelet Case Study. In 2020 28th European Signal Processing Conference (EUSIPCO) (pp. 2200-2204). IEEE.
- Legros, Q., & Fourer, D. (2021, August). A Novel Pseudo-Bayesian Approach for Robust Multi-Ridge Detection and Mode Retrieval. In 2021 29th European Signal Processing Conference (EUSIPCO) (pp. 1925-1929). IEEE.
- Legros, Q., Fourer, D., Meignen, S., & Colominas, M. A. (2022). Instantaneous Frequency Estimation In Multi-Component Signals Using Stochastic EM Algorithm. arXiv preprint arXiv :2203.16334.
- Fourer, D. (2022, September). Séparation de sources harmoniques/percussives utilisant des estimateurs locaux de modulation linéaire AM-FM. In GRETSI 2022.

To appear :

- Legros, Q., & Fourer, Time-Frequency Ridge Estimation of Multi-Component Signals using Sparse Modeling of Signal Innovation.
- Legros, Q., & Fourer, Pseudo-Bayesian Approach for Robust Mode Detection and Extraction based on the STFT

Plan

HVS detection from EEG signals using recursive synchrosqueezing transform

- Partners : Paris, Tsing Hua University (Taiwan)
- Involved Tasks : T1.3 and T4.1 : multivariate SST and application to the study of ECG and EEG signals

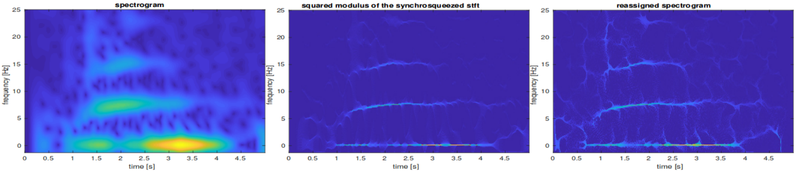
Contributions

- A new HVS method designed to ECG signals
- Combines recursive synchrosqueezing with a detector
- Application on real-world signals

Motivation

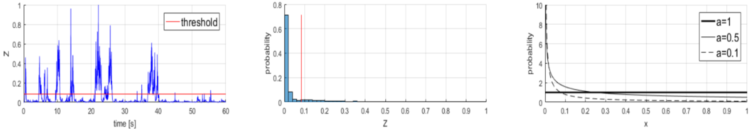
- Fast Prediction of Parkinson high voltage spindles (HVS)

Numerical results



(a) spectrogram (b) synchrosqueezed STFT (c) reassigned spectrogram

FIGURE 1 – Comparisons of the different recursive TFRs computed for an EEG signal with a HVS.



(a) saliency function (b) histogram (c) beta law distribution

FIGURE 2 – Resulting saliency function (a) and its histogram (b) computed from the recursive synchrosqueezed STFT of an EEG signal. Its probability density function can be compared to a beta distribution (c) with parameter $a \in \{0.1, 0.5, 1\}$.

Application of the recursive and fft-based reassignment method to radar signal analysis

- Partners : Paris, Warsaw Univ. (Poland)
- Involved Tasks : T4.3 : software development

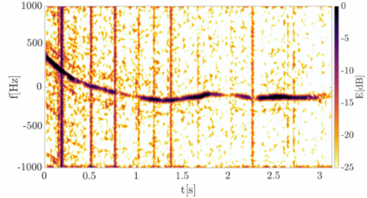
Contributions

- A complexity comparison between classical and recursive reassignment
- Application to radar signals

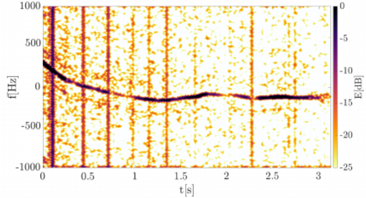
Motivation

- Future application to radar signal processing

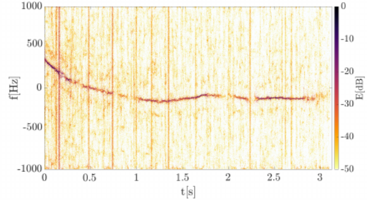
Numerical results



(c) Recursive spectrogram of the echo originating from a drone

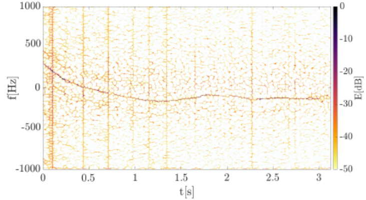


(c) FFT-based spectrogram of the echo originating from a drone



(d) Recursive reassigned spectrogram of the echo originating from a drone

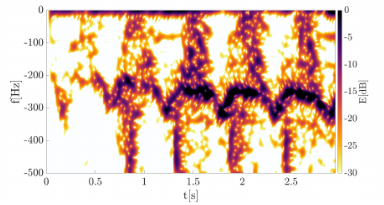
Fig. 2: Results for the recursive method.



