

Explorer la musique à l'âge numérique

Hugues Vinet

En cinquante ans, le développement de l'informatique et des technologies numériques a suscité un bouleversement radical des pratiques musicales, de la création à la diffusion en passant par la lutherie, l'édition, la production, et bientôt les nouveaux modes d'écoute instrumentée. Après une présentation des principaux jalons de cette histoire récente, à laquelle j'ai eu la chance de participer au cours des vingt dernières années¹, je proposerai une mise en perspective des mutations en cours à partir des dernières avancées des recherches scientifiques et technologiques appliquées à la musique. Mon propos visera en particulier, sur la base d'une synthèse des différentes approches existantes et d'exemples issus des recherches de l'Ircam, à montrer en quoi les technologies numériques ouvrent de nouvelles heuristiques en matière de représentation et de manipulation du musical, aussi bien pour les créateurs que pour le public mélomane.

Synthèse et traitement du signal audionumérique

Les premières synthèses audionumériques, conduites par Max Mathews aux Bell Labs à la fin des années cinquante, sont quasi-contemporaines de l'apparition de l'informatique et des premiers convertisseurs numérique-analogique. Initialement limitées à la synthèse d'oscillateurs périodiques élémentaires, les possibilités de traitement sonore vont être progressivement étendues à tous types de sons à partir des années 60 et 70, avec le développement des théories du traitement du signal numérique. Celles-ci représentent, par rapport aux techniques analogiques, un saut qualitatif, et ouvrent de nombreux types de transformations sonores irréalisables auparavant : transposition, *time-stretching*², micro-montage, hybridation entre sons, etc .

L'informatique introduit également un nouveau paradigme d'utilisation : la programmation. Ainsi, dès l'origine, les premiers logiciels de synthèse musicale se présentent-ils sous la forme de *langages musicaux*, offrant à l'utilisateur musiciens des fonctions de programmation relatives aux différentes étapes de création d'une séquence sonore, structurées selon les métaphores de la lutherie et de la composition : construction d'un « instrument » sous la forme d'un algorithme de synthèse par assemblage de fonctions élémentaires (oscillateurs, filtres, retards, etc.), puis constitution d'une « partition » pour cet instrument par spécification de la variation des paramètres d'entrée de l'algorithme au cours du temps.

Une autre évolution importante des technologies de traitement du signal concerne la possibilité du traitement sonore en *temps réel*, dont des centres liés à la création contemporaine comme l'Ircam ont été les pionniers : alors qu'il fallait autrefois attendre plusieurs heures pour le calcul de quelques secondes de son, l'électronique devient un instrument à part entière, qui dialogue avec les interprètes en direct pendant le concert selon des modes d'interaction conçus par le compositeur. Le logiciel Max, langage graphique pour la réalisation d'algorithmes de traitement d'informations musicales en temps réel, constitue l'aboutissement actuel de ces travaux. Jusqu'à présent réservée aux initiés, la programmation est ainsi rendue accessible, par l'intermédiaire d'interfaces graphiques intuitives, à des catégories d'utilisateurs de plus en plus larges.

¹ En tant que responsable des recherches et développements de deux institutions dédiées à la création musicale contemporaine : le Groupe de recherches musicales de l'Ina, puis l'Ircam.

² Variation de vitesse de lecture sans modification de hauteur.

MIDI et le Home studio

La norme MIDI, introduite en 1983, propose pour la première fois un codage numérique des informations musicales, sur la base d'une représentation simplifiée du jeu au clavier : hauteur de note au demi-ton, vitesse (vitesses d'enfoncement des touches), etc. Elle s'impose rapidement comme standard d'interconnexion entre instruments de musique électroniques, en introduisant dès lors une dissociation, caractéristique de la lutherie électronique, entre dispositifs de commande gestuelle et synthétiseurs. L'essor simultané de la microinformatique va cristalliser une nouvelle forme de pratiques de composition amateur : le *Home studio* (studio personnel). L'ordinateur, désormais accessible au plus grand nombre, y est le centre d'un dispositif d'instruments de musique interconnectés en MIDI, et dont il gère à la fois les réglages et configurations, mais également le pilotage à travers des logiciels de composition d'un nouveau type : les séquenceurs. Dans la même mouvance, le bouleversement introduit dans les métiers de l'édition à la fin des années 80 par le développement des techniques de publication assistée par ordinateur s'applique également à la musique : l'édition des partitions et leur impression de qualité professionnelle devient bientôt accessible au plus grand nombre et s'intègre au sein du home studio comme une des fonctions disponibles.

