

# VISUALISATION DU JEU INSTRUMENTAL

C. Daudin, D. Fober, S. Letz, Y. Orlarey, Y. Chapuis

{daudin,fober,letz,orlarey,chapuis}@grame.fr

Grame - Centre national de création musicale

## RÉSUMÉ

Nous abordons la visualisation du jeu instrumental sous l'angle de la représentation objective du son, avec pour objectif de rendre apparentes des caractéristiques du signal en termes d'interprétation musicale. Cet article expose les différentes stratégies mises en oeuvre pour exploiter voire amplifier la complémentarité des perceptions visuelles et sonores, avec un point de vue particulier sur les applications dans le domaine de la pédagogie de la pratique instrumentale.

## 1. INTRODUCTION

La visualisation du sonore est un champ d'investigation déjà ancien. Dès les années 1940, issu de travaux dans le domaine de l'acoustique, le *sonographe* est créé par les laboratoires Bell; depuis, le *sonagramme*<sup>1</sup> a évolué mais est resté l'une des représentations les plus importantes pour étudier visuellement la composition spectrale du son. Avec l'avènement de l'ordinateur et du son numérique, les visualisations se sont multipliées, principalement pour des besoins d'analyse acoustique, mais le champ d'investigation s'est rapidement élargi aux exigences de domaines très variés. Dès le début des années 90, l'*acousmographe*, développé par le GRM [6] [3], propose des outils de représentation du sonore, dédiés à la transcription des musiques électroacoustiques. Très rapidement, les domaines de l'analyse pour extraction d'information ou recherche de similarités, vont également proposer des représentations nouvelles du sonore [2] [9]. Les logiciels de montage, d'édition ou de contrôle du son feront également appel à des représentations graphiques [1], pouvant servir les besoins d'interfaces utilisateur, voire de support à la *navigation* dans le sonore [10]. Des applications aussi diverses que la visualisation du son pour les sourds [7] se pencheront sur la question de la *visibilité* du sonore.

Loin d'être exhaustives, les quelques références ci-dessus montrent cependant la richesse et la diversité des approches dans le domaine de la visualisation du sonore. Le travail que nous présentons, à la fois se nourrit des représentations existantes, et s'en différencie en ce sens que nous nous attachons plus par-

ticulièrement au jeu instrumental, c'est à dire à visualiser les caractéristiques du sonore en terme d'interprétation musicale (articulation, justesse, nuances...). Dans ce domaine particulier, il existe également des travaux proposant des représentations graphique au titre d'aide à l'écoute ou à l'analyse [5] [4], ou encore à vocation pédagogique [8]. Notre approche s'en différencie en ce sens qu'elle s'applique à des signaux audios, qu'elle est *temps différé*, qu'elle est conçue pour coexister avec l'espace temporel linéaire de la notation musicale, et qu'elle n'est pas basée sur une analyse du signal ayant pour but d'en extraire des informations symboliques de plus haut niveau.

Nous nous intéressons aux problèmes suivants : comment afficher les données de base du signal pour obtenir des représentations claires, précises et intuitives du jeu instrumental? Quelles caractéristiques du jeu sont alors visualisables et dans quelle mesure le sont-elles? Pour chaque caractéristique, quelles sont les représentations graphiques les plus fidèles à la perception auditive?

Le travail que nous présentons a été réalisé dans le cadre du projet VEMUS<sup>2</sup>, c'est à dire dans un contexte pédagogique centré sur la pratique instrumentale. Il se limite à des instruments monodiques. Nous décrivons tout d'abord les stratégies adoptées pour représenter les caractéristiques de base du jeu instrumental : hauteur, justesse, intensité, articulations, timbre. Nous présenterons ensuite les résultats obtenus en utilisant les dérivées des valeurs du signal. Nous montrerons comment la combinaison des représentations permet d'augmenter leur lisibilité. Enfin, nous exposerons comment la comparaison de courbes graphiques en permet une meilleure compréhension, et comment elle peut servir un but pédagogique tout en s'affranchissant de la problématique de la qualité du jeu instrumental.

## Préambule

Les données que nous utiliserons dans ces représentations sont directement extraites du signal. On dispose ainsi :

- de la fréquence fondamentale du signal ;
- de la valeur efficace du signal (en anglais : RMS, *Root Mean Square*). Cette valeur est couramment

<sup>1</sup> ou spectrogramme

<sup>2</sup> IST-27952

utilisée en acoustique pour représenter l'intensité ;

- du spectre du signal ;
- des amplitudes des harmoniques, extraites du spectre.

## 2. LA HAUTEUR

Un problème simple pour la représentation de la hauteur est de savoir si l'on doit tracer une courbe ou des points. La réponse apparaît rapidement : relier 2 points est, sémantiquement, une interpolation. Or, la hauteur n'est pas une valeur continue dans le jeu instrumental (du moins, pas sur les instruments qui nous intéressent). On représentera donc chaque valeur uniquement par un point.

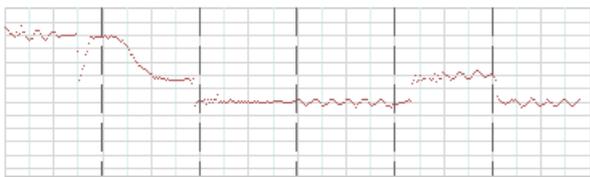
De plus, les données sont issues de systèmes de "pitch-tracking" (détection de la hauteur) qui sont - à l'heure actuelle - imparfaits : ne pas relier les points entre eux minimise l'impact graphique d'éventuelles valeurs erronées (qui seraient, en liant les points, représentées par des pics).

