

Navigation dans les bases de données musicales

Hugues Vinet¹

Les rebondissements de l'affaire Napster l'auront montré : la généralisation en cours des technologies numériques est susceptible d'ébranler, sinon de remettre en cause les formes existantes de diffusion de la musique. Ces évolutions concrétisent l'émergence d'une nouvelle ère de la reproductibilité et induisent potentiellement le déploiement d'une nouvelle génération de dispositifs techniques de médiation, qui sont à même de bouleverser notre rapport à l'écoute musicale (Stiegler, 2002). D'une part en autorisant tous types de représentation du contenu musical, non plus limitées au codage audionumérique de son enregistrement, mais combinant différentes descriptions et médias documentaires associés (partition, représentations graphiques, commentaires textuels, etc.). D'autre part en structurant ces représentations, en particulier par l'utilisation de bases de données et en organisant leur manipulation interactive sous la forme d'applications intégrant des interfaces homme-machine adaptées. La constitution d'un nouveau *système technique de diffusion de la musique*, qui serait aux réseaux numériques ce que la radio est à l'analogique hertzien, est certes encore loin d'être stabilisée et de proposer des alternatives viables à l'édition traditionnelle sous forme de disques compacts. Compte tenu des enjeux économiques considérables qu'elles représentent, les problématiques liées à la recherche d'informations musicales font toutefois l'objet d'une mobilisation importante de la part de différentes communautés scientifiques, dont la mise en place depuis 2000, du colloque international ISMIR (*International Conference on Music Information Retrieval*), marque la

¹Directeur scientifique de l'Institut de recherche et coordination acoustique/ musique (Ircam).

structuration. Certains de ces travaux, réalisés en particulier dans le cadre de projets pilotes aux échelles nationale et européenne, se traduisent d'ores et déjà par la réalisation de prototypes d'applications proposant de nouvelles heuristiques en matière de navigation par le contenu musical et sonore. Après un exposé des questions d'ordre scientifique soulevées par la mise en œuvre de ce type d'applications, nous présenterons une synthèse des approches et solutions, représentatives de l'état de l'art actuel de la recherche, retenues dans le cadre de trois cas concrets : la gestion d'échantillons sonores, les bases de données d'enregistrements musicaux et la présentation hypermédia d'œuvres musicales.

Description des contenus musicaux

La question de la navigation peut ici se ramener, à un premier stade d'élaboration conceptuelle, à celle de la description des contenus musicaux et sonores, relevant elle-même de la problématique plus générale de l'analyse. Cette approche, promue dans le cadre du processus de normalisation MPEG-7, vise la manipulation des contenus audiovisuels à travers l'utilisation de données intermédiaires, dites *métadonnées*, constituant des représentations adaptées à l'application visée. On peut distinguer dans ce cadre les descriptions de haut niveau, relatives aux structures de connaissance constitutives de l'application, de descriptions de bas niveau, extraites automatiquement à partir des signaux audionumériques et éventuellement cachées à l'utilisateur. L'intérêt de l'extraction automatique est de limiter tant que faire se peut la saisie d'information manuelle. La méthodologie nécessaire à la réalisation de telles applications combine donc idéalement :

- une approche *top-down*, visant l'élaboration de structures de description de haut niveau des contenus musicaux qui rendent compte à la fois des fonctionnalités cibles et des processus cognitifs relatifs à l'écoute musicale
- une approche *bottom-up*, concernant l'utilisation de techniques de traitement de signal pour l'analyse automatique et le calcul des descripteurs de bas niveau
- la *mise en correspondance* entre les deux types de représentation à travers l'utilisation de méthodes d'apprentissage automatique, devant permettre in fine l'inférence automatique des structures de haut niveau.