Du fait de l'accroissement de la puissance des microordinateurs, permettant non plus seulement le traitement d'informations MIDI, mais également des signaux sonores, le phénomène des home studios va se propager aux studios de production professionnels et marquer la fin des techniques analogiques. Au-delà du saut qualitatif provoqué par l'audio numérique, l'un des principaux facteurs technologiques de rupture introduits par l'utilisation de l'ordinateur concerne le stockage des enregistrements sonores, qui passe de supports linéaires (la bande magnétique) à des supports rotatifs (disque dur, CD), voire à des mémoire vives, permettant un accès quasi-instantané, sans déroulement, à tout instant du son. Cette nouvelle possibilité suscite la réalisation d'applications de montage et de mixage *virtuels* : l'opération, *non destructive*, porte désormais sur une représentation du matériau et non plus sur son support physique, le résultat sonore étant immédiatement disponible à l'écoute.

Du signal aux modèles physiques

Les recherches menées en acoustique sur le fonctionnement des sources sonores, et notamment des instruments de musique, aboutissent à la réalisation de modèles informatiques de synthèse, dits *modèles physiques*. Ceux-ci se distinguent des modèles de signaux dans la mesure où ils s'attachent à décrire les causes physiques de production des sons, et non à en modéliser les effets par des formalismes mathématiques issus des théories du signal. Les sons ainsi produits ont la richesse de sons naturels. De plus, les paramètres de ces modèles correspondent à des grandeurs physiques (dimensions et matériaux d'objets modélisés, force ou pression exercées, etc.) et revêtent une signification plus immédiate pour l'utilisateur musicien. Enfin, le fonctionnement de l'instrument y est décrit comme un système dynamique, permettant une simulation réaliste de phénomènes complexes tels que les transitions entre notes, les bifurcations ou les passages à l'octave dans les instruments à vent. Nécessitant une puissance de calcul importante, ainsi que la mise en œuvre de dispositifs de commande élaborés, ces modèles, jusqu'à présent réservés aux centres de recherche, trouvent aujourd'hui une diffusion plus large, notamment sous la forme de logiciels musicaux. Une des applications les plus abouties de cette technique est le logiciel Modalys de l'Ircam dont la fonction relève de la *lutherie virtuelle* : Modalys permet la « construction » d'instruments virtuels par assemblage de structures physiques élémentaires dont l'ensemble des caractéristiques physiques sont paramétrables : cordes, tuyaux, plaques, membranes, anches simples et doubles, embouchure de trompette, etc. Le formalisme utilisé autorise la création de structures éventuellement irréalisables, tels que des hybrides entre une membrane et un

tube, mais dont le son produit garde une vraisemblance physique. La lutherie informatique, du moins dans ses possibilités sonores, peut ainsi être considérée comme une extension de la lutherie traditionnelle, dans la mesure où tous types d'instruments existants peuvent être simulés. En pratique, si les sources sonores peuvent être ainsi modélisées dans l'ordinateur, le problème de leur contrôle, nécessitant l'apprentissage de l'expertise d'un interprète, reste entier. Une résolution possible passe par la réalisation de dispositifs de captation gestuelle adaptés, effectuant non seulement une analyse fine des mouvements de l'interprète, mais également, dans le cas le plus complet, un retour d'effort lui fournissant les informations de proprioception simulant son contact avec la structure physique virtuelle. L'analyse et la modélisation des informations gestuelles, issues de dispositifs de capture du mouvement, constitue aujourd'hui une nouvelle direction de recherche, visant notamment à définir des représentations pertinentes des informations gestuelles et les mettre en œuvre dans des applications informatiques pour la création et l'interprétation.

Une autre application musicale de la modélisation acoustique concerne la *spatialisation* des sons, c'est-à-dire la reproduction ou la simulation de sources sonores localisées dans un espace doté de caractéristiques acoustiques particulières. Les applications en production sonore ont longtemps reposé sur des systèmes de captation et de restitution stéréophoniques, l'effet de spatialisation étant simulé à l'aide de techniques de traitement de signal : panoramique³ pour la position des sources, réverbération artificielle pour l'effet de salle. Pour une salle de géométrie donnée, revêtue de matériaux connus, la modélisation acoustique permet, là encore, une simulation plus précise et réaliste, mais est mieux adaptée à des configurations statiques (telles qu'en acoustique architecturale) qu'à des applications de spatialisation visant la simulation de sources en mouvement. Les travaux les plus avancés dans ce domaine portent sur une description perceptive de l'effet spatial (relative à la qualité acoustique de la salle, aux positions des sources,...) dans le cadre de systèmes, tels le Spatialisateur de l'Ircam, qui adaptent automatiquement cette spécification spatiale au dispositif de restitution utilisé : stéréophonie, Dolby Surround, haut-parleurs multiples, casque, etc. Parmi les techniques de diffusion les plus perfectionnées, destinées à supplanter à terme les haut-parleurs dans certaines configurations, on peut mentionner la *Wavefield synthesis*, ou holographie acoustique : des panneaux dotés de multiples sources élémentaires entourent l'auditoire et recréent un champ acoustique tridimensionnel homogène pour tous les auditeurs. La précision croissante de ces dispositifs de rendu ouvre un nouveau champ expressif lié à la conception spatiale du son, qui dépasse l'expertise traditionnelle des ingénieurs du son en stéréophonie, reposant sur des représentations numériques combinant données physiques, traitement du signal et descriptions symboliques des scènes sonores.

Composition assistée par ordinateur

La diffusion, au cours des années 80, du protocole MIDI et des logiciels d'édition de partitions, avait marqué le développement de méthodes de *représentation symbolique* des informations musicales dans l'ordinateur, sous différentes formes spécialisées, selon qu'elles concernent le codage du jeu instrumental ou des partitions. Les applications de ces représentations symboliques vont alors être généralisées dans la réalisation, au sein de centres de recherche liés à la création contemporaine, de logiciels dits de *Composition assistée par ordinateur*, tels que Patchwork puis OpenMusic, développés à l'Ircam : les fonctions de programmation qu'ils proposent, appliquées ici aux structures musicales (et non plus au signal sonore), permettent aux compositeurs le calcul de matériaux complexes (agrégats de hauteurs, rythmes, timbres,...) destinés à les assister dans leur travail d'écriture orchestrale. Les avancées récentes de ces travaux convergent également vers le contrôle de synthétiseurs

³ Application de gains différents sur les canaux gauche et droite.

sonores directement à partir de structures musicales symboliques, ainsi que vers la mise en œuvre de méthodes d'analyse musicale automatisée, qui ouvrent de nouvelles possibilités aux musicologues et à l'explicitation des œuvres.

Métadonnées et descriptions des contenus musicaux

Une étape ultérieure en matière de représentation des informations musicales est franchie avec la réalisation de bases de données d'enregistrements musicaux dotées de fonctions de *recherche par contenu*, qui seraient au contenu musical ce que les moteurs de recherche sur le Web sont au texte. Les recherches dans ce domaine, qui représentent des enjeux économiques considérables, impliquant une réorganisation potentielle du système actuel de distribution de la musique enregistrée, ont contribué à fédérer, depuis les années 2000, d'importantes communautés de chercheurs. Les problèmes abordés consistent d'une part à constituer des structures pertinentes de description des informations musicales en fonction des applications visées, d'autre part à réaliser des procédures d'analyse, effectuant autant que possible automatiquement l'extraction des informations nécessaires à partir des enregistrements ou de la partition. (cf encadré [Navigation dans les bases de données musicales](#)). Des recherches avancées visent également le développement de techniques de *synthèse et de traitement sonore par le contenu*, à travers lesquelles l'utilisateur musicien spécifie le résultat sonore attendu selon un ensemble d'attributs qualitatifs, et non plus en détaillant les méthodes et algorithmes explicites de production du son : « son plus métallique, moins sourd, plus granuleux, avec moins d'attaque, ... ». L'ensemble de ces applications, relevant de fonctions de gestion des informations musicales dites *par le contenu*, partagent un mode d'accès commun aux contenus musicaux, dans lequel ceux-ci ne sont pas manipulés directement, mais à travers des structures de descriptions intermédiaires, dites *métadonnées*, qui formalisent les connaissances nécessaires à l'application. Ces structures constituent un type spécifique de représentations informatiques de la musique et se présentent généralement sous une forme textuelle associée à des catégories, taxonomies, réseaux sémantiques et autres descriptions qualitatives. Convenons de les appeler, dans une référence strictement étymologique, *représentations cognitives*. De par leur nature contextuelle, les connaissances utiles pour une application informatique donnée dépendent notamment de la compétence de l'utilisateur, des corpus manipulés, des fonctions visées, et se présentent généralement non pas sous la forme de structures figées, mais de bases de connaissances évolutives et personnalisables. Dans le contexte de nouvelles applications de distribution de la musique enregistrée par les réseaux, cette possibilité d'adaptation personnalisée de l'offre représente un attrait commercial important.

