SUIVI DU FONDAMENTAL DE SIGNAUX MUSICAUX À HAUTE RÉOLUTION TEMPORELLE

V. Gibiat, P. Jardin

► To cite this version:

V. Gibiat, P. Jardin. SUIVI DU FONDAMENTAL DE SIGNAUX MUSICAUX À HAUTE RÉ-
SOLUTION TEMPORELLE. Journal de Physique IV Proceedings, 1992, 02 (C1), pp.C1-97-C1-100.
10.1051/jp4:1992117 . jpa-00251130

HAL Id: jpa-00251130

<https://hal.science/jpa-00251130>

Submitted on 1 Jan 1992

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.

SUIVI DU FONDAMENTAL DE SIGNAUX MUSICAUX À HAUTE RÉOLUTION TEMPORELLE

V. GIBIAT et P. JARDIN*

LOA/ESPCI, 10 rue Vauquelin, F-75005 Paris, France

*ESIEE, Cité Descartes, BP. 99, F-93162 Noisy le Grand cedex, France

Résumé: le suivi du fondamental très utilisé dans le domaine du traitement de la parole reste en comparaison un outil relativement peu employé dans le cadre de l'acoustique musicale. Hormis quelques applications spécifiques, son emploi n'est souvent que le résultat d'une analyse du spectre de puissance à court terme. L'Analyse Spectrale Différentielle déjà testée en particulier dans le traitement de la parole nous a semblée autoriser une bonne précision temporelle et fréquentielle. Cette technique simple, permet d'obtenir avec des moyens légers une information précise sur l'évolution des fréquences dans le signal. Nous avons ainsi obtenu sur des signaux courts d'instruments à vent des informations sur l'évolution de la hauteur émise lors de l'attaque et pendant le son tenu ainsi que pour des variations de l'émission de pianissimo à fortissimo. On peut également, sur des phrases musicales longues, détecter très précisément la note jouée ainsi que l'écart en cents avec le tempérament égal, réalisant ainsi un détecteur de mélodie relativement simple.

Abstract: pitch extraction is a common tool in speech processing, but seems less used in musical acoustics. It is often obtained through short time power spectra which does not allow precise information. Differential Spectral Analysis allows an easy extraction of the main frequencies present in a signal.

We have obtained upon short signals (1 second) from woodwinds, a great deal of informations on pitch variations. One also can, on long duration musical sentences, determine very precisely the note played and the discrepancy in cents with the equal tempered pitch.

I) INTRODUCTION.

Le suivi du fondamental, (détection de la mélodie), est une technique très employée en parole où sa connaissance joue un rôle important. Dans la domaine de l'acoustique musicale, la difficulté majeure vient du fait que cette mélodie évolue sur plusieurs octaves, entre 50Hz. et plus de 4000Hz. et que la précision requise est très importante (moins de 10 cents dans toute la gamme soit environ .1 Hz dans le grave). Les techniques existantes de suivi de la mélodie sont soit trop imprécises pour donner des indications quand au fonctionnement de l'instrument [1], soit trop lourdes en calcul pour être d'un accès pratique [2]. Les dispositifs commerciaux d'accord des instruments permettent eux de répondre aux conditions de précision et de souplesse, mais sont beaucoup trop lents pour permettre une étude sur les parties transitoires de signaux. Leur temps de réponse ne permettant pas de suivi de mélodies variant rapidement.

Nous proposons d'utiliser l'Analyse spectrale Différentielle pour mesurer la fréquence fondamentale de signaux musicaux. Cette méthode déjà employée dans des domaines variés,[3,4], pourrait être une alternative intéressante à des techniques lourdes comme le vocoder de phase.

II) ANALYSE SPECTRALE DIFFÉRENTIELLE

L'Analyse spectrale Différentielle (ASD) peut se définir comme une transformation de Fourier glissante à origine de temps fixe et utilisant une fenêtre de pondération rectangulaire. Sa définition mathématique est:

$$F(n, t) = \frac{1}{T} \int_{t-\frac{T}{2}}^{t+\frac{T}{2}} f(t') \exp(-2i\pi \frac{n}{T} t') dt' \quad (1)$$

alors que l'expression normale d'une transformation de Fourier glissante serait:

$$F(n, t) = \frac{1}{T} \int_{t-\frac{T}{2}}^{t+\frac{T}{2}} f(t') \text{Rect}^T(t' - t) \exp(-2i\pi \frac{n}{T} (t' - t)) dt'. \quad (2)$$