Le premier problème à résoudre, qui se heurte à la limitation de nos connaissances en matière de cognition musicale, porte donc sur la conception de structures de description de haut niveau. L'expression musicale repose certes sur des théories bien établies, qui en définissent les éléments constitutifs de base (hauteurs, rythmes, nuances, etc.) et, dans certaines limites, les règles de composition (harmonie et contrepoint). Les recherches et développements, effectués dans des institutions comme l'Ircam en rapport avec la création contemporaine, aboutissent à la formalisation informatique de ces structures musicales dans le cadre d'environnements de

composition assistée par ordinateur, sous forme de représentations symboliques, orientées-objet, décrivant les informations contenues dans la partition de manière plus complète que ne le permet le standard MIDI (Lerdahl *et al.*, 1985), (Assayag *et al.*, 1999). Ces représentations ne peuvent pour autant être exploitées directement dans les applications de navigation, qui nécessitent une approche de plus haut niveau. En effet, ce n'est pas au niveau du matériau à la base de l'écriture musicale (la note) qu'une structuration de l'espace musical peut être constituée, mais à un niveau structurel plus élevé, tant dans la dimension temporelle (le *motif* rythmique ou mélodique, les *mouvements* de la pièce) que celle de la superposition (harmonie, timbre). De plus, une telle description à partir de la notation ne prend pas en compte les caractéristiques acoustiques de la musique, alors même que le « son » d'un groupe de musique, d'un orchestre symphonique ou d'un ensemble baroque, notion qu'il conviendrait de préciser, ne correspond pas moins à une référence cognitive relativement partagée dans l'écoute musicale (Delalande, 2001). La réalisation d'applications de navigation nécessite donc la constitution de descriptions ad-hoc, qui consisteront nécessairement en une synthèse, une réduction de l'information musicale de départ, et seront conçues en adéquation avec un contexte applicatif donné. Il convient donc d'explicitier la nature des fonctions associées à de telles interfaces.

Fonction et modes des interfaces de navigation musicale

Dans le contexte des applications informatiques pour la création musicale, la prise en compte des caractéristiques relatives aux interfaces homme-ordinateur nous a conduit à définir une typologie distinguant fonctions opératoires, conceptuelles et heuristiques (Vinet, 1999). Les fonctions de type *opératoire* peuvent se résumer à la production ou au traitement d'informations. La fonction *conceptuelle* concerne l'utilisation de l'ordinateur en tant qu'outil de représentation, de prothèse à la pensée en quelque sorte. La fonction *heuristique* considère l'ordinateur comme outil de créativité, susceptible de produire des résultats inattendus mais pertinents, éventuellement sur la base d'une spécification incomplète de l'utilisateur. Les interfaces de navigation relèvent manifestement des deux dernières catégories, dans la mesure où elles mettent en œuvre des espaces de représentation et permettent de les parcourir pour en découvrir les informations. Dans ce cadre, une question centrale concerne l'adéquation entre l'espace de représentation proposé et l'organisation des connaissances propres à l'utilisateur, ou, éventuellement, la possibilité donnée à celui-ci d'infléchir l'espace de représentation en fonction de sa propre conception. L'interface de navigation revêt ici une double fonction heuristique : la recherche de sons, mais aussi, implicitement, la découverte, à travers ses modalités interactives, des structures de représentation du sonore qu'elle met en œuvre.

En ce qui concerne les applications de navigation dans les bases de données musicales, deux modes de navigation complémentaires peuvent être envisagés d'emblée : la navigation *entre sons différents*, et la navigation *à l'intérieur d'un son*, dans son déploiement au cours du temps. Ces deux aspects de la navigation rendent compte de fonctions d'applications différentes. Le premier cas concerne la mise en œuvre de procédures de recherche dans des bases de données sonores ou musicales. Le second cas concerne le développement d'interfaces, tirant parti de l'utilisation conjointe de plusieurs modalités, notamment auditives et visuelles, pour offrir une présentation élaborée d'un morceau de musique associant de manière coordonnée différents types de représentations et documents autour de l'œuvre.

