



Laboratoire d'ElectroMagnétisme et d'Acoustique

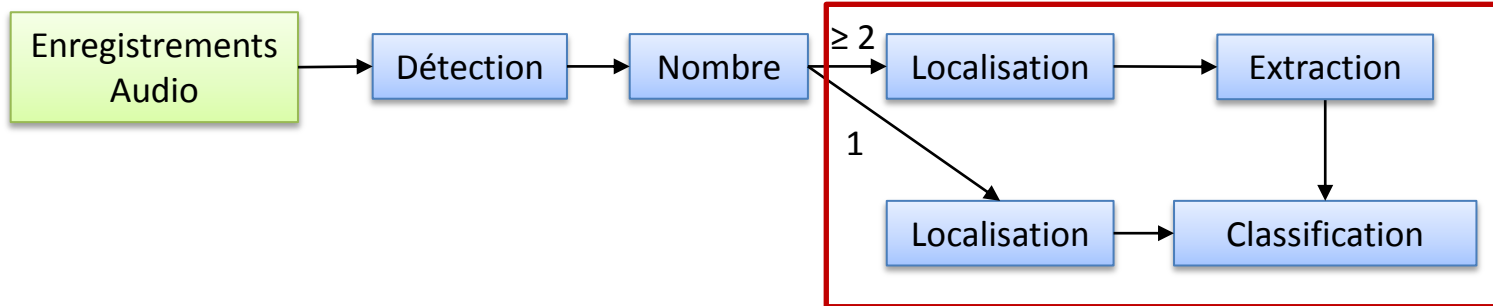
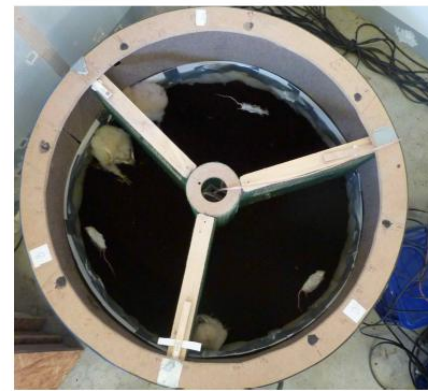
Localisation de sources, extraction et caractérisation de signaux acoustiques: application aux cris de chouettes effraies en milieu confiné

**Dorian Cazau (Stagiaire M1 - UPMC),
Patrick Marmaroli (EPFL), Xavier Falourd (EPFL),
Charlène Ruppli (UNIL), Amélie Reiss (UNIL),
Hervé Lissek (EPFL), Alexandre Roulin (UNIL)**



Contexte et objectifs

- ▶ **Contexte biologique: quémande alimentaire des oisillons**
 - But: comprendre mécanismes de la négociation.
 - Organisation, dynamique des échanges ?
- ▶ **Objectif 1: banc d'essai existant à instrumentaliser**
 - Choix des capteurs, géométrie
- ▶ **Objectif 2: algorithme de post-traitement**
 - Traitement automatique des campagnes de mesures
 - Quelle chouette vocalise ? Quel type de cri ? Durée ?