L'information de hauteur peut être considérée, en termes musicaux, sous deux angles :

- la hauteur des notes jouées et leurs relations, ou hauteur *mélodique* ;
- la justesse.

### 2.1. Hauteur mélodique

On représente ici les notes sur une échelle de demitons, obtenant une représentation similaire à la portée musicale traditionnelle. Pour représenter des interprétations dont le registre est étendu, on choisira un niveau de détail faible pour limiter l'espace graphique consommé.

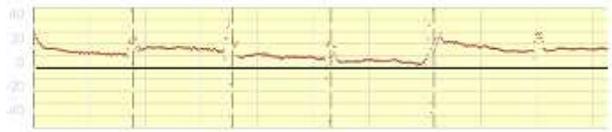


**Figure 1.** Courbe de hauteur, graduée en 1/2 tons. Vièle Erhu, musique chinoise.

### 2.2. Justesse

La justesse fait référence ici à la capacité de l'interprète à respecter un tempérament donné. Pour cela, nous utilisons une échelle de cents (centième de 1/2 ton) pour représenter la différence entre les hauteurs effectives et théoriques de la note. Estimer la hauteur théorique d'une note pose à la fois le problème du diapason et du tempérament. Pour un tempérament

donné, on peut adopter différentes stratégies pour estimer la hauteur théorique :

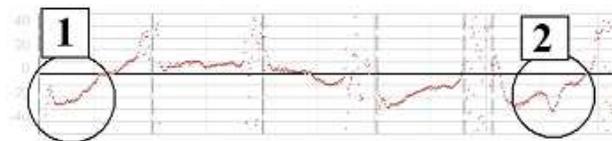


**Figure 2.** Courbe de justesse. L'interprète joue trop haut.

1. elle peut être le "1/2 ton le plus proche" d'après un diapason *fixe*, qui peut être le "La 440" ou un accord précisé par l'instrumentiste. C'est ce qui est fait sur la figure 2 ;
2. on peut calculer le diapason a posteriori, en choisissant la valeur qui minimise l'écart type de la justesse sur toute l'interprétation ;
3. on peut aussi, par exemple, fixer le diapason à partir des premières notes, pour détecter si l'interprétation dévie par la suite.

Il est intéressant de s'interroger sur l'attitude de notre oreille par rapport à la justesse : est-ce qu'elle fonctionne selon l'un des schémas précédents (1, 2 ou 3) ? pour un instrument soliste, est-elle capable de procéder à une réévaluation du diapason ?

Par ailleurs, la courbe de justesse utilisant une échelle détaillée, les moindres variations de hauteur  $y$  sont amplifiées (figure 3). Contrairement aux représentations temps-réel de type "accordeur électronique", le tracé dans le temps permet alors de visualiser clairement des problèmes d'intonation ou de tenue de note (figure 3). Pour représenter plus spécifiquement la justesse, en ignorant ces instabilités locales, on pourrait penser à un lissage des valeurs, pour ne représenter que la hauteur *globalement perçue* de la note.



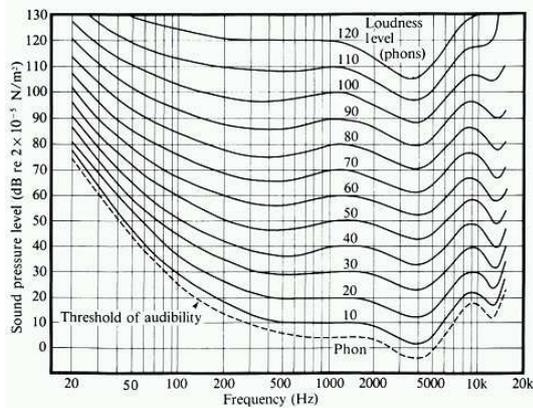
**Figure 3.** Courbe de justesse. Le niveau de zoom permet de voir des variations fines de la hauteur : 1. défaut d'intonation ; 2. légère instabilité.

## 3. L'INTENSITÉ & LES ARTICULATIONS

On utilisera les mêmes valeurs RMS du signal pour représenter deux caractéristiques du jeu musical distinctes : l'intensité (c'est à dire les nuances, la dynamique) et les articulations.

### 3.1. Intensité

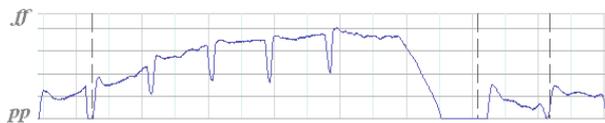
Il faut noter que, pour avoir une représentation de l'intensité fidèle à la perception, il serait judicieux d'utiliser des valeurs tenant compte des variations de la sensibilité de l'oreille selon la fréquence du son (figure 4). Toutefois, nous nous limiterons ici à la représentation des simples valeurs RMS.



**Figure 4.** Courbes d'isophonie de Fletcher-Munson. A chaque courbe correspond un niveau de perception constant : l'oreille est moins sensible dans les basses et hautes fréquences.