La différence essentielle tient à l'emploi d'une exponentielle d'analyse qui se "déroule" en même temps que le signal. Ceci permet d'estimer très simplement l'évolution des coefficients $F(n, t)$ grâce à l'expression:

$$F(n, t + dt) - F(n, t) = \frac{1}{T} (f(t + dt) - f(t)) \exp(-2i\pi \frac{ndt}{T}) \quad (3)$$

qui sous forme discrète devient:

$$F_{n,k} - F_{n,k-1} = \frac{1}{N} (f_k - f_{k-N}) \exp - 2i\pi \frac{nk}{N} \quad (4)$$

avec $T = Ndt$ et $dt = 1/f_e$ où f_e est la fréquence d'échantillonnage.

On montre facilement que dans le cas d'un signal quasimonochromatique $x(t)$, tel que $f(t)$ soit le signal analytique associé à $x(t)$, les expressions (3) et (4) permettent l'accès rigoureux à la fréquence instantanée au sens de Ville, comme:

$$\nu(t) = 1/2\pi \cdot \arg(F(n, t + dt) - F(n, t)). \quad (5)$$

Si $f(t)$ est un signal réel alors le résultat de l'expression (5) oscille à deux fois la valeur de $\nu(t)$ [5]. La valeur exacte est donc facilement rétablie par filtrage passe bas de l'expression (3) ou (4).

Lors de l'étude d'un signal à structure harmonique ou quasi harmonique, l'ASD se comporte comme un filtre numérique en peigne particulier dont la fonction de transfert est esquissée sur la figure 1. On peut parler de bandes d'analyses bien que chacune d'entre elle porte la contribution de l'ensemble du contenu fréquentiel du signal par convolution avec la transformée de Fourier de la fenêtre rectangulaire. L'opération de filtrage passe bas indispensable pour la détermination de la fréquence instantanée dans chaque bande permet également de limiter les interférences provenant des lobes secondaires du sinus cardinal.

L'ASD permet donc de suivre, moyennant le choix initial d'un fenêtre d'analyse de largeur T , l'évolution fréquentielle des composantes, éventuellement non harmoniques, contenues dans un signal réel. Bien entendu comme toutes les méthodes d'analyses temps fréquences, elle reste liée à une relation d'incertitude qui ici apparaît clairement dans l'impossibilité de séparer par l'analyse des composantes donc l'écart dans l'espace des fréquences serait inférieur à $1/T$.

III) APPLICATION AU SUIVI DE FONDAMENTAL.

L'idée d'utiliser l'ASD au suivi du fondamental de signaux divers n'est pas en soi nouvelle. Cette technique a déjà été éprouvée sur des signaux de parole [5]. Dans le cadre des signaux musicaux cette utilisation est en revanche nouvelle. L'intérêt de ces signaux réside dans leur structure quasi harmonique et dans la présence d'un fondamental physique souvent d'énergie importante. Dans les cas où ce fondamental est de plus faible énergie qu'un des autres constituants du signal, il reste néanmoins d'amplitude relativement importante, alors que le fondamental du signal de parole lui est toujours beaucoup plus faible.

Le principal problème rencontré dans l'utilisation de l'ASD reste le choix de la fenêtre d'analyse. Celle ci impose une connaissance a priori du résultat qui en limite l'intérêt. Malgré tout le choix de cette fenêtre est suffisamment lâche pour que l'on puisse presque la choisir "à l'aveugle". Les signaux musicaux ont leurs fondamentaux qui varient de manière importante dans une zone de fréquences comprises entre moins de 50 Hz. et plus de 4000 Hz. Nous avons choisi de fixer arbitrairement une fenêtre d'analyse à $1/T = 100 \text{ Hz.}$ et de filtrer le résultat de l'équation (5) par un filtre passe bas numérique de type Butterworth d'ordre 4. De cette façon tous les signaux de fondamentaux situés entre 0 Hz et 200 Hz sont détectés dans le premier lobe de la transformée de Fourier de la fenêtre d'analyse tandis que les fondamentaux plus aigus interviennent à travers leurs contributions dans les lobes secondaires. La figure 2 illustre cette situation. Sur un signal de clarinette (une suite de notes du grave à l'aigu) on voit que l'ASD suit correctement le fondamental jusqu'à ce que celui ci soit trop haut. On conjugue ici les inconvénients d'un fondamental de trop faible amplitude et d'un trop grand éloignement de la fenêtre d'analyse.

figure 2

Nous avons pallié cet inconvénient en effectuant l'analyse sur plusieurs bandes successives. La fréquence calculée se situant dans ou hors de la bande (coefficient n), il est simple de choisir la bande dans laquelle l'énergie extraite est maximum et donc d'obtenir la meilleure précision sur la fréquence instantanée. Dans cette situation l'ASD remplit parfaitement son rôle et la précision obtenue permet de réaliser un véritable suivi de mélodie, mais également de contrôler la "justesse" avec laquelle cette mélodie est donnée et les éventuelles micro variations que l'on peut rencontrer (micromélodies c.f. figure 3).

figure 3

IV) RESULTATS SUR DES SIGNAUX TRANSITOIRES.

La précision obtenue sur des suivis de mélodie nous a conduit à envisager un autre type d'applications. L'étude fine de l'évolution de la hauteur dans les signaux transitoires n'ayant à notre connaissance pas été faite, et ayant remarqué dans des conditions particulières des variations notables de hauteurs dans des signaux en cours de jeu, nous nous sommes attachés à l'étude de la micro mélodie liée à l'attaque des instruments à anches simples. L'étude porte alors sur des signaux courts (moins de 1 seconde) et pour lesquels on cherche à déterminer les éventuels écarts à un tempérament égal. La figure 4 présente un exemple de ces résultats. L'attaque est toujours le siège d'une variation de hauteur rapide très souvent dans le sens de la baisse. Ce qui revient à dire que l'instrument commence son oscillation à une fréquence supérieure à celle de la partie tenue. Le choix de la largeur de fenêtre T est ici choisi en fonction de la note jouée. La précision du résultat ne sera pratiquement pas alors affectée par l'écart (faible) de $1/T$ avec la fréquence vraie du signal.

figure 4

V) VARIATIONS DE HAUTEURS DANS LES NUANCES DE JEU.

Le précédent paragraphe se réfère plus au fonctionnement physique de l'instrument, mais le jeu instrumental est aussi fait de variations de nuances et d'intensités. La clarinette a cette particularité que ses notes graves sont de hauteurs d'autant plus basses que la note est jouée forte. Cet effet est très facile à mesurer avec un accordeur électronique. On peut également se servir de l'ASD pour suivre la variation de fréquence sur une note. Ceci permet de vérifier cet effet lié au fonctionnement non linéaire de l'instrument et en même temps de confronter l'ASD à un autre système de mesure de hauteur. La précision de la mesure est telle que nous avons pu mettre en évidence cet effet d'abaissement avec l'augmentation de la nuance de jeu pour l'ensemble des notes des registres grave médium et clairon de l'instrument. L'explication de ce phénomène comme du à l'inharmonicité forte des pics d'impédances de la clarinette dans le grave ne peut donc pas être invoquée comme seule cause. Les mesures ont bien entendu été réalisées en éliminant toute variation autre que la pression d'alimentation en utilisant une embouchure semi-artificielle où le musicien n'intervient pas au niveau de l'embouchure mais uniquement pour fournir la pression requise à l'émission sonore.

VI) CONCLUSION.

Nous avons utilisé l'ASD comme un outil simple et robuste. Ses performances et sa légèreté en font un bon outil pour toutes les applications de suivi d'une ou même plusieurs fréquences dans un signal où la structure est proche de celle d'un signal harmonique. La précision est au moins aussi bonne en fréquence que celle d'un accordeur électronique (quelques cents), mais sa réponse temporelle est en revanche beaucoup plus rapide. Ceci permet de suivre les évolutions de mélodies et de micromélodies. Si on ajuste la largeur de la fenêtre la méthode réagit alors avec la même vitesse temporelle et s'y ajoute une précision en fréquence largement meilleure que 5 cents associée à une perte de souplesse.

REFERENCES:

- [1] J. BROWN, J. Acoust. Soc. Am. 89 (1991) p425.
- [2] J.L. FLANAGAN "Speech Analysis Synthesis and Perception" §8-114 p378 SPRINGER Ed. Berlin 1972.
- [3] V. GIBIAT, P. JARDIN, F. WU Acustica 63 (1987) p 90.
- [4] L. VIGNAUD "Etude des transitoires d'attaque du violon par des méthodes temps fréquence" Mémoire de magistère de l'ENS de Lyon (non publié) 1991.
- [5] P. JARDIN "Evaluation des performances d'une technique de Fourier glissant (Analyse Spectrale Différentielle) pour le traitement des signaux de parole. thèse de doctorat de troisième cycle Université Paris 6 1984.