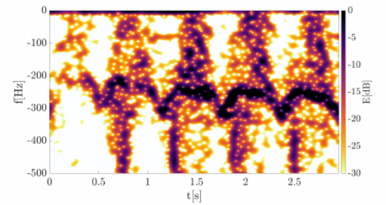
(d) FFT-based reassigned spectrogram of the echo originating from a drone

Fig. 3: Results for the classical FFT-based method.

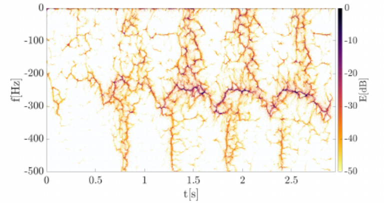
Numerical results



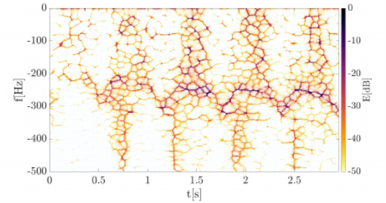
(a) Recursive spectrogram of the echo originating from a walking human



(a) FFT-based spectrogram of the echo originating from a walking human



(b) Recursive reassigned spectrogram of the echo originating from a walking human



(b) FFT-based reassigned spectrogram of the echo originating from a walking human

Smart Beekeeping based on TF analysis and deep learning

- Partners : Paris
- Involved Tasks : T4.2 : application of SST to audio signals

Contributions

- Combine TF analysis and deep neural networks for audio classification
- Application to beehive signal analysis for bee queen detection [Orlowska, Fourer, 21]
- Analysis of piping signals [Fourer, Orlowska, 22]

Motivation

- Predicting the health state of a beehive (smart Beekeeping)

Numerical results : piping signals analysis

Table 4: Experiment 3: Simultaneously Detection and classification classification comparative results.

Method	Feat. dimension	Label	Recall	Precision	F - score	Accuracy
TTB+SVM	164	Tooting	0.88	0.78	0.83	0.82
		Quacking	0.03	0.12	0.05	
		Non-piping	0.99	0.89	0.94	
1D-CNN	11,025	Tooting	0.93	0.84	0.88	0.85
		Quacking	0.10	0.54	0.16	
		Non-piping	0.99	0.86	0.92	
MFCC+CNN	17×47	Tooting	0.88	0.81	0.84	0.84
		Quacking	0.18	0.45	0.26	
		Non-piping	**0.99**	**0.90**	**0.95**	
STFT+CNN	512×42	Tooting	**0.94**	**0.97**	**0.95**	**0.91**
		Quacking	**0.50**	**0.76**	**0.60**	
		Non-piping	**0.99**	0.89	0.94	

Numerical results

Table 4 Comparison of the classification results in Experiment 2 (4-fold hive-independent cross-validation).

Method	Features	Label	Precision	Recall	F - score	Accuracy
MFCCs+CNN [11]	20x44	Queen	0.36	0.44	0.40	0.31
		No queen	0.22	0.16	0.19	
STFT+CNN	513×44	Queen	0.77	0.76	0.66	0.55
		No queen	0.33	0.20	0.33	
CQT+CNN	513×44	Queen	0.10	0.07	0.08	0.25
		No queen	0.32	0.41	0.36	
mean-CQT+CNN	27x44	Queen	0.25	0.11	0.16	0.38
		No queen	0.41	0.65	0.50	
mean-STFT+CNN	27x44	Queen	0.71	0.86	0.78	0.75
		No queen	0.81	0.64	0.71	
mean-STFT+CNN+DA	27×44	Queen	**0.96**	**0.99**	**0.96**	**0.96**
		No queen	**0.99**	**0.94**	**0.96**	