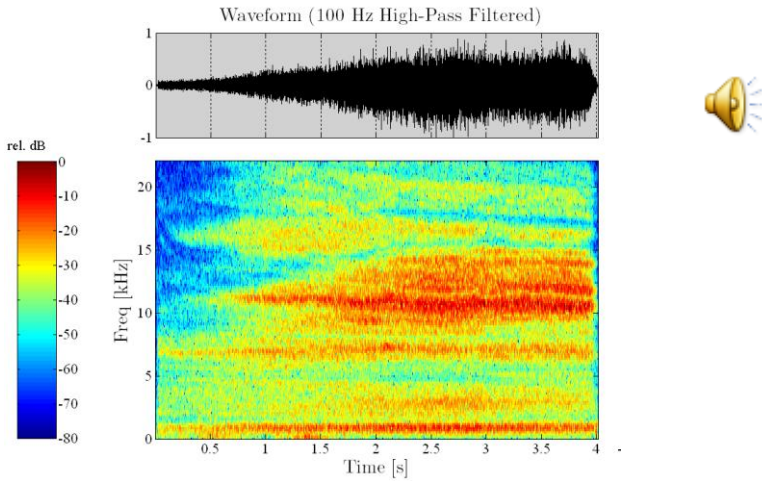


Chouette Effraie: *Tyto alba*

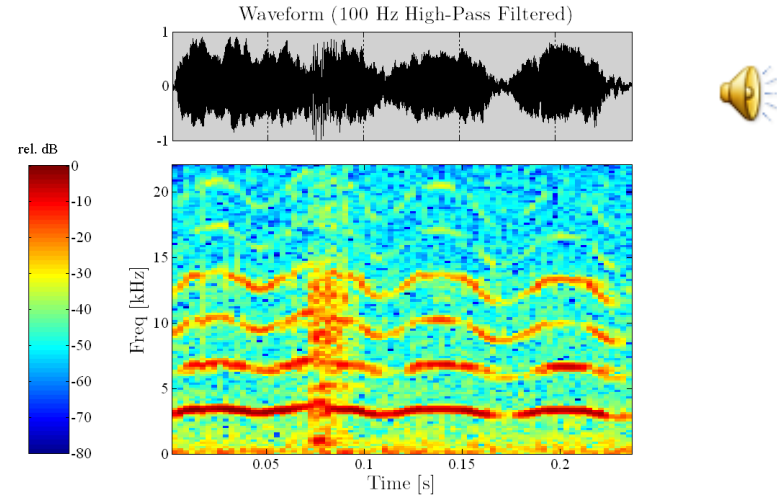


Contexte, exemples de vocalisations

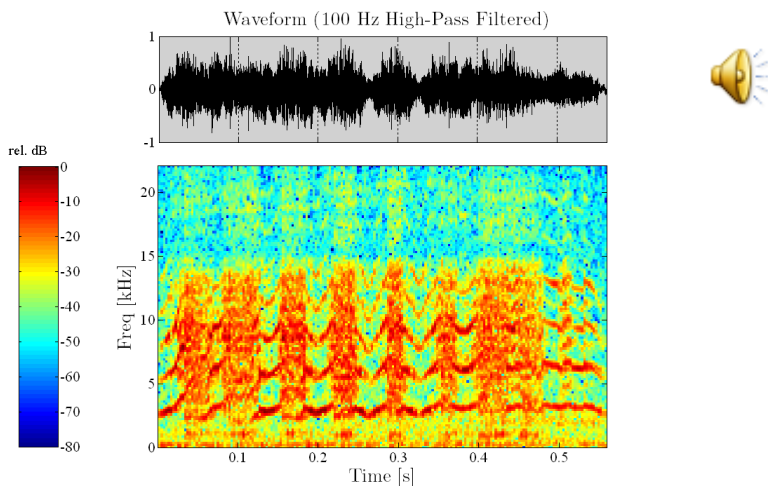
Chuintement anti-prédateur



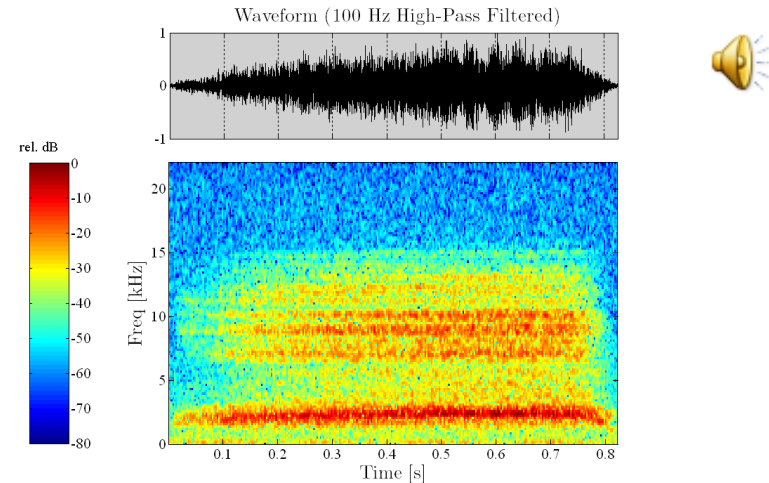
Gazouilli A



Gazouilli B

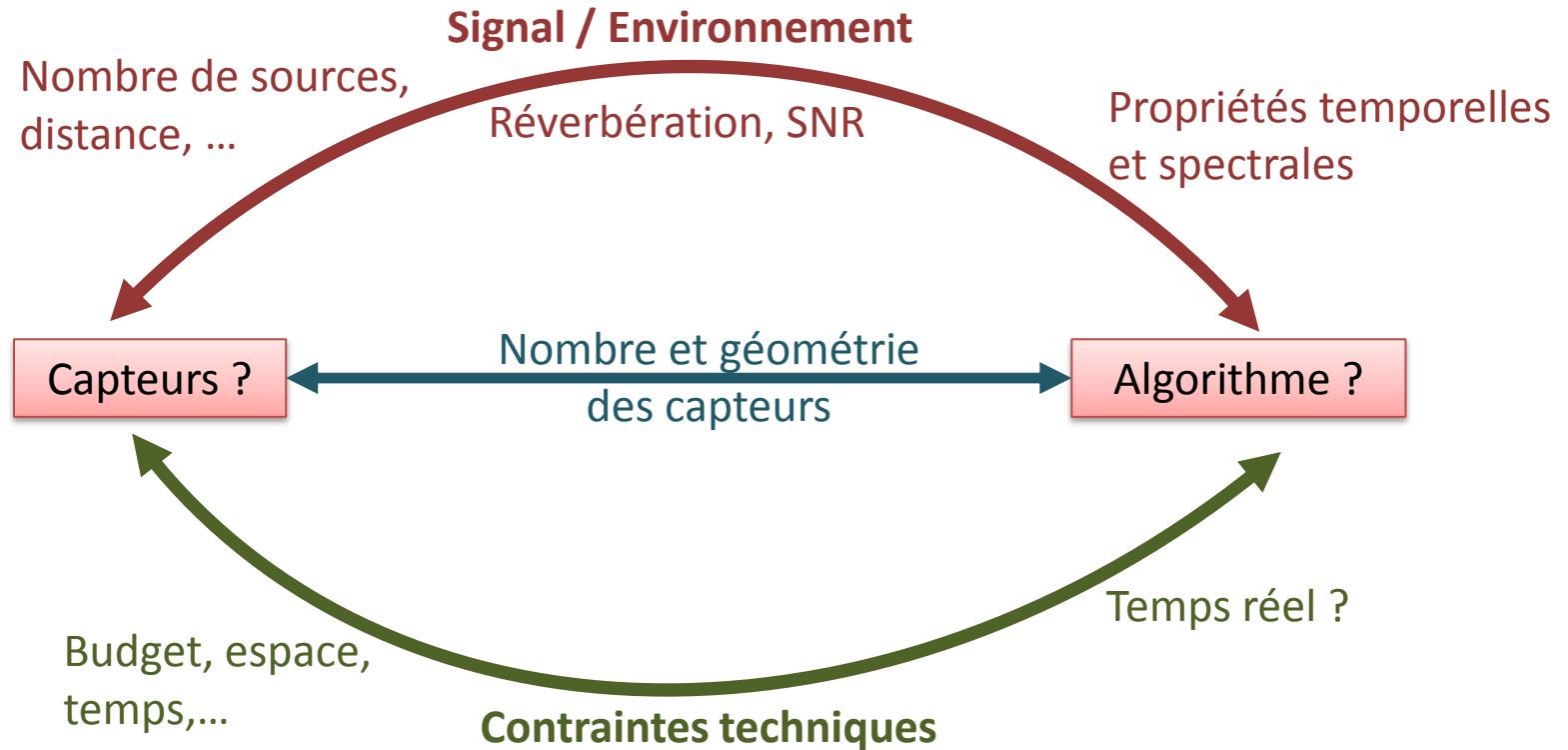


Cri de négociation





Instrumentalisation du banc d'essai



Position des capteurs

1) Solide de Platon autour de la source

Equality holds if and only if

1. $\sum_{i=1}^N \mathbf{g}_i = 0$, with

$$\mathbf{g}_i = \frac{\mathbf{x}_i - \mathbf{x}_s}{|\mathbf{x}_i - \mathbf{x}_s|_2} \quad (3.151)$$

coordonnées capteur i
coordonnées source s

2. the matrix $\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_N]$, satisfies $\mathbf{G}\mathbf{G}^T = (N/p)\mathbf{I}$, i.e. \mathbf{G} has orthogonal row vectors with equal row norm.

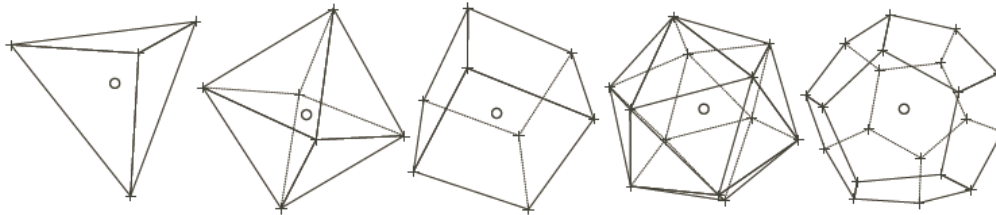


Figure 3.9: Tetrahedron, Octahedron, Hexahedron (cube), Dodecahedron, and Icosahedron

Source: Neering2009
D mo: Yang2005

Position des capteurs

1) Solide de Platon autour de la source

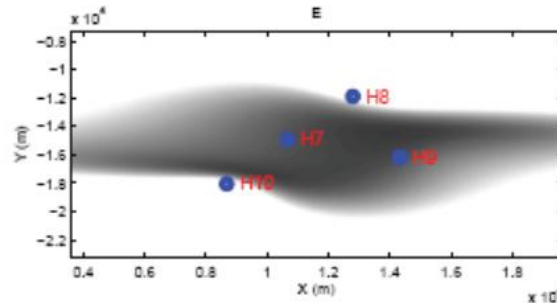
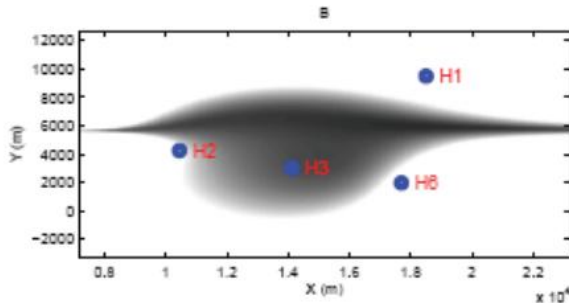
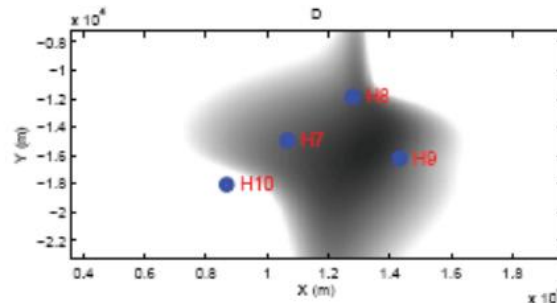
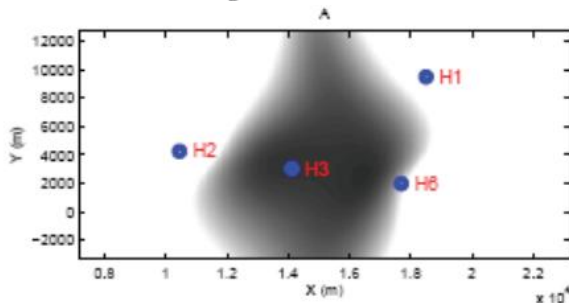
Equality holds if and only if

$$1. \sum_{i=1}^N \mathbf{g}_i = 0, \text{ with}$$

$$\mathbf{g}_i = \frac{\mathbf{x}_i - \mathbf{x}_s}{|\mathbf{x}_i - \mathbf{x}_s|_2}. \quad (3.151)$$

↗ coordonnées capteur i
 ↘ coordonnées source s

2. the matrix $\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_N]$, satisfies $\mathbf{G}\mathbf{G}^T = (N/p)\mathbf{I}$, i.e. \mathbf{G} has orthogonal row vectors with equal row norm.



Source: Benard-Caudal2009

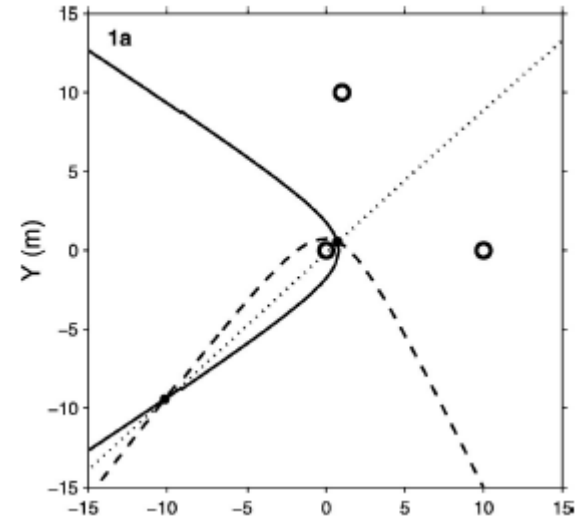
■ ■ ■ ■ ■ Nombre de capteurs

1) Solide de Platon autour de la source

2) Nombre de capteurs:

-localisation **2D**: minimum **4** capteurs

-localisation **3D**: minimum **5** capteurs

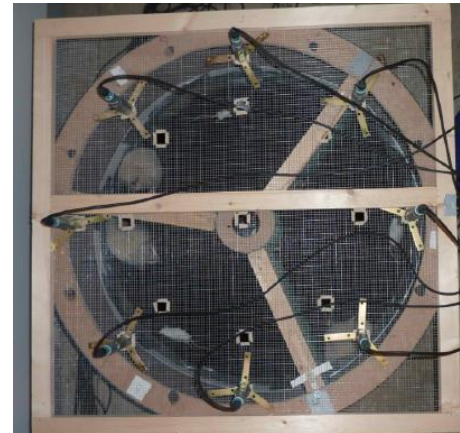
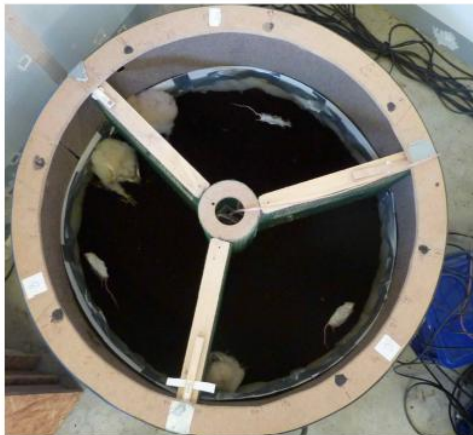