Il est communément admis, depuis les expériences du Dr. Gustav Fechner (1860), que la perception est le logarithme de l'excitation. On représentera donc l'intensité sur une échelle logarithmique (c'est à dire en décibels).

Il est important de fixer, en dB, une limite inférieure de représentation. On pourrait utiliser le seuil de perception de l'oreille. Or, il est inutile de représenter le bruit ambiant. On propose donc une valeur qui correspond à la nuance "pianissimo". Par extension, on graduera l'échelle logarithmique de pianissimo à fortissimo (figure 5). La correspondance entre une valeur en dB et une nuance musicale est spécifique à chaque instrument : on effectuera les calibrages correspondants.



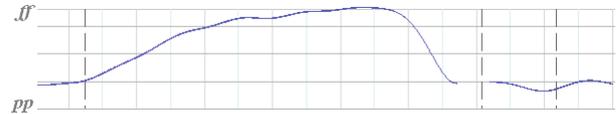
**Figure 5.** Courbes d'intensité. Interprétation d'un crescendo suivi d'un piano.

Un instrumentiste n'envisage généralement pas l'intensité d'une manière aussi proche des données physiques ; il aura certainement une vision des nuances au niveau de la phrase mélodique et de sa dynamique globale, de son mouvement. Pour cela, nous

proposons une représentation lissée, qui nécessite un traitement simple des données, par exemple :

$$x_i = \frac{1}{2k+1} \cdot \sum_{j=i-k}^{i+k} x_j$$

Un exemple de représentation qui peut ainsi être obtenue est donné par la figure 6, qui représente la même interprétation que la figure 5.

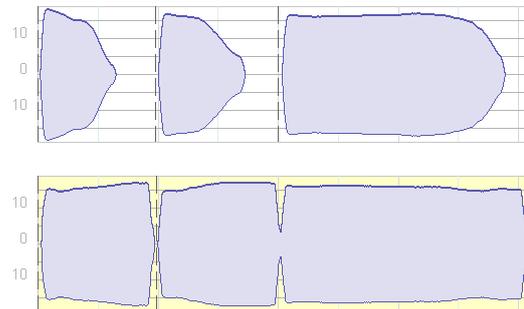


**Figure 6.** Courbe d'intensité lissée.

### 3.2. Articulations

Pour représenter les articulations, on propose une représentation dont le symbolisme graphique nous semble plus adapté ; cette représentation est une simple enveloppe symétrique.

La figure 7 donne un exemple de cette représentation, en montrant deux interprétations : trois notes jouées staccato (en haut) et jouées liées (en bas). Les notes sont alors comme des *bulles sonores*, collées ou détachées.



**Figure 7.** Courbes d'articulation. En haut, une interprétation *staccato* ; en bas, *legato*.

Précisons que l'on est bien conscient que les informations données sur l'articulation par les valeurs d'intensité sont incomplètes, voire grossières ; par exemple, pour les instruments à anche, l'articulation peut varier du son "T" au son "D" : ces subtilités ne sont pas visibles ici. Elles pourraient en revanche l'être sur un spectrogramme. On étudiera dans le futur les possibilités de fournir des représentations plus fines, prenant en compte les spécificités de chaque instrument.

## 4. LE TIMBRE

Avant toute chose, il est important de répéter qu'il n'est pas question ici de représenter graphiquement

des données résultant de modèles estimatoires complexes (de type caractérisation automatique du timbre par des descripteurs linguistiques), mais simplement de proposer des représentations utilisant les données brutes du spectre.

#### 4.1. Remarques sur le spectrogramme

La grande diffusion du spectrogramme n'empêche pas l'existence de certains problèmes, qu'il est intéressant de soulever.

Tout d'abord revient la question d'une représentation tenant compte de la variation de la sensibilité de l'oreille selon la fréquence (revoir la figure 4). En effet, si dans une situation d'analyse physique du son (par exemple en lutherie) on veut avoir une représentation fidèle à la *réalité physique*, on peut se demander s'il ne serait pas plus adapté, dans un contexte de représentation fidèle à la *perception auditive*, de représenter les valeurs de sonie (intensité perçue).

On pourrait arguer que notre oreille n'atteint pas le niveau de précision d'un spectrogramme et que, dès lors, parler de spectrogramme *fidèle à la perception* n'a pas de sens. Des recherches approfondies seraient nécessaires pour le déterminer.

Ensuite se pose la question de l'échelle de couleurs. En effet, les spectrogrammes utilisent généralement différentes couleurs pour *faire ressortir* des valeurs. Cela permet par exemple de représenter plus nettement les "raies harmoniques" : au delà d'une certaine valeur, on utilisera une couleur beaucoup plus vive.

Mais ces échelles de couleur *non-linéaires* introduisent de fait un biais dans la représentation : les valeurs ne sont plus représentées avec une importance graphique qui leur est proportionnelle. Ce biais risque alors de donner des représentations qui induisent en erreur sur la réalité des données, comme l'illustre la figure 8. L'exemple que nous donnons ici est extrême car, en pratique, les échelles de couleurs sont choisies pour limiter le biais.

Il faudrait donc n'utiliser que des échelles de couleurs linéaires, comme des niveaux de gris évoluant linéairement du blanc au noir. Mais paradoxalement, de tels spectrogrammes sont considérés comme moins éloquents, moins lisibles, alors qu'ils donnent une représentation plus fidèle à la réalité des données.

