

Images et sons dans les hypermédias

De la correspondance à la fusion

Hervé Zénouda (Paris le 08/09/2005)

H2PTM05, 'Créer, jouer, échanger : expériences de réseaux', Novembre 2005

Résumé :

Nous chercherons à souligner dans cet article les continuités et les ruptures qu'apporte le numérique dans le rapport qui se crée entre les images et les sons. Depuis l'influence mutuelle des arts graphiques et des arts sonores en peinture jusqu'aux hypermédias aujourd'hui, on peut repérer une dynamique de convergence toujours plus affirmée des deux médias.

Si le numérique, avec son codage discret, est un moment fort de cette trajectoire, il engendre différents modes de communication spécifiques : interactivité, générativité, simulation, qui chacun, à sa manière, modifie les rapports entre les deux modalités, comme autant d'étapes vers une fusion des conditions de production.

L'évolution de cette relation concerne autant le concepteur (qui manipule une 'matière audiovisuelle' toujours plus riche) que l'utilisateur (apparition de nouvelles applications, évolution des fonctions de recherche documentaire ...) et modifie la nature même du média : plus vraiment une image et plus vraiment un son, cette fusion engendre l'apparition d'un nouvel objet audiovisuel composite, structuré et complexe.

Mots-clefs :

Sons et images numériques, interactivité, générativité, simulation, convergence, synesthésie, mapping

Abstract :

The goal of this article is to underline the continuities and breaks, the digital technology brought to the relationship between pictures and sounds. From the mutual influence of graphic and sound arts on painting, to the hypermedia today, we can easily spot an always-stronger convergence dynamic between these two media. Digital technology, with its discrete common encoding, is a strong point of this evolution and generates different specific ways of communicating: interactivity, generativity, and simulation. Each of them, in its own way, changes the relationships between the two representations, as different steps towards a fusion of the production conditions.

This relationship's evolution concerns as much the designer (who manipulates an always-richer "audiovisual material") as the user (new applications appear, documentary search functions evolve...) and modifies the nature of the media itself : not fully a picture nor a sound anymore, this fusion creates a new composite audiovisual object, structured and complex.

Key words :

Digital sounds and pictures, interactivity, generativity, simulation, convergence, synesthesia, mapping.

"En affichant une volonté conceptuelle et expérimentale sans autre borne que celle du calcul et de la description logique, le numérique introduit donc en cascade, dans la pensée et la pratique de ce qu'on appelle "forme" ou "signe", des remaniements essentiels ; ainsi, grâce au traitement informationnel, nous continuerons de voir de l'image, à entendre du son, à toucher de la matière ; mais ce stade morphologique de l'information n'existe que produit "en bout de chaîne" ; la reconnaissance formelle, sensible, n'aura donc lieu qu'à titre d'événement terminal, présupposant, en amont de la forme ou de l'information sensorielle, le processus complexe d'une morphogenèse dont la "matière" n'est autre que celle du nombre (l'élément 0/1 et ses combinaisons) qui seul a présidé à sa possible existence."

Alain Renaud Alain ("L'image sans gravité, la forme image aux risques de l'information" in "Revue d'esthétique", 1994)

Cherchant à resituer dans la continuité des médias précédant la relation image/son, telle qu'on peut l'analyser aujourd'hui dans les hypermédias, nous regarderons de manière synthétique l'influence de la musique sur la peinture ainsi que la cohabitation du sonore et du visuel sur un même support-film avec le cinéma.

Nous essaierons ainsi de montrer ce mouvement de convergence qui va d'un rapport de correspondance à une fusion des deux médias.

Si le numérique, avec son codage discret, est une étape particulièrement marquante dans cette tendance à la fusion (création d'un "substrat" numérique commun à toutes les modalités), c'est le programme, cette "interface langage", qui permet de passer du niveau "physique" au niveau "cognitif" du numérique avec des niveaux de représentations de plus en plus abstraits.

Les différents modes de communication que permettent les hypermédias : interactivité, générativité et simulation, qui chacun, à leur manière, modifient les rapports entre les deux modalités, sont autant d'étapes vers une fusion des conditions de production : la constitution d'un langage commun en amont des médias, pour mettre en place une situation d'émergence uni média.

Un média qui n'est plus vraiment une image ni plus vraiment un son, mais un nouvel objet audiovisuel programmable.

I/ Quelques repères historiques :

Nous allons regarder succinctement différentes étapes de cette dynamique de convergence avec :

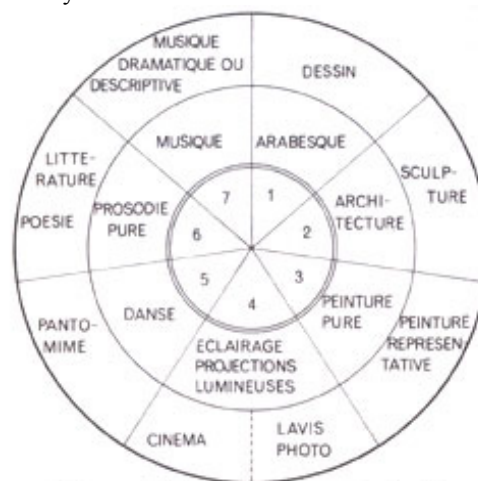
- . L'esthétique comparée d'Etienne Souriau, première approche synthétique des différents arts.
- . Les correspondances entre arts graphiques et arts sonores en peinture.
- . Le cinéma narratif comme premier support de synchronisation image/son et le cinéma expérimental comme recherche d'une convergence synesthésique entre les deux modalités.