Source: Spiesberger2001

Nombre de capteurs

- 1) Solide de Platon autour de la source
- 2) Nombre de capteurs:
 - localisation **2D**: minimum **4** capteurs
 - localisation **3D**: minimum **5** capteurs

micros : AKG Perception 170
carte son : Presonus Firestudio et Digimax FS (microphone preamplifier)
ordi : Mac Pro RAID Card ; 2 x 2.26 GHz Quad Core; Intel Xeon
(2 disques durs en raid) - (Mac OSX)

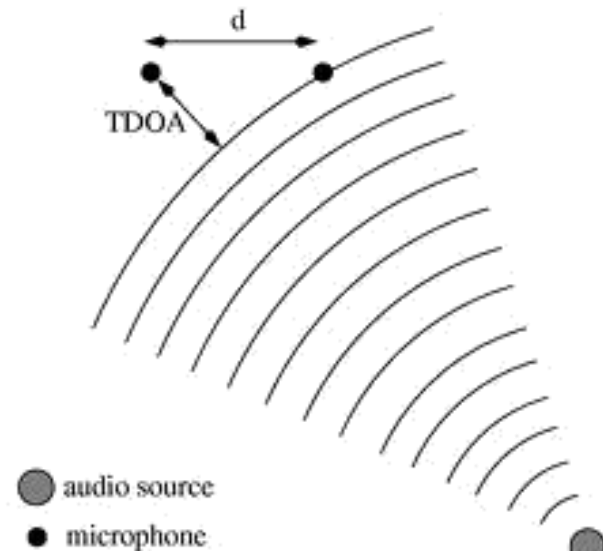
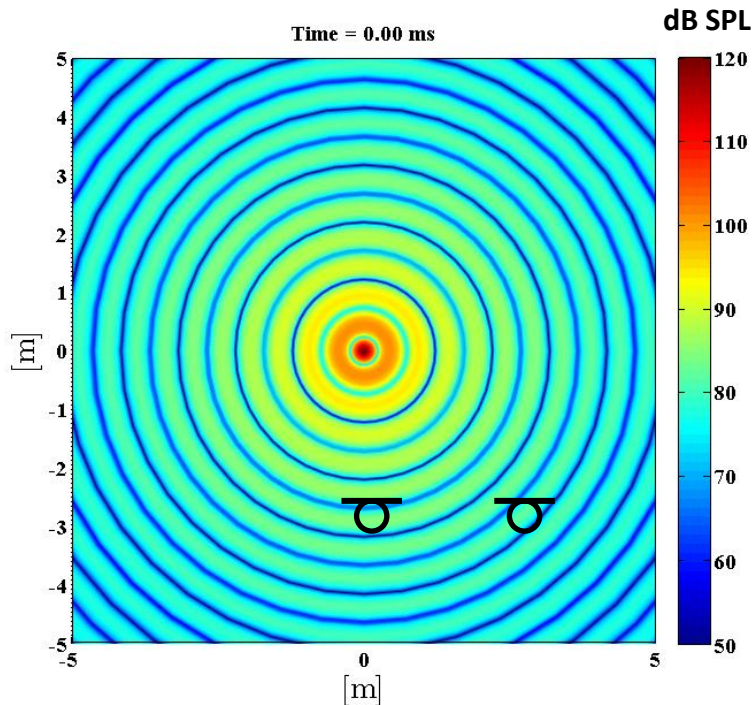


Localisation: choix algorithmique

Amplitude Vs.

Phase

**TDOA-based method: Time
Difference Of Arrival**



Localisation: choix algorithmique

Amplitude Vs.

Phase

TDOA-based method: Time Difference Of Arrival

Subspace-based methods

Capon spectrum (Hawkes et al., 1998)
MUSIC (Schmidt, 1986)
ESPRIT (Wong et al. 1997)
TF-MUSIC (Zhang et al, 2001)
Wideband-MUSIC

Active

RADAR
SONAR

Analytic

Intersection of
hyperboloids

Others

Coherence test
MAMDF
Minimum variance (Nehorai et. al, 1994)
Propagator method (Tayem et al, 2005)
...

Active / Passive

Single / Multiple Sources

2D / 3D

More / Less sources than sensors

Narrowband / Wideband

Reverberant / Clean conditions

Overlapped / Sparse sources

Spatial Maximum Likelihood

SRP-PHAT

Beamforming

Accumulated correlation

Hemisphere sampling

Localisation : approche hyperbolique

Idée: estimer la différence de temps d'arrivée de l'onde sonore entre plusieurs capteurs, puis en déduire son origine par triangulation.

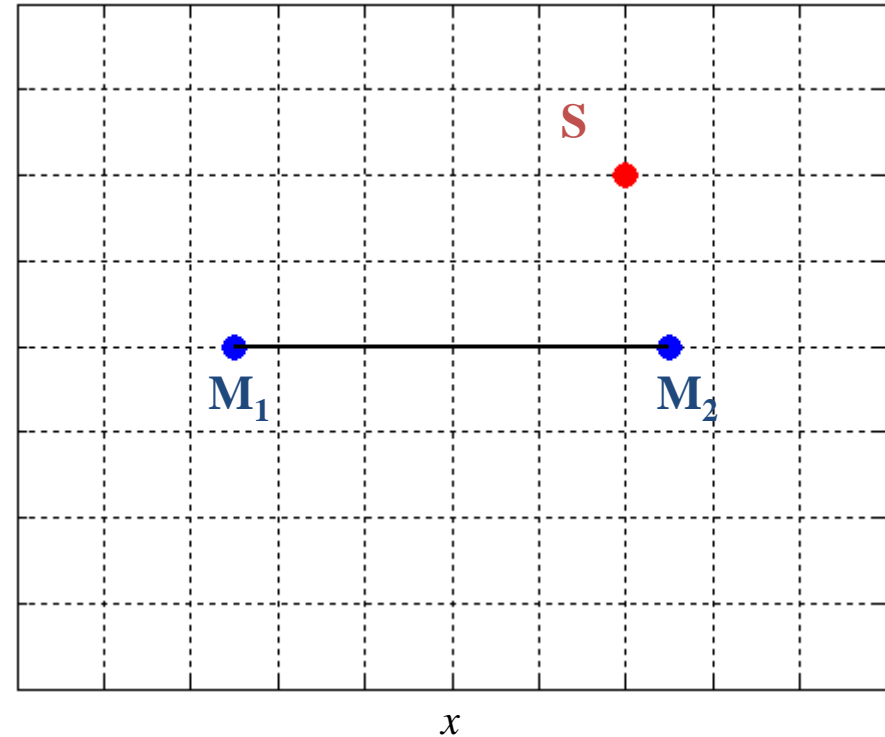
Modèle

$$p_i(t) = s \left(t - \frac{d(S, M_i)}{c} \right) + n_i(t)$$

Différence de temps d'arrivée (TDOA)

$$\tau_{12} = \frac{d(S, M_2) - d(S, M_1)}{c}$$

y



Localisation : approche hyperbolique

Idée: estimer la différence de temps d'arrivée de l'onde sonore entre plusieurs capteurs, puis en déduire son origine par triangulation.

Modèle

$$p_i(t) = s \left(t - \frac{d(S, M_i)}{c} \right) + n_i(t)$$

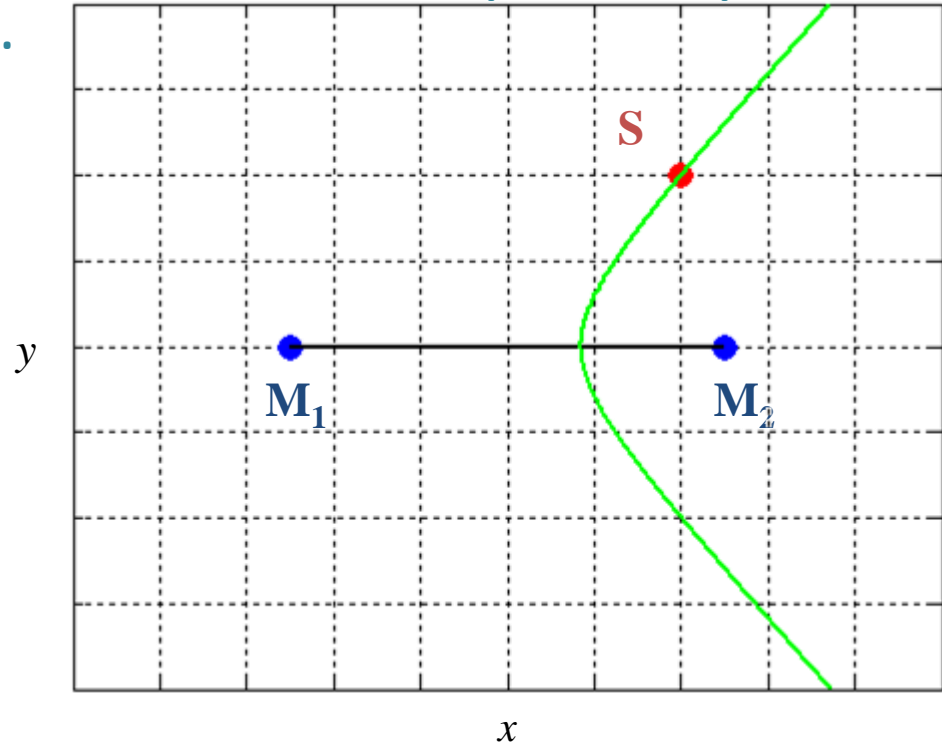
Différence de temps d'arrivée (TDOA)