Le spectrogramme comportant une grande quantité d'informations, il peut être intéressant de disposer de représentations plus lisibles, ciblant certaines caractéristiques. Nous proposons ici des représentations qui ne montrent que les valeurs des harmoniques. Toute valeur non-harmonique est ignorée.

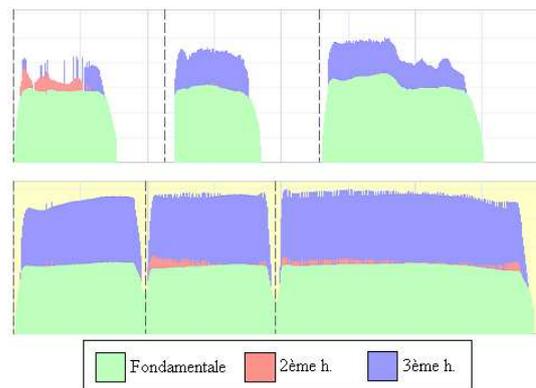
#### 4.2. Représentations en empilement d'harmoniques

La figure 9 est un exemple de représentation des harmoniques, en *empilement d'enveloppes*. Dans ce type

de représentation :

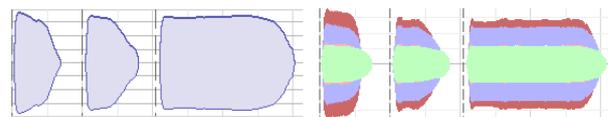
- chaque harmonique est représentée par une enveloppe d'une épaisseur proportionnelle à l'amplitude de l'harmonique ;
- les enveloppes sont empilées les unes sur les autres, par ordre croissant d'harmonique ;
- les enveloppes sont de couleurs différentes, une légende permettant de connaître l'harmonique représentée par chaque enveloppe.

Un exemple du type d'information qui peut être apporté par cette représentation est donné dans la figure 9, où l'on voit clairement la faiblesse de la 3ème harmonique chez le clarinettiste débutant. Cette faiblesse se traduit par un timbre pauvre, "nasillard". On note par ailleurs l'absence, dans les deux interprétations, de la 2ème harmonique (l'octave), absence caractéristique du timbre de la clarinette.



**Figure 9.** Courbes d'harmoniques empilées. Deux interprétations : en haut, timbre nasillard ; en bas, timbre chaud.

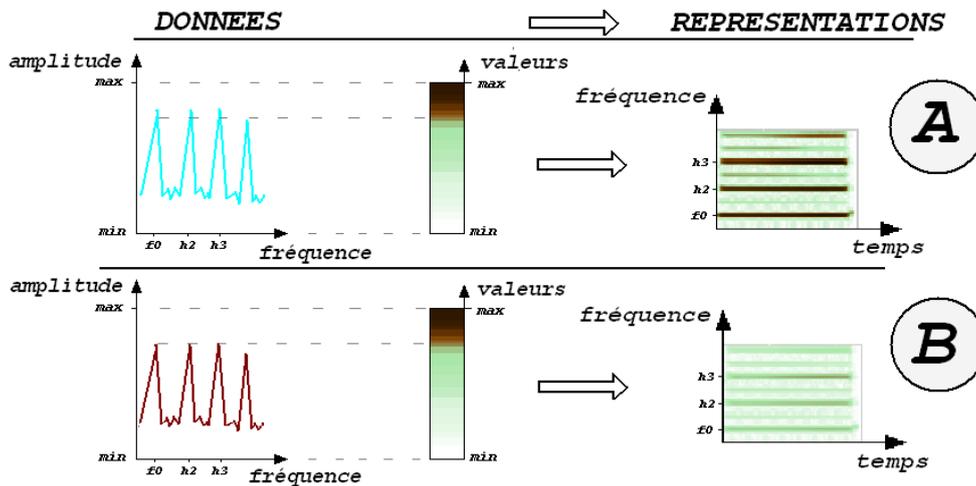
Sur la figure 10, on peut remarquer que ces enveloppes d'harmoniques empilées, représentées de manière symétrique, font apparaître les mêmes informations d'articulations que la courbe d'articulation utilisant les valeurs RMS.



**Figure 10.** Pour une même interprétation : à droite, courbe d'harmoniques empilées ; à gauche, courbe d'articulation. Il y a une grande ressemblance dans l'enveloppe globale.

##### 4.2.1. Epaisseur des enveloppes

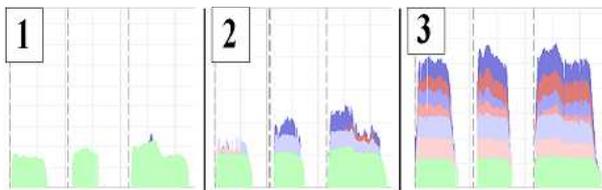
Comme pour la représentation de l'intensité, pour avoir une représentation fidèle à la perception, il faut utiliser une échelle logarithmique. Mais le cas des



**Figure 8.** Exemple de biais introduit dans le spectrogramme par une mauvaise échelle de couleur : les données A, en haut, sont très légèrement supérieures aux données B, en bas ; la différence entre leurs représentations respectives est donc disproportionnée.

harmoniques est plus complexe que celui de l'intensité : il est moins trivial de déterminer un seuil de représentation pour les harmoniques que pour une intensité globale, puisqu'on ne peut alors plus se baser sur le seuil de perception.