1/ Etienne Souriau et l'Esthétique comparée :

En 1947, Etienne Souriau publie son livre "La correspondance des arts" ([SOU 69]) où il développe l'idée que toute œuvre s'exprime sous quatre principaux niveaux d'existence :

- . Une *existence physique* représentée par la matérialité de l'œuvre.
- . Une *existence phénoménale* où s'exprime son caractère sensible. Les matériaux sensibles sont constitués non pas de sensations mais "d'essences pures, de qualités génériques et absolues" qu'il nomme "qualias sensibles". Il énumère sept registres de qualias (Lignes, Volumes, Couleurs, Luminosités, Mouvements, Sons articulés, Sons musicaux).
- . Une *existence "réique" ou chosale* où s'exprime la capacité de l'œuvre à "représenter à l'esprit des choses". Il différencie à ce niveau deux approches : les arts *représentatifs* ou de *second degré* (peinture, sculpture, ...) et les arts non représentatifs (musique ...) appelés *présentatifs* ou de *premier degré*.
- . Une existence transcendante, qui appelle à un dépassement de l'œuvre "vers le haut" à l'aide du symbolisme.

Fort de ces sept registres de qualia et de la distinction entre arts du premier et du second degré, Souriau nous propose un schéma synthétique de son système des beaux-arts :



1, Lignes; 2, Volumes; 3, Couleurs; 4, Luminosités; 5, Mouvements; 6, Sons articulés; 7, Sons musicaux.

1/ Lignes, 2/Volumes, 3/Couleurs, 4/ Luminosités, 5/ Mouvements, 6/ Sons articulés, 7/ Sons musicaux

Le premier cercle intérieur comprend les arts du premier degré (*présentatifs*), le cercle extérieur ceux du second degré (*représentatifs*).

Malgré le caractère arbitraire ou partiel de cette classification (et les difficultés d'intégrer les arts de synthèse : le cinéma serait ici uniquement muet !), cette approche synthétique de l'ensemble des arts permet de mettre en évidence des correspondances intéressantes en reliant les arts sur l'axe des qualia ou sur celui des degrés.

Le schéma ne prend néanmoins en compte que les plans du phénoménal et du réique, laissant de côté les correspondances pouvant se faire sur les deux plans manquants : celui de la matérialité et celui de la transcendance.

2/ Les rapports musique/peinture : Jacques Parrat et Jean Yves Bosseur

L'histoire des relations et des influences mutuelles entre les arts graphiques et les arts sonores est bien trop longue et riche pour tenter ici de circonscrire la question.

Nous pouvons toutefois y voir les bases de la correspondance entre le sonore et le visuel.

Jacques Parrat ([PAR N.C.]) propose trois niveaux de correspondances entre les deux arts :

a) Les fondements rationnels :

À partir de trois catégories de structures initiales : les *structures algébriques* (applications d'opérations directes et réversibles), les *structures d'ordre* (relation de ressemblance et réciprocity) et les *structures topologiques* (relations de voisinage, limite, continuité) ; les différents langages vont se développer suivant leurs propres fonctions. Les langages esthétiques seront sous la dépendance des phénomènes perceptifs et des sens convoqués (musique : oreille et temps, peinture : œil et espace).

Le langage musical fondé sur des opérations logico-mathématiques (augmentation, diminution, renversement, rétrogradation ...) appliquées à des éléments simples (notes, accords, intervalles, durées) montre la relation directe entre la musique et les nombres, la vocation naturelle de la musique vers l'abstraction.

Le langage pictural, pendant des siècles, va se développer dans une toute autre direction, celle de la représentation (réalisme, connaissance du monde extérieur, fonctions de communication...).

b) Les fondements sensoriels :

L'œuvre d'art, avant d'être perçue par l'esprit comme une construction, est d'abord saisie par les sens. À ce niveau de perception, les correspondances se tissent sur deux plans distincts : le cadre perceptif (le temps et l'espace) et l'objet matériel (la forme, le matériau).

Le temps et l'espace :

Si l'acceptation de la musique comme art du temps et la peinture comme art de l'espace ne peut être réfutée, chacune des pratiques va s'intéresser à représenter la dimension qui lui manque :

En musique, l'espace est présent à différents niveaux :

- . par l'utilisation de l'écriture qui est une représentation spatiale de la musique qui permet l'élaboration de constructions et de formes,
- . avec la dimension verticale (harmonique) de la musique qui associe simultanément des "hauteurs" de notes distinctes et crée ainsi un espace,
- . l'espace "interne" de la musique créée par l'orchestration et la gestion des différents plans sonores des timbres et des dynamiques,
- . l'espace "externe" de la musique représenté par l'espace acoustique du lieu de concert ou de diffusion de la musique (voir les techniques de multi-diffusion de la musique acousmatique).

En peinture, la temporalité est surtout liée au regard, au temps de "lecture" de l'œuvre. Si l'œuvre picturale s'offre de façon simultanée, l'analyse, le jugement, la pensée intègre la notion de temps. Le jeu sur les nuances de couleurs ainsi que sur les masquages permet au peintre d'offrir des niveaux de lecture avec des éléments directement repérables et d'autres nécessitant un "temps" de lecture.