$$\tau_{12} = \frac{d(S, M_2) - d(S, M_1)}{c}$$

1 TDOA -> 1 Hyperbole ou Isodiachron
(Spiesberger2004)

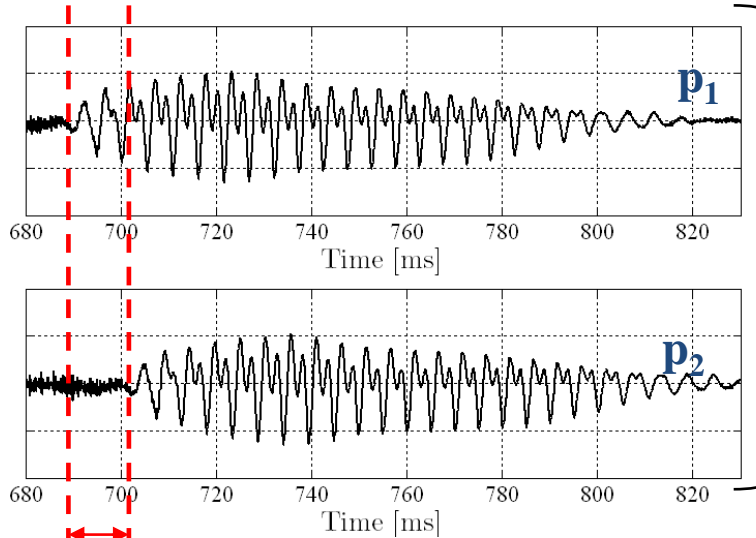
$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad \text{avec} \quad a = \frac{c\tau_{12}}{2}$$

$$b = \sqrt{\left(\frac{d(M_1, M_2)}{2} \right)^2 - a^2}$$

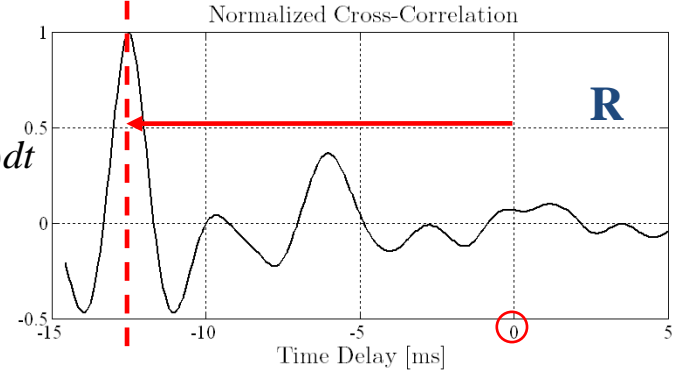


Localisation : approche hyperbolique

Estimation TDOA: fonction d'inter-corrélation



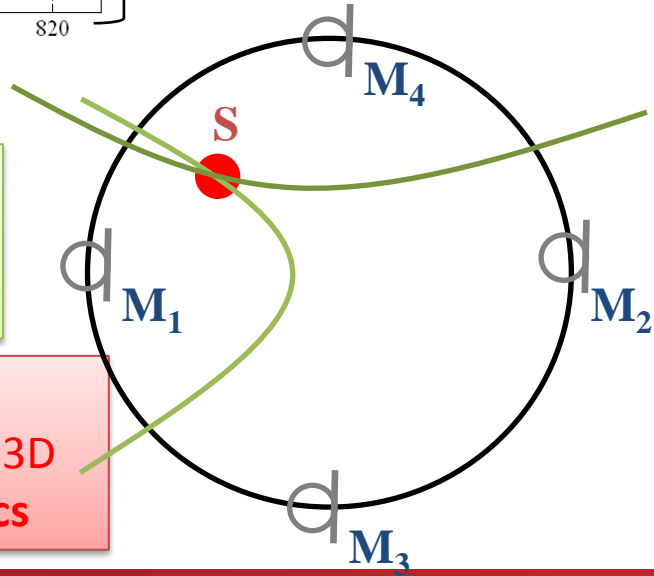
$$R(\tau) = \int_{-\infty}^{+\infty} p_1(t) p_2(t - \tau) dt$$
$$\hat{\tau}_{12} = \arg \max_{\tau} R_{\tau}$$



τ_{12}

Avantages:
- rapide
- Précis en champ proche

Inconvénients:
- Analytiquement complexe en 3D
- Suppose une recherche de pics



Il faut faciliter la recherche de pics !

■ ■ ■ ■ ■ Pondération PHAT

Estimation TDOA: fonction d'inter-corrélation généralisée

$$R(\tau) = \int_{-\infty}^{+\infty} p_1(t) p_2(t - \tau) dt$$

$$R(\tau) = \int_{-\infty}^{+\infty} P_1(f) P_2^*(f) e^{2j\pi f\tau} df$$

$$R(\tau) = \int_{-\infty}^{+\infty} \psi(f) P_1(f) P_2^*(f) e^{2j\pi f\tau} df$$

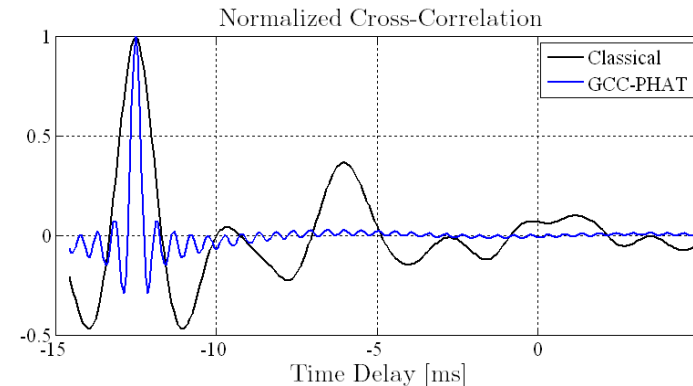
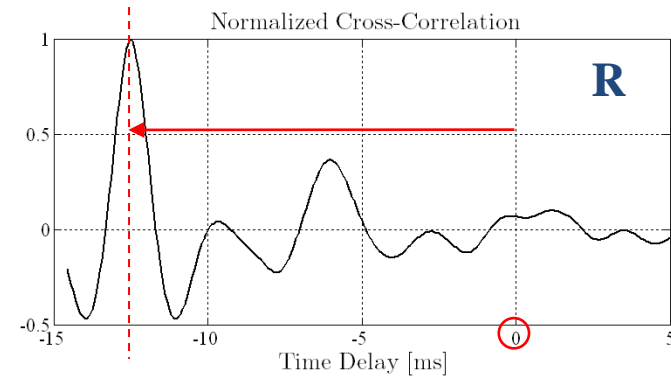
$$\psi(f) = \frac{1}{|P_1(f) P_2^*(f)|}$$



Fourier Domain

Generalized Cross Correlation

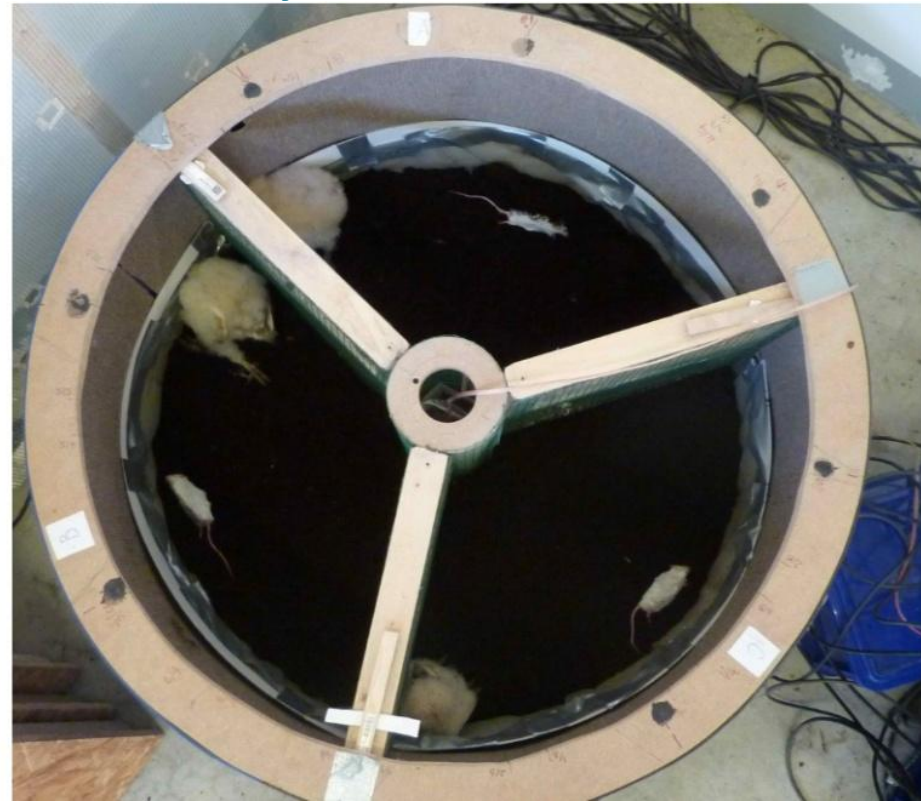
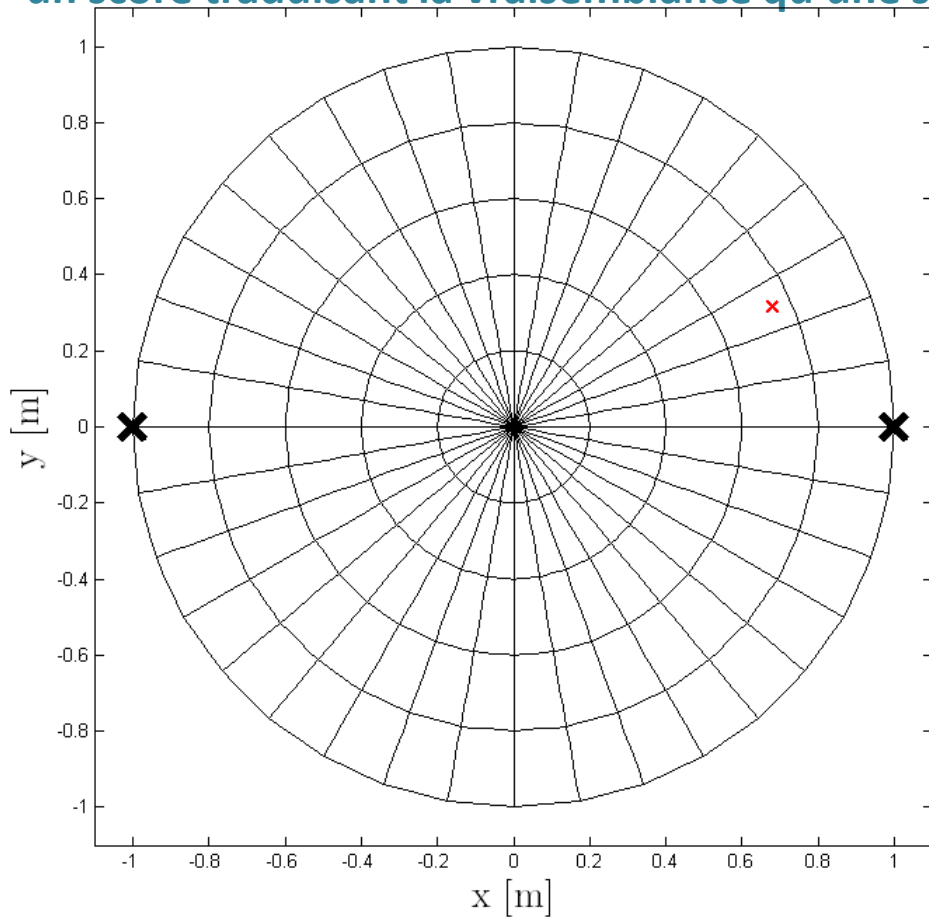
PHase Transform ponderation (GCC-PHAT), Knapp1976



Vraisemblance spatiale (SRP-PHAT)



Idée: quadriller une zone de recherche en petites parcelles puis attribuer à chaque parcelle un score traduisant la vraisemblance qu'une source sonore s'y trouve.

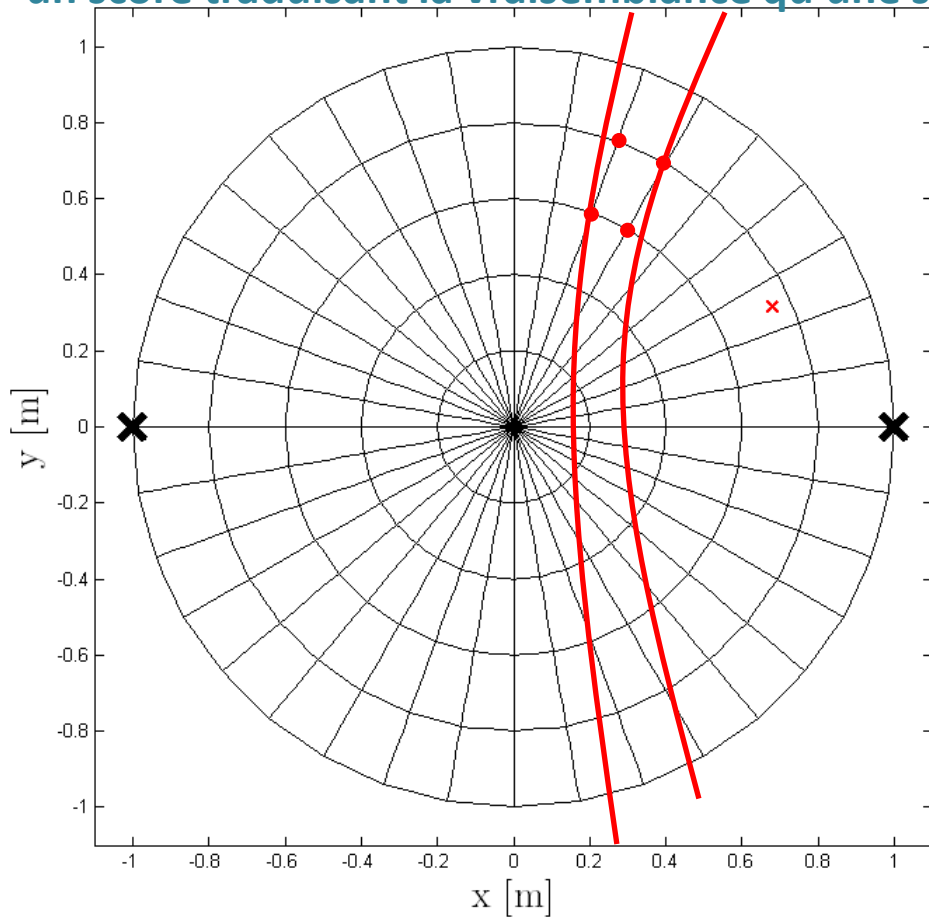




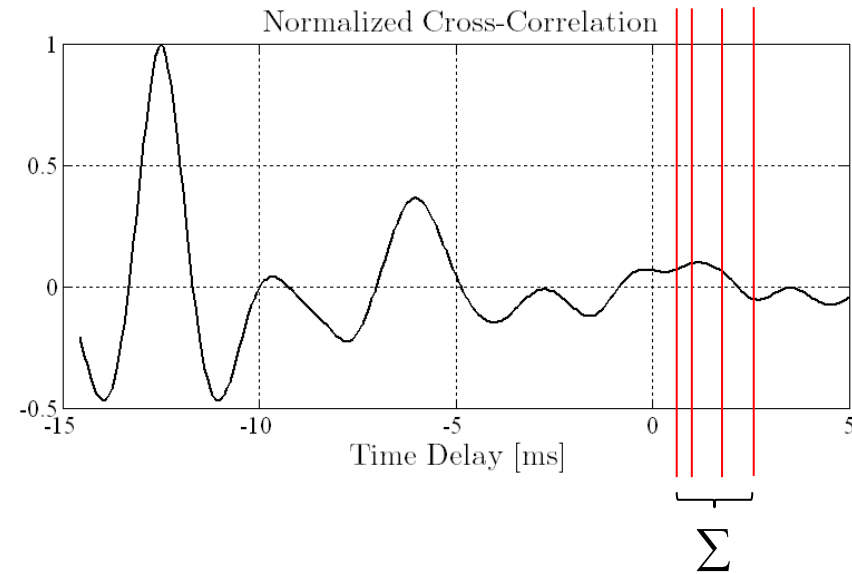
Vraisemblance spatiale (SRP-PHAT)



Idée: quadriller une zone de recherche en petites parcelles puis attribuer à chaque parcelle un score traduisant la vraisemblance qu'une source sonore s'y trouve.

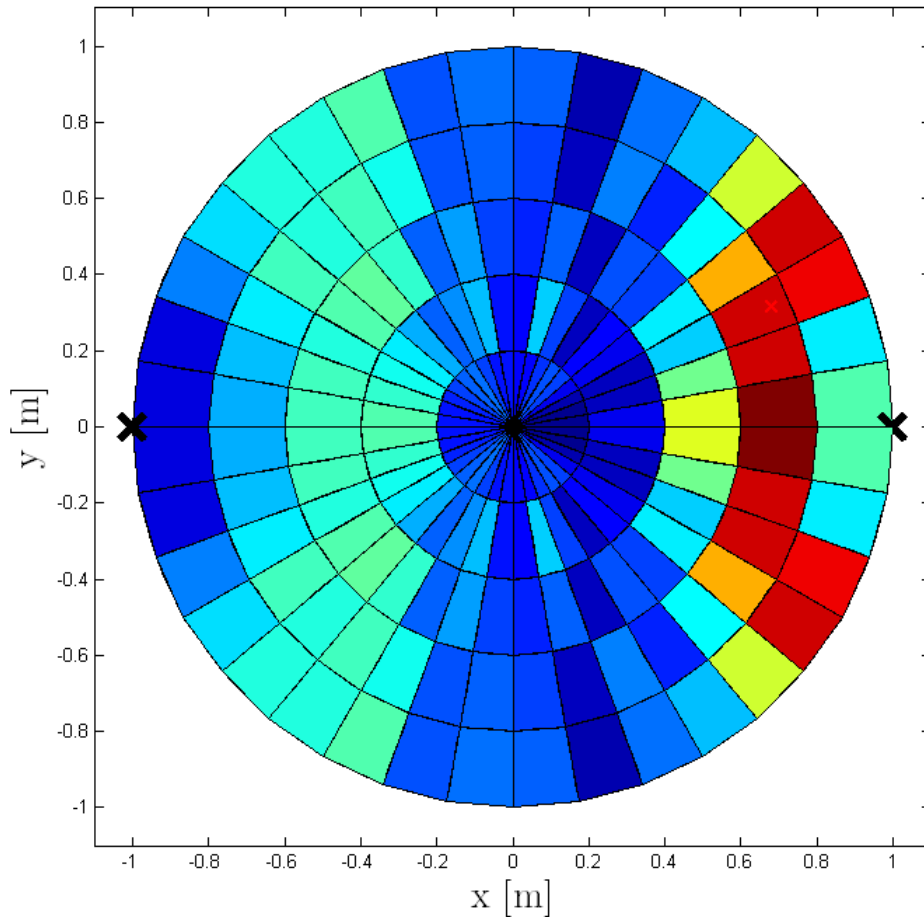


M paires de microphones ($m = 1, \dots, M$)
N cellules ($n = 1, \dots, N$)

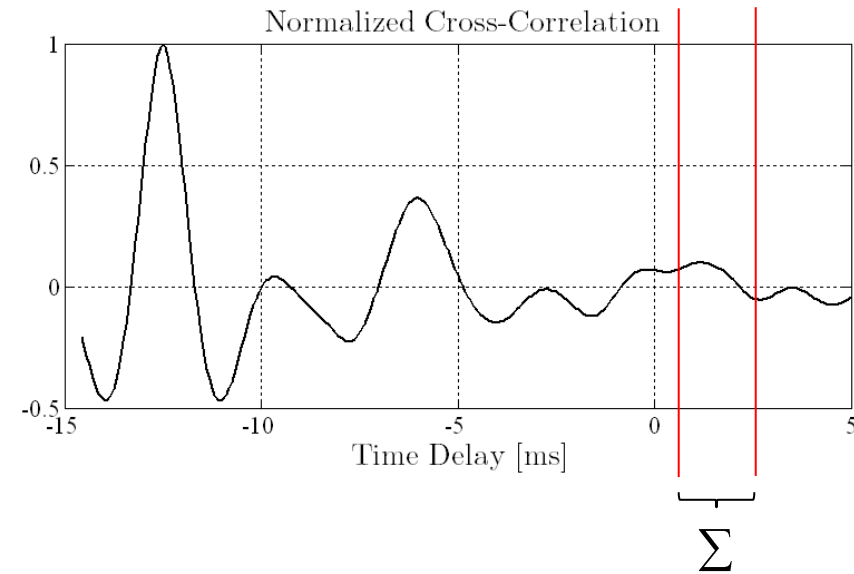


Vraisemblance spatiale (SRP-PHAT)

Autres noms dans la littérature: « Global Coherence Field » ou « Matched Field Processing »



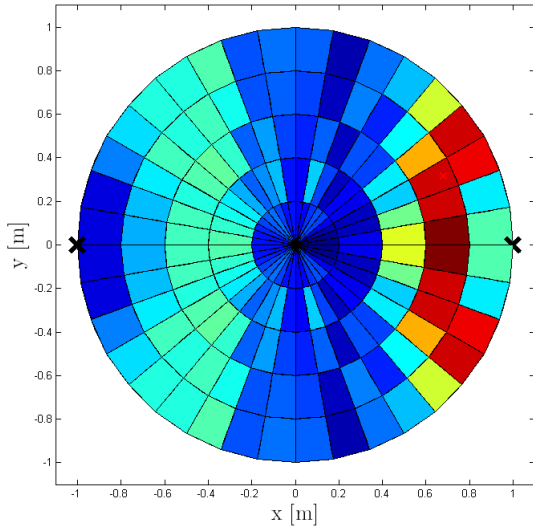
M paires de microphones ($m = 1, \dots, M$)
 N cellules ($n = 1, \dots, N$)



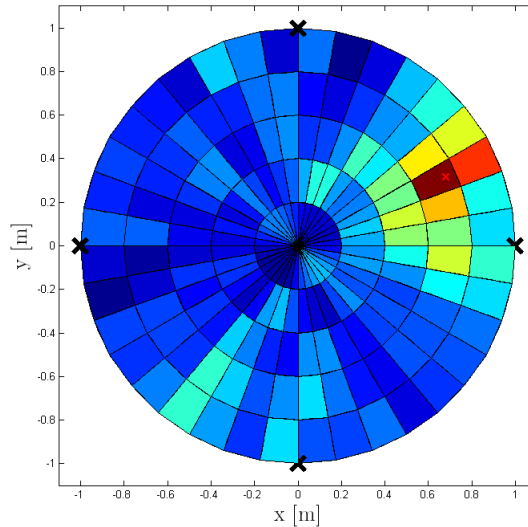
$$GCF(n) = \sum_{m=1}^M \sum_{k=k_n^{\min}}^{k_n^{\max}} R_m(k)$$

Localisation: méthode SRP-PHAT (3/3)

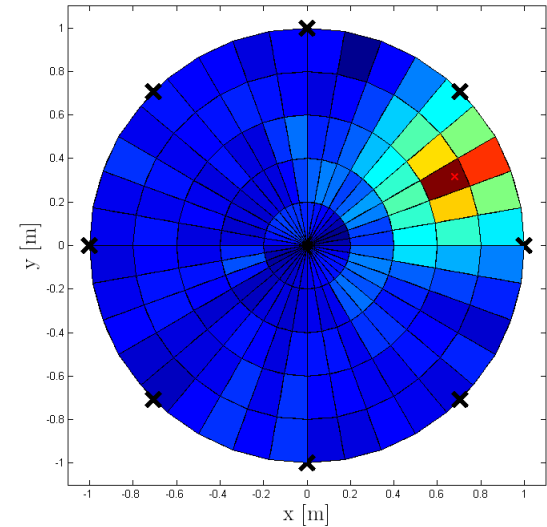
Plus de paires = moins d'ambiguïtés (on bénéficie de la redondance d'information)



2 microphones = 1 paire



4 microphones = 6 paires



8 microphones = 28 paires

Algorithme de localisation:

SRP-PHAT

- Recherche grossière
- Contraint la recherche de pics

HYPERBOLE

- Localisation précise

Extraction

Algorithme d'extraction:

BEAMFORMING

DENOISING

-Réhausse le signal d'intérêt
-Permet de modéliser le bruit

-Sépare signal du bruit

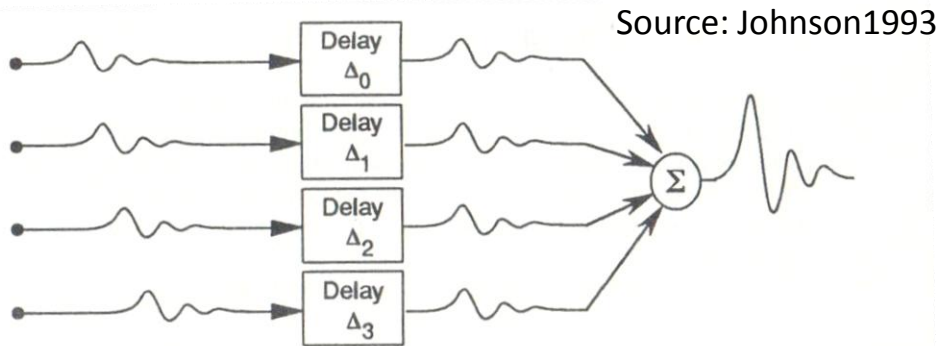
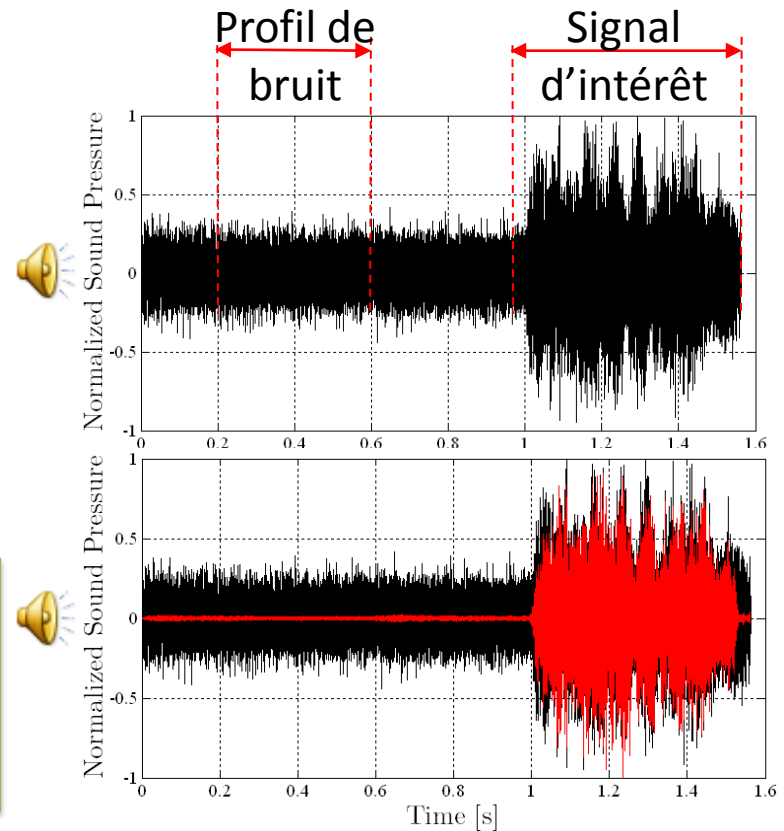


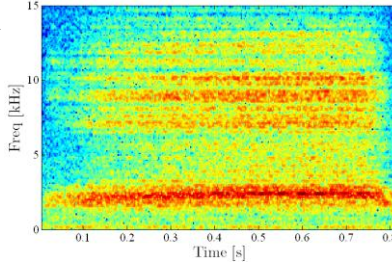
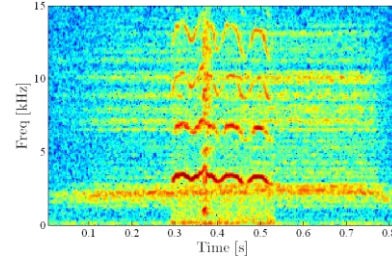
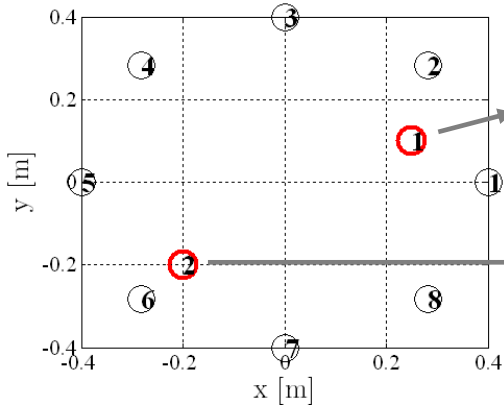
Figure 4.4 Adjusting the signal processing delays Δ_m to reinforce the signal is sometimes called *stacking*. As shown, the maximum-energy choice for Δ_m aligns the sensor outputs so that the sensors' waveforms resemble the signal waveform stacked on itself.

Beamforming:
VanVeen1988

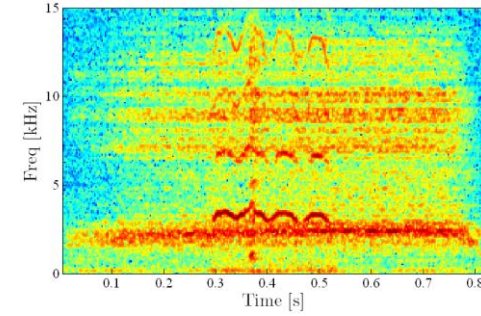
Denoising:
Soustraction spectrale (Boll1979)
Filtrage de Wiener (Wiener1949)
Log-STSA (Ephraim1985)



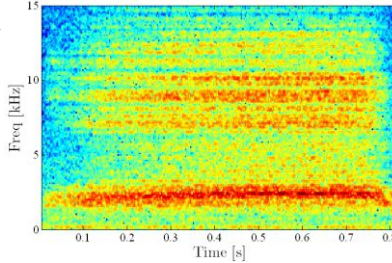
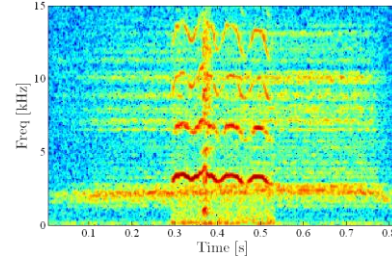
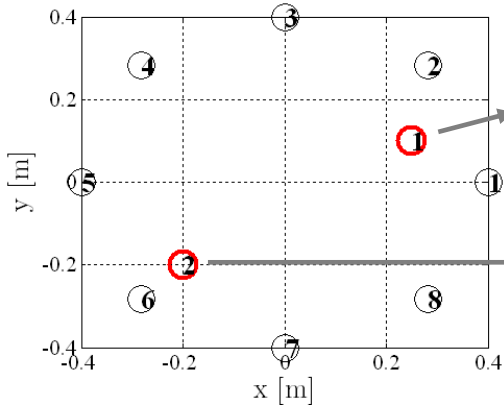
Extraction: simulation



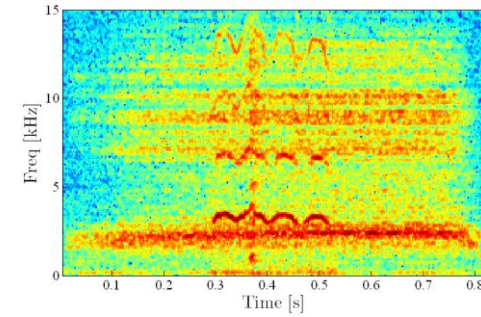
Sur les microphones 1 à 8:



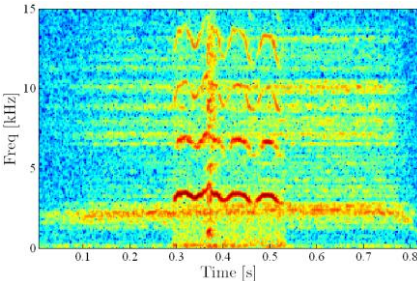
Extraction: simulation



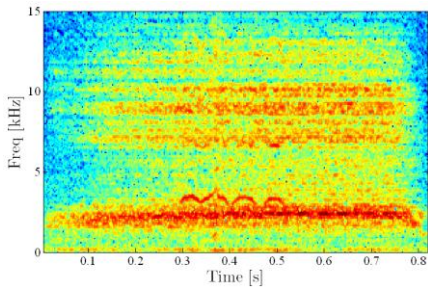
Sur les microphones 1 à 8:



BF sur source 1



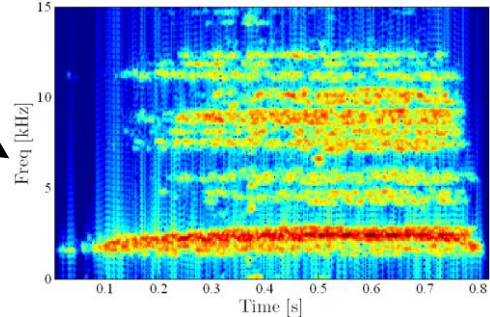
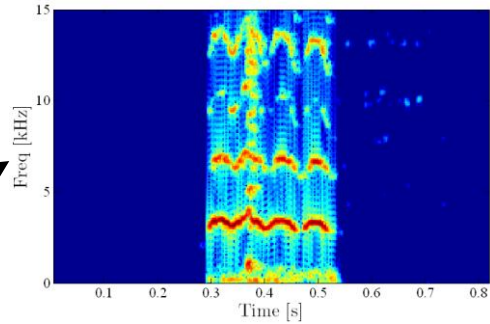
BF sur source 2



Signal d'intérêt

Denoising

Profil de bruit





Classification

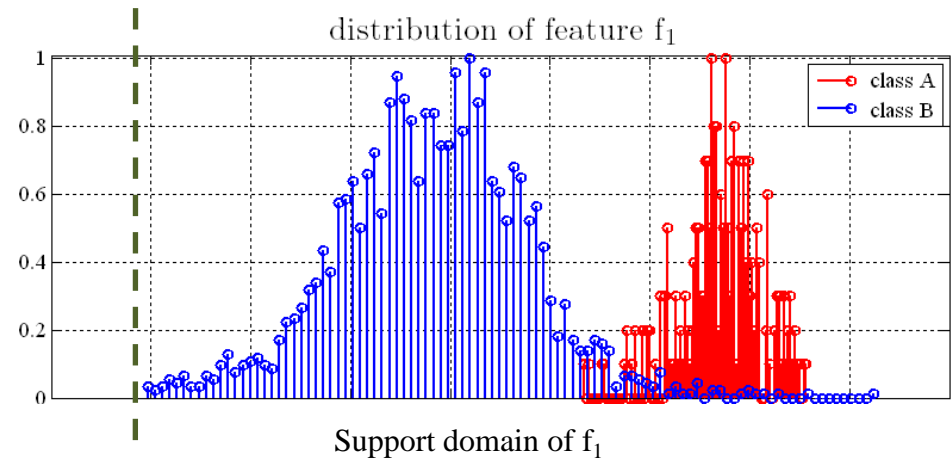
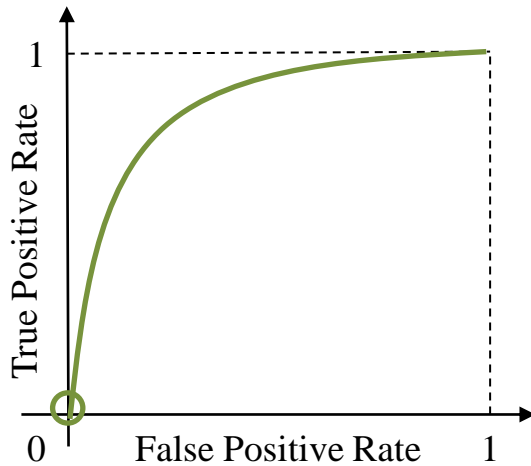


Figure 1: Basic structure of a sound classification system comprising feature extraction, classification, and postprocessing.

Source: Allegro2001

Classificateur: classification par seuillage sur histogramme

Choix du seuil: Analyse ROC



Classification

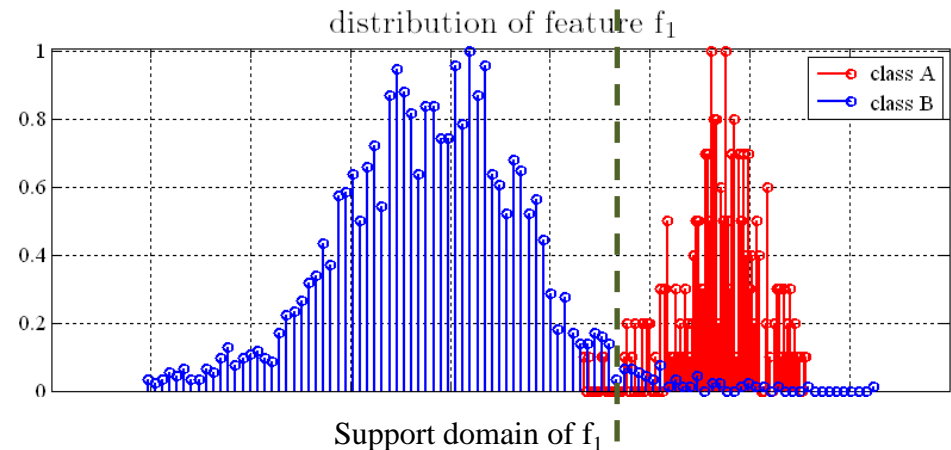
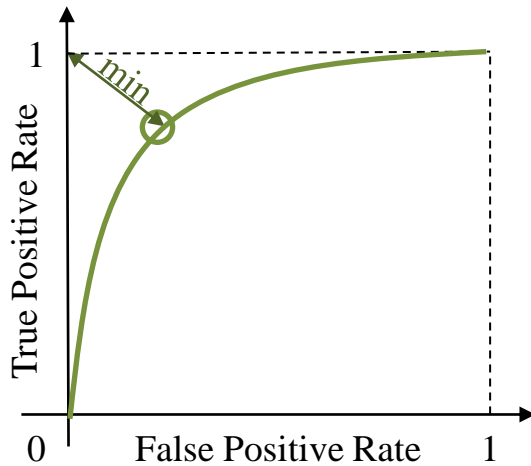


Figure 1: Basic structure of a sound classification system comprising feature extraction, classification, and postprocessing.

Source: Allegro2001

Classificateur: classification par seuillage sur histogramme

Choix du seuil: Analyse ROC



Classification

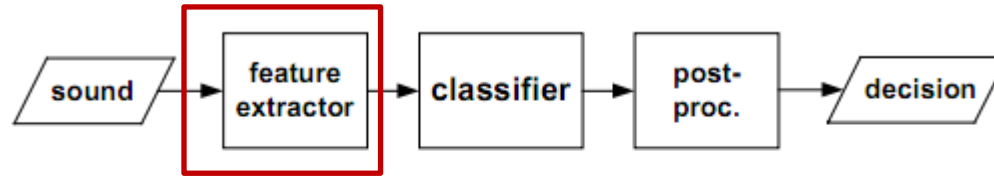


Figure 1: Basic structure of a sound classification system comprising feature extraction, classification, and postprocessing.

Source: Allegro2001

Descripteurs acoustiques:

- Leq: Niveau Sonore Equivalent (dB)
- Fleq: Niveau Sonore Equivalent sur une certaine bande de fréquences (dB)
- CGT: Centre de Gravité Temporel (s)
- HAR: Harmonicité (coefficient)
- ROF: Point de Roll-Off Fréquentiel (Hz)
- GCS: Centre de Gravité Spectral (Hz)
- ... Peeters2004

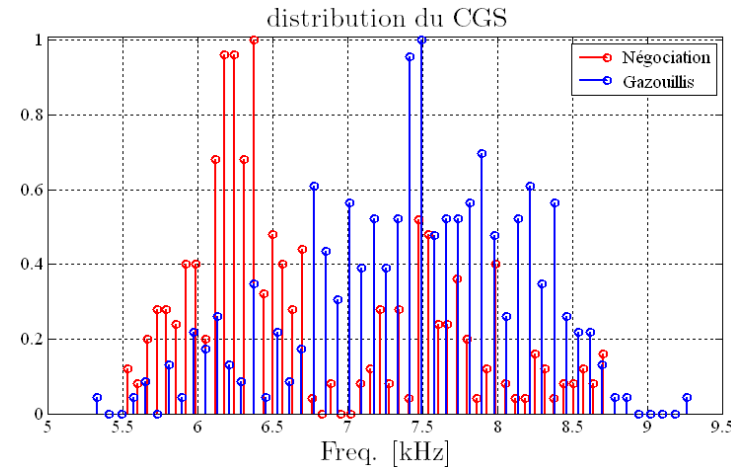
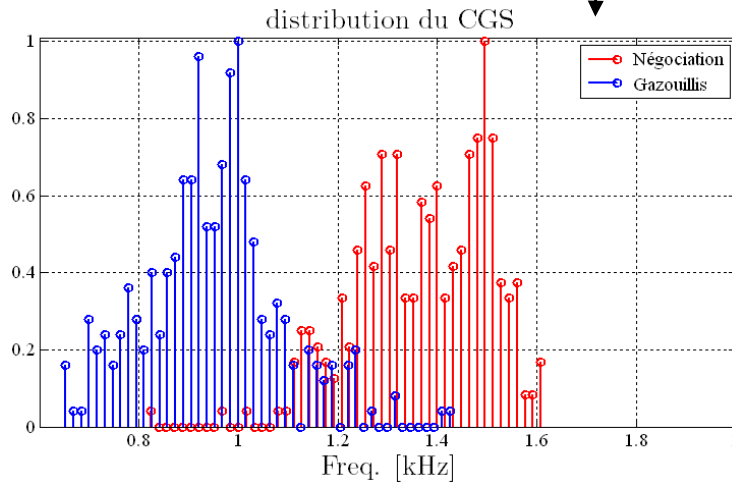
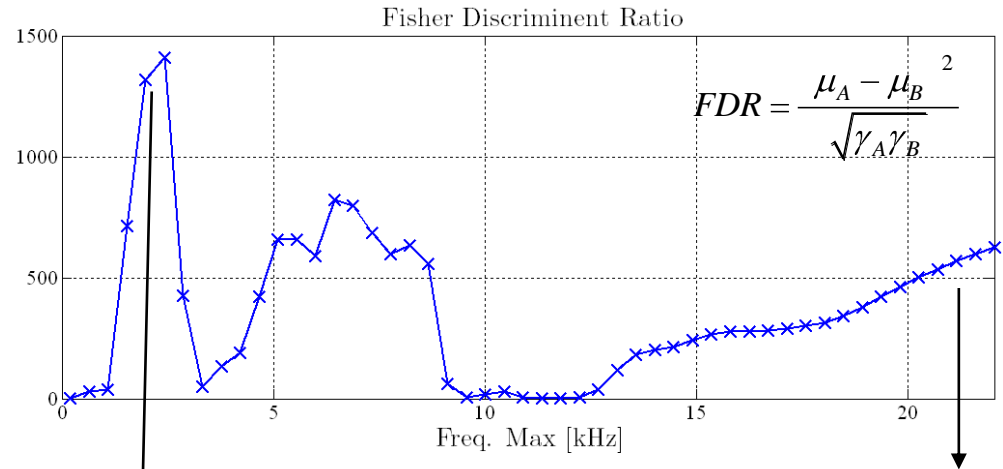
Optimiser les paramètres en utilisant le Fisher Discriminant Ratio (FDR)

Classification

Optimiser les paramètres (par ex. *Fisher Discriminant Ratio* ou *Kullback-Leibler Divergence*)

$$GCS(P(f)) = \sum_{f=f_{min}}^{f_{max}} \frac{f \times P(f)}{P(f)}$$

= barycentre du spectre



Conclusion

- ▶ Un banc d' à été instrumentalisé et des méthodes ont été développées pour
 - Localiser: localisation par SRP-PHAT puis par intersection d'hyperboles
 - Extraction: association du Beamforming et du Denoising
 - Classification: par seuillage

- ▶ **A venir:**
 - Qualifier et améliorer la détection sur les campagnes de mesures.
 - Classification plus « intelligente » (plus en plus efficace au fur et à mesures des enregistrements).
 - Etablir des statistiques pertinentes sur la communication entre chouettes.

Bibliography

Bir2011: Acoustic source direction by hemisphere sampling, Stanley T. Birchfield and Daniel Kahn Gillmor

Spiesberger2001: Hyperbolic location errors due to insufficient numbers of receivers, John L. Spiesberger, JASA 2001

Neering2009: Optimization and estimation techniques for passive acoustic source localization, phd Thesis, 2009

Benard-Caudal2009: Highly defined whale group tracking by passive acoustic Stochastic Matched Filter, F. Benard-Caudal & H. Glotin, 2009

Yang2005: Cramer-Rao Bound and Optimum Sensor Array for Source Localization from Time Differences of Arrival, B. Yang and J. Scheuing

Spiesberger2004: Geometry of locating sounds from differences in travel time: Isodiachrons, John L. Spiesberger, JASA 2004

Johnson1993: Array signal processing: concepts and techniques, book, Don H. Johnson and Dan. E. Dudgeon 1993, Prentice Hall PTR

Boll1979: Suppression of Acoustic Noise in Speech Using Spectral Subtraction, 1979, Steven F. Boll

Ephraim1985: Speech Enhancement Using a Minimum Mean-Square Error Log-Spectral Amplitude Estimator, 1985, Ephraim and Malah

Wiener1949: Extrapolation, Interpolation, and Smoothing of Stationary Time Series. N. Wiener, New York: Wiley 1949

Allegro2001: Automatic Sound Classification Inspired by Auditory Scene Analysis, Silvia Allegro, Michael Büchler and Stefan Launer, 2001

Peeters2004: A Large Set Of Audio Features For Sound Description (similarity and classification) in the CUIDADO project, G. Peeters, 2004

Knapp1976: The generalized correlation method for estimation of time delay, C. Knapp and G. Carter, 1976

VanVeen1988: Beamforming: a versatile approach to spatial filtering, Van Deen and Buckley, 1988



**Merci pour votre
attention !!**

**N'ayez pas peur
de poser vos
questions...**