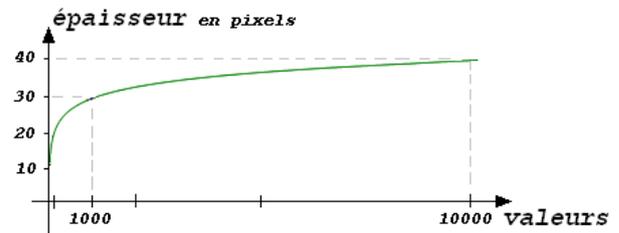
La figure 11 illustre l'importance du choix du seuil, en montrant que les harmoniques apparaissent et disparaissent selon que leur valeur est respectivement supérieure ou inférieure au seuil. On croirait que le timbre de l'image 3 est plus riche que celui de l'image 1, alors que c'est le même : seul le seuil change.



**Figure 11.** Courbes d'harmoniques empilées. 3 représentations avec des seuils logarithmiques de plus en plus bas et donc permissifs.

On pourrait conclure qu'il faut montrer le plus d'harmoniques possible, ne rien cacher, et donc choisir un seuil très bas. Or, ce serait un très mauvais choix. L'exemple suivant illustre le problème : on représente des valeurs comprises entre 0 et 10000 sur une échelle logarithmique, dans un rectangle haut de 40 pixels. L'échelle est donc la suivante :  
 $\text{épaisseur en pixel} = 10 * \log_{10}(\text{valeur})$

On voit rapidement sur la figure 12 que 90% des valeurs seront représentées par une épaisseur comprise entre 30 et 40 pixels : la discrimination graphique des valeurs est très faible. Il faut donc choisir un seuil de représentation (dans notre exemple, on



**Figure 12.** Exemple d'échelle graphique :  $\text{épaisseur} = 10 * \log_{10}(\text{valeur})$

prendrait 1000) pour utiliser au mieux l'espace graphique disponible.

Au lieu d'utiliser un simple seuil binaire *représenter / ignorer*, on pourrait utiliser des fonctions non-linéaires : en dessous du seuil, on représenterait tout de même l'harmonique, mais avec une épaisseur moindre qu'au dessus du seuil. Ce problème rejoint finalement celui des échelles de couleur non-linéaires des spectrogrammes (figure 13).

#### 4.3. Représentations en empilement de proportions des harmoniques

On peut aussi représenter la *proportion* des harmoniques dans le son total. L'information d'intensité absolue est alors supprimée : on ne montre que le timbre. Cette représentation peut être, selon les cas, plus éloquent que la représentation absolue : si l'on voit sur la figure 9 que le rapport "fondamentale/3e harmonique" n'est pas le même dans les deux interprétations, c'est beaucoup plus clair sur la figure 14.

Un problème avec ces représentations en enveloppes empilées est que seule la couleur différencie

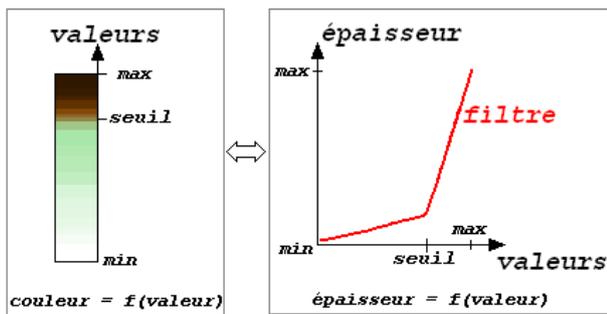


Figure 13. Analogie entre une échelle de couleurs non-linéaire sur un spectrogramme et une fonction non-linéaire valeur  $\rightarrow$  épaisseur.

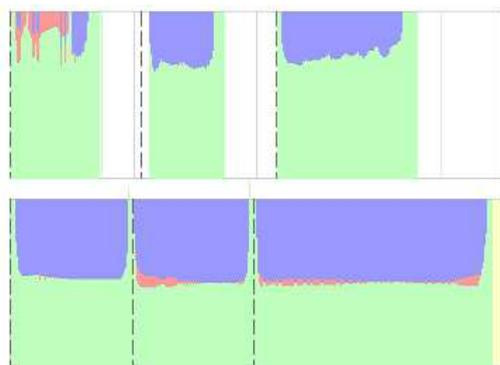


Figure 14. Empilement de proportions d'harmoniques. L'équilibre fondamentale/3e harmonique est différent entre les 2 interprétations.

les harmoniques les unes des autres ; par exemple, si la 10<sup>me</sup> harmonique a les même valeurs que la 2<sup>me</sup>, elles auront des enveloppes de formes identiques - seule la couleur changera. La figure 15 illustre ce problème.

#### 4.4. Spectrogrammes symbolisés

On propose par ailleurs trois représentations voisines du spectrogramme, inspirées de son allure générale, mais encore une fois, seules les valeurs d'harmonique sont représentées. On obtient alors des spectrogrammes *symbolisés, iconifiés* ; par rapport au spectrogramme classique, leur moins grande densité d'information en augmente la lisibilité.