D'autre part, avec des écoles comme le cubisme ou le futurisme, le temps est directement représenté dans la peinture en faisant cohabiter différents points de vue représentant le volume (cubisme) ou le mouvement (futurisme).

La forme, le matériau :

- . *Convergence phénoménologique :*
 - . Intensité du son et "valeurs" plastiques
 - . Timbres et Couleurs
 - . Densités et textures sonores et picturales

. *Convergence kinesthésique* :

Importance du "geste", du "trait" dans la perception de la peinture comme de la musique. Energie du corps laissant une trace dans l'œuvre.

"Au-delà de la perception du visible et du sonore, il existe, semble-t-il, une perception globale, originelle, indifférenciée, de nature musculaire et kinesthésique, ressentie charnellement par le corps tout entier ... Peinture et musique sont ainsi originellement et organiquement liées au corps, au geste."¹

c) Les fondements irrationnels

Analysant les capacités de la musique à provoquer des émotions, Jacques Parrat propose trois types de signes expressifs sur lesquels se fondent les pouvoirs d'expression musicale :

- . Les signes expressifs d'origine organique (1)
- . Les signes expressifs d'origine structurelle (2)
- . Les signes expressifs par association (3)

Qu'on peut appliquer à la musique :

- . la manière dont l'élément brut, charnel (cri, gémissement, sanglot ..) va être représenté et exprimé par le langage musical (1),
- . leur potentiel expressif ne vient plus des signes organiques préexistant à la musique mais à la structure même du langage utilisé (langage tonal pour la période classique) (2),
- . musique descriptive ou imitative et plus récemment, l'utilisation de sons concrets (3).

Ainsi qu'à la peinture :

- . expression des sentiments humains dans la peinture (1),
- . expressivité liée à la composition des couleurs et de la géométrie (symétrie proportions ...) (2),
- . fonctions de représentation de la peinture (3).

d) Les correspondances structurelles :

Jean Yves Bosseur ([BOS 98]) propose un niveau de correspondance supplémentaire, celui du niveau structurel : La liberté prise par la peinture du XX ième siècle par rapport à sa fonction de représentation, va pousser de nombreux peintres à chercher dans la musique (art abstrait par définition) des modèles de structuration (cf : la distinction que fait Souriau entre les arts du premier et du second degré). Les techniques comme le canon, l'imitation, l'augmentation, la diminution, le renversement appliqués aux transformations de motifs plastiques permettent à Klee, par exemple, de maîtriser la forme générale et la représentation du temps dans ses tableaux. Avec la notion de "peinture polyphonique", Klee définira l'existence de plusieurs thèmes graphiques qui se développent simultanément (pour plus de détails sur les influences de la musique sur le travail de Klee ainsi que l'impact de sa peinture sur des compositeurs comme Pierre Boulez, voir la thèse de Ekaterina Kouprovskaja-Bruggeman [KOU 03]).

3/ Cinéma narratif et cinéma expérimental

On retrouve l'équivalent de l'opposition figuratif- abstrait au cinéma avec l'opposition narratif-expérimental. De la même manière que pour la peinture, quand le cinéma cherche à s'émanciper du narratif, il va chercher du côté de la musique d'autres modèles que le roman ou le théâtre pour se structurer.

Premier support de *synchronisation* de l'image et du son, le cinéma pose des questions sur la convergence des médias (l'émergence d'un discours polyphonique multi sensoriel), sur les enjeux de l'inter sensorialité (les notions de "valeur ajoutée" et "d'indices matérialisants" développées par Michel Chion [CHI 03]), sur l'élaboration d'un vocabulaire de la relation de l'image animée et du son (fonction du son dans les raccords de plans, jeux sur les champs : in/hors/off, notion de point de vue et point d'écoute ...).

D'autre part, l'arrivée du parlant a eu comme effet de stabiliser la vitesse d'enregistrement et de lecture des images "permettant à cet art cinématographique (fixant le mouvement), de devenir également un art "chronographique" (fixant le temps)" (Pratiques n°14 Muriel Caron [PRA 03])

Deke Dusinberre ([BEA 86]) souligne l'importance de la fixation du temps dans le rapprochement de la musique et du cinéma "De surcroît, la musique comme le cinéma, se déroulent dans le temps et c'est le principe du rythme, d'une durée rythmée, qui est la base de son esthétique, et qui représente l'affinité la plus marquante entre ces deux arts temporels."

¹ Jacques Parrat, Des relations entre la peinture et la musique dans l'art contemporain, Grenoble, Editions Z'édicions, date n.c. p.72

On peut noter que cette référence à la musique apparaît très tôt dans l'histoire du cinéma (bien avant le cinéma sonore) avec des réalisateurs comme Louis Delluc ou Germaine Dulac (les "symphonies visuelles" ou le cinéma "pur")².

Le cinéma expérimental va développer ainsi des analogies avec la musique sur trois principaux niveaux (dont les deux premiers retrouvent les niveaux de correspondances repérés dans la relation musique/peinture) :

- . Synesthésie entre le visuel et le sonore avec la "musique audiovisuelle" ou "colorée" (James et John Withney, Mac Laren, Len Lye, Oskar Fischinger ...).
- . Modèle d'organisation de la forme et de la durée (Yann Beauvais, Jean Michel Bouhours ...).