L'espace de représentation, de la partition aux modèles

Les synthèses qui précèdent ont mis en évidence quatre types principaux, distincts et complémentaires, de représentations numériques des informations musicales. Historiquement, les représentations de type *signal* et *symbolique* ont été présentes dès les débuts de l'informatique musicale et sont aujourd'hui largement mises en œuvre, sous des formes diverses, dans les applications commerciales. L'utilisation des représentations de types *physique* et *cognitif* sont plus récentes et leur débouchés émergents. L'objet des normes et standards techniques a suivi cette évolution : après le codage des signaux audio numériques (AES-EBU, etc.) et MIDI, les normes audiovisuelles MPEG1 (notamment à travers le format mp3) et MPEG2 sont restées au niveau du signal en se concentrant sur la compression des données, puis sont complétées par MPEG4, qui aborde la description de scènes sonores composites incluant des représentations de types signal, symbolique et physique. Les dernières normes en date, MPEG7 et MPEG21, se concentrent quasi-exclusivement sur le

niveau cognitif, visant la réalisation d'applications de gestion des informations musicales par le contenu.

Il est aisé de montrer que ces quatre types de représentations s'organisent en *niveaux d'abstraction* croissants de description du phénomène musical, du physique au cognitif en passant par le signal puis le symbolique. Du point de vue de la théorie de l'information, cet ordre rend également compte de quantités d'information⁴ décroissantes. L'opération de conversion d'une représentation à celle d'un niveau supérieur correspond donc à une extraction d'information, et relève du processus générique d'analyse. Par exemple, le passage du signal au symbolique (de l'enregistrement à la partition) correspond-il à une double discrétisation, à la fois sur l'échelle des temps (temps continu à temps pulsé) et des valeurs prises par les grandeurs décrites (e.g. passage du continuum des fréquences aux échelles de hauteurs). Inversement, le processus de synthèse correspond-il à une génération d'information, à partir de données discrètes, par activation de connaissances implicites propres au modèle de synthèse utilisé (e.g. modèle physique, détenant le fonctionnement d'un instrument donné). Or, la création artistique, vue sous l'angle de la technique, peut être ramenée à la synthèse, en tant que processus de concrétisation et de spécification, partant de conceptions abstraites et de formes potentielles et aboutissant, à travers des opérations et choix successifs, à une œuvre dont l'ensemble des caractéristiques matérielles sont déterminées. En musique, ce processus de concrétisation s'effectue traditionnellement selon deux phases successives, la première, incombant au compositeur, visant la réalisation d'une représentation de type symbolique (la partition), la seconde, de l'ordre de l'interprétation, réalisant une actualisation, en tant que grandeurs physiques, du potentiel qu'elle contient. L'utilisation de l'informatique bouleverse cette division du travail obligée à travers la médiation de la partition, permettant par exemple au compositeur d'être son propre luthier et interprète ou à l'instrumentiste de concevoir des dispositifs interactifs pour l'improvisation. Plus encore, les fonctions introduites par la programmation permettent au créateur de réaliser son propre *modèle génératif*. Ce concept, que je propose comme extension de celui de modèle descriptif en épistémologie, rend compte des différents modes d'utilisation de l'ordinateur dans la création musicale, ainsi que de la nature du transfert conceptuel, au-delà de la notion générale de technique, entre recherche scientifique et création artistique. L'objet du modèle génératif est de produire des matériaux musicaux (échantillons de signal, notation, etc.) comme résultat d'un calcul algorithmique. Les paramètres d'entrée du modèle représentent l'espace de variation possible, chaque instance de matériau produit étant fonction d'une combinaison particulière de ces données d'entrée. Cet espace, programmé par le créateur pour décrire les variations qu'il souhaite apporter au matériau, définit une heuristique propre au modèle, en particulier par extrapolation (voisinage d'une position définie) et interpolation (trajet d'un point à un autre). Il généralise ainsi l'espace de manipulation de la partition, qui reposait sur les catégories traditionnelles de la théorie musicale⁵ et de l'instrumentarium, qu'il étend respectivement du point de vue des structures de représentation, désormais configurables selon une infinité de systèmes propres aux différentes approches esthétiques, et de la lutherie électronique comme algorithmique de synthèse.