Python codes

https://github.com/agniorlowska/beequeen_prediction
https://fourer.fr/dcase22

Speech Emotion Recognition using TF Analysis

- Partners : Paris
- Involved Tasks : T4.2 : application of SST to audio signals

Contributions

- Combine TF analysis and deep neural networks for audio classification
- Efficient data augmentation technique for audio classification
- Investigation of several deep CNN architectures originally designed for image classification
- Application to speech signal for emotion recognition [Xia, Fourer, 21]

Motivation

- Speech Emotion Recognition

Main Idea

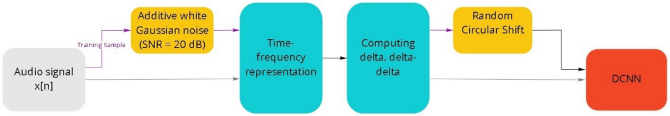
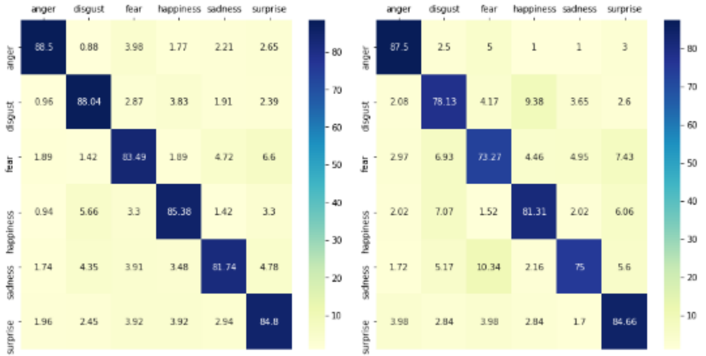


Figure 1: *Overview of the proposed approach.*

Python codes

https://github.com/llnanis/SER-RCS

Numerical results

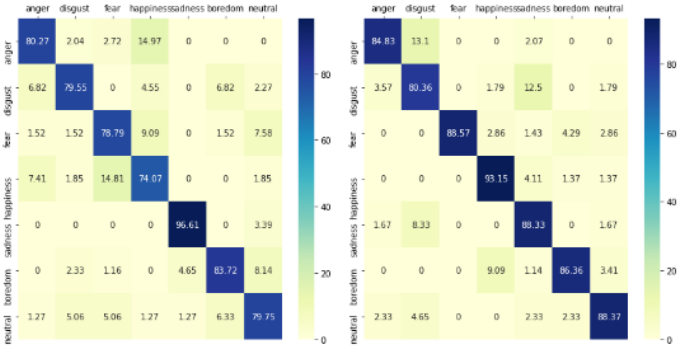


(a) proposed, STFT-Alex+RCS41 (Acc. 85.33%)

(b) DCNN-DTPM [12] (Acc. 79.25%)

Figure 4: *eNTERFACE05 confusion matrices obtained using our proposed method STFT-Alexnet + RCS41 (a) and DCNN-DTPM [12].*

Numerical results



(a) proposed, STFT-Alex+RCS19 (Acc. 81.82%)

(b) DCNN-DTPM [12] (Acc. 87.31%)

Figure 5: *EMO-DB confusion matrices obtained using our proposed method STFT-Alexnet + RCS19 (a) and DCNN-DTPM [12].*

Objective 4 : Applications and Software developments

- Souriau, R., Fourer, D., Chen, H., Lerbet, J., Maaref, H., & Vigneron, V. (2019, August). High-Voltage Spindles detection from EEG signals using recursive synchrosqueezing transform. In GRETSI.
- Abratkiewicz, K., Samczynski, P., & Fourer, D. (2020, September). A Comparison of the Recursive and FFT-based Reassignment Methods in micro-Doppler Analysis. In 2020 IEEE Radar Conference (RadarConf20) (pp. 1-6). IEEE.
- Orlowska, A., Fourer, D., Gavini, J. P., & Cassou-Ribehart, D. (2022). Honey Bee Queen Presence Detection from Audio Field Recordings using Summarized Spectrogram and Convolutional Neural Networks. In International Conference on Intelligent Systems Design and Applications (pp. 83-92). Springer, Cham.
- Xia, S., Fourer, D., Audin, L., Rouas, J. L., & Shochi, T. (2022, May). Speech Emotion Recognition using Time-frequency Random Circular Shift and Deep Neural Networks. In Speech Prosody 2022.

To appear :

- D. Fourer and A. Orlowska, Detection and Identification of Beehive Piping Audio Signals. Proc. DCASE 2022. Nancy, France. **(accepted for publication)**

Plan

On the use of concentrated TFR with DNN

- Partners : Paris, Nantes
- Involved Tasks : T4.3 and T2.2 : development of original DNN-based signal processing tools

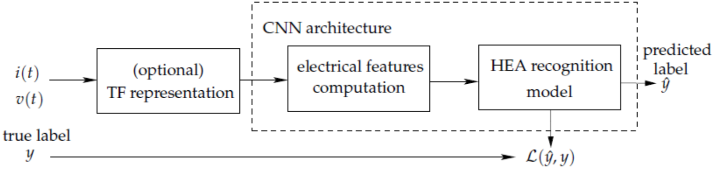
Contributions

- First study on the relevance of concentrated TFR applied to DNN-based signal classification
- Combination of 2D-CNN architecture with reassigned spectrogram and synchrosqueezed STFT
- Application to non-intrusive load monitoring (NILM)

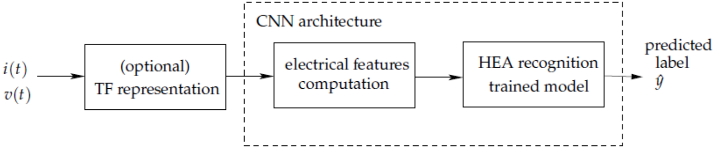
Motivation

- Improving the accuracy of existing NILM methods
- Investigating the relevance of concentrated TFR for signal classification

Main Idea



(a) training step



(b) testing step

Numerical results

Table 1. Comparative results (in percentage) of the different HEA recognition methods applied to the PLAID dataset. The window parameter L is empirically chosen to provide the best results.

	Acc	F_M	Rec	Pre
P, Q + Random Forest [15,19]	97.8	97.7	97.6	97.9
STFT (L = 60, single-input CNN)	87.1	87.2	87.3	88.4
STFT (L = 600, CNN with two channels)	97.7	97.5	97.5	97.9
STFT (L = 600, CNN concatenated)	95.6	95.7	95.5	96.1
Synchrosqueezing (L = 600, single-input CNN)	91.9	92.1	92.4	93.1
Synchrosqueezing (L = 60, CNN with two channels)	85.4	85.0	85.4	86.1
Synchrosqueezing (L = 60, CNN concatenated)	87.2	87.3	87.4	87.9
Time-reassigned synchrosqueezing (L = 60, single-input CNN)	85.8	86.1	86.4	85.9
Time-reassigned synchrosqueezing (L = 60, CNN with two channels)	91.4	91.2	90.9	92.1
Time-reassigned synchrosqueezing (L = 60, CNN concatenated)	92.3	92.3	92.4	91.9
Reassigned spectrogram (L = 600, single-input CNN)	74.4	75.0	74.1	77.3

DNN-based reassignment

- Partners : Paris
- Involved Tasks : T2.2 : development of original DNN-based signal processing tools

Contributions

- Improving the readability of a TFR
- Comparison of classical and DNN-based reassignment methods

Motivation

- Improving the robustness to noise with a data-driven approach
- Reassignment can be viewed as an image post-processing operation

Main Idea

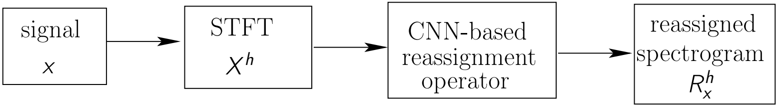


Figure : Proposed DNN-based reassignment method

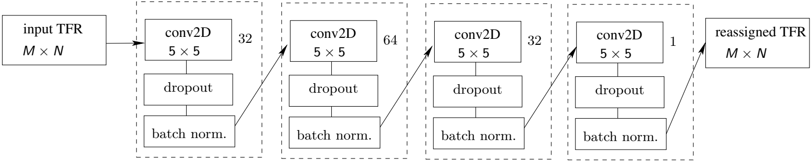


Figure : Proposed architecture based on 2D CNN.

- Uses 2D convolutional neurons with a 5×5 kernel
- Activation function : REctified Linear Unit (RELU)
- Dropout : Randomly discard 10% of the computed coefficients
- Optimizer : RMSProp [1]

1. Bengio, Yoshua. "Rmsprop and equilibrated adaptive learning rates for nonconvex optimi-

Unitary test on a sinusoidal signal 1/2

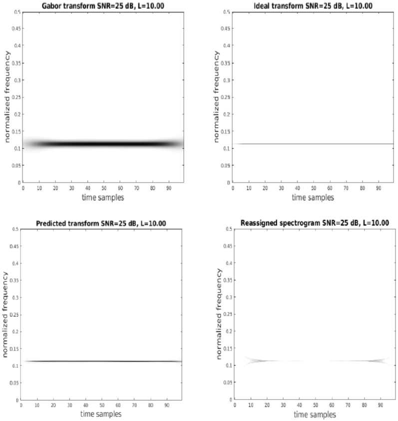


Figure : Comparison between $|X^h|$, ITFR, DNN1 estimation and classical reassigned spectrogram.

Unitary test on a sinusoidal signal 2/2

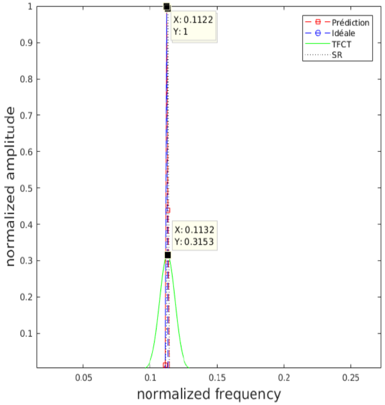


Figure : Comparison between $|X^h|$, ITFR, DNN1 estimation and classical reassigned spectrogram at a given time instant $n = 50$.

Unitary test on an impulse signal 1/2

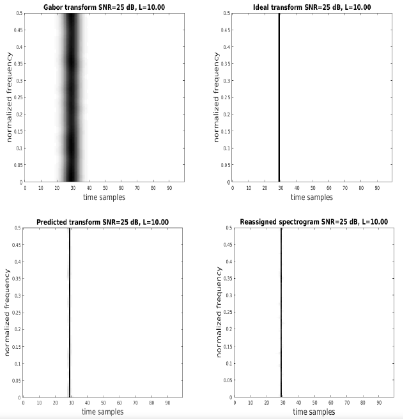


Figure : Comparison between $|X^h|$, ITFR, DNN2 estimation and classical reassigned spectrogram.

Unitary test on an impulse signal 2/2

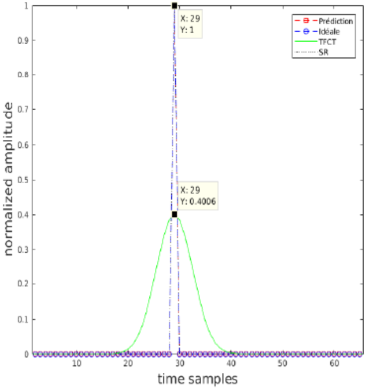


Figure : Comparison between $|X^h|$, ITFR, DNN2 estimation and classical reassigned spectrogram at a given normalized frequency $\lambda = 0.3$.

Noisy signal (SNR=5dB)

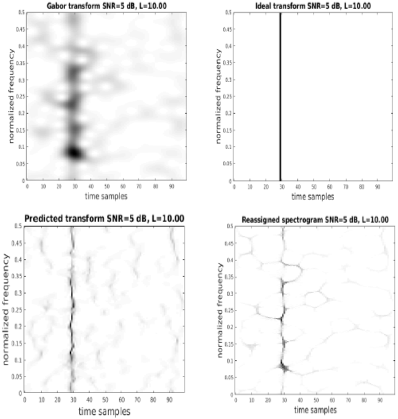


Figure : Comparison between $|X^h|$, ITFR, DNN2 estimation and classical reassigned spectrogram.

Noisy signal (SNR=5dB)

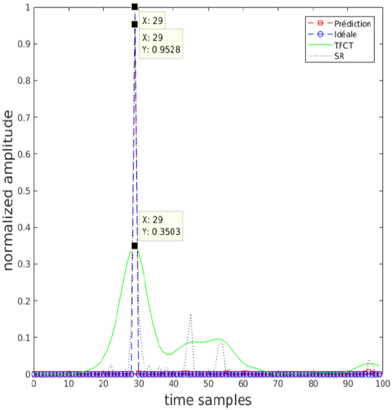


Figure : Comparison between $|X^h|$, ITFR, DNN2 estimation and classical reassigned spectrogram at a given normalized frequency $\lambda = 0.3$.