Et un troisième niveau plus spécifiquement lié au support film comme succession d'images fixes :

- . Modèle d'organisation non plus du contenu des images mais de leur rapport entre elles (l'organisation des photogrammes) faisant du cinéma non plus une machine à représenter mais une machine perceptive ("cinéma comme machine à rythme visuel") : Le cinéma graphique ou structurel (Jean Michel Bouhours, Robert Breer, Peter Kubelka ...).

Yann Beauvais souligne l'association note/photogramme de ce type de cinéma :

"Nombre de films travaillèrent ces rythmes photogrammiques selon des systèmes métriques plus ou moins définis, rigides ou non. Cette appréhension d'une rythmique assigne d'une certaine manière une équivalence entre la note et le photogramme, comme on peut le voir dans "Jum Jum" de Werner Nekes, où il utilise des techniques d'associations comme on peut les retrouver en musique comme par exemple : l'inversion, le renversement, l'inversion rétrograde ... Ce détour par l'assimilation du photogramme à la note présuppose une détermination de l'organisation des parties dans l'ensemble, d'où le recours à une partition comme système d'écriture pré-filmique. Cette notation est ce qui lie le plus le cinéma et la musique via le graphisme."³

4/ Un historique rapide de la convergence image et son

Nous avons vu dans ces quelques lignes, l'ébauche d'une longue histoire d'un rapprochement entre le visuel et le sonore qui débute par un jeu d'influences et de correspondances entre deux pratiques séparées (la peinture et la musique). Dans ce cadre, de nombreux systèmes de correspondances ont été tentés pour relier les éléments primaires d'une pratique : couleurs, formes, lignes et courbes, valeurs de couleur pour la peinture, avec ceux d'une autre : hauteurs, intervalles, timbres, intensité pour la musique (cf : Kandinsky, Blanc-Gatti, Scriabine ...). On a vu que la recherche de modèles d'organisation commun a été aussi un fil conducteur important de l'évolution parallèle des deux arts (cf : Klee ...).

Le cinéma, support de synchronisation des deux modalités, permet d'exprimer une réalité multi-sensorielle unique. La fixation du temps étant le point de correspondance le plus fort entre la musique et le cinéma, que l'on doit considérer comme objets temporels.

Le cinéma narratif a assujéti ces différents éléments constitutifs au récit. On a pu ainsi développer un riche langage audiovisuel fait de synchronisations et de contrepoints sensoriels et d'espaces virtuels multiples (in, hors, off).

Le cinéma expérimental, de son côté, a renoué avec les aspirations de la peinture en cherchant à développer un langage synesthésique et sensoriel en dehors du récit : l'abstraction comme correspondance forte avec la musique. Le cinéma expérimental a ainsi poussé loin la convergence des deux sens et a préparé, d'une certaine manière, la notion de multimédia et l'avènement du numérique.

Si dans le cinéma, c'est la gestion du temps et la synchronisation qui instaure l'effet de synchrèse (terme créé par Michel Chion en contractant les termes synchronisation et synthèse), on reste dans une situation où deux flux distincts s'influencent mutuellement.

Avec le numérique, la fusion des deux modalités va s'opérer à plusieurs niveaux : au niveau du codage (codage binaire commun aux différents médias), et au niveau des différentes structures langagières d'organisation de la substance audiovisuelle (méta-data, méta-langages, unités sémiotiques diverses ...) qui s'imbriquent les unes les autres en différentes couches.

² Sur l'influence de la musique sur le cinéma muet et la consécration du parlant comme victoire du modèle narratif, voir le livre de Nouredine Ghali, *L'avant garde cinématographique en France dans les années 20*, Editions Paris expérimental, Paris, 1995 ([GHA 95])

³ Yann Beauvais, "Notations" dans *Musique Film*, Editions La cinémathèque française, Paris, 1986 ([BEA 86])

II/ Sons et images dans les hypermédias :

Le codage commun aux différents médias ainsi que les capacités de calcul et de manipulations interactives du numérique, accentue de manière inédite la tendance à la convergence du visuel et du sonore initiée par les médias précédents.

5/ De nouvelles situations de communication

Les hypermédias, qui sont à la fois des systèmes de *représentation* et de *production* d'images et de sons, instaurent de nouvelles situations de communication comme *l'interactivité, la générativité ou la simulation* :

- . Avec *l'interactivité*, le contrôle est donné à l'utilisateur, le système réagissant à ses manipulations dans un cadre défini par le concepteur.
- . Avec la *générativité*, le système va créer des objets multimédias autonomes indépendamment des actions de l'utilisateur.
- . Avec la *simulation*, il s'agira de créer un monde ou un "être" artificiel en amont de toute représentation médiatique.

Chacune de ces situations engendre des relations particulières entre les images et les sons et sont autant d'étapes vers la fusion des modalités.

Mais si l'on peut dissocier, en théorie, ces différentes approches, de nombreux systèmes hypermédias proposeront des solutions qui hybrident ces différents modes de communication avec aux deux extrêmes des systèmes interactifs non génératifs et des simulations ou des systèmes génératifs non interactifs (systèmes autonomes).

Interactivité et générativité :

Face à une réalisation hypermédia "classique" (interactive), l'utilisateur entre en contact avec, non plus un objet fixe, mais une œuvre à multiples facettes où le parcours entre ses éléments est soumis aux décisions du récepteur.

