



HAL
open science

Sculpture temps-fréquence des sons par modification de leur transformée en ondelettes

Florent Jaillet, Bruno Torrèsani

► **To cite this version:**

Florent Jaillet, Bruno Torrèsani. Sculpture temps-fréquence des sons par modification de leur transformée en ondelettes. 6e Congrès Français d'Acoustique, Société Française d'Acoustique, Apr 2002, Lille, France. pp.455-458. hal-04125089

HAL Id: hal-04125089

<https://hal.science/hal-04125089>

Submitted on 12 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Sculpture temps-fréquence des sons par modification de leur transformée en ondelettes

Florent Jaillet^{1,2} et Bruno Torrèsani²

¹GENESIS, Bâtiment Beltram, Domaine du Petit Arbois,
BP 69, 13545 Aix en Provence Cedex 4, France

²LATP, CMI, Université de Provence, 39 rue F. Joliot-Curie, 13453 Marseille cedex 13, France

RESUME

Le vocodeur de phase est un outil d'analyse/modification/synthèse des sons basé sur la transformée de Fourier à court terme qui est couramment utilisé pour le traitement localisé à la fois en temps et en fréquence des sons. Il est utile, pour avoir une analyse mieux adaptée à certains types de sons ou pour tenir compte des données psychoacoustiques concernant la perception auditive, d'utiliser d'autres types de transformées pour les mêmes applications, telles que la transformée en ondelettes.

En conséquence, nous avons étudié les aspects théoriques permettant d'effectuer le même type de traitement avec la transformée en ondelettes qu'avec le vocodeur de phase.

De plus, une discrétisation inhabituelle de la transformée en ondelettes est proposée pour une implantation numérique. Cette discrétisation est basée sur la constitution de paquets d'ondelettes dans l'espace de Fourier. Elle permet de construire une analyse par un banc de filtres dont les largeurs de bande sont données par une fonction quelconque de la fréquence centrale. Ceci permet en particulier de construire une analyse selon l'échelle des Barks.

Un exemple d'application de la transformée en ondelettes sur un bruit de moteur diesel est également présenté.

I. INTRODUCTION

Dans cette étude, on souhaite construire une transformée temps-fréquence utilisable dans un outil informatique permettant à l'utilisateur de modifier localement les caractéristiques fréquentielles du son par action graphique sur une représentation de type sonagramme.

Pour cette application, la possibilité de reconstruction du signal à partir de sa transformée est particulièrement importante, car on souhaite agir sur le son par modification des coefficients de sa transformée. En particulier, dans les régions temporelles où aucune modification n'est apportée, le signal ne doit pas être modifié. Il est donc nécessaire de conserver une relation de reconstruction exacte.

De plus, les modifications étant déterminées graphiquement sur une représentation sonographique de la transformée temps-fréquence, celle-ci doit bien se prêter à la représentation graphique et à l'analyse visuelle.

I.1. Vocodeur de phase

La transformée temps-fréquence couramment utilisée pour ce type d'application dans les domaines du traitement de la parole et des signaux musicaux est la transformée de Fourier à court terme. La chaîne d'analyse/modification/synthèse associée est appelée "vocodeur de phase" [1] [2].

Pour un signal x , la transformée de Fourier à court terme est définie par :

$$F_x(\tau, \nu) = \int_{-\infty}^{+\infty} x(t)h^*(t - \tau)e^{-i2\pi\nu t} dt,$$

avec h la fenêtre d'analyse vérifiant $\int_{-\infty}^{+\infty} |h(t)|^2 dt = 1$.

On peut reconstruire le signal à partir de sa transformée suivant la relation d'inversion :

$$x(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F_x(\tau, \nu)h(t - \tau)e^{i2\pi\nu t} d\tau d\nu.$$

Cette transformée peut s'interpréter comme une décomposition sur une famille d'atomes temps-fréquence $h_{\tau\nu}(t) = h(t - \tau)e^{i2\pi\nu t}$ obtenus par décalage temporel et modulation de la fenêtre d'analyse h . Tous ces atomes ont en conséquence la même concentration dans le plan temps-fréquence et la transformée aura une résolution temporelle et fréquentielle identique pour tous les points.

Un des intérêts de la transformée de Fourier à court terme est la possibilité d'implantation informatique efficace grâce à l'utilisation des algorithmes rapides de calcul de la transformée de Fourier discrète.

I.2. Intérêt d'autres représentations

La transformée de Fourier à court terme n'est pas bien adaptée à la représentation de certains types de sons, notamment ceux qui sont fortement non-stationnaires, comme dans les zones transitoires (par exemple les attaques ou les sons impulsifs). Pour ces types de sons, l'approche temps-échelle utilisée dans la transformée en ondelettes est mieux adaptée.

De plus, en considérant les résultats obtenus dans les études sur la perception auditive et en particulier sur les bandes critiques de l'oreille [6], on observe qu'au dessus de 500Hz, le fonctionnement de l'oreille peut-être interprété comme un filtrage à fac-

teur de qualité (rapport de la largeur de bande sur la fréquence centrale du filtre) constant. Ainsi la transformée en ondelettes (qui réalise une analyse à facteur de qualité constant) est mieux adaptée que la transformée de Fourier à court terme (qui réalise une analyse à largeur de bande constante) lorsque l'on souhaite reproduire ce fonctionnement.

Pour ces raisons, il est intéressant d'étudier les possibilités qu'offre la transformée en ondelettes pour des applications de type vocodeur de phase.

II. TRANSFORMÉE EN ONDELETTES

II.1. Définition

Dans la suite, compte tenu du fait que l'on analyse des fonctions réelles, on se restreint aux valeurs positives de fréquence. Pour cela, on associe de façon bijective à chaque signal f de $L^2(\mathbf{R})$ à valeurs réelles le signal analytique correspondant z_f tel que $\hat{z}_f(\nu) = 0$ pour $\nu < 0$ et $\hat{z}_f(\nu) = 2 * \hat{f}(\nu)$ pour $\nu \geq 0$, en notant \hat{g} la transformée de Fourier d'une fonction g .

Les définitions suivantes sont en conséquence données pour des signaux tels que leur transformée de Fourier soit nulle pour les valeurs négatives de la fréquence.

La transformée en ondelettes d'un signal x est définie par :

$$T_x(b, a) = \frac{1}{a} \int_{-\infty}^{+\infty} x(t) h^* \left(\frac{t-b}{a} \right) dt$$

avec h l'ondelette mère telle que $\int_0^{+\infty} |\hat{h}(\nu)|^2 \frac{d\nu}{\nu} = 1$. Le signal peut-être reconstruit à partir de sa transformée en utilisant la formule d'inversion :

$$x(t) = \int_{-\infty}^{+\infty} \int_0^{+\infty} T_x(b, a) h \left(\frac{t-b}{a} \right) \frac{db da}{a^2}. \quad (1)$$

La transformée en ondelettes peut être interprétée comme une décomposition du signal sur la famille d'atomes $h_{ba}(t) = \frac{1}{a} h \left(\frac{t-b}{a} \right)$ obtenus par translation temporelle d'une durée b et dilatation temporelle d'un facteur d'échelle a de l'ondelette mère. La longueur des supports temporels des atomes sera donc dépendante de la valeur du facteur d'échelle a .

La transformée de Fourier de l'atome h_{ba} est $\hat{h}_{ba}(\nu) = e^{-i2\pi\nu b} \hat{h}(a\nu)$. Si l'ondelette mère a une transformée de Fourier bien localisée autour d'une fréquence centrale ν_0 avec une largeur de bande $\Delta\nu$, l'atome h_{ba} aura une transformée de Fourier localisée autour de la fréquence centrale $\frac{\nu_0}{a}$ et de largeur de bande $\frac{\Delta\nu}{a}$.

En conséquence, la transformée en ondelettes peut être interprétée comme une analyse temps-fréquence, la fréquence centrale d'un atome étant directement déterminée par le coefficient d'échelle correspondant. De plus, il apparaît que la résolution de l'analyse varie avec la fréquence. La résolution fréquentielle

diminue lorsque la fréquence augmente, alors que la résolution temporelle augmente.

II.2. Discrétisation

La discrétisation naturelle de la transformée en ondelettes tient compte de l'évolution de la résolution temporelle et fréquentielle des atomes en fonction du facteur d'échelle. Pour cela, on calcule la transformée pour des valeurs de b et a de la forme $b = nb_0 a_0^m$ et $a = a_0^m$, avec $n, m \in \mathbf{Z}$, $b_0 > 0$ et $a_0 > 0$.

Il est démontré dans [3] que la famille $(h_{a_0^m, nb_0 a_0^m})$ d'atomes forme un repère (frame en anglais) de $L^2(\mathbf{R})$ pour des valeurs de b_0 et a_0 permettant une discrétisation suffisamment fine.

Pour conserver une reconstruction exacte à partir de ce repère, il faut alors introduire le repère dual qui permet la reconstruction en remplaçant les intégrales de la formule d'inversion (1) par des sommes discrètes. Cependant, le calcul des fonctions duales constitue un problème complexe, et rend l'utilisation de cette discrétisation peu souple pour notre application.

De plus cette discrétisation ne conserve pas l'invariance par translation temporelle de la transformée, qui est souhaitable pour une bonne analyse et une représentation graphique intuitive.

D'autres approches de discrétisation permettent de conserver cette invariance, comme les ondelettes dyadiques [4] ou la décomposition de Littlewood-Paley [5].

Avec les discrétisations habituelles, un problème se pose pour les composantes les plus graves qui correspondent aux valeurs du facteur d'échelle tendant vers $+\infty$. En effet, en limitant nécessairement ces valeurs pour une utilisation informatique, on perd les composantes les plus graves, ce qui fausse la reconstruction.

Une approche plus souple basée sur les paquets d'ondelettes et assurant une reconstruction exacte a été utilisée dans cette étude.

III. PAQUETS D'ONDELETTES

III.1. Définition

Pour une suite (a_j) strictement décroissante de nombres réels indexée sur \mathbf{Z} , c'est à dire des valeurs telles que $a_{j+1} < a_j$, vérifiant $\lim_{j \rightarrow -\infty} a_j = +\infty$ et $\lim_{j \rightarrow +\infty} a_j = 0$, on définit les paquets d'ondelettes (également appelés ondelettes intégrées) H_j comme des fonctions telles que :

$$|\hat{H}_j(\nu)|^2 = \int_{a_{j+1}}^{a_j} |\hat{h}(a\nu)|^2 \frac{da}{a}.$$

Il faut noter que cette construction n'implique pas de contrainte sur la phase des paquets d'ondelettes, ce qui laisse une liberté supplémentaire de choix dans les fonctions que l'on utilisera.

L'ensemble des paquets d'ondelettes permet de construire une partition unitaire de l'espace Fourier, c'est à dire :

$$\sum_{j=-\infty}^{+\infty} |\hat{H}_j(\nu)|^2 = 1.$$

En introduisant une fonction d'échelle ϕ telle que $|\hat{\phi}(\nu)|^2 = \int_1^{+\infty} |\hat{h}(u\nu)|^2 \frac{du}{u}$, on peut écrire

$$|\hat{\Phi}_{j_0}(\nu)|^2 + \sum_{j=j_0}^{+\infty} |\hat{H}_j(\nu)|^2 = \sum_{j=-\infty}^{+\infty} |\hat{H}_j(\nu)|^2 = 1,$$

avec Φ telle que $\hat{\Phi}_j(\nu) = \hat{\phi}(a_j\nu)$, et les paquets d'ondelettes vont permettre la reconstruction exacte du signal, selon la relation :

$$x = \int_{-\infty}^{+\infty} \langle x, \Phi_{j_0}^b \rangle \Phi_{j_0}^b + \sum_{j=j_0}^{+\infty} \int_{-\infty}^{+\infty} \langle x, H_j^b \rangle H_j^b db$$

avec Φ_j^b et H_j^b telles que $\Phi_j^b(t) = \Phi_j(t-b)$ et $H_j^b(t) = H_j(t-b)$.

Les paquets d'ondelettes peuvent se déduire de la fonction d'échelle par la relation

$$|\hat{H}_j(\nu)|^2 = |\hat{\Phi}_{j+1}(\nu)|^2 - |\hat{\Phi}_j(\nu)|^2,$$

et seule la connaissance de la fonction d'échelle sera en conséquence nécessaire pour construire les paquets d'ondelettes.

Remarque : La constitution des paquets d'ondelettes permet d'obtenir une transformée pour laquelle le facteur d'échelle est discrétisé, mais les paquets d'ondelettes restent des fonctions de la variable temporelle continue. Cependant, dans le cas où l'ondelette mère est à bande limitée, les paquets d'ondelettes seront également à bande limitée et il sera possible d'utiliser les résultats courants sur l'échantillonnage pour construire une version à temps discret de la transformée.

Dans le cas général où l'ondelette n'est plus à bande limitée, on montre qu'il est possible de construire des repères (frames) à partir des paquets d'ondelettes en discrétisant la variable temporelle.

III.2. Calcul des moments

Les caractéristiques du paquet d'ondelettes H_j varient en fonction des bornes d'intégration a_{j+1} et a_j . Il est souhaitable notamment de pouvoir connaître la fréquence centrale et la largeur de bande du paquet d'ondelettes en fonction de celles de l'ondelette mère h et des bornes d'intégration.