Multi-component noisy signal 1/2

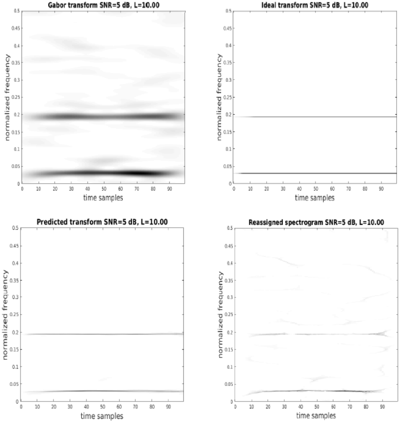


Figure : Comparison between $|X^h|$, ITFR, DNN2 estimation and classical reassigned spectrogram.

Multi-component noisy signal 2/2

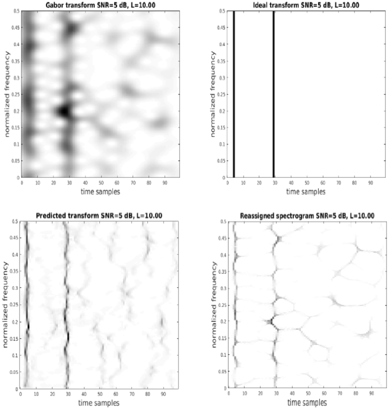


Figure : Comparison between $|X^h|$, ITFR, DNN2 estimation and classical reassigned spectrogram.

Overlapping components 1/2

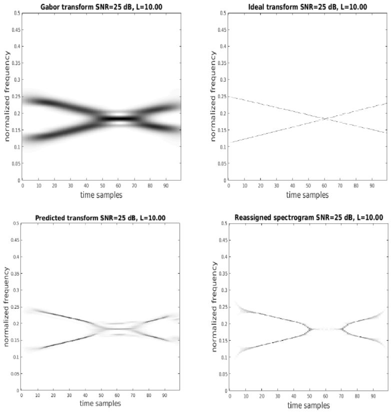


Figure : Comparison between $|X^h|$, ITFR, DNN2 estimation and classical reassigned spectrogram.

Overlapping components 2/2

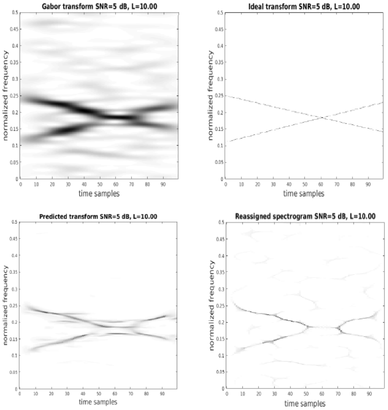


Figure : Comparison between $|X^h|$, ITFR, DNN2 estimation and classical reassigned spectrogram.

Objective 2 : Improving signal representations using data-driven

- Houidi, S., Fourer, D., & Auger, F. (2020). On the use of concentrated time-frequency representations as input to a deep convolutional neural network : Application to non intrusive load monitoring. Entropy, 22(9), 911.

In preparation :

- D.Fourer, Q. Legros and F. Auger. Improving the readability of Time-frequency and Time-Scale Representations using Deep Convolutional Neural Networks

Plan

Outlook

2022

- oct (1 month) : EM-based IF and CR estimation methods
- nov-dec (2 months) : Finalizing and submitting work on DNN-based reassignment

With extension for Quentin (6 extra months)

- jan-mar (3 months) :
  - Mode-extraction DNN-based methods (image-based segmentation using advanced neural network architectures)
  - TFR information estimation with F. Auger (Nantes)
- apr-jun (3 months) : Adaptive representations learning (investigation based on recurrent CNN with constraint)

Other ideas :

- spectrogram/scalogram phase reconstruction
- spectrogram/scalogram zeros/peaks statistical analysis
- overlapping modes reconstruction

Future work directions / Ideas for future projects and collaborations

- A public Benchmarking for mode extraction, denoising and IF/CR estimations methods (with Juan)
- TFR structure information extraction from a machine learning point of view (with Francois)
- New academic project proposal on signal processing for smart beekeeping (with all bees friends)
- Green IA : using TF analysis for computing low-dimension signal models for efficient computation, embedded systems and energy saving
- ...