



La musique en images

CHRISTOPH PÖPPE

Comment former des images à partir des sons.

Comment Transformer les sons en images ? La recherche d'une telle opération est utile : un coup d'œil suffirait pour savoir si une émission sonore est intentionnelle ou aléatoire. Gordon Monro, à l'Université de Sydney, et Jeff Pressing, à l'Institut de psychologie de Melbourne, explorent ce type de visualisations.

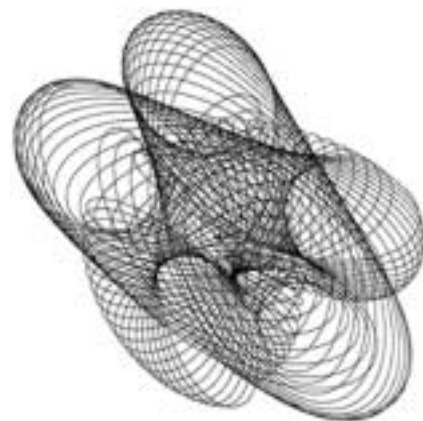
Ils ont notamment utilisé les données fournies par les systèmes d'enregistrement de disques compacts : environ 44 000 fois par seconde, la tension électrique formée à partir du son, dans le microphone, est mémorisée ; on transforme ainsi un signal continu en une suite de valeurs, qui sont stockées sous forme codée sur les disques compacts.

Désignons ces valeurs par $s_1, s_2, s_3, \dots$ et portons dans un plan muni d'un repère les points (s_i, s_{i+1}) . Normalement, on utilise les repères cartésiens pour représenter les fonctions : aux variables x de l'axe des abscisses, les fonctions associent des valeurs y repérées sur l'axe des ordonnées. Dans l'étude du son, on ne cherche pas si une fonction f associe chaque valeur à la précédente ($s_{i+1} = f(s_i)$), mais on construit un graphe en portant des points d'abscisse s_i et d'ordonnée s_{i+1} dans un plan comme si une telle fonction existait. Ce mode de représentation

fait ainsi apparaître d'éventuelles relations (des «corrélations») entre une valeur et celle qui la précède immédiatement, autrement dit entre la suite des valeurs et une version d'elle-même décalée dans le temps, d'où le nom de «diagramme d'autocorrélation» pour l'image que forment les innombrables points qui décrivent un signal sonore.

Quelle image obtient-on par ce procédé ? Sans précautions particulières, l'image semble insignifiante, car, avec une fréquence d'échantillonnage de 44 000 par seconde, deux valeurs successives sont si proches que rien n'apparaît : les valeurs s_i et s_{i+1} sont quasi identiques, et la représentation graphique correspondante est très proche de la droite $x = y$, la bissectrice du premier quadrant.

L'image devient plus intéressante quand on utilise un écart supérieur entre les valeurs représentées, tout en conservant tous les points de l'échantillonnage : on porte par exemple les points (s_i, s_{i+10}) au lieu des points (s_i, s_{i+1}) . D'ailleurs, au lieu de 10, on peut choisir n'importe quel écart de temps k afin d'obtenir une image aussi instructive que possible. Avec un pas k égal à 10 et un son de fréquence égale à 1 000 hertz (ce qui correspond à une note aiguë de soprano), les deux valeurs à représenter



1. Image d'un accord de trois sons purs : do, sol et mi.

sont espacées d'un quart de période environ, ce qui est optimal.

Pour une oscillation sinusoïdale, l'image ainsi obtenue est une ellipse (lorsque l'écart de temps est exactement d'un quart de période, c'est même un cercle). D'autre part, tout signal périodique peut être décrit comme une somme de sinusoïdes de fréquences multiples de la fréquence minimale, dite fondamentale. Lorsque la première oscillation harmonique (une octave au-dessus du son fondamental, c'est-à-dire à une fréquence double de celle du fondamental) s'ajoute au fondamental avec une intensité inférieure, l'ellipse se déforme jusqu'à former des nœuds (voir l'encadré ci-contre). Sur l'image, un déphasage entre les deux oscillations apparaît alors qu'il est inaudible. Des harmoniques supérieures engendrent de remarquables motifs en forme de fleurs, lorsque l'on choisit un écart de temps approprié, et un accord à trois notes donne une courbe compliquée (voir la figure 1). Un son de flûte aigu et très fort donne un champ gris avec une structure à peine reconnaissable (voir la figure 3).

Dans quelles conditions les points forment-ils des courbes reconnaissables ? Supposons qu'existe réellement une fonction f qui corresponde à la relation $s_{i+1} = f(s_i)$. Si seule la première valeur s_1 de la suite est fixée, toutes les autres valeurs seraient déterminées par la fonction f : $s_2 = f(s_1)$, $s_3 = f(s_2)$, $s_4 = f(s_3)$, etc. La fonction f est ce que l'on nomme une fonction d'itération, et le système dyna-

Images d'autocorrélation obtenues pour des superpositions d'oscillations sinusoïdales pures. À gauche, les images sont celles que l'on obtient pour une note sinusoïdale pure à laquelle on a ajouté la note à l'octave supérieure, deux fois moins intense ; les deux courbes ne diffèrent que par le déphasage des deux notes. À droite, l'image est celle d'une note sinusoïdale avec plusieurs harmoniques (le nombre m indique le rapport entre les fréquences de l'oscillation harmonique et de l'oscillation de base). Les écarts de temps (α indique l'écart en fraction de période) sont choisis de façon à obtenir une image la plus symétrique possible.

