

Maltraitement du signal

Partie 1 : prolégomènes

Gaël Mahé

Université Paris Cité

octobre 2022

Plan

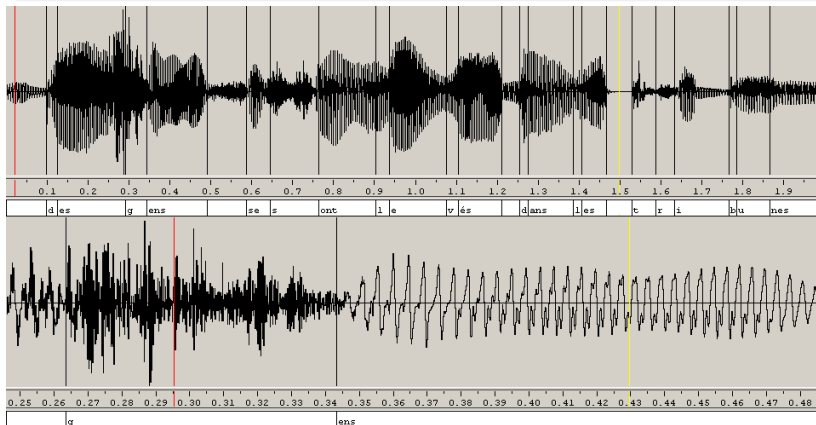
- 1 Représentations des signaux audio
- 2 L'oreille, récepteur imparfait
- 3 Maltraiter le signal
- 4 Quelques notions sur la qualité audio

Plan

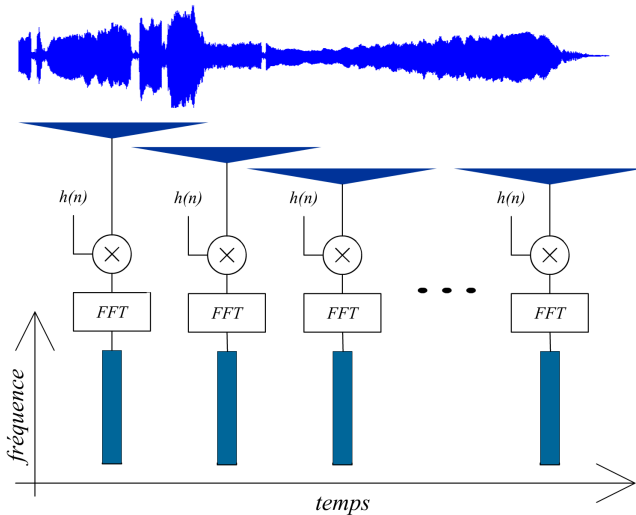
- 1 Représentations des signaux audio
- 2 L'oreille, récepteur imparfait
- 3 Maltraiter le signal
- 4 Quelques notions sur la qualité audio

Stationnarité locale

- Les signaux audio sont **non-stationnaires**
- mais peuvent être considérés **localement** stationnaires (sur quelques dizaines de ms).



Transformée de Fourier à Court Terme

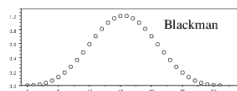
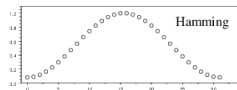
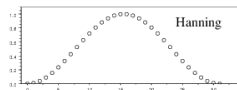
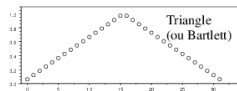
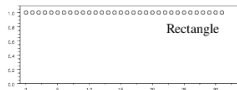


Rappels sur l'analyse spectrale numérique

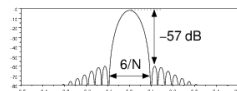
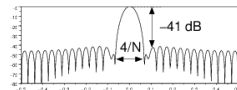
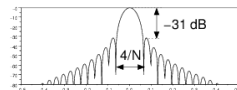
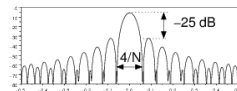
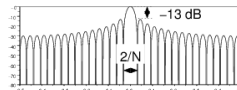
- Spectre de l'observation : $X'(f) = X(f) * H(f)$
avec $|H(f)| = |\text{sinc}(\pi Nf)|$ pour une fenêtre rectangulaire
- TFD \rightarrow on n'observe que des **échantillons** du spectre !
- Amélioration visualisation via **zéropadding** :
 - Compléter $(x(n))_{0 \leq n \leq N-1}$ par N' zéros
 - TFD de x complété
 - ▶ Spectre (TFTD) de x reste le même : $X(f) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi nf}$
 - ▶ Mais spectre discrétisé (TFD) sur $N + N'$ points, espacés de $\frac{1}{N+N'}$
- Reste problème de la **résolution**
 - **fréquentielle** : écart fréquentiel minimal entre 2 raies
 - **en amplitude** : écart d'amplitude maximal entre 2 raies

Fenêtres d'analyse et leurs spectres d'amplitude.

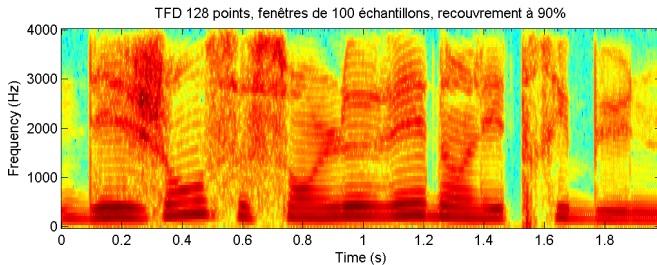
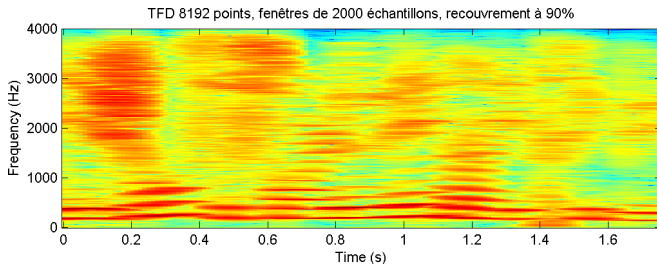
Fenêtres



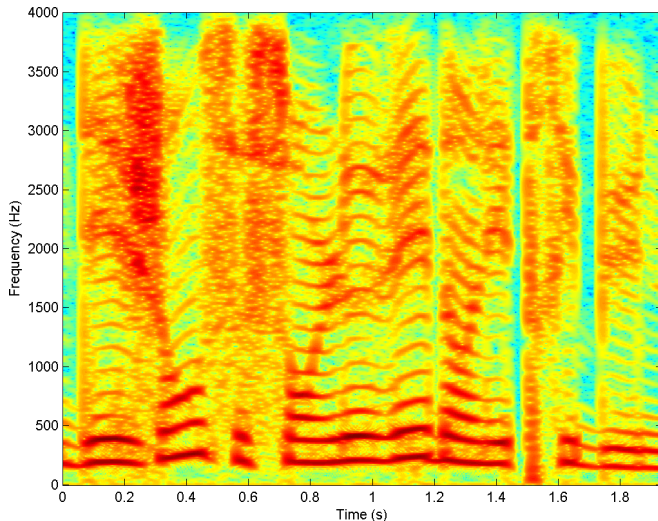
Spectres d'amplitude (en dB)



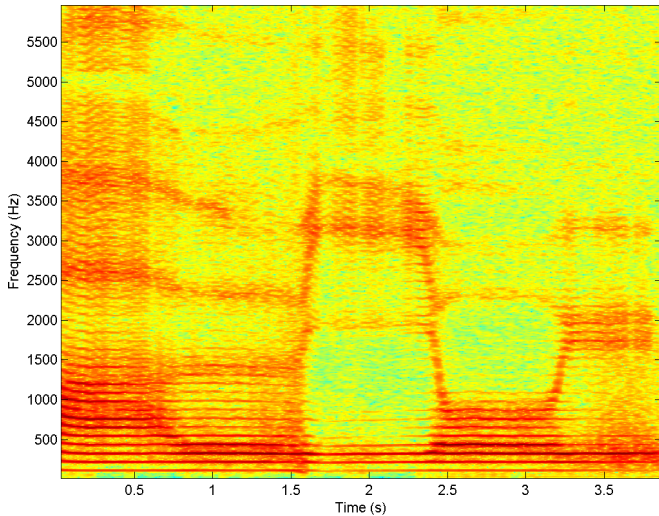
Attention aux paramètres !



Compromis résolution temporelle / fréquentielle



Voyelles : "a e i o u"



Autres représentations

- Modified Discrete Fourier Transform (MDCT)
- Ondelettes
- Décompositions *ad hoc* :
 - Modèle auto-régressif (AR)
 - Bancs de filtres simulant les filtres cochléaires
 - Décomposition enveloppe / structure temporelle fine

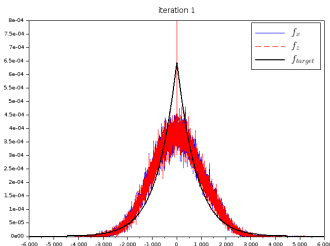
Histogramme

Les échantillons temporels $x(n)$ et temps-fréquence $|X(t,f)|$ suivent une loi gaussienne généralisée :

$$f(x) = \frac{\beta}{2\alpha\Gamma(1/\beta)} \exp\left(-\left(\frac{|x-\mu|}{\alpha}\right)^\beta\right), \quad \text{avec :}$$

- α = facteur d'échelle (lié à la variance)
- β = facteur de forme

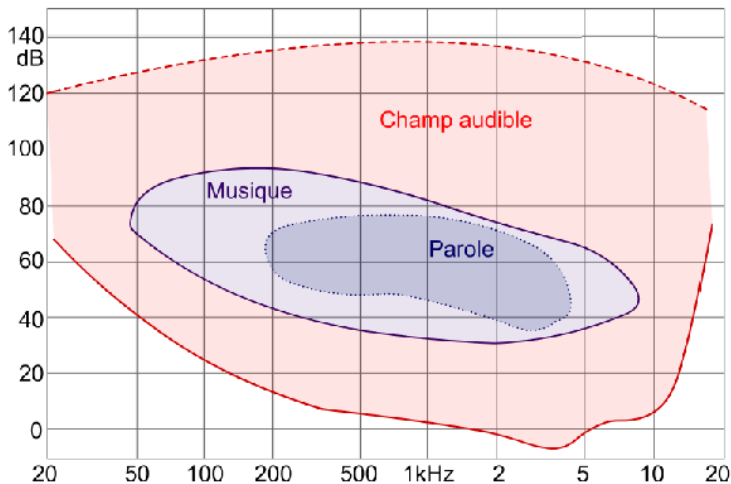
Exemple : histogramme (éch. temporels) de 10s de piano à 32kHz :



Plan

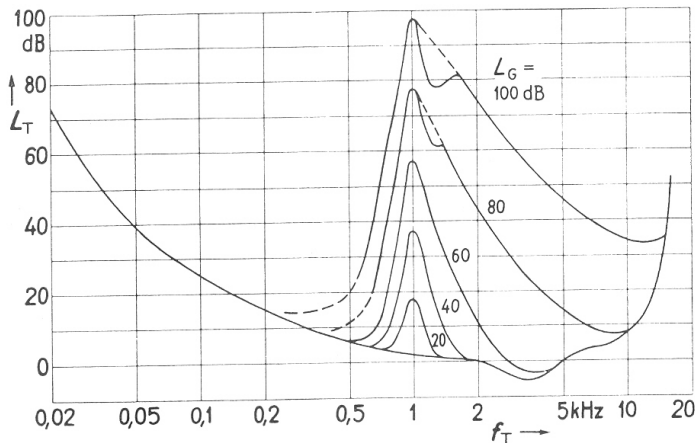
- 1 Représentations des signaux audio
- 2 L'oreille, récepteur imparfait**
- 3 Maltraiter le signal
- 4 Quelques notions sur la qualité audio

Le champ audible



Masquage simultané (fréquentiel)

Seuil d'audition en présence d'un bruit de bande étroite (160 kHz) de fréquence centrale 1 kHz :



Masquage simultané (2)

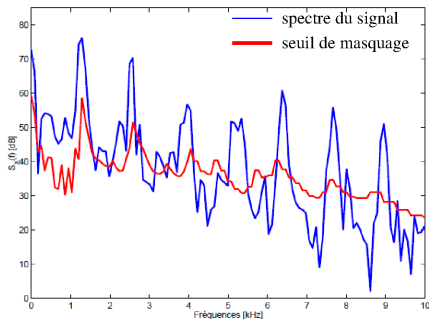
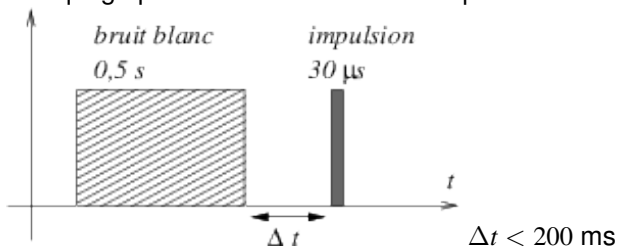


Figure – Seuil de masquage d'un son complexe.

Le signal masque tout autre signal
dont le spectre est $<$ seuil de masquage.

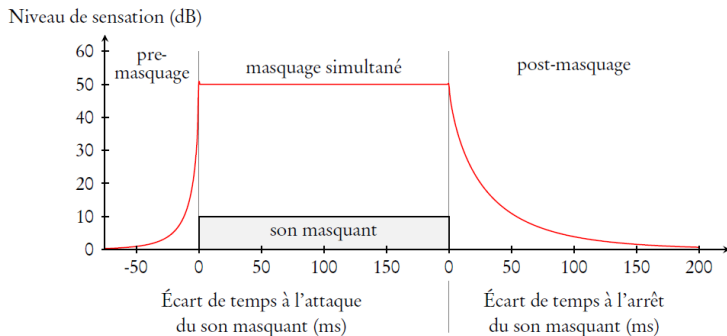
Masquage temporel

- Masquage postérieur : le 1er son masque le 2e



- Masquage antérieur : le 2e son masque le 1er
 $\Delta t < 40$ ms

Masquage : en résumé



Plan

- 1 Représentations des signaux audio
- 2 L'oreille, récepteur imparfait
- 3 Maltraiter le signal**
- 4 Quelques notions sur la qualité audio

On peut donc modifier imperceptiblement le signal

- **Comment ?**

- en ajoutant du bruit
- en supprimant des composantes inaudibles

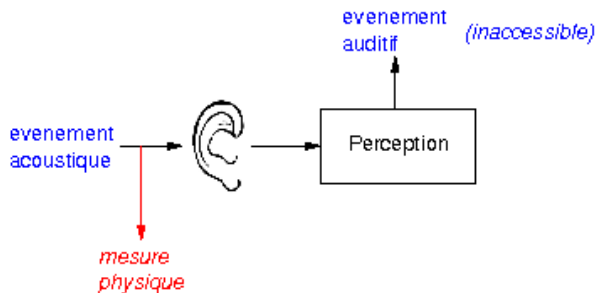
- **Pour quoi ?**

- Compression
- Tatouage (watermarking)
- Conditionnement

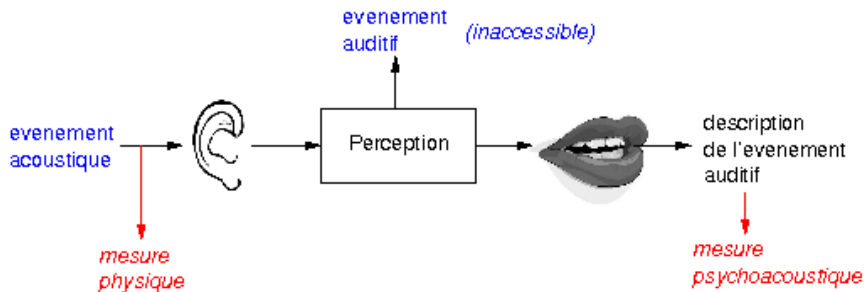
Plan

- 1 Représentations des signaux audio
- 2 L'oreille, récepteur imparfait
- 3 Maltraiter le signal
- 4 Quelques notions sur la qualité audio**

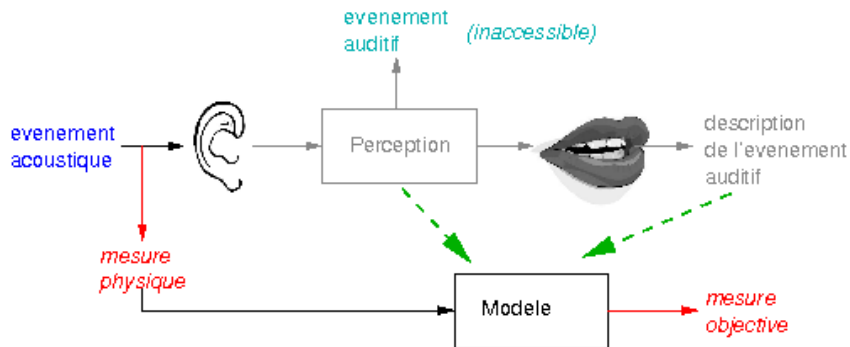
Comment mesurer la qualité audio ?



Comment mesurer la qualité audio ?



Comment mesurer la qualité audio ?



Modèles instrumentaux paramétriques

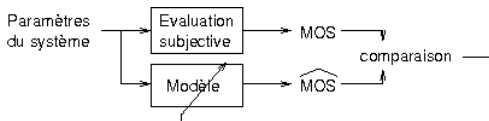
Principe

Paramètres physiques du système → note MOS

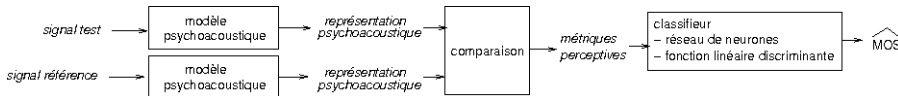
Paramètres physiques = délai, écho, bruit, atténuation...

Ex : norme ITU-T G.107

Apprentissage du modèle :



Modèles instrumentaux fondés sur les signaux



Ex :

- Pour la parole : PESQ (Perceptual Evaluation of Speech Quality) = ITU-T P.862
- Pour l'audio : PEAQ (Perceptual Evaluation of Audio Quality) = ITU-R BS.1387