Face à un hypermédia mettant en œuvre des fonctions génératives, l'interacteur entre en communication avec un "modèle" informatique. Le concepteur du système travaille ainsi sur un aspect de la représentation bien différent de l'approche traditionnelle des médias linéaires. Il ne travaille plus sur la réalisation d'un objet fini (ou semi-ouvert) mais sur un modèle qui intègre les modalités de production d'une multitude de "possibles" ainsi que l'évolution temporelle de ces modalités. On n'est donc plus dans le régime de la trace mais bien dans celui de la production d'un réel multiple, d'une infinité de résultats toujours différents.

Générativité et simulation :

La distinction entre générativité et simulation est bien plus fine à définir puisque tout système génératif est une simulation d'un processus régi par un certain nombre de règles et que toute simulation s'actualise par des processus génératifs. Si cette proximité est vraie d'un point de vue technique (du point de vue de la production) elle est à nuancer du point de vue de la réception. La distinction de ces deux modes est bien à chercher du côté de la situation de communication mise en place. Une situation de dialogue avec un être ou un monde simulé est une situation fort différente de manipuler des jeux de règles d'un générateur de texte ou de musique.

6/ Deux approches pour lire la relation image/son

Pour mieux distinguer nos trois modes de communication, nous utiliserons deux notions :

- . celle de cause (développée par Roland Cahen ou Antoine Schmitt),
- . celle de mapping 'a-musicologique' (développée par Alain Bonardi et Francis Rousseaux).

La première met l'accent sur la situation de communication du point de vue de la réception, la seconde sur les impératifs de production (même si ce mapping n'est pas sans répercussions sur la perception de l'utilisateur).

Une "Esthétique de la cause" ?

Roland Cahen (compositeur) propose ([CAH 00] de diriger la réflexion vers les notions de cause et d'action : "Il s'agit d'une distinction de causalité. C'est l'idée de cause exclusive. La notion de cause serait donc à la fois au cœur de la genèse lorsqu'elle est exclusive et de l'interaction dans le cas d'une relation de cause à effet."

La différence entre l'interactivité et la générativité serait donc à trouver dans la cause de l'événement : dans les systèmes génératifs la cause est cachée dans le cœur du programme, dans le cas de l'interactivité, elle est toujours lisible par l'interacteur car directement liée à son intervention.

Roland Cahen utilise la notion de cause en la démarquant de la notion schaefferienne (la cause comme source de production du son) :

"La musique interactive se préoccupe des causes, non plus pour analyser le phénomène sonore pour lui-même, mais pour organiser des situations relationnelles entre un effet ; un son et sa cause ; l'événement qui le produit".

D'où l'importance primordiale d'une certaine lisibilité de la chaîne action-réaction d'une situation interactive.

Cette notion de cause, nous la retrouvons dans les approches de simulation de la synthèse par modèles physiques (où ce n'est plus le son qui est simulé mais son origine, sa cause) ou dans le travail d'artistes-programmeurs comme Antoine Schmitt ([SCH 02]) qui développe un discours sur "L' esthétisme de la cause" en parlant de ses travaux d'art programmé tourné vers la simulation d'êtres artificiels (l'enjeu de ces œuvres est la découverte progressive du comportement à la source de toute expression médiatisée du programme).

Cette approche peut nous aider à redéfinir nos trois modes de communication en soulignant que quand la cause :

- . pointe vers l'interacteur, on se trouve plutôt dans une situation d'*interactivité*,
- . pointe vers les règles du programme, on se trouve plutôt dans une situation de *générativité*,
- . pointe vers le monde ou l'être simulé, on se trouve plutôt dans une situation de *simulation*.

Importance du "mapping" :

L'analyse d'hypermédias mettant en oeuvre des interactions multi-modales, souligne la position centrale d'un "mapping" signifiant (table de correspondances entre les différents modes : image, son, geste) dans le contexte de l'oeuvre.

À la différence de l'informatique musicale où le rapport image/son est codifié de manière stricte : soit l'image est une représentation scientifique du son (forme d'onde ...) soit elle en est une représentation symbolique (solfèges divers), les œuvres hypermédias proposent des représentations plus libres et subjectives où le choix même du type de "mapping" est souvent le cœur même de l'oeuvre. Bonardi et Rousseaux ([ROU 03] [BON 03]) proposent de nommer ces types de réalisations : "systèmes a-musicologiques" pour définir des systèmes qui proposent de manipuler de la matière sonore à partir d'une représentation visuelle qui ne fait pas référence aux critères classiques de la musicologie (solfège, formes musicales ...).

Ces tables de correspondance ("mapping") se définissent différemment selon que l'on se trouve dans les différentes situations de communication décrites plus haut :

. Dans le cas de l'interactivité, les concepteurs sont face à deux objets distincts (les images et les sons) et doivent trouver des points de correspondances entre les deux (voir pour exemple les travaux exemplaires de Nicolas Clauss, Jean Birgé et Frédéric Durieux sur les sites : www.flyingpuppet.com et www.lecielestbleu.com). Cette situation rentre bien dans la catégorie des "mapping a-musicologiques" décrit par Rousseaux et Bonardi.