Le numérique comme restauration du continuum technique des représentations

Au-delà des apports de l'ordinateur à la création, la démocratisation de l'Internet et des réseaux numériques ouvre de nouveaux modes de diffusion et d'accès aux contenus musicaux numérisés. En effet, le médium de l'*album*, collection de morceaux enregistrés, qui structure tout le système actuel de distribution de la musique, est limitatif à la fois du point de vue de la

⁴ Nombres de bits nécessaires au codage de l'information.

⁵ Hauteurs tempérées, métrique et rythmes, nuances, etc.

nature des données transmises (le signal de l'enregistrement) et des possibilités d'interfaces d'accès et de manipulation de ces données. Les réseaux numériques autorisent toutes formes de diffusion combinant différentes représentations des contenus musicaux sur les quatre niveaux définis plus haut, ainsi que des données documentaires associées (textes, photos, vidéos d'interviews, etc.). Ainsi, le contenu musical peut être transmis non plus uniquement comme un bloc homogène, déjà composé et mixé, mais comme l'articulation structurelle d'éléments de matériau accessibles à la fois individuellement et en contexte (mélodies, rythmes, couleurs harmoniques,...), tant dans la dimension temporelle (organisation dans le temps des différents matériaux) que de la superposition et de la spatialité polyphoniques (notamment : possibilité d'écoute indépendante des différentes voies). Cette « décomposition » structurée de l'œuvre peut résulter d'une analyse musicologique faite a posteriori, voire même d'un accès direct aux modèles et structures de représentations mises en œuvre par les compositeurs dans les outils de production. C'est en ce sens que la diffusion numérique restaure potentiellement le continuum des représentations, des outils de création aux nouveaux instruments d'écoute. L'enjeu, en matière de diffusion culturelle, est important : une telle médiation analytique des œuvres est susceptible de favoriser leur appropriation par le public mélomane à travers sa familiarisation avec les éléments et structures les constituant, en vue d'une meilleure intégration synthétique de leur fonction musicale en contexte, voire à travers des outils de re-composition permettant d'expérimenter des possibilités de variation et de réarrangement des matériaux. Cette dernière possibilité ouvre également la voie à de nouvelles formes de création spécialement conçues en fonction du caractère interactif des interfaces d'accès. Tels sont notamment les objectifs poursuivis par le projet européen SemanticHIFI (**voir encadré**). L'intérêt, évident pour la musique contemporaine, porteuse d'une complexité la rendant difficile d'accès, existe également pour les musiques dites populaires, dont les constituants structurels gagneraient à être explicités, ne serait-ce que pour être objectivés et démystifiés. Un tel projet, réalisable dès aujourd'hui techniquement, implique surtout une modification radicale de la chaîne technologique de production, de diffusion, d'accès et de manipulation des représentations numériques du musical, préservant autant que possible toute la richesse et la variété de leurs formes existantes et à venir.

Le projet SemanticHIFI (encadré)

Le projet SemanticHIFI, soutenu par la Commission européenne (programme IST) et démarré début 2004 pour une durée de trois ans, vise la préfiguration de la chaîne Hi-fi de demain. Il regroupe à cet effet un consortium européen, coordonné par l'Ircam, associant industriels (Sony Europe, Native Instruments, leader mondial des logiciels pour DJ) et laboratoires de recherche (Fraunhofer IDMT, Sony-CSL, UPF à Barcelone, Université Ben Gourion en Israël, Ircam).

Les principales limites des chaînes Hi-fi traditionnelles résultent d'abord de celles des supports de diffusion de la musique, qui se présentent essentiellement sous forme d'enregistrements stéréophoniques. Les possibilités de manipulation liées à ce type de support d'informations musicales se résument ainsi en quelques fonctions : sélection du morceau, lecture, arrêt, commande de volume, etc. L'objectif de SemanticHIFI est d'expérimenter des modes d'interaction plus riches, fondés sur des représentations plus élaborées des contenus musicaux. Ainsi, la chaîne Hi-fi de demain sera dotée de capacités de stockage permettant la gestion de dizaines de milliers de morceaux, et SemanticHIFI prévoit la réalisation d'outils de classification personnalisée, qui, analyseront automatiquement les contenus musicaux, autorisant différentes heuristiques de navigation entre morceaux (voir encadré). Une autre fonction prévue concerne la navigation à l'intérieur d'un morceau, selon plusieurs modes complémentaires : d'une part la capacité de se déplacer virtuellement entre les différents instruments ou voies de polyphonie et de réaliser son propre mixage assisté, à l'aide d'interfaces apportant une nouvelle dimension au rendu spatial tridimensionnel dans le contexte domestique (voir figure XXX) ; d'autre part, à partir d'un découpage automatisé dans le temps des différentes parties constituant le morceau, la possibilité de donner à voir ces structures internes et de passer instantanément de l'une à l'autre (voir figure YYY). La production d'œuvres spécialement conçues pour ces dispositifs de présentation analytique et de navigation fait également partie de l'expérimentation à mener. D'autres fonctions plus ludiques, de l'ordre de l'interprétation (accompagnement automatique, chœur virtuel, chef d'orchestre, etc.) sont également prévues, ainsi que la possibilité de créer de petites compositions par assemblage de fragments musicaux (fonction DJ). Enfin, la chaîne Hi-fi de demain sera connectée à l'Internet et dotée de fonctions d'accès à des sites distants, mais aussi de partage, auprès de communautés de pairs, du travail d'indexation, de spatialisation, d'interprétation et de composition réalisé grâce à ces nouvelles fonctions conférant un rôle plus actif à l'utilisateur mélomane.

Navigation dans les bases de données musicales

Les bases de données de morceaux de musique enregistrée tendent à se généraliser, qu'elles soient sous la forme de serveurs centralisés (portail d'achat sur le Web) ou de réseaux d'ordinateurs interconnectés en *peer-to-peer* (Kazaa, etc.), voire de logiciels pour microordinateurs (logiciel iTunes d'Apple). Afin de faciliter la recherche parmi les très nombreux items accessibles tout en favorisant la découverte de nouveaux morceaux, différents modes de description sont proposés, dépassant les informations éditoriales traditionnelles (titres, noms des compositeurs et interprètes,...). La plupart des systèmes reposent en particulier sur des taxonomies de genres musicaux (jazz, country, classique,...), généralement incompatibles entre elles, chacune ayant ses spécificités, mais restant nécessaires pour restreindre le champ de recherche. L'utilisation d'autres modes de description, plus objectifs, a fait l'objet de travaux récents portant en particulier sur le tempo, l'intensité sonore ou la couleur orchestrale, l'avantage de tels descripteurs étant qu'il peuvent être calculés automatiquement à partir des enregistrements, évitant une indexation manuelle. Un autre mode d'association entre morceaux (filtrage collaboratif) consiste à observer les comportements des usagers et à dégager des classes d'utilisateurs ayant des goûts proches, le système proposant alors des morceaux choisis par les autres usagers de la même classe. La combinaison de ces différents critères objectifs et subjectifs permet la constitution de *mesures de similarité* composites entre morceaux et ouvrent des heuristiques de *recherche par l'exemple* : à partir d'un morceau donné comme point de départ par l'utilisateur, le système peut fournir des listes de morceaux proches selon telle ou telle combinaison de critères de similarité, produisant des découvertes insolites, mais gardant toujours une cohérence d'appariement. Au-delà de ces critères de description globaux, les systèmes les plus perfectionnés reposent sur une analyse des structures internes du morceau. Ainsi, dans le cas de mélodies simples (chansons), les méthodes de *recherche par chantonnement* obtiennent un bon taux de fiabilité et de robustesse... aux prestations approximatives de chant d'un thème connu de la part de l'utilisateur. Des recherches plus avancées, mais encore expérimentales, extraient automatiquement les motifs représentatifs d'œuvres polyphoniques et les associent selon des critères de similarité combinant mélodie, harmonie, rythme, etc. La complexité qui résulte des structures ainsi extraites est toutefois à la hauteur de la celle du contenu musical lui-même, le problème étant de dégager un niveau de description à la fois représentatif et suffisamment simplifié pour fournir des interfaces de navigation lisibles et opérationnelles.

Légende figure XXX : Logiciel MusicSpace : interface de navigation entre les différentes parties instrumentales d'un morceau, l'effet de spatialisation simulant le déplacement de l'auditeur au sein de l'orchestre. Un système de contraintes programmables restreint les variations possibles pour préserver la cohérence du résultat sonore (mixage assisté). Crédit : Olivier Delerue, François Pachet, ©Sony-CSL.

Légende figure YYY: Production d'un résumé automatique à partir d'une analyse de signal, par visualisation de la succession des principaux états constituant la structure du morceau. Crédit : Geoffroy Peeters, ©Ircam.