Les descriptions de haut niveau à concevoir dans les deux cas sont différentes. Le premier amène la constitution d'espaces dans lesquels chaque son est un point, ou tout du moins une entité close. Les descriptions associées aux sons se réfèrent à la globalité de leur déroulement temporel, afin de dégager des critères pertinents de mise en relation entre sons et donc de structuration de l'espace de navigation : ségrégation, similarité, ordre, appartenance à des catégories communes ou distinctes, etc. En particulier, lorsque tous les descripteurs sont exprimés sous forme numérique, un son est caractérisé par un point dans un espace de dimension N et le recours à des techniques d'analyse en composantes principales permet de réduire celui-ci à un espace de dimension 2 ou 3, utilisable dans une interface graphique de navigation. Le second cas induit l'extraction de structures internes et leur mise en relation, cette fois à l'intérieur du son lui-même, de sorte que l'espace de navigation résulte, selon une projection synthétique, de l'espace musical propre au morceau. Cela est particulièrement vrai du déploiement temporel de celui-ci, généralement figuré selon des représentations graphiques que nous avons qualifiées de *représentations temporelles*, dans l'établissement d'une typologie des interfaces homme-machine propres aux logiciels musicaux (Vinet, 1999a), consistant en une *spatialisation de leur dimension temporelle*.

En pratique, les deux cas ne sont pas disjoints, des situations hybrides pouvant être rencontrées, par exemple en organisant la navigation dans un corpus de morceaux en fonction de structures motiviques qu'ils contiennent.

Il s'ensuit de ces différents aspects que, dans le cas général, la mise en place d'un espace de navigation musicale résulte d'un processus d'*analyse/synthèse*, selon lequel les caractéristiques descriptives pertinentes sont extraites du matériau de départ, puis synthétisées en vue de la constitution d'un espace « navigable ». Nous développons dans la suite la manière dont ces considérations d'ordre général s'instancient dans le cadre des trois types d'applications visées.

Gestion d'échantillons sonores

L'usage d'échantillons, sons élémentaires contenant des événements simples (tels qu'une note isolée jouée sur un instrument, un cri d'animal, ou un bruit d'automobile) s'est généralisé dans différents contextes de production musicale et sonore en studio. Leur mode de diffusion était jusqu'à présent limité aux cédéroms, interdisant à l'utilisateur un accès global à l'ensemble de ses sons, ou l'obligeant à en assurer lui-même la gestion sur son disque dur d'ordinateur. Les technologies des réseaux et des bases de données rendent possible la réalisation d'applications de gestion globale d'échantillons, en particulier sous la forme de serveurs en ligne, permettant l'accès et le partage de sons entre différents utilisateurs. Leur réalisation nécessite la mise en œuvre d'un système de classement permettant de retrouver facilement les sons déjà présents, mais définissant aussi explicitement la manière de compléter la base existante en rajoutant de nouveaux sons. Les besoins relatifs à de telles applications ne sauraient toutefois se limiter à cette fonction de classement², mais relèvent également d'une fonction de type heuristique. En effet, dans un contexte de production ou de création sonore, il ne s'agit pas seulement d'accéder efficacement aux sons connus, mais aussi de découvrir des sons inattendus à partir d'une spécification non explicite ou incomplète, éventuellement métaphorique. D'un point de vue scientifique, l'intérêt de travailler sur les échantillons en vue de l'obtention de structures de descriptions adaptées tient d'abord à leur relative simplicité : il est plus facile d'extraire automatiquement, par l'utilisation de techniques de traitement de signal, les principales caractéristiques d'un son isolé que d'un morceau de musique enregistré. Il n'en reste pas moins à définir les structures de descriptions de haut niveau pertinentes du point de vue de l'application.