. Dans le cas des systèmes génératifs, cette correspondance se fait plutôt par le biais d'un métalangage de communication entre les différents générateurs de médias. Chaque générateur dégage des structures (durées, ordre de thèmes ...) pouvant être utiles aux autres générateurs (voir pour exemple le spectacle de Jean Pierre Balpe et Jaccopo Badoni-Schilingui "Trois mythologies et un poète aveugle" 1997).

Le "mapping" se fait ici via une étape de traduction dans une structure langagière.

. Dans le cas des réalisations mettant en oeuvre des systèmes de simulation, un niveau d'abstraction et de fusion est franchi. Il ne s'agit plus de coordonner plusieurs générateurs obéissant à des règles spécifiques mais d'avoir un seul objet programmé (un être ou un monde simulé) qui s'exprime par des générateurs de médias différents. On pourrait parler alors d'un langage en amont des médias, d'un langage "a-média" ([BOU 02]) Les représentations émergentes proviennent dans ce cas d'une même cause : l'objet simulé.

Ainsi, nous pouvons résumer l'impact des différentes approches de communication sur la relation image/son :

- . Interactivité Æ "Mapping a-musicologique"
- . Générativité Æ Métalangage
- . Simulation Æ Langage en amont des médias (langage "a-média")

Du point de vue de la réception, dans l'interactivité la sensation de fusion est obtenue grâce à l'accentuation d'effets de synchronisations entre les modalités (synchronisation déjà largement utilisée par le cinéma).

La générativité accentue l'expérience d'émergence médiatique d'un flot de données toujours neuf. La fusion dépendra alors de la lisibilité et de la cohérence des règles de production des différentes modalités

Dans la simulation, la cohérence du monde ou de l'être simulé est garante de l'effet de fusion des modalités.

7/ Vers une syntaxe commune aux différentes modalités (en amont et en surplomb des médias)

Plusieurs recherches parallèles vont dans le sens de l'élaboration de descriptions communes aux différentes modalités. Nous avons relevé ici quatre domaines qui privilégient des structures langagières soit pour produire des médias (1, 2, 3), c'est-à-dire 'en amont', soit pour les décrire (3, 4), c'est-à-dire 'en surplomb'.

Certaines parties de ces syntaxes sont liées à la modalité concernée, d'autres sont communes aux modalités en jeu : c'est-à-dire '*a-média*'.

Enfin, on peut souligner que ces syntaxes peuvent être constituées d'un langage (1,3), d'unités sémiotiques (2) ou de méta-données (4)

1) La synthèse sonore par modèle physique

Cette approche simule le comportement physique réel d'un objet auquel on injecte de l'énergie. Ici la modélisation ne cherche pas à simuler un certain résultat sonore mais plutôt à modéliser la cause de cet effet sonore. Cette approche prend ainsi le contre-pied des méthodes de synthèse classiques qui partent pour la plupart du son et de sa description. La synthèse par modèle physique réintroduit l'événement-source au cœur de la création sonore.

En partant de cette approche de modèles physiques, le laboratoire ACROE (Grenoble) a élaboré un langage "Cordis-Anima" qui permet, à partir d'une même description, de produire à la fois du son et de l'animation 3 D.

2) Une approche originale représentée par le groupe "Transitoire Observable" réunissant autour de Philippe Bootz, Alexandre Gherban et Tibor Papp des artistes/chercheurs (dont Jean Pierre Balpe ou Antoine Schmitt).

Ce groupe développe des stratégies variées pour élaborer des hypermédias d'un genre nouveau : non plus "orientés médias" mais plutôt "orientés programme". Dans ce nouveau type de production, le sens doit se construire au niveau du code, les différentes stratégies de représentation (image, son, texte) étant soumises aux décisions du programme en temps réel.

Pour cela, l'on doit pouvoir décrire un langage en amont des médias avec une sémantique et des règles de production commune.

Une des pistes de recherche du groupe est une généralisation des U.S.T. musicales au domaine visuel. Les Unités Sémiotiques Temporelles proposent un découpage en unités significatives, prenant comme unité de mesure le plus petit mouvement énergétique faisant sens sur le plan temporel (comme la chute, l'élan ...). Cette unité de sens permettrait d'identifier des schémas universels applicables aussi bien au sonore qu'au visuel.

3) Les interfaces multi-modales :

Ce domaine applicatif qui cherche à enrichir le dialogue homme/machine par l'utilisation de différents modes dans l'échange d'informations souligne la nécessité de différencier, pour un message donné, son contenu sémantique de sa forme de représentation. Cette différenciation permet de faciliter la traduction du message d'un mode à l'autre.

Les interfaces multi-modales s'intéressent autant à la production qu'à la réception d'un message. Du point de vue de la machine, la réception d'un message multimodal suppose une étape de 'fusion' des différentes modalités dans un langage commun (le CMR : Common Meaning Représentation). A l'inverse, en situation de production de message, une étape de 'fission' de l'information en plusieurs modalités est nécessaire.

4) L'évolution de la norme Mpeg (Mpeg4, Mpeg7 et Mpeg21) :

L'évolution de Mpeg fait passer la norme de "simples" algorithmes de compression (Mpeg1 et 2) à une norme complexe qui intègre :

- . des découpages en objets séparés et une représentation objet qui permet une certaine interactivité avec ces objets différenciés (Mpeg4),
- . des analyses et des descriptions de contenus (mpeg7),
- . la prise en compte des notions de droits d'auteurs et de sécurité (Mpeg21).

Du côté de l'analyse du signal, ces nouvelles normes sont capables de scruter les contenus vidéos et sonores de manières de plus en plus fines et d'en comprendre certains éléments constitutifs (séparation du fond et des personnages par exemple pour la vidéo dans la norme Mpeg4, reconnaissance des formes musicales : mélodie, tonalité, rythme pour le son dans la norme Mpeg7 ...). Cette tendance à l'analyse du contenu des médias est associée à l'élaboration de descripteurs, de plus en plus sophistiqués, qui définissent de véritables contextes de connaissances (les ontologies d'un domaine) pour permettre des inférences sur ces descripteurs, des raisonnements automatiques sur ces connaissances. Ces techniques sont de plus en plus souvent réversibles et permettent à partir de descripteurs et de réseaux de connaissances de produire des médias.

8/ De nouvelles relations avec l'utilisateur

Ces différentes directions qui privilégient les méta-données et les structures langagières pour la description ou la production de médias permettent des interactions nouvelles avec ces objets.

Ces nouvelles applications centrées sur le contenu et la connaissance permettent :

- . des recherches par similarité (“trouve moi une forme ou une mélodie qui ressemble à ...”), par éléments constitutifs (“trouve moi une forme ou une mélodie qui contient ...”) par raisonnement (“trouve moi une forme ou une mélodie qui est influencée par ...”),
- . l’apparition d’outils d’aide à la création d’un genre nouveau (véritables assistants créatifs qui proposent des solutions qui nécessitent un raisonnement, des connaissances, des profils stylistiques ...).

Par exemple, les futurs systèmes d’écoute qui favorisent une ‘écoute active’, proposent à l’auditeur d’accéder et de manipuler la structure même de l’œuvre (ce que Hughes Vinet [VIN 04], appelle le *modèle génératif* du compositeur) permettant ainsi des parcours multiples, des mixages ou des orchestrations personnalisés de l’œuvre. Ce qui est transmis ici n’est plus ni le son ni la partition mais les outils et les modèles même de production.

Outre l’intérêt pédagogique de ce type de système (où l’œuvre et l’analyse de l’œuvre ne font qu’un), on voit bien que l’on se dirige vers une maléabilité toujours plus grande du flux audiovisuel. Ce n’est plus en effet uniquement l’accès au contenu qui est personnalisé mais bien la ‘substance audiovisuelle’ elle-même qui s’adapte et réagit dans un dialogue toujours plus étroit avec l’utilisateur.

9/ Vers un nouvel objet audio-visuel

La tendance à la fusion permet d’entrevoir l’émergence d’un nouvel objet audiovisuel :

Cette fusion s’effectue sur deux plans parallèles :

- . fusion des modalités,
- . fusion du sensible et du calcul.

Fusion des modalités :

On l’a vue, la fusion des modalités, avec le numérique, va s’opérer au moins sur deux niveaux :

- . au niveau du codage commun qui permet d’associer les paramètres spécifiques d’un média à ceux d’un autre, et d’accentuer ainsi leur proximité. Ce niveau de fusion prolonge et accentue les effets de synchrèse et de valeur ajoutée qu’à repéré Michel Chion au cinéma,
- . au niveau de différentes structures langagières qui décrivent et produisent la substance audiovisuelle (méta-data, méta-langages, unités sémiotiques diverses ...).

Fusion du sensible et du calcul :

Lev Manovitch, dans son livre ‘The language of new media’ [MAN 01], décrit les hypermédias comme ‘cinématographique en surface, digital dans sa structure et computationnel dans sa logique’⁴.

Cette convergence de l’audiovisuel et de l’informatique font cohabiter deux logiques (Manovitch parle de ‘couche culturelle’ et de ‘couche computationnelle’⁵) qui appelle une nouvelle approche de l’étude des médias qui va chercher ses modèles dans le domaine de l’informatique⁶.

Il semble que l’impact de cette convergence n’est pas encore évaluée à sa juste mesure en terme de modification de la notion même de média. On continue, en effet, à parler d’image ou de son comme entité homogène alors que l’on est de plus en plus confronté à des objets composites, structurés et complexes.

Ce nouvel objet audiovisuel pourrait être caractérisé par un enchevêtrement subtil entre ses différentes composantes : l’image, le son, l’interaction permise entre l’utilisateur, son comportement. Ce comportement inclus la capacité de mémorisation et de transformation dans le temps de toutes ces dimensions (réflexivité).

En tant qu’objets programmables, ils subissent directement l’influence des différents paradigmes de programmation (par objets, par agents, par réseau de neurones ...) et l’on peut ainsi considérer les hypermédias, qui émergent aujourd’hui, comme une forme médiatique issue du paradigme de la programmation orientée objet.

⁴ ‘To summarize, the visual culture of a computer age is cinematographic in its appearance, digital on the level of its material and computational (i.e. software driven) in its logic.’ (‘The language of new media’, p.180)

⁵ ‘Because new media is created on computers, distributed via computers and stored and archived via computers, the logic of a computer can be expected to significantly influence the traditional cultural logic of media ; that is, we may expect that the computer layer affect the cultural layer.’ (Idem, p.46)

⁶ ‘New media may look like media, but this is only the surface. ... To understand the logic of new media, we need to turn to computer science. It is there that we may expect to find the new terms, categories, and operations that characterize media that became programmable. From media studies, we move to something that can be called ‘software studies’, from media theory to software theory. (Idem, p.47-48)

Conclusion :

On a vu dans ces quelques pages comment le numérique représente une rupture dans la relation image/son (modalités de communication spécifiques, calcul, description langagière ...) et s'inscrit, en même temps, dans une logique historique de cette relation (peinture, cinéma, cinéma expérimental).

Cette logique peut être résumée comme une dynamique de convergence qui va de la correspondance à la fusion.

Cette fusion des modalités s'effectue sur plusieurs plans :

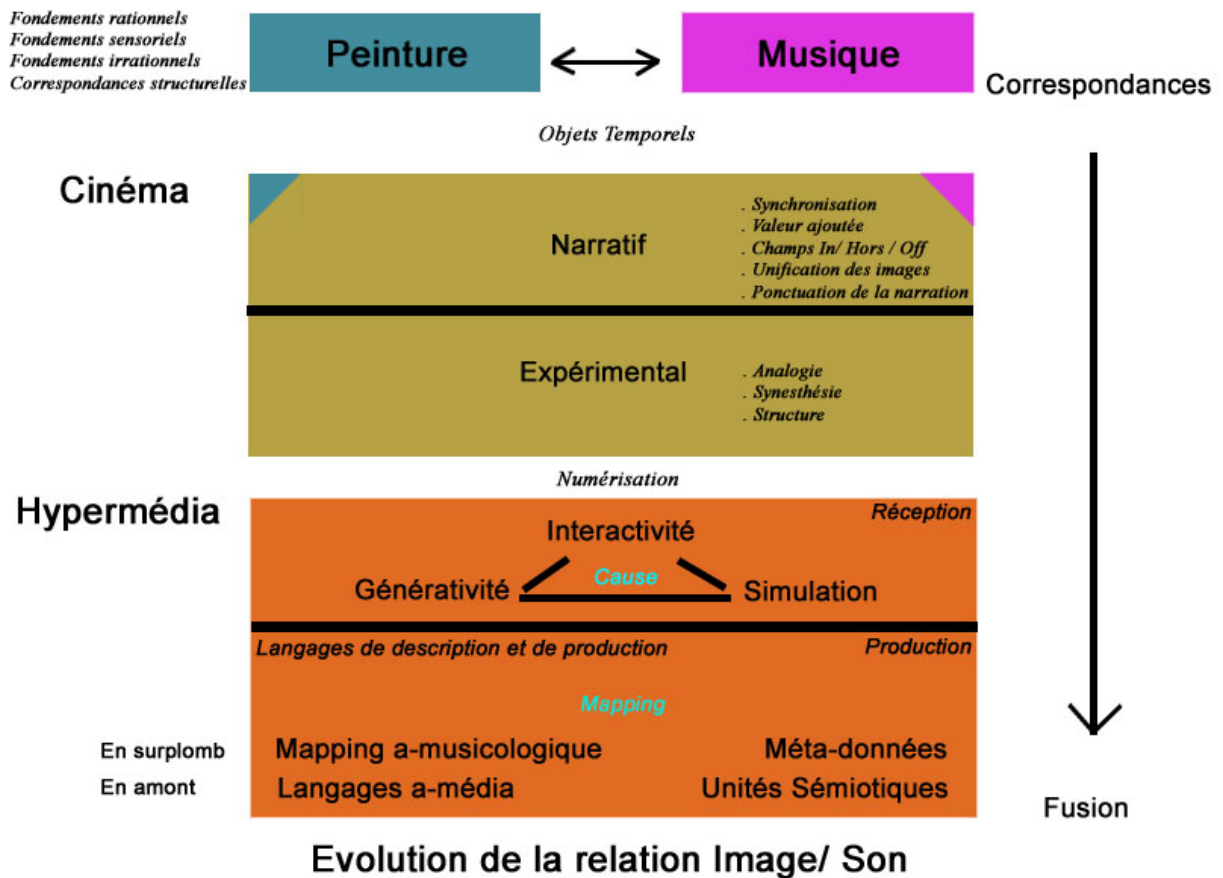
- . au niveau du codage numérique qui accentue leur proximité,
- . au niveau des structures langagières de description et de production.

D'autre part, un mouvement parallèle tend à fusionner la logique audiovisuelle (la 'couche culturelle') avec la logique informatique (la 'couche computationnelle').

L'addition de ces deux logiques de fusions modifie en profondeur la notion même de média pour faire émerger un nouvel objet audiovisuel programmable qui n'est plus vraiment une image, une vidéo ou un son mais un objet audiovisuel composite, structuré et complexe.

Ce changement de statut a des répercussions à la fois techniques et esthétiques et concerne autant la production que la réception (comme on a pu le voir succinctement dans les exemples d'applications possibles)

Attaquée de toutes parts (par l'analyse du signal, par les méta-données, par les métalangages ...), qu'elle soit produite (par des langages 'en amont' des médias et par du calcul) ou bien décrite (par des méta-données 'en surplomb' des médias), la 'substance audiovisuelle' a tendance à se dissoudre ainsi progressivement dans du langage.



BIBLIOGRAPHIE SELECTIVE

Esthétique comparée :

[SOