

Deep learning par la pratique

Partie 1 : Deep learning appliqué au son



Présenté par **Morgan Gautherot**

Comprendre une onde sonore



Partie 1 : Deep learning appliqué au son



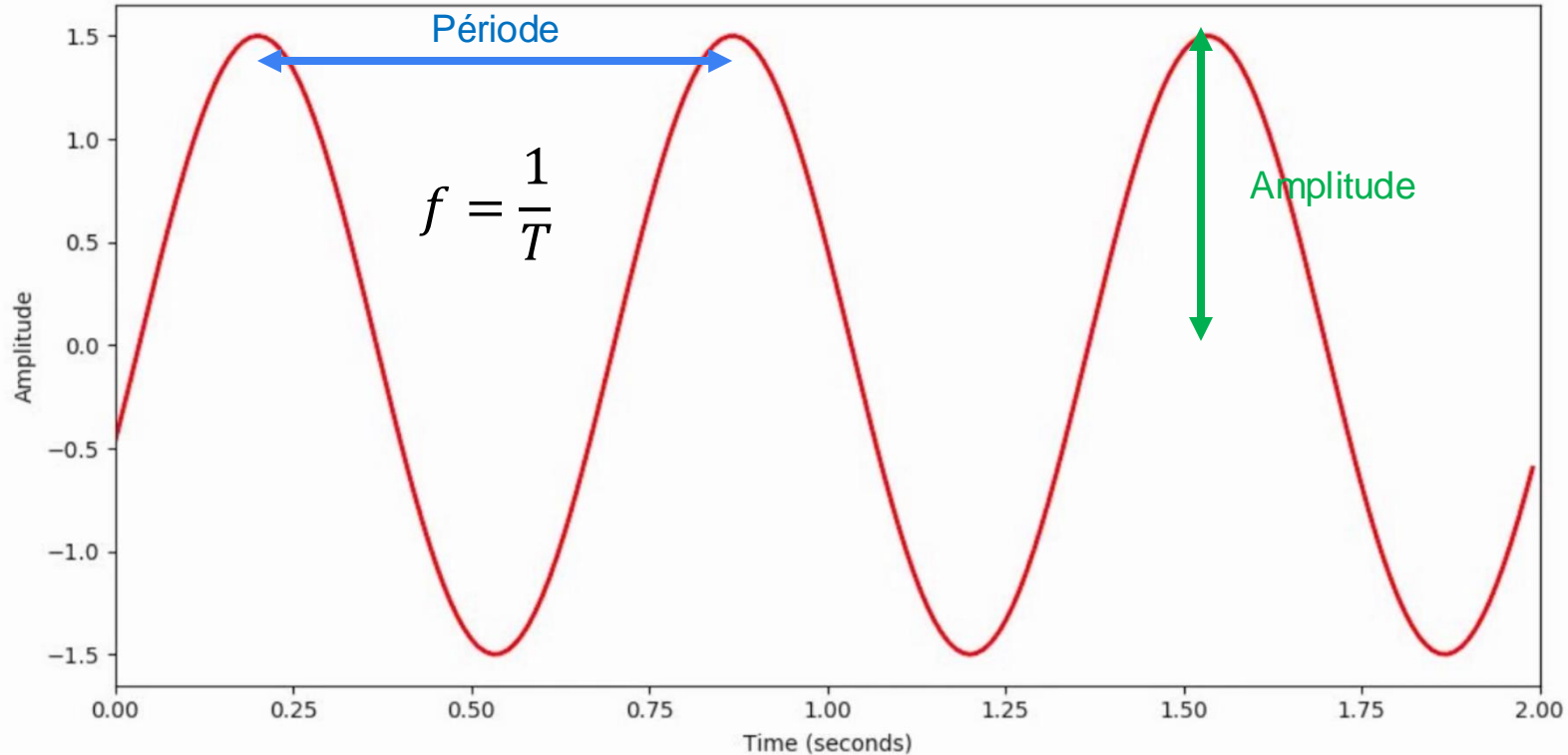
Le son

- Produit par la vibration d'un objet
- Les vibrations déterminent l'oscillation des molécules d'air
- L'alternance de la prédiction de l'air provoque une onde

$$y(t) = A \cdot \sin(2\pi f t + \varphi)$$

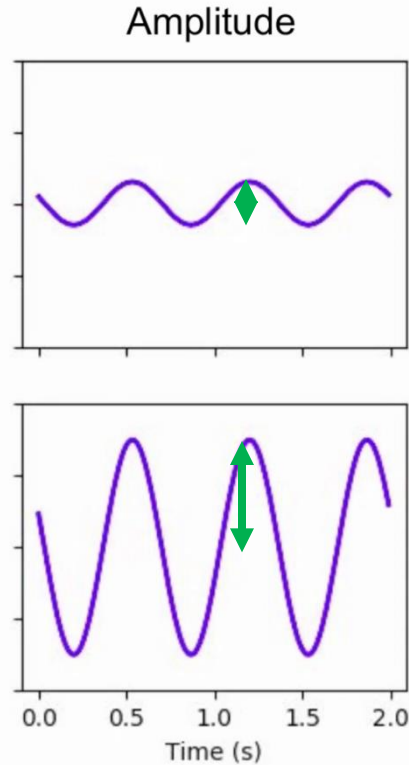
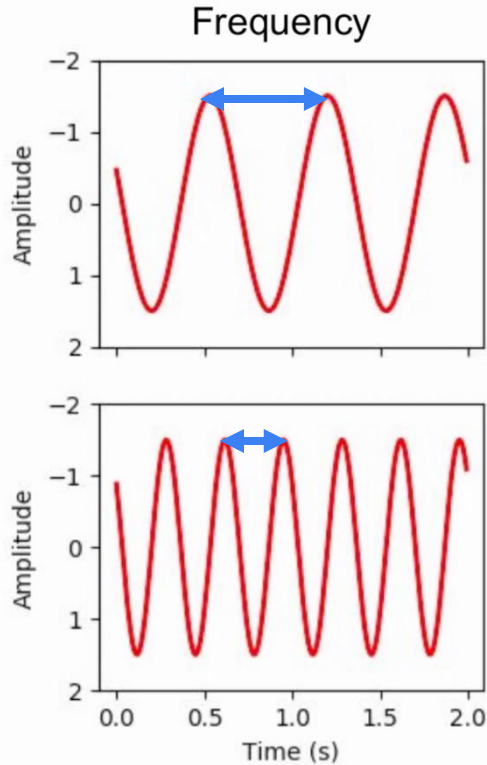


Onde





Fréquence/hauteur et amplitude/intensité

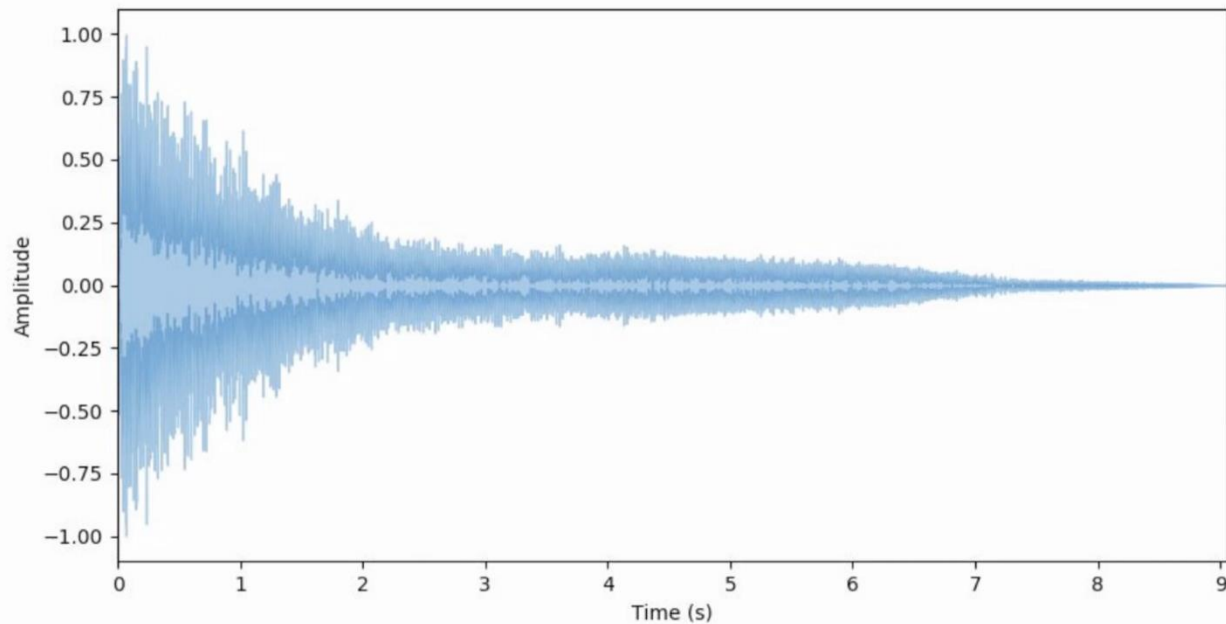


Plus grande fréquence -> son plus aigu

Plus grande amplitude -> Volume plus important



Une onde réel (note de piano)





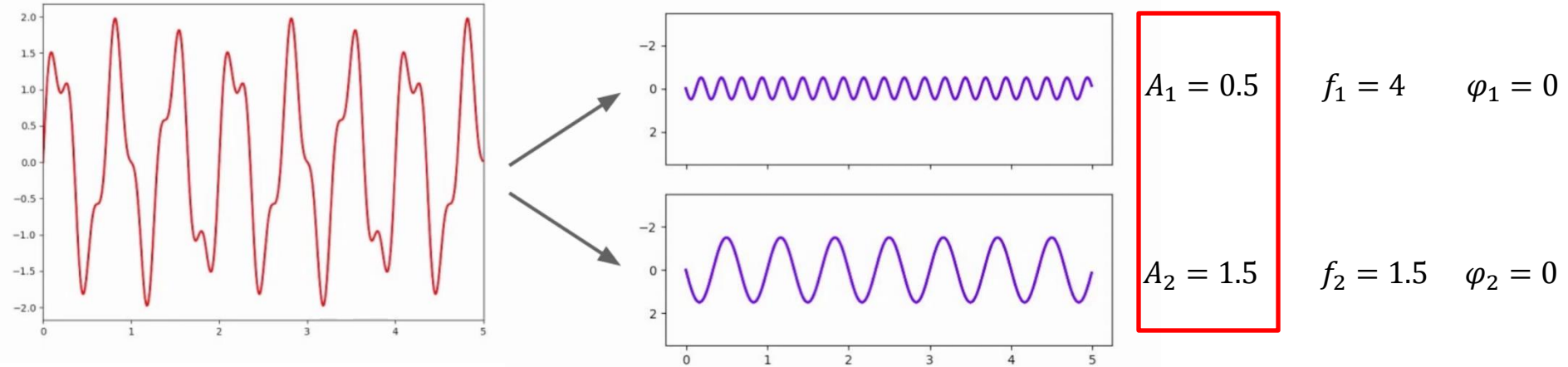
La transformation de Fourier

Décomposer un son périodique complexe en une somme d'ondes sinusoïdales oscillant à différentes fréquences.



La transformation de Fourier

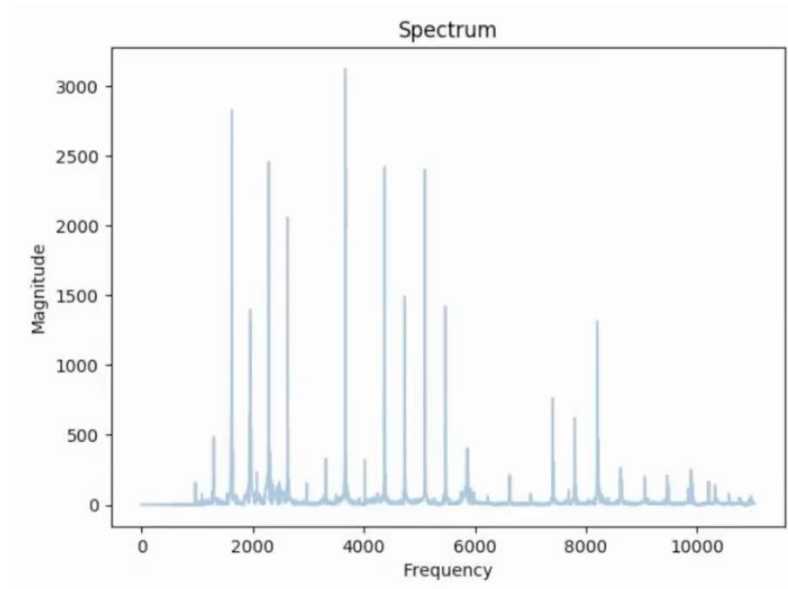
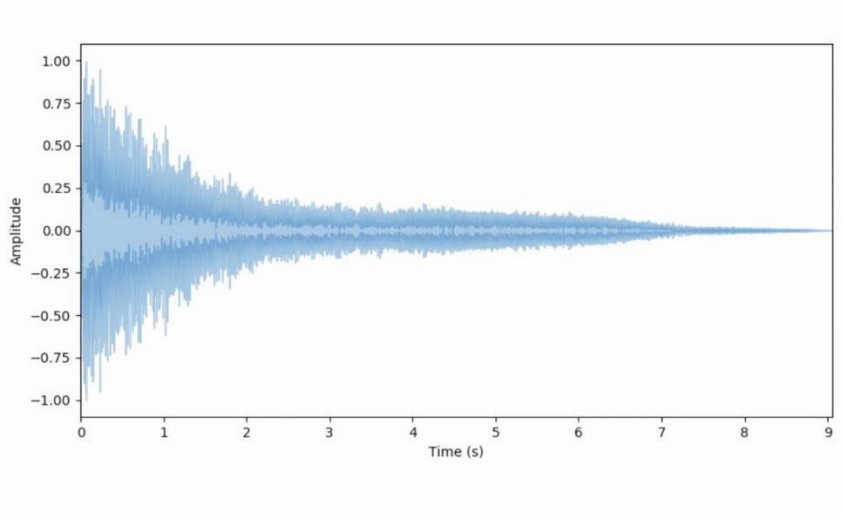
Plus l'amplitude est élevée plus le son contribue à la complexité du son d'origine



$$S = A_1 \sin(2\pi f_1 t + \varphi_1) + A_2 \sin(2\pi f_2 t + \varphi_2)$$



La transformation de Fourier d'une onde réelle



Perte d'information sur le temps

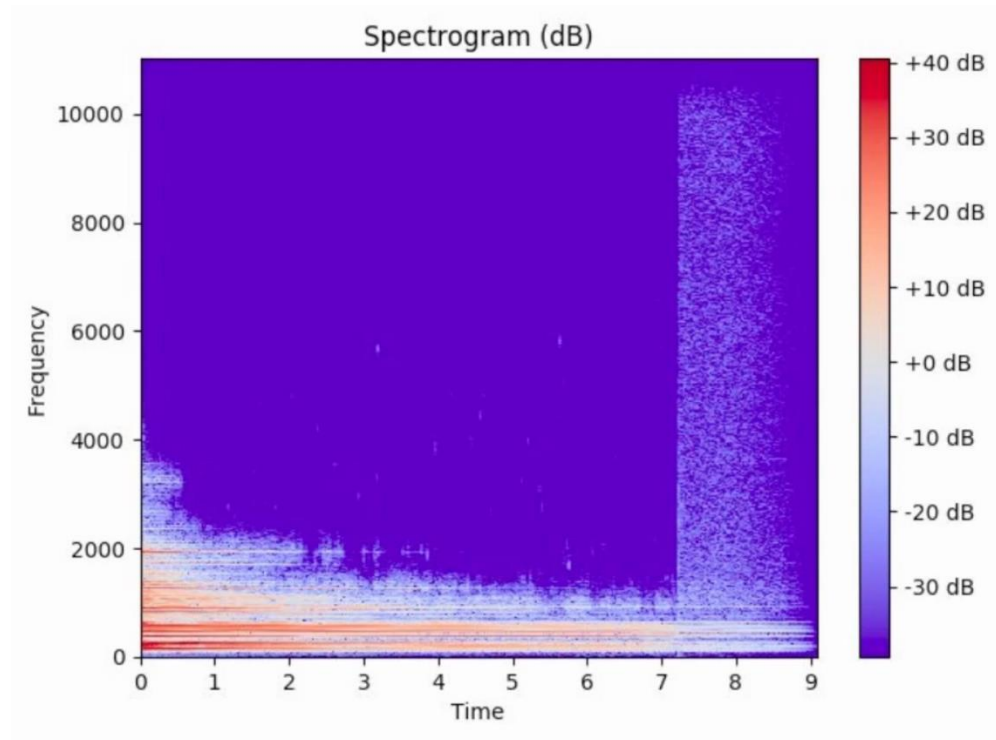


Transformée de Fourier à temps court (STFT)

- Calcul de plusieurs FFT à différents intervalles
- Préserve les informations temporelles
- Taille de la fenêtre de temps fixe
- Donne un spectrogramme (temps + fréquence + amplitude)

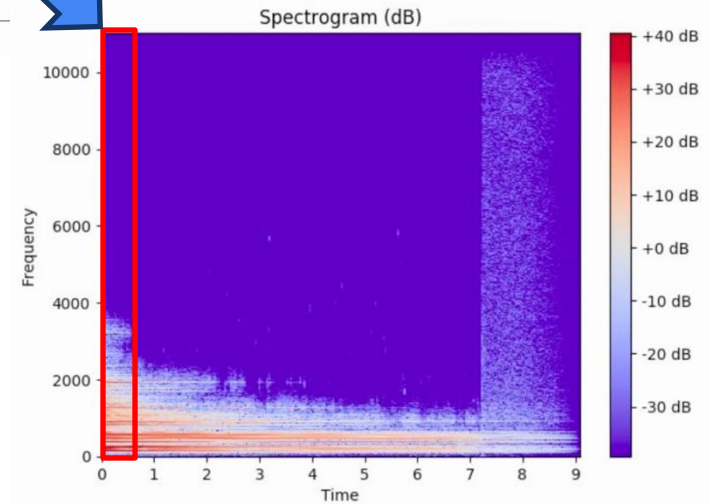
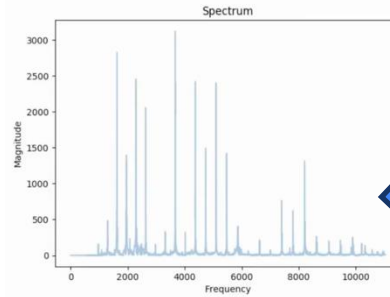
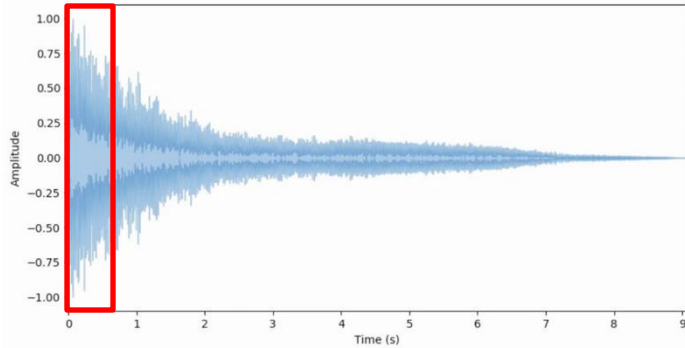


Transformée de Fourier à temps court (STFT)



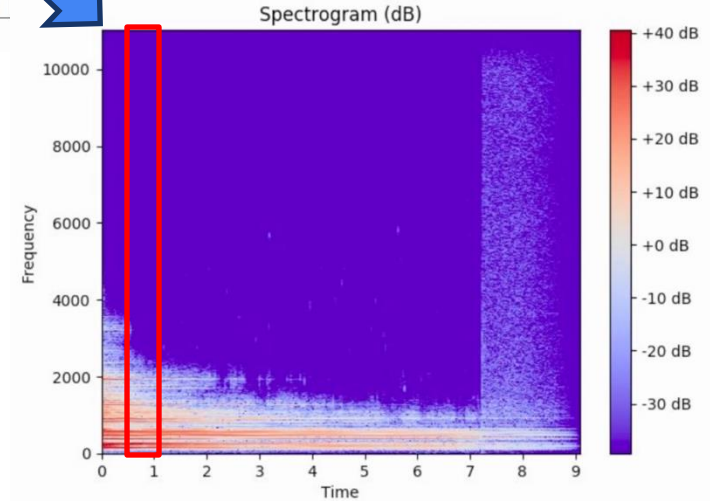
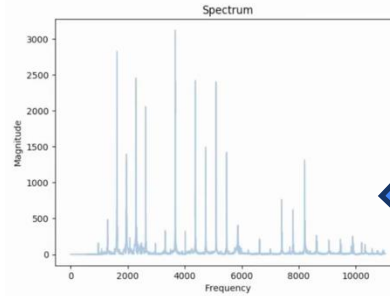
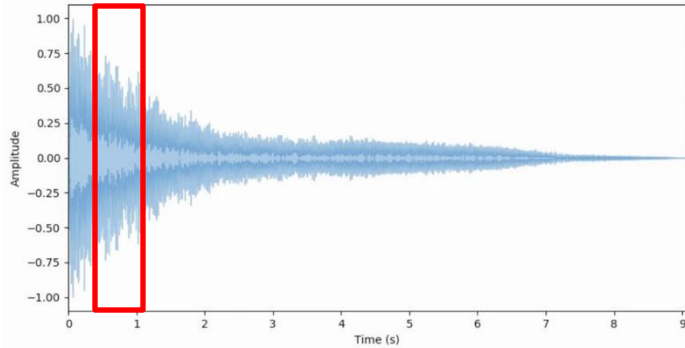


D'une onde au spectrogramme



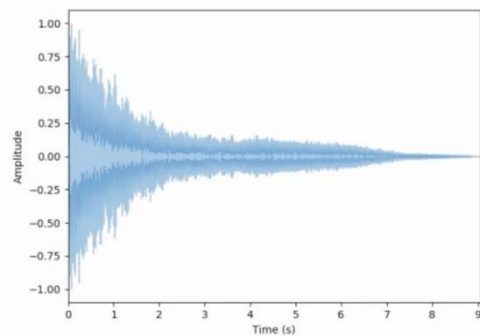


D'une onde au spectrogramme

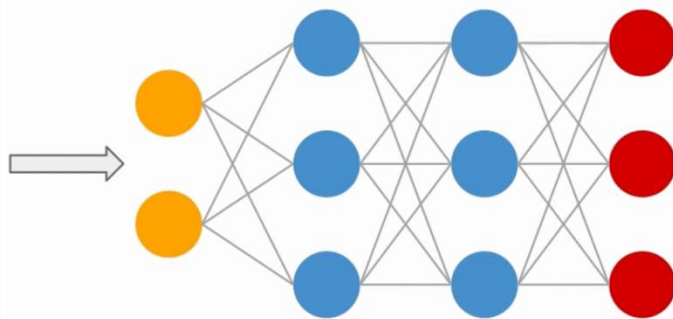
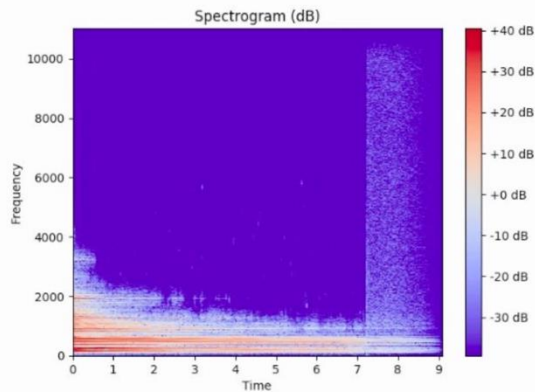




Pourquoi c'est important en deep learning ?



STFT





Mel frequency Cepstral Coefficients (MFCCs)

- Capture des aspects de timbres/textures du son
- Caractéristiques du domaine des fréquences
- Approche du système auditif humain
- 12 à 40 coefficients
- Calculés à chaque fenêtre de temps



D'une onde au MFCCs

