Filtering out time-frequency areas using Gabor multipliers

A.Marina Krémé [1,2], Valentin Emiya[2] , Caroline Chaux [1] and Bruno Torrésani [1]

[1] *Institut de Mathématiques de Marseille, Aix Marseille Univ, CNRS, France.*
[2] *Laboratoire d'Informatique et des Systèmes, Aix Marseille Univ, CNRS, France.*

ICASSP2020 Barcelona

Institut de Mathématiques de Marseille

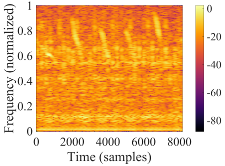
I2M - UMR CNRS 7373
Aix-Marseille Université • CNRS • École Centrale de Marseille

LIS
LABORATOIRE D'INFORMATIQUE & SYSTÈMES

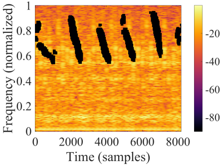
Outline

Outline

Problem overview and baseline solutions

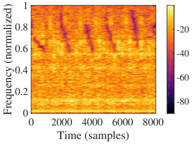


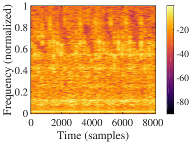
Obs: target + perturbation

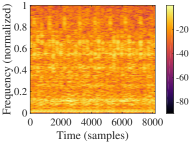


Ω : masked area

Goal: filter out signal perturbation and minimize the artefacts







ZerVal method:
fill Ω with zeros
Audible artefact

RandVal method:
interp in Ω + random phases [ANSYS]

RedEnerg
Our method

Problem of filtering out time-frequency areas formulation

Time-frequency transform: $\mathcal{T} : \mathbb{R}^L \longrightarrow \mathbb{R}^{M \times N}$

The perturbation signal $\mathbf{x}_{per}$ is well localized in a region Ω and the target signal $\mathbf{x}_{ref}$ is everywhere in TF plane

Observations: $\mathbf{x}_0 = \mathbf{x}_{ref} + \mathbf{x}_{per}$

Problem of filtering out time-frequency areas formulation

$$\mathbf{x}^*_\lambda = \underset{\mathbf{x} \in \mathbb{C}^L}{argmin} \underbrace{\|\mathcal{T}\mathbf{x} - \mathcal{T}\mathbf{x_0}\|^2_{\bar{\Omega}}}_{data\ fitting} + \lambda \underbrace{\|\mathcal{T}\mathbf{x}\|^2_{\Omega}}_{energy\ within\ \Omega}, \quad \lambda > 0,$$

$$where \quad \|\mathbf{y}\|^2_{\Omega} := \sum_{k \in \Omega} |\mathbf{y}[k]|^2.$$

Outline

Gabor transform

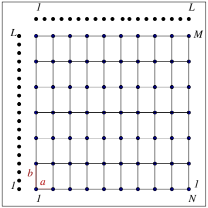


Figure: Lattice: $\Lambda = \mathbb{Z}_M \times \mathbb{Z}_N$ with $N = L/a$ and $M = L/b$

For a window $\mathbf{g} \in \mathbb{R}^L$, the Gabor atom $\mathbf{g}_{mn}$ at TF point $(m, n) \in \Lambda$:

$$\mathbf{g}_{mn}[l] = \mathbf{g}[l - na]e^{2i\pi mbl/M}, \forall l \in \mathbb{Z}_L.$$

Gabor transform:

$$\mathcal{V}_{\mathbf{g}}\mathbf{x}[m, n] = \langle \mathbf{x}, \mathbf{g}_{mn} \rangle = \sum_{l=0}^{L-1} \mathbf{x}[l]\mathbf{g}[l - na]e^{-2i\pi mbl/M}.$$

Gabor multiplier

For suitably chosen $\mathbf{g}$ and Λ (i.e. generating a tight Parseval frame) then, $\forall \mathbf{x} \in \mathbb{C}^L$, we have:

$$\mathbf{x} = \sum_{m,n} \mathcal{V}_{\mathbf{g}}\mathbf{x}[m,n]\mathbf{g}_{mn} = \sum_{m,n} \langle \mathbf{x}, \mathbf{g}_{mn} \rangle \mathbf{g}_{mn} \ .$$

