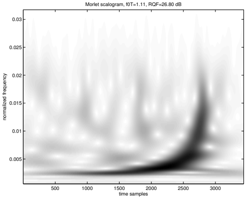
Improving the Readability of Time-frequency Representations using Deep Neural Networks

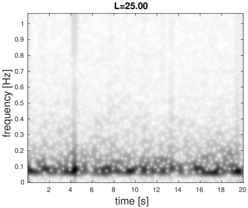
D. Fourer, F. Auger, S. Houidi, M. Sebag, H. Maaref, V. Vigneron

31 janvier 2020

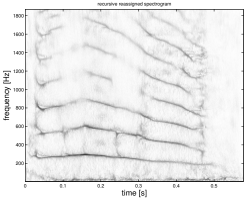
Motivation : non stationary multicomponent signals are everywhere...



GW150914



Draupner wave



Cello

ASTRES project (ANR-13-BS03-0002) 2013-2017

Consolidate, unify, extend and apply Reassignment, Synchosqueezing and EMD methods for :

- Computing efficient and meaningful signal representations
- Disentangling the elementary components of multicomponent signals
- Developing and applying tools for high-level processing (e.g. signal restoration, information extraction, modeling and regression, etc.)

The new ASCETE project

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 tasks :

- 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

The reassignment method [Kodera *et al.* 1978] [Auger & Flandrin 1995]

TF reassignment improves the energy concentration (readability) of any bilinear distribution by reassigning its energy to new locations closer to real signal support.

Considering a time-frequency representation (TFR) of a signal x expressed in terms of the Wigner-Ville distribution as :

$$\mathrm{TFR}_x(t,\omega) = \iint_{\mathbb{R}^2} \mathrm{WV}_x(\tau,\Omega)\Phi(t-\tau,\omega-\Omega)d\tau d\Omega$$

Method description

- Computation of the reassignment operators :

$$\hat{t}(t,\omega) = \frac{\int_{\mathbb{R}^2} \tau \mathrm{WV}_x(\tau,\Omega)\Phi(t-\tau,\omega-\Omega)d\tau d\Omega}{\int_{\mathbb{R}^2} \mathrm{WV}_x(\tau,\Omega)\Phi(t-\tau,\omega-\Omega)d\tau d\Omega} \quad (1)$$

$$\hat{\omega}(t,\omega) = \frac{\int_{\mathbb{R}^2} \Omega \mathrm{WV}_x(\tau,\Omega)\Phi(t-\tau,\omega-\Omega)d\tau d\Omega}{\int_{\mathbb{R}^2} \mathrm{WV}_x(\tau,\Omega)\Phi(t-\tau,\omega-\Omega)d\tau d\Omega}. \quad (2)$$

- Computation of the reassigned time-frequency representation :

$$\mathrm{RTFR}_x(t,\omega) = \int_{\mathbb{R}^2} \mathrm{TFR}_x(\tau,\Omega)\delta(t-\hat{t}(\tau,\Omega))\delta(\omega-\hat{\omega}(\tau,\Omega))\, d\tau d\Omega \quad (3)$$

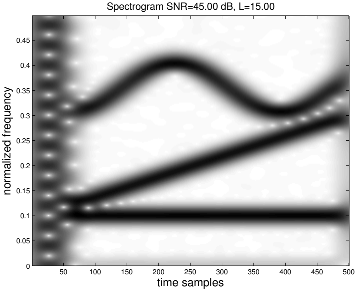
Example : the reassigned spectrogram

$X^h(t,\omega) = \int_{\mathbb{R}} x(\tau) h(\tau - t)^* \, e^{-j\omega\tau} \, d\tau$ **being the STFT of a signal** x **using a differentiable analysis window** h.

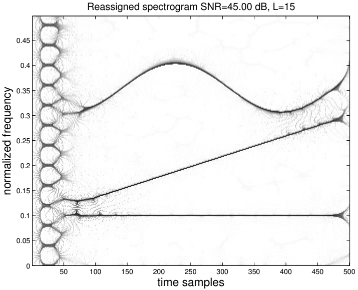
$$\hat{t}(t,\omega) = -\frac{\partial \Phi_x^h}{\partial \omega}(t,\omega) = t + \mathrm{Re}\left(\frac{X^{Th}(t,\omega)}{X^h(t,\omega)}\right) \quad \text{, with } Th(t) = t\, h(t) \tag{4}$$

$$\hat{\omega}(t,\omega) = \omega + \frac{\partial \Phi_x^h}{\partial t}(t,\omega) = \omega + \mathrm{Im}\left(\frac{X^{Dh}(t,\omega)}{X^h(t,\omega)}\right) \quad \text{, with } Dh(t) = \frac{dh}{dt}(t) \tag{5}$$

$$\mathrm{R}_x(t,\omega) = \iint_{\mathbb{R}^2} \left| X^h(\tau,\Omega) \right|^2 \delta(t - \hat{t}(\tau,\Omega))\delta(\omega - \hat{\omega}(\tau,\Omega)) \, d\tau \frac{d\Omega}{2\pi} \tag{6}$$



(a) $|X^h(t,\omega)|^2$



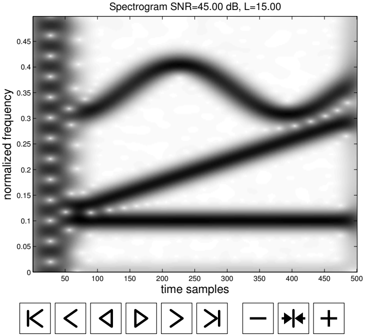
(b) $|R_x^h(t,\omega)|^2$

Synchrosqueezing

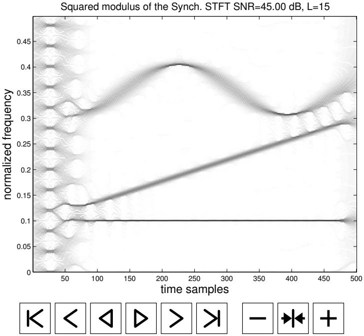
Can be viewed as a particular reassignment method which allows to compute sharpen and reversible TFRs [Daubechies 1996, 2011] [Thakur 2011]. Computation of the synchrosqueezed STFT and of its signal reconstruction formula :

$$\mathbf{S}_{\boldsymbol{x}}(\boldsymbol{t}, \omega) = \frac{\mathbf{1}}{\boldsymbol{h}(\mathbf{0})} \int_{\mathbb{R}} \boldsymbol{X}^{\boldsymbol{h}}(\boldsymbol{t}, \Omega)\delta(\omega - \hat{\omega}(\boldsymbol{t}, \Omega)) \frac{\boldsymbol{d}\Omega}{\mathbf{2}\pi} \quad (7)$$

$$\hat{\boldsymbol{x}}(\boldsymbol{t}) = \int_{\mathbf{supp}_{\Omega}(\boldsymbol{x})} \mathbf{S}_{\boldsymbol{x}}(\boldsymbol{t}, \Omega)\boldsymbol{d}\Omega \quad (8)$$



(c) $|X^{\boldsymbol{h}}(t, \omega)|^2$



(d) $|S_{\boldsymbol{x}}(t, \omega)|^2$

Recent advances for reassignment/synchrosqueezing

- Levenberg-Marquardt reassignment [Auger, et al., 2012]
- Second-order vertical and oblique synchrosqueezing [Oberlin, et al., 2015]
- Recursive Levenberg-Marquardt reassignment and synchrosqueezing [Fourer, et al., 2016]
- Chirp demodulation [Meignen et al., 2017]
- Higher-order synchrosqueezing [Pham, Meignen et al., 2017]
- Synchro-extracting transform [Yu et al., 2017]
- Horizontal (time-reassigned) synchrosqueezing [He et al., 2019]
- High-order chirp demodulation [Pham, Meignen et al., 2019]
- Second-order horizontal synchrosqueezing [Fourer, Auger, 2019]
- ...