La première est un empilement de bandelettes, une par harmonique, où l'intensité des harmoniques est, comme pour le spectrogramme, représentée par une couleur (figure 16, à gauche).

La deuxième utilise l'autre symbolisme du spectrogramme, à savoir l'épaisseur des raies harmoniques. On représente donc parallèlement, une enveloppe pour chaque harmonique (figure 16, à droite).

La troisième est la combinaison des deux précé-

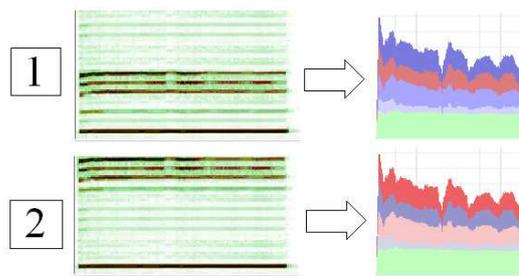


Figure 15. Deux spectrogrammes et leurs représentations en enveloppes harmoniques empilées correspondantes : deux spectres très différents sont représentés par le même empilement d'enveloppes, à leurs couleurs près.

entes ; l'information y est donc représentée de manière redondante, par la couleur *et* par l'épaisseur (figure 16, au centre).

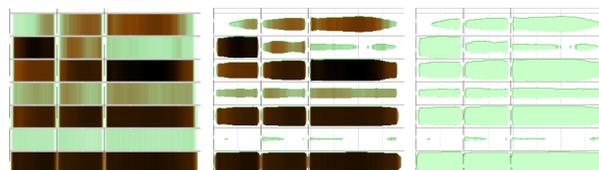


Figure 16. Spectrogrammes symbolisés : à gauche en termes de couleurs ; à droite en termes d'épaisseur. Au centre, leur combinaison.

## 5. REPRÉSENTATIONS DES DÉRIVÉES

Nous présentons ici des propositions utilisant la dérivée de la hauteur et de l'intensité ; la dérivée du spectre ne donne pas de résultats exploitables - pour le moment. La représentation des dérivées des données de base est, à notre connaissance, inédite.

### 5.1. Dérivée de la hauteur

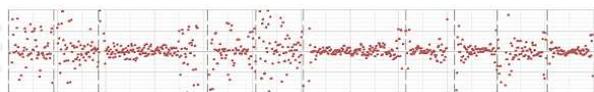
La dérivée de la hauteur va permettre de visualiser, naturellement, les variations de hauteur ; il s'agit d'expérimenter dans quelle mesure elle permet de visualiser des éléments comme la stabilité de la tenue d'une note, ou les vibrati.

L'efficacité de la représentation proposée ici (figure 17) repose sur la combinaison de deux choix graphiques :

- une échelle détaillée, par exemple un axe de -10 cents à +10 cents, pour faire apparaître même les variations de hauteurs les plus fines ;
- l'utilisation de petites tâches pour chaque valeur.

Ainsi, un timbre instable et "tremblant", entraînant des variations chaotiques de la dérivée sera repré-

senté par un "nuage" de tâches centrées autour de zéro. Le nuage sera plus ou moins étendu (en hauteur) selon l'amplitude du tremblement.



**Figure 17.** Dérivée de la hauteur. Echelle [-5 cents, +5 cents].

On peut comparer cette représentation avec la représentation de la justesse que nous avons présentée plus tôt. Contrairement à celle-ci, la représentation de la dérivée s'affranchit du problème de la "note théorique" en comparant chaque hauteur à la hauteur précédente ; c'est donc une représentation de la *stabilité*, et non de la justesse.

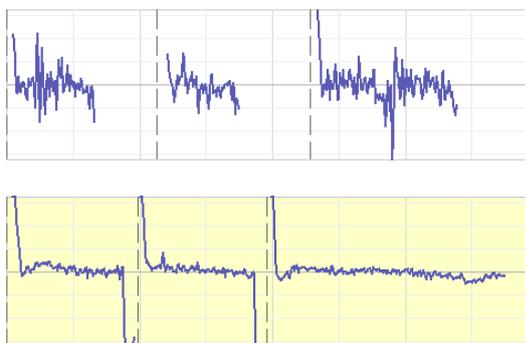
Il est intéressant de comparer par ailleurs les représentations respectives du vibrato par chacune des courbes. Sur l'image 1 de la figure 18, le vibrato final n'est pas parfait, "propre" : si la courbe de justesse nous permet de voir précisément les détails de ce vibrato, la courbe de dérivée se "contente" de le représenter par un nuage.

## 5.2. Dérivée de l'intensité

La dérivée de l'énergie permet de visualiser les variations d'énergie.

La figure 19 donne un exemple de graphique obtenu. Les attaques sont représentées par des pics vers le haut. Les fins de notes sèches sont des pics vers le bas, alors que les decrescendi donnent des traits qui s'arrêtent, sans pic vers le bas final.

Comme pour la représentation de la dérivée de la hauteur, on choisit une échelle détaillée pour visualiser les variations fines. On obtient un graphisme proche d'un *sismogramme*, où l'on peut visualiser les tremblements de l'intensité.



**Figure 19.** Courbe de dérivées des valeurs RMS. En haut, clarinettiste débutant : l'intensité est instable dans la tenue des notes ; en bas, clarinettiste expérimenté : tenue de note stable.

## 6. REPRÉSENTATIONS COMBINÉES

Les représentations combinées ont pour but de représenter différentes données en un unique graphique. Leur intérêt réside donc dans leur concision, qui permet d'avoir une vue plus globale de l'interprétation.

### 6.1. Combinaison hauteur et intensité

Nous proposons ici une représentation combinée de la hauteur et de l'intensité. Elle consistera simplement en une enveloppe pour l'intensité dessinée le long de la courbe de hauteur. Elle permet d'avoir une vue globale de l'interprétation, en traçant, en un seul objet graphique, l'évolution de ces deux informations centrales que sont la hauteur et l'intensité. Les éléments musicaux caractérisés par la relation hauteur - intensité sont ainsi représentés, comme l'illustre la figure 20 : la trille comme bref décalage de la hauteur avec intensité continue, ou un vibrato decrescendo classique en fin de phrase.

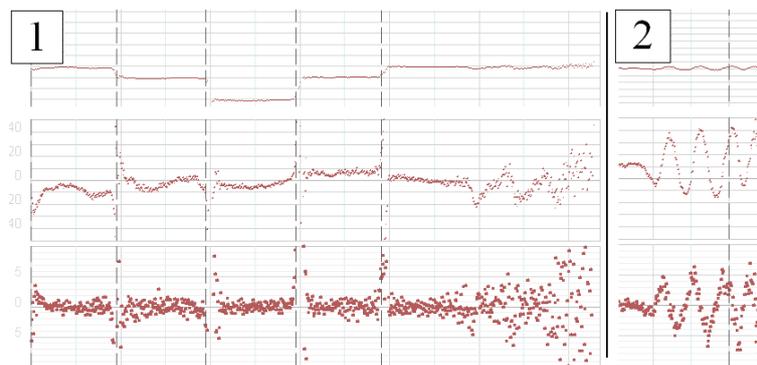


**Figure 20.** Combinaison hauteur & intensité. Phrase avec trille et vibrato.

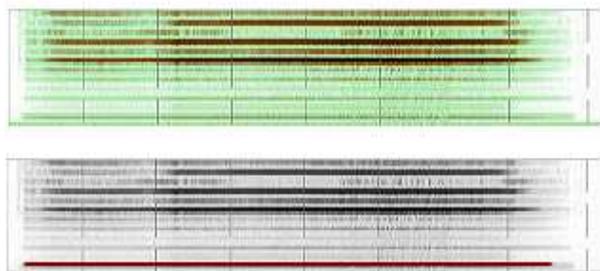
### 6.2. Combinaison hauteur et contenu spectral

Nous proposons deux approches de cette combinaison :

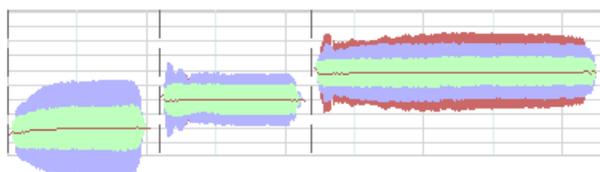
- la première est de simplement représenter une courbe de hauteur en Hertz en surimpression sur le spectrogramme. On obtient ainsi un spectrogramme amélioré où l'on repère clairement quelle "raie" est la fondamentale, ce qui n'est pas nécessairement évident quand on n'est pas habitué à cette représentation (figure 21).
- la deuxième approche est du même type que la combinaison "hauteur et intensité". L'enveloppe d'intensité est remplacée par une enveloppe représentant le contenu spectral du son. L'enveloppe de contenu spectral peut alors prendre les apparences diverses que nous avons énoncées plus haut. La figure 22 est un exemple, qui permet de voir (dans une certaine mesure) la hauteur, les articulations et le timbre.



**Figure 18.** En haut, courbe de hauteur mélodique ; au milieu, courbe de justesse ; en bas, courbe de dérivée de la hauteur.



**Figure 21.** Où est la fondamentale sur le spectrogramme du haut ? Celui du bas ajoute la courbe de hauteur qui marque cette fondamentale, pour la repérer facilement.



**Figure 22.** Courbe symétrique d'enveloppes harmoniques empilées combinée avec une courbe de hauteur.

## 7. LA COURBE GRAPHIQUE COMME MODÈLE

On aura constaté en lisant cet article que l'on compare souvent les mêmes représentations d'interprétations différentes, ceci pour une raison simple : c'est à sa capacité à discriminer graphiquement deux interprétations différentes que l'on estime la qualité d'une représentation ; la différence de perception entre les deux représentations graphiques doit de plus être proportionnelle à celle entre les interprétations.

Dans le contexte du projet VEMUS, la comparaison des représentations présente un grand intérêt pédagogique : tout comme l'élève dispose d'un *modèle sonore* de la partition à travailler, enregistré par un

professeur, il peut également se référer aux représentations graphiques de ce modèle sonore et les comparer à celles de son propre jeu instrumental (figure 23). Le système de représentation peut ainsi s'affranchir de tout critère qualitatif : pas besoin de se "risquer" à qualifier, à évaluer automatiquement les interprétations de l'élève ; le modèle graphique de l'interprétation du professeur fait référence.

## 8. PERSPECTIVES

On se penchera à l'avenir sur les différents problèmes soulevés tout au long de cet article, comme la prise en compte de la variation de la sensibilité de l'oreille avec la fréquence, ou les différents problèmes de calibrage des échelles logarithmiques ; on cherchera aussi des représentations *fin*es de l'articulation, tenant compte des spécificités de chaque catégorie d'instrument.

On peut penser à de nouvelles représentations, comme le timbre dans un espace de couleur, par exemple de type HSV (*Hue - Saturation - Value*, teinte - saturation - luminosité).

Par ailleurs, nous comptons proposer une courbe de déviation temporelle, c'est à dire qui permette de visualiser les variations de l'interprétation par rapport à un métronome de référence.

Mais l'avenir concret de ces représentations passe par le projet VEMUS. Dans ce cadre, il s'agira de tester concrètement leur validité. Cela se fera de différentes manières :

- via des questionnaires d'évaluation interactifs, où l'on demande d'associer la bonne courbe avec la bonne interprétation ;
- une phase de test d'un prototype de VEMUS sera menée dans des écoles de musique : les professeurs et les élèves retourneront alors leurs impressions vis-à-vis des différentes représentations. Cela permettra d'en valider certaines, d'en améliorer d'autres et de laisser de côté celles jugées peu pertinentes.

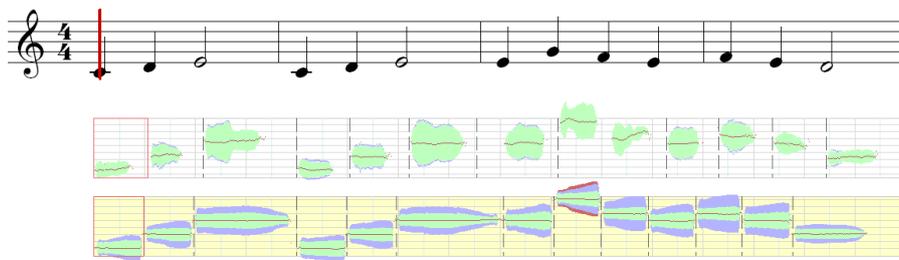


Figure 23. Comparaison de la courbe du professeur et de l'élève dans la partition étendue de VEMUS.

## 9. CONCLUSION

Nous avons proposé une palette de représentations, plus ou moins innovantes, toutes orientées vers la visualisation de caractéristiques du jeu instrumental.

Certaines caractéristiques, comme la justesse, sont représentées avec une précision qui dépasse celle de notre oreille ; en revanche, des caractéristiques plus complexes comme les articulations ou le timbre sont représentées de manière incomplète. Il devient difficile de les faire apparaître en utilisant directement les données de base ; des systèmes plus "intelligents" seraient peut-être nécessaires.

Avoir une perception objective de son propre jeu est très difficile pour un instrumentiste ; on espère que les représentations présentées ici, qui viennent s'ajouter à celles déjà existantes, seront un outil efficace pour le développement de cette oreille extérieure, pour que l'instrumentiste puisse bientôt s'affranchir de toutes ces représentations. Hors cadre pédagogique, ces représentations pourront être utilisées à des fins d'analyse du jeu instrumental.

## 10. REFERENCES

- [1] Jean Bresson and Carlos Agon. Sdif sound description data representation and manipulation in computer assisted composition. In *ICMC'04 : Proceedings of the International Computer Music Conference*, pages 520–527. ICMA, 2004.
- [2] Jonathan Foote. Visualizing music and audio using self-similarity. In *MULTIMEDIA '99 : Proceedings of the seventh ACM international conference on Multimedia (Part 1)*, pages 77–80. ACM Press, 1999.
- [3] Yann Geslin and Adrien Lefevre. Sound and musical representation : the acousmographe software. In *ICMC'04 : Proceedings of the International Computer Music Conference*, pages 285–289. ICMA, 2004.
- [4] Rumi Hiraga and Noriyuki Matsuda. Visualization of music performance as an aid to listener's comprehension. In *Proceedings of the working conference on Advanced visual interfaces*, pages 103–106, May 2004.
- [5] Rumi Hiraga, Reiko Mizaki, and Issei Fujishiro. Performance visualization : a new challenge to music through visualization. In *Proceedings of the tenth ACM international conference on Multimedia*, pages 239–242, December 2002.
- [6] Olivier Koechlin and Hughes Vinet. The acousmographe, a macintosh software for the graphical representation of sounds. In *Proceedings of the International Computer Music Conference (ICMC'91)*, pages 586–588. ICMA, 1991.
- [7] F. Wai ling Ho-Ching, Jennifer Mankoff, and James A. Landay. Can you see what i hear? the design and evaluation of a peripheral sound display for the deaf. In *CHI '03 : Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 161–168, New York, NY, USA, 2003. ACM Press.
- [8] Philip McLeod and Geoff Wyvill. Visualization of musical pitch. In *Computer Graphics International 2003*, 2003.
- [9] Geoffroy Peeters, Amaury La Burthe, and Xavier Rodet. Toward automatic music audio summary generation from signal analysis. In *ISMIR'02 : Proceedings of the International Conference on Music Information Retrieval*, pages 94–100. IRCAM, 2002.
- [10] Anne Sedes, Benoît Courribet, and Jean-Baptiste Thiébaud. Visualization of sound as a control interface. In *Proceedings of the 7th Int. Conference on Digital Audio Effects (DAFx'04)*, pages 390–394, 2004.