Gabor multiplier is similarly defined via a transfer function also called mask $\mathbf{m}$ in the time-frequency domain, i.e,

Definition

The Gabor multiplier associated to $(\mathbf{g}, \Lambda)$ with mask $\mathbf{m} \in \mathbb{R}^{M \times N}$ is defined by:

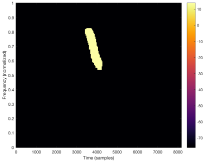
$$\mathfrak{M}_{\mathbf{m}}\mathbf{x} = \sum_{m,n} \mathbf{m}[m,n] \langle \mathbf{x}, \mathbf{g}_{mn} \rangle \mathbf{g}_{mn} \ ,$$

where $\mathbf{m}$ denotes also the operator of pointwise multiplication by $\mathbf{m}$.

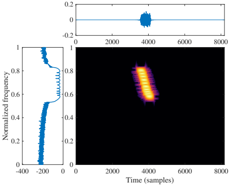
Main property [H. G. Feichtinger and K. Nowak, 2002.]

If **m** is real-valued then $\mathfrak{M}_{\mathbf{m}}$ is Hermitian. Then there is an orthonormal basis of $\mathbb{C}^L$ formed by $\mathfrak{M}_{\mathbf{m}}$ eigenvectors.

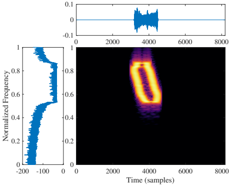
Illustration: eigenvectors associated with large eigenvalues tend to follow TF localization properties of the mask.



(a) mask



(b) Large eigenvalue



(c) Small eigenvalue

Outline

TF areas filtering: normal equations

Objective function:

$$f(\mathbf{x}) = \|\mathcal{V}_{\mathbf{g}}\mathbf{x} - \mathcal{V}_{\mathbf{g}}\mathbf{x_0}\|_{\overline{\Omega}}^2 + \lambda\|\mathcal{V}_{\mathbf{g}}\mathbf{x}\|_{\Omega}^2$$

Normal equations:

$$(\mathfrak{M}_{\overline{\Omega}} + \lambda\mathfrak{M}_{\Omega})\mathbf{x} = \mathfrak{M}_{\overline{\Omega}}\mathbf{x_0}$$
$$[I + (\lambda - 1)\mathfrak{M}_{\Omega}]\,\mathbf{x} = (I - \mathfrak{M}_{\Omega})\mathbf{x_0}$$

$[I + (\lambda - 1)\mathfrak{M}_{\Omega}]$ is invertible for $\lambda > 0$ due to:

If $\mathbf{m} \in \mathbb{C}^{M\times N}$, then $\mathfrak{M}_{\mathbf{m}}$ defines a bounded operator with operator norm $\|\mathfrak{M}_{\mathbf{m}}\|_{op} \leq C\|\mathbf{m}\|_{\infty}$, where C is a constant. In particular, if $\mathbf{g}$ and Λ generate a Parseval frame, then $\|\mathfrak{M}_{\mathbf{m}}\|_{op} \leq \|\mathbf{m}\|_{\infty}$.

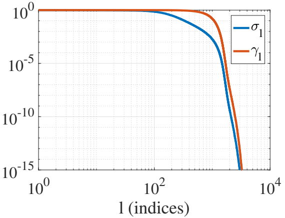
TF areas filtering: analytical solution

$\mathfrak{M}_\Omega$ is diagonalizable, so there is a unitary matrix U and a diagonal matrix $D = \text{diag}(\sigma_1, ..., \sigma_L)$, $\sigma_1 \geq \cdots \geq \sigma_L$ such that $\mathfrak{M}_\Omega = UDU^{-1}$.

Analytical solution

$$\mathbf{x}_\lambda^* = \mathbf{x_0} - U\,\text{diag}\underbrace{\left(\frac{\lambda\sigma_l}{1-(1-\lambda)\sigma_l}\right)}_{\gamma_l, l=1,\cdots,L} U^{-1}\mathbf{x_0},\ \lambda > 0$$

Typical behavior of eigenvalues ($\lambda = 0.1$): $0 \leq \sigma_l, \gamma_l \leq 1$



Efficient tuning of the hyperparameter

Computational issues

1. Most costly part: diagonalization of $\mathfrak{M}_\Omega$. It does not depend on λ and is done only once ☺.
2. Low rank approximations can be used to reduce computing time ☺.

Tuning of the regularization parameter λ

- λ can be tuned to match a prescribed target energy $E = \|\mathcal{V}_\mathbf{g}\mathbf{x}_\lambda^*\|_\Omega^2$.
- We set $E = \|\mathcal{V}_\mathbf{g}\mathbf{x}_\lambda^*\|_{\Omega'}^2$ from a region Ω' similar to Ω and not affected by the perturbation.

Outline

Problem data

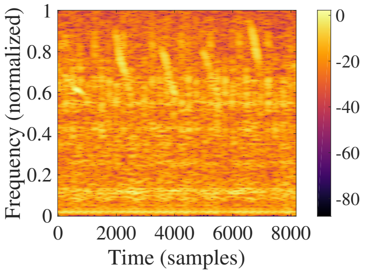
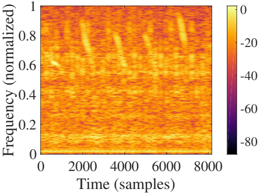


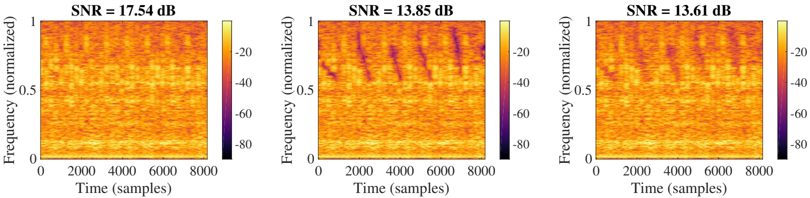
Figure: Observations (left), Mask (right)

- signal length: $L = 8192$
- sampling frequency: $fs = 8000$ Hz
- $\mathbf{x}_{ref}$: wide-band engine sound and $\mathbf{x}_{per}$: chirp-like bird song
- $\mathbf{g}$: hanning window of length 128 $(16ms)$, hop size $a = 32$ $(4ms)$, and frequency shift $b = 512$ $(16Hz)$.

Comparative reconstruction results



$$SNR(\mathbf{x}_{ref}, \mathbf{x}) = 20 \log 10 \left(\frac{\|\mathbf{x}_{ref}\|}{\|\mathbf{x}_{ref} - \mathbf{x}\|} \right)$$



(a) RedEnerg (b) ZerVal (c) RandVal

SNRs for several targets and perturbations

	Target			
	car	train	aircraft	**perturbation**
RedEnerg	**25.05**	**23.88**	**26.80**	Beeps
ZerVal	24.23	16.52	24.73	
RandVal	24.90	16.15	24.73	
RedEnerg	**18.88**	**22.49**	**20.72**	Finger snap
ZerVal	17.18	20.70	16.59	
RandVal	16.92	16.17	15.93	
RedEnerg	**21.67**	**20.97**	**17.96**	Clicks
ZerVal	17.61	13.98	12.33	
RandVal	10.57	5.60	11.07	
RedEnerg	**18.01**	**21.52**	**20.04**	Birdsong
ZerVal	17.4	21.52	19.71	
RandVal	17.13	21.28	19.56	

Outline

Conclusion

- We have addressed the problem of estimating a target signal, with no assumption on its contents, when perturbated by an additive signal that is well located in a region Ω of the TF plane.
- We have proposed an optimization problem in which the energy in Ω is controlled.
- We have used the nice properties of the Gabor multipliers for its analytical and numerical resolution.
- The proposed method outperforms some industrial baseline systems in terms of reconstruction SNR.
- Ongoing work: efficient algorithms for speeding up computations.

ama-marina.kreme@univ-amu.fr

Thank you for your attention!