The (already outdated) TFTB and ASTRES toolbox

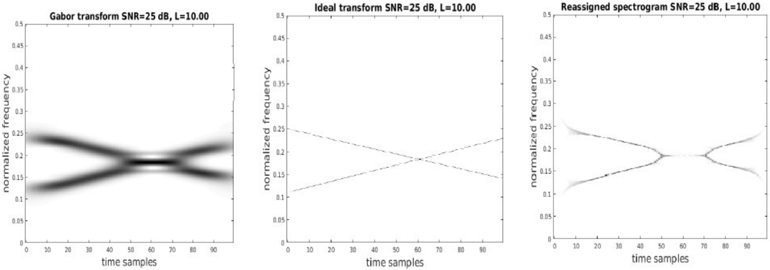
- *F. Auger, P. Flandrin, P. Gonçalvès and O. Lemoine. "Time-frequency toolbox, CNRS/Rice University, France.", 1995.*

  `http://tftb.nongnu.org`

- *D. Fourer, J. Harmouche, J. Schmitt, T. Oberlin, S. Meignen, F. Auger and P. Flandrin. The ASTRES Toolbox for Mode Extraction of Non-Stationary Multicomponent Signals. Proc. EUSIPCO 2017, Aug. 2017. Kos Island, Greece.*

  `https://github.com/dfourer/ASTRES_toolbox`

Limitations and current challenges



- Robustness to noise
- Overlapping components
- Efficient Ridge estimation and mode extraction
- Optimal hyperparameters tuning
- Generalization to multidimensional signals
- Perfect signal reconstruction from (synchrosqueezed or not) TFRs
- Phase retrieval from real-valued TFRs (eg. spectrogram, scalogram, etc.)
- ...

New proposed approach

Principle : regression of the reassignment operation using Convolutional Neural Networks (CNN)

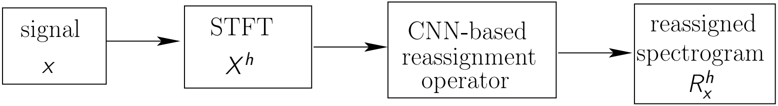


Figure : Proposed DNN-based reassignment method

Motivation

- Reassignment can be viewed as an image post-processing operation
- Capability to synthesize the ideal time-frequency representation of a given signal model allowing to generate a simulated training datasets
- Consideration of noisy signals to improve the robustness of the trained model

Deep neural network architecture

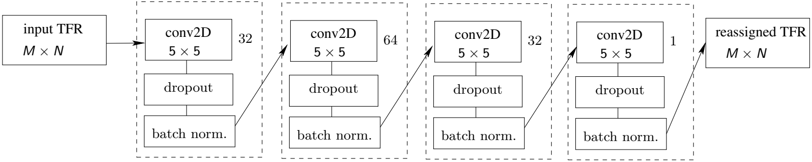


Figure : Proposed architecture based on 2D CNN.

- 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 optimization." corr abs/1502.04390 (2015).

Sinusoidal signal model

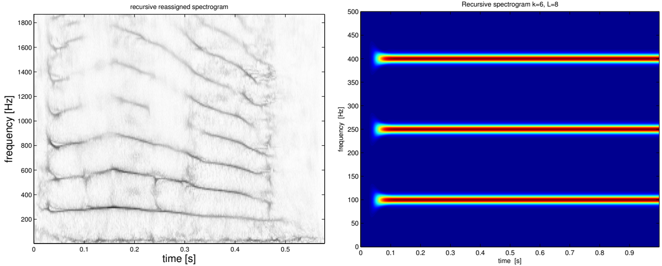
- Time-domain expression :

$$x(t) = a_x(t)\, \mathbf{e}^{j(\phi_{\mathbf{x}}+\omega_{\mathbf{x}}(t)t)} \tag{9}$$

with $a_x(t)$, $\omega_x(t)$ the instantaneous amplitude and instantaneous frequency. ϕ_x is the initial phase.

- Ideal time-frequency representation :

$$\text{ITFR}_x(t,\omega) = \begin{cases} a_x(t) & \text{if } \omega = \omega_x(t), \\ 0 & \text{otherwise} \end{cases} \tag{10}$$



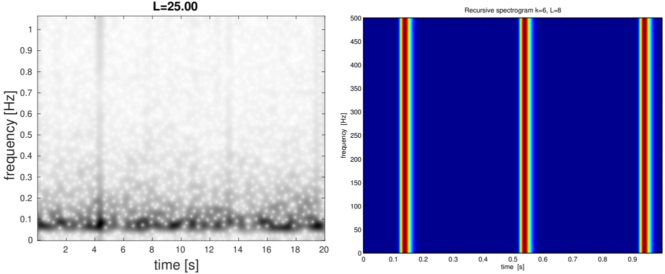
Impulse signal model

- Time-domain expression :

$$x(t) = a_x \delta(t - t_0) \tag{11}$$

- Ideal time-frequency representation :

$$\text{ITFR}_x(t, \omega) = \begin{cases} a_x & \text{if } t = t_0, \forall \omega \\ 0 & \text{otherwise} \end{cases} \tag{12}$$



Mixture signal model :

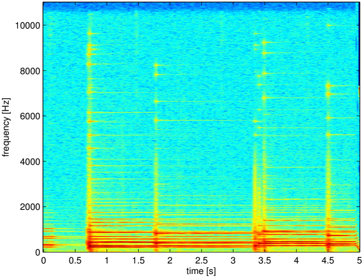
- Time-domain expression :

$$x(t) = \sum_{i=1}^{I} s_i(t) \tag{13}$$

where s_i denotes an elementary signal component.

- Ideal time-frequency representation :

$$\text{ITFR}_x(t,\omega) = \sum_{i=1}^{I} \text{ITFR}_{s_i}(t,\omega) \tag{14}$$



Implementation

Training

- We compute for each signal x its magnitude $|X^h|$ (X^h being the STFT of x) and its ideal time-frequency representation Y.
- We train the CNN to minimize the Mean Squared Error (MSE) between the estimated TFR $\hat{Y}$ and the ideal TFR :

$$\mathcal{L}(Y, \hat{Y}) = \frac{1}{NM} \sum_{n=1}^{N} \sum_{m=1}^{M} |Y[n, m] - \hat{Y}[n, m]|^2 \tag{15}$$

DNN-based reassignment

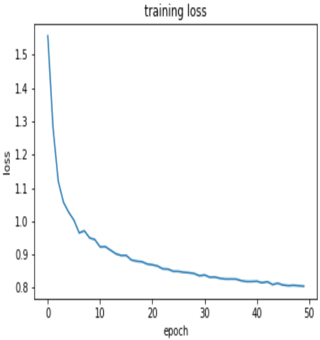
- The trained DNN operator is applied on $|X^h|$ to estimate $\hat{Y}$.

Experiment

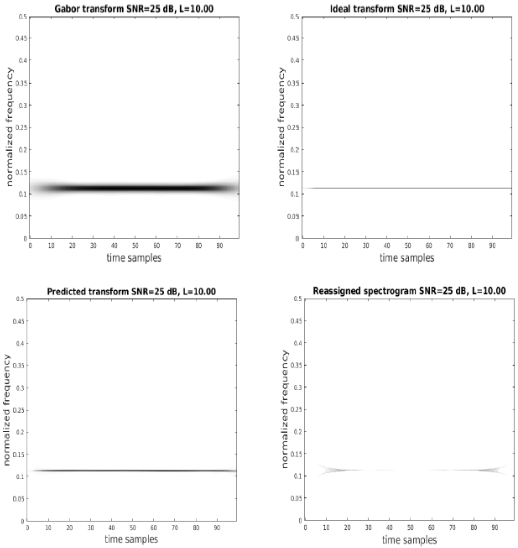
We consider 3 datasets made of 3,000 randomly (uniformly sampled) generated multicomponent signals ($I \in [1;10]$) merged with a white Gaussian noise (SNR $\in [5, 2545]dB$) :

- **Sinusoidal dataset** such as $x(t) = \sum_{i=1}^{I} \exp\left(\sum_{p=0}^{P} c_p t^p\right)$ with $c_p \in \mathbb{C}$ and $P \leq 2$
- **Impulse dataset** such as $x(t) = \sum_{i=1}^{I} a_i \delta(t - t_i)$ with $a_i, t_i \in \mathbb{R}^+$
- **Sinusoid + impulses dataset** : merging of the two previously proposed datasets

The 3 datasets lead to 3 distinct DNN models : DNN1, DNN2 and DNN3.



Unitary test on a sinusoidal signal 1/2



Unitary test on a sinusoidal signal 2/2

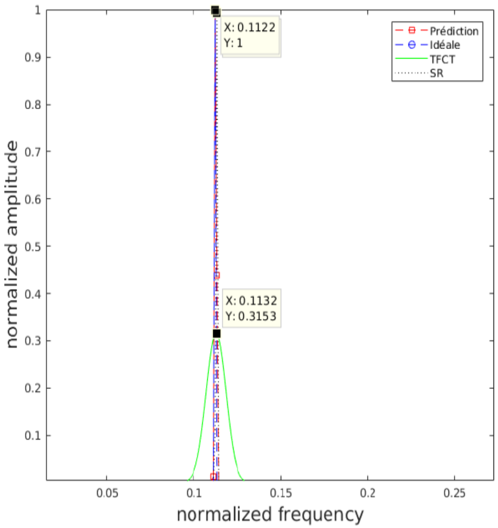


Figure: Comparison between $|X^h|$, ITFR, DNN1 estimation and classical reassigned

Unitary test on an impulse signal 1/2

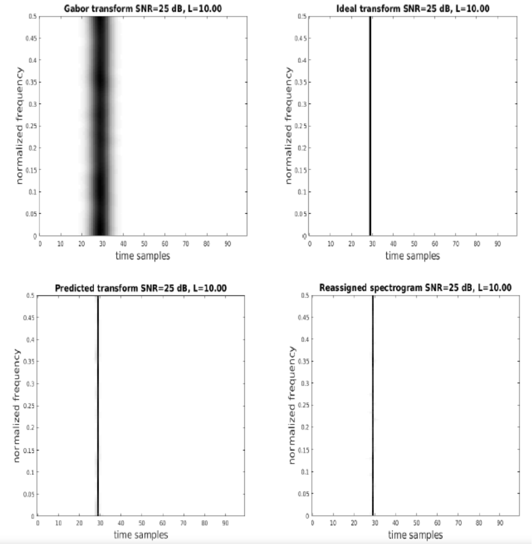


Figure : Comparison between $|X^h|$, ITFR, DNN2 estimation and classical reassigned

Unitary test on an impulse signal 2/2

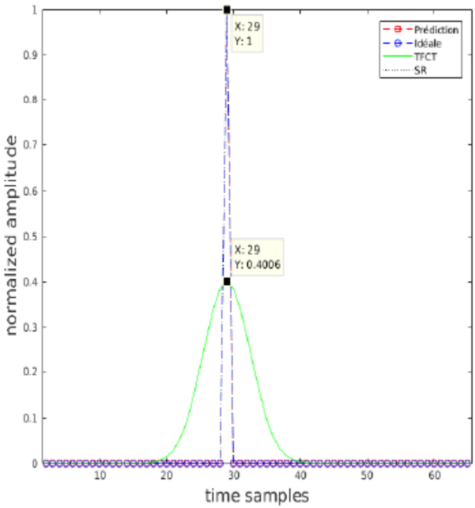
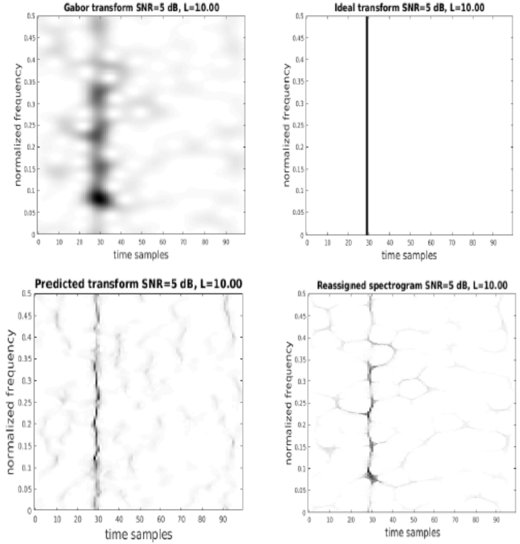


Figure : Comparison between $|X^h|$, ITFR, DNN2 estimation and classical reassigned

Noisy signal (SNR=5dB)



Noisy signal (SNR=5dB)

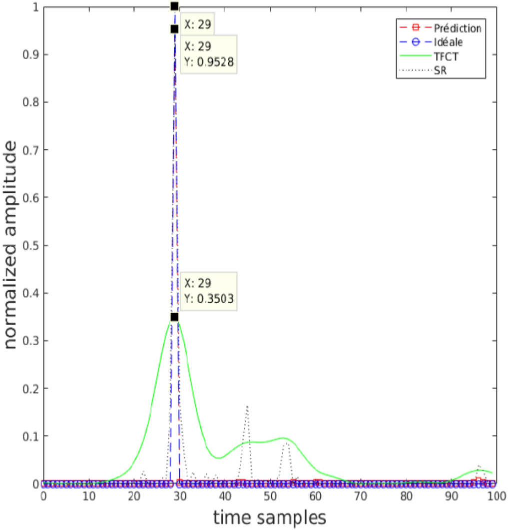
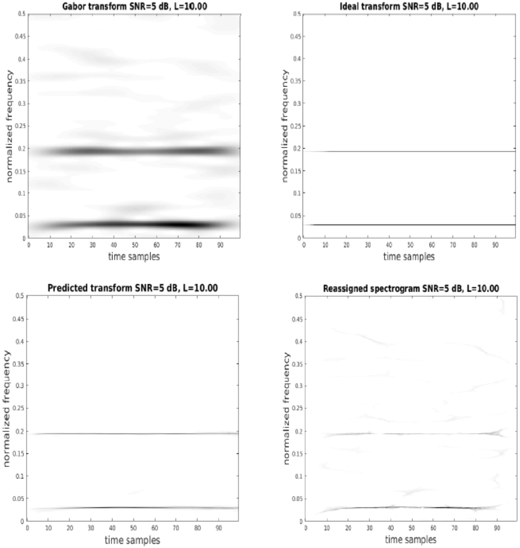


Figure: Comparison between $|X^h|$, ITER, DNN2 estimation and ideal signal

Multi-component noisy signal 1/2



Multi-component noisy signal 2/2

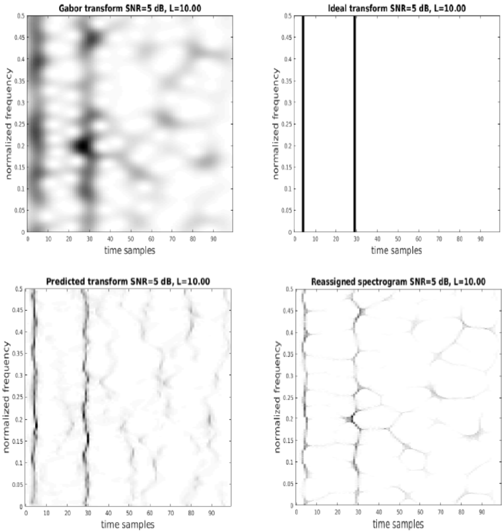


Figure: Comparison between $|X^h|$, ITER, DNN2 estimation and classical reassigned

Overlapping components 1/2

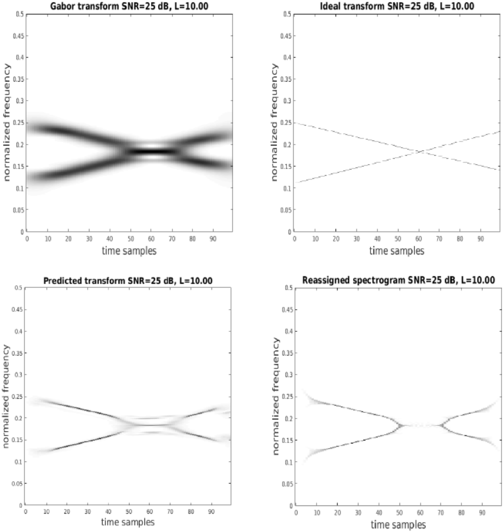
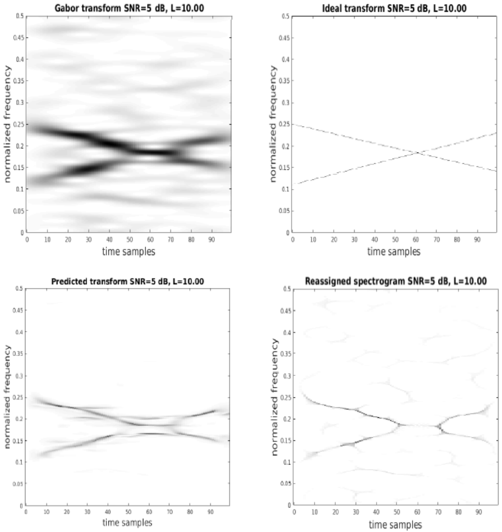


Figure : Comparison between $|X^h|$, ITFR, DNN2 estimation and classical reassigned

Overlapping components 2/2



Summary

Contributions

- A new proposed DNN-based reassignment operator
- A slight improvement in presence of noise when compared to classical reassignment
- Deals with overlapping components

Limitations

- Computed TFRs are non invertible (for the moment)
- Reassignment operators parameters (IF and IG_d) are not explicited
- Requires more investigation on real-world signals

Future work directions

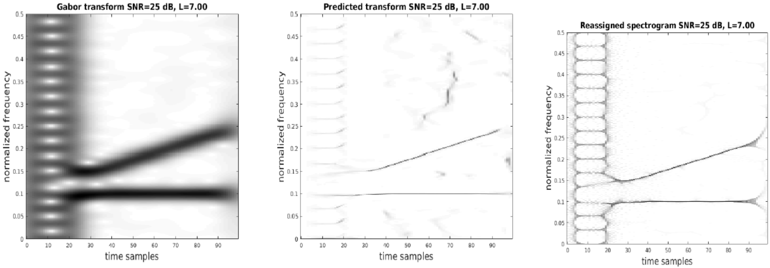


Figure : Comparison between DNN3 and classical reassignment.

- Increasing the database size and the number of DNN parameters
- Applications on real-world data (audio, biomedicine, etc.)
- A more complete comparative investigation and evaluation
- Modification of the architecture to estimate the reassignment operators (IF and IG_d)
- Application to DNN-based ridge estimation