Le premier mode de description des sons qui s'impose s'attache aux conditions de leur production, notamment en rapport avec une classification des sources sonores. Cette approche est d'autant plus pertinente d'un point de vue cognitif qu'elle rend compte d'une des fonctions probablement les plus anciennes de la perception auditive à des fins de survie : la reconnaissance de sources sonores dans l'environnement, en particulier lorsque celles-ci sont situées hors du champ de vision. L'utilisation à cet effet d'une taxonomie arborescente, organisant les différents types de sources selon un niveau de spécification croissant, est aisément concevable et est, de fait, largement adoptée dans les applications existantes, d'autant qu'elle est compatible avec les systèmes de gestion de fichiers existants et ne nécessite pas l'utilisation de bases de données. Celles-ci deviennent utiles si l'on souhaite s'affranchir des structures arborescentes et autoriser un parcours, une topologie de navigation plus ouverte. Cette approche a été développée

² Cette fonction correspondant au type conceptuel défini plus haut, le système de classement fournissant le principe d'organisation de l'espace, la topologie de navigation entre sons.

dans le projet Studio en ligne (Ballet *et al.*, 1999), proposant une base de données de 117 000 échantillons instrumentaux, décrits à partir de leurs modes de production. L'interface de sélection de Studio en ligne permet de spécifier un son en commençant par l'une quelconque des rubriques disponibles. Par exemple, le choix initial du mode de jeu *pizzicato* restreint automatiquement les possibilités ultérieures de sélection d'instruments à ceux compatibles avec lui. Plus généralement, les systèmes de classement des sources sous forme arborescente présentent des limitations drastiques à l'utilisation, en particulier si l'on privilégie une approche heuristique. Ainsi, on s'attendrait à ce qu'un son d'avertisseur, qui est certes associé à la catégorie voiture ou véhicule, le soit aussi directement à celle des sons urbains, etc. Les recherches effectuées dans le cadre du projet Ecrins (*Environnement de Classification et de Recherche Intelligente de Sons*), (Geslin *et al.*, 2002), proposent une solution à ce problème en autorisant l'établissement de liens symboliques entre tout son et toute catégorie, la sémantique étant spécifiée manuellement, ou produite à partir d'un processus d'apprentissage automatique.

Un autre mode de description des sons, également développé dans Ecrins, s'attache cette fois à une approche morphologique du phénomène sonore, dans la lignée des travaux de Pierre Schaeffer (Schaeffer, 1966). Celui-ci développe en particulier le concept d'*écoute réduite*, en référence directe à la réduction phénoménologie de Husserl, selon lequel un son peut être écouté en tant que tel, dans son propre déploiement, en faisant abstraction de ses causes de production. Dans le cas des échantillons sonores, ceux-ci peuvent ainsi être caractérisés selon leurs profils dynamique (intensité), de hauteur, leur grain, etc. L'intérêt d'une telle description est qu'elle rend compte de caractéristiques objectivables du son, pouvant être automatiquement extraites par analyse du signal.

Ces deux types d'approches, parmi d'autres, sont représentatives de l'opportunité de décrire les sons selon des points de vue complémentaires (causal, morphologique, etc.), correspondant chacun à une logique, un mode de classification autonome. La navigation peut ainsi s'effectuer à l'intérieur d'un mode donné, mais aussi en combinant les caractéristiques de plusieurs points de vue (recherche des sons de violon ayant tel profil de hauteur par exemple). De plus, chaque mode de classification, qui tend à réaliser une partition, au sens mathématique, de l'ensemble des sons, peut, du moins théoriquement, faire l'objet de fonctions d'apprentissage, de sorte que tout nouveau son introduit dans le système soit, par l'utilisation de techniques de classification, automatiquement rangé dans une classe unique pour chaque mode. Dans Ecrins, la possibilité est en outre donnée à l'utilisateur de créer ses propres modes de classification et de lancer les procédures d'apprentissage correspondantes. En pratique, les méthodes actuelles de classification automatique des sons, qui présentent de bonnes performances dans des contextes réduits, sont sensibles à la définition des corpus d'apprentissage, et doivent être davantage considérées comme des aides à la

classification manuelle que comme des fonctions entièrement automatisables.

Une autre heuristique de navigation, complémentaire aux précédentes et expérimentée dans le cadre des projets Studio en ligne, Cuidado (*Content based User Interfaces for Digital Audio Databases Online*) (Vinet et al., 2002) et du produit SoundFisher³, concerne la *recherche par l'exemple* : les sons sont recherchés par similarité avec un son de référence donné par l'utilisateur. Celui-ci peut être issu d'un fichier d'échantillon sélectionné par l'utilisateur, mais aussi produit par lui à l'aide d'un microphone : cette modalité d'entrée, complémentaire à une description textuelle, est particulièrement pertinente lorsqu'elle est soumise à des professionnels exercés, dont le perfectionnement des capacités auditives s'est accompagné manifestement de compétences motrices de production sonore. La mise en œuvre de critères de similarité perceptive utilisables à cet effet tire parti de travaux effectués en psychologie expérimentale sur la perception du timbre musical, aboutissant à la constitution d'*espaces perceptifs de timbres*, dans lesquels sont organisés les sons. Chaque axe de cet espace correspond à un critère perceptif indépendant, corrélé à certaines caractéristiques acoustiques du son (brillance, attaque, etc.). La similarité entre deux sons s'y mesure par le calcul d'une distance euclidienne sur les différentes dimensions perceptives, avec la possibilité pour l'utilisateur de ne retenir que certains critères de ressemblance par rapport au son d'exemple.

Bases de données d'enregistrements musicaux

Ce type d'application concerne également la navigation entre sons enregistrés, mais relatifs cette fois à des morceaux de musique et non des sons isolés. L'avènement de l'Internet et de techniques de compression sonore, conçues dans le cadre de la norme MPEG, rend possible la constitution de nouveaux circuits de diffusion de la musique enregistrée, notamment à partir de serveurs en ligne. Même si la viabilité économique et juridique de tels services reste encore à démontrer, ceux-ci présentent un certain nombre d'avantages à la fois du point de vue des consommateurs et des « majors », détentrices des droits des œuvres : accès rapide aux contenus, information personnalisée, valorisation à long terme de fonds dont la période de commercialisation est révolue (Top50), promotion d'artistes inconnus en fonction des goûts des auditeurs, etc.

Nous ne traiterons pas ici des techniques et formats documentaires pour la gestion de bases de données musicales, mais concentrerons notre propos sur les problématiques de description des œuvres proposant des heuristiques originales de navigation. Le critère de classement le plus évident, utilisé par la plupart des sites existants⁴ repose sur des taxonomies

³ <http://www.soundfisher.com/>

⁴ Voir par exemple <http://amazon.com> et <http://moodlogic.com>

des genres musicaux. La constitution de description de haut niveau pertinentes se heurte cependant à la difficulté d'en automatiser l'extraction. Une approche largement répandue consiste à isoler des critères de similarité musicale d'ordre culturel à partir de mesures de co-occurrence : des morceaux ou genres musicaux seront d'autant plus proches culturellement qu'on en trouvera de nombreuses références juxtaposées, en particulier sur le Web. Une autre approche consiste cette fois à extraire ces critères non de statistiques d'usage, mais directement partir du contenu musical, avec la difficulté de mise en œuvre de l'analyse de signaux audio numériques aussi complexes que des enregistrements polyphoniques. Les tentatives effectuées jusqu'à présent d'extraction des rythmes à partir des signaux et de corrélation avec les genres musicaux n'ont pas à notre connaissance donné de résultats probants. L'approche développée par différentes équipes de recherche, en particulier dans le cadre du projet Cuidado, porte sur des descripteurs plus simples de tempo, d'intensité subjective ou de surface musicale, qui rendent bien compte de caractéristiques pertinentes d'un point de vue cognitif et utilisables à des fins de classification.

Une autre voie de description des contenus musicaux part non pas des signaux audio, mais de représentations symboliques proches de la partition, sous forme de fichiers MIDI, disponibles en particulier pour une partie importante du corpus classique, pour en extraire des structures représentatives des formes musicales : rythmes, profils mélodiques, harmonie, etc. Une heuristique assez répandue de recherche par l'exemple prend ici la forme de *recherche par chantonnement*, où le système, à partir de l'analyse d'un signal audio produit par l'utilisateur à l'aide d'un micro, recherche les mélodies proches, selon une métrique de similarité suffisamment permissive pour autoriser des interprétations approximatives. La problématique générale de sélection de structures musicales de haut niveau adaptées, et de critères de mise en relation de morceaux de musique sur cette base, reste cependant encore largement inexplorée.

Une fois ces critères de description de haut niveau constitués, se pose le problème de l'interface de navigation proprement dite. La plupart des systèmes existants se limitent à des fonctions de recherche par multi-sélection à partir de listes de termes. Quelques interfaces proposent des modes de navigations graphiques entre genres, par exemple sur la base du concept d'arbres hyperboliques développé à Xerox PARC.⁵ Une heuristique originale concerne la génération automatique de séquences musicales, en réponse à une requête de l'utilisateur (Pachet *et al*, 1999). En effet, les technologies de distribution électronique de la musique s'affranchissent de l'album, support traditionnel de diffusion, pour se concentrer sur le titre, le morceau de musique, qui devient potentiellement l'item à la base de l'échange commercial. Du point de vue du consommateur mélomane, rares sont cependant les situations dans lesquelles un morceau est écouté

⁵ <http://www.eicvest.no/cuidado/styles.html>

isolément : l'intérêt de programmes de radio ou de concert bien conçus réside aussi dans l'impression d'ensemble donnée par la succession des œuvres. Ce constat est mis à profit, en particulier dans le cadre de Cuidado, pour proposer la génération automatique de séquences musicales, spécifiées à partir d'un jeu de contraintes pouvant porter sur des caractéristiques statistiques de l'ensemble de la séquence, ainsi que sur son évolution au cours du temps. Cette heuristique constitue une forme de navigation originale, dans laquelle l'application fournit un cheminement, une *trajectoire* dans l'espace de navigation spécifiée par ses caractéristiques globales.

Présentation hypermédia d'œuvres musicales

Le type d'application envisagé ici concerne l'utilisation de technologies hypermédia à des fins de navigation à l'intérieur d'une œuvre musicale. Les dispositifs actuels de médiation de la musique, encore universellement en usage dans les chaînes Hifi, sont limités à des interfaces élémentaires sous la forme de panneaux de contrôle Lecture/Arrêt/Pause, la commande de volume, etc. Cette limitation en terme d'interactivité tient d'abord à celle du codage des informations musicales, restreint à leur enregistrement sous forme audionumérique. L'utilisation de descriptions plus élaborées du contenu musical, faisant apparaître les structures internes de l'œuvre, permet d'envisager des formes de présentation plus poussées, tirant parti des technologies hypermédia existantes. La réalisation de telles interfaces, susceptibles de supplanter à terme les formes actuelles de médiation des œuvres enregistrées, s'inscrit également dans un objectif de diffusion culturelle, en particulier pour les musiques dites savantes, visant à favoriser la compréhension des œuvres et à offrir des supports, en particulier visuels, à l'écoute.

Des recherches récentes, réalisées en particulier dans le cadre de Cuidado, aboutissent à la production automatique d'un résumé de l'œuvre, par des procédés de segmentation, de comparaison et de visualisation des segments ainsi obtenus. Il est ainsi possible de mettre en évidence la structure globale du morceau et produire des icônes graphiques correspondant aux différents segments, dont la ressemblance visuelle rend compte de la similarité de leurs contenus sonores. Certaines réalisations multimédia recourent à l'utilisation de sonagrammes, représentations graphiques du spectre du signal sonore au cours du temps, comme support visuel à l'écoute⁶. Au-delà de ces procédés automatisés, la constitution de descriptions adéquates des structures internes de l'œuvre relève plus généralement d'un travail d'analyse musicale, dans lequel l'intervention humaine est essentielle au choix à la fois des stratégies et points de vue d'analyse, mais aussi des formes, notamment graphiques, adéquates à leur médiation. Le logiciel Acousmographe a été conçu dans cet esprit pour la

⁶ <http://www.38rugissants.com/Site 2000/>

production de représentations graphiques synchrones au déroulement sonore, à partir d'un repérage temps-fréquence sous forme de sonagrammes (Koechlin *et al.*, 1991). Il a été abondamment utilisé pour l'analyse d'œuvres contemporaines (Besson *et al.*, 1995). Une approche systématique de la production hypermédia autour d'œuvres musicales, telle qu'elle est en particulier envisagée à l'Ircam, passe la constitution d'outils d'analyse, assistant le musicologue pour la mise en évidence de régularités dans l'œuvre, et d'outils auteur dotés de fonctions spécifiques, telles que des possibilités de synchronisation et d'hyperliens entre différents types de documents (enregistrements sonores, partitions et autres représentations graphiques, annotation textuelles, etc.) organisés en base de données.

Conclusion

Cette synthèse que nous avons proposée des approches existantes en matière d'interfaces de navigation dans des corpus musicaux rend compte de recherches récentes ou en cours, issues pour l'essentiel de laboratoires spécialisés, qui ne constituent que les prémices d'un domaine voué à un développement important. Notre propos s'est concentré sur la structuration de l'espace de navigation à partir de l'obtention de descriptions adaptées du contenu musical, principal verrou scientifique identifié, dont la résolution appelle un questionnement interdisciplinaire en matière de cognition musicale. Même si ces travaux sont représentatifs de l'état de l'art scientifique et technologique, ils ne font qu'illustrer des potentialités dont l'actualisation effective, s'inscrivant dans le déploiement d'un système technique généralisé de diffusion de la musique, passe par la mise à l'épreuve de conditions de faisabilité d'ordre juridique, économique et, finalement, d'usage.

Bibliographie

- Assayag G., « Computer-Assisted Composition at IRCAM : From PatchWork to OpenMusic », *Computer Music Journal*, Vol 23, N°3, MIT Press, 1999 , p. 59-72.
- Ballet G., Borghesi R., Hoffmann P., Lévy F., « Studio Online 3.0: An Internet "Killer Application" for Remote Access to IRCAM Sounds and Processing tools », *Actes des Journées d'Informatique Musicale (JIM'99)*, 1999
- Disponible en ligne à :
http://www.ai.univ-paris8.fr/~jim99/actes_html/BalletJIM99.htm
- Besson D., *Les musicographies*, cédérom, Institut national de l'audiovisuel, 1995.
- Delalande F., *Le son des musiques, entre technologie et esthétique*, Paris, Editions INA-Buchet/Chastel, 2001.
- Geslin Y, Mullon M., Jacob M., « Ecrins, an audio content description environment for sound samples », *Actes du colloque International Computer Music Conference (ICMC02)*, International Computer Music Association, San Francisco, 2002.
- Koechlin O, Vinet H., « The Acousmographe, a Macintosh software for the graphical representation of sounds », *actes du colloque International Computer Music Conference (ICMC91)*, International Computer Music Association, San Francisco, 1991.

Lerdahl F., Jackendoff R., *A Generative theory of tonal music*, Editions MIT Press, 1985.

Pachet F., Roy P., Cazaly, D. « A Combinatorial approach to content-based music selection », *Actes du colloque IEEE International Conference on Multimedia Computing and Systems*, Florence, Italie, Vol. 1, 1999, p.457-462.

Schaeffer P., *Traité des objets musicaux, essai inter disciplines*, Seuil, 1966

Stiegler B., « La numérisation du son », *Culture et recherche*, N°91-92, Ministère de la culture et de la communication, 2002, p.3-6.

Disponible en ligne à :

http://www.culture.gouv.fr/culture/editions/r-cr/cr91_92.pdf

Vinet H., «Introduction », In *Interfaces hommes-machine et création musicale*, éditeurs Hugues Vinet et François Delalande, Editions Hermes Science, 1999, p.9-14.

Vinet H., «Concepts d'interfaces graphiques pour la production musicale et sonore», In *Interfaces hommes-machine et création musicale*, éditeurs Hugues Vinet et François Delalande, Editions Hermes Science, 1999, p.97-121.

Vinet H., Herrera P., Pachet F., « The CUIDADO project », *Actes du colloque International Conference on Music Information Retrieval (ISMIR02)*, Ircam, Paris, 2002.