Un des intérêts de la discrétisation avec les paquets d'ondelettes est qu'il est possible de calculer ces paramètres explicitement, ce qui permettra de connaître et contrôler précisément le comportement de la version discrétisée de la transformée utilisée.

Pour cela, en considérant que pour une fonction g , la fonction $\rho_g(\nu) = \frac{|\hat{g}(\nu)|^2}{\int_0^{+\infty} |\hat{g}(\nu)|^2 d\nu}$ représente la densité répartissant l'énergie de la fonction g selon les fréquences, on va définir la fréquence centrale et la largeur de bande en fonction des moments de cette densité.

Le moment d'ordre n de ρ_g étant défini par :

$$\langle \nu^n \rangle_g = \int_0^{+\infty} \nu^n \rho_g(\nu) d\nu = \frac{\int_0^{+\infty} \nu^n |\hat{g}(\nu)|^2 d\nu}{\int_0^{+\infty} |\hat{g}(\nu)|^2 d\nu},$$

on peut définir la fréquence centrale f_g de g comme la moyenne de ρ_g , $f_g = \langle \nu \rangle_g$, et la largeur de bande Δ_g comme l'écart type de ρ_g qui vérifie la relation $\Delta_g^2 = \langle \nu^2 \rangle_g - \langle \nu \rangle_g^2$.

Pour le paquet d'ondelettes H_j , on obtient :

$$\langle \nu^n \rangle_{H_j} = \frac{a_j^{n+1} - a_{j+1}^{n+1}}{(n+1)(a_j - a_{j+1})a_j^n a_{j+1}^n} \langle \nu^n \rangle_h.$$

La fréquence centrale du paquet d'ondelettes peut s'exprimer en fonction de celle de l'ondelette mère :

$$f_{H_j} = \frac{a_j + a_{j+1}}{2a_j a_{j+1}} f_h,$$

et sa largeur de bande vérifie :

$$\Delta_{H_j}^2 = \frac{a_j^2 + a_j a_{j+1} + a_{j+1}^2}{3a_j^2 a_{j+1}^2} \Delta_h^2 + \frac{(a_j - a_{j+1})^2}{12a_j^2 a_{j+1}^2} f_h^2.$$

En prenant pour a_j des valeurs de la forme $a_j = \lambda^j$ avec $0 < \lambda < 1$, on obtient : $f_{H_j} = \frac{1+\lambda}{2\lambda^{j+1}} f_h$ et

$$\Delta_{H_j}^2 = \frac{1+\lambda+\lambda^2}{3\lambda^{2(j+1)}} \Delta_h^2 + \frac{(1-\lambda)^2}{12\lambda^{2(j+1)}} f_h^2.$$

Ces formules permettent de bien apprécier le comportement de la transformée construite et également de choisir les valeurs de discrétisation en fonction des caractéristiques que l'on souhaite donner à la transformée.

III.3. Transformée sur l'échelle des Barks

En choisissant les a_j de façon non régulière, il est possible de faire varier la largeur de bande des paquets d'ondelettes selon une fonction donnée de leur fréquence centrale. Ainsi, en choisissant des valeurs appropriées, chaque paquet peut avoir une largeur de bande égale à la largeur de la bande critique de l'oreille dans la zone de fréquence correspondante. Pour cela on utilise l'échelle des Barks définie dans [6], qui est telle qu'un accroissement de fréquence égal à la largeur de bande critique correspond à une augmentation de 1 Bark.

La valeur en Barks en fonction de la fréquence en Hz peut être déterminée par la relation donnée dans [7]

$$z(f) = \frac{26.81f}{1960+f} - 0.53.$$

Le choix de valeurs de a_j telles que

$$z(f_{H_j} + \frac{\Delta_{H_j}}{2}) - z(f_{H_j} - \frac{\Delta_{H_j}}{2}) = 1 \text{ Bark}$$

permet donc d'aboutir à une transformée ayant les caractéristiques souhaitées.

Pour cela, on fixe une valeur a_0 telle que la fréquence de coupure de la fonction passe-bas $\hat{\Phi}_0$ soit la fréquence à partir de laquelle on souhaite effectuer l'analyse sur l'échelle de Bark, par exemple 20Hz, puis on choisit de proche en proche la valeur a_{j+1} permettant d'obtenir une largeur de 1 Bark avec la valeur de a_j qui vient d'être fixée.

IV. EXEMPLE D'APPLICATION : RALENTI DE MOTEUR DIESEL

Des exemples sonores correspondant à cette partie sont disponibles à l'adresse internet www.cmi.univ-mrs.fr/~jaillet/cfa02.

La figure 1 montre le module de la transformée en ondelettes d'un extrait de son de moteur diesel. On voit apparaître de manière cyclique des pics d'énergie, surtout visibles en hautes fréquences, correspondant au bruit du culbuteur. En sélectionnant graphiquement ces régions de forte énergie, et en modifiant leur niveau, on peut simuler l'effet qu'auraient des modifications mécaniques qui influeraient sur le niveau du bruit du culbuteur. On peut ainsi par exemple mesurer l'influence de ce bruit sur la gène perçue.

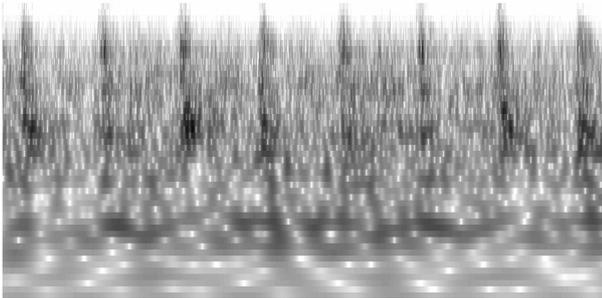


Figure 1: Module de la transformée en ondelette d'un bruit de moteur (ondelette de Morlet).

La figure 2 montre la forme d'onde du signal original et la forme d'onde du signal pour lequel on a diminué le niveau des bruits du culbuteur.

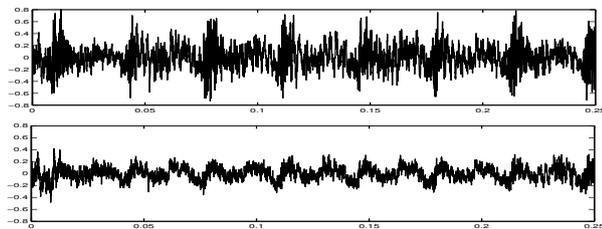


Figure 2: Forme d'onde originale en haut et forme d'onde traitée en bas.

V. CONCLUSION

L'utilisation de la transformée en ondelettes pour des applications de type vocodeur de phase a été étudiée et une discrétisation appropriée utilisant les paquets d'ondelettes a été proposée, permettant une implantation informatique de la transformée.

Cet outil a montré son intérêt pratique sur le traitement de certains types de sons.

La souplesse d'utilisation des paquets d'ondelettes a permis de construire une transformée suivant l'échelle des Barks.

De plus, cette souplesse ouvre d'intéressantes perspectives. En effet, plutôt que de l'utiliser pour adapter la transformée à une échelle, il est possible de l'utiliser pour rendre la transformée adaptative, c'est à dire choisir la forme des paquets en fonction du signal, de façon à optimiser la représentation selon des critères donnés.

Ces critères pourraient être des critères de concentration énergétique de la transformée, du même type que ceux utilisés dans les choix de meilleure base [8], ou encore liés aux connaissances a priori sur la structure du signal.

REFERENCES

1. J. Moorer, *The use of the phase vocoder in computer music applications*, Journal of the audio Engineering Society 26(1), 42-45, (1978)
2. M. Dolson, *The phase vocoder : a tutorial*, Computer Music journal 10(4), 14-27, (1986)
3. I. Daubechies, *The wavelet transform, time-frequency localization and signal analysis*, IEEE Transactions on information theory 36(5), 961-1005, (September 1990)
4. S. Mallat, *Une Exploration des Signaux en Ondelettes*, Editions de l'Ecole Polytechnique, Ellipses diffusion,(2000)
5. M. Duval-Destin, M.A. Muschietti, B. Torrèsani, *Continuous wavelet decompositions, multiresolution, and contrast analysis*, SIAM J. Math. Anal. 24(3), (may 1993)
6. E. Zwicker, R. Feldkeller, *Psychoacoustique, l'oreille récepteur d'information*, Collection Technique des Télécommunications, Masson, (1981)
7. C. D'Alessandro et D. Beautemps, *Représentation et modification du signal de parole par transformée en ondelettes utilisant des contraintes auditives*, Note et documents LIMSI 90-10 (Octobre 1990)
8. H. Šikić, M. Wickerhauser, *Information cost functions*, Applied and Computational Harmonic Analysis 11(2), 147-166, (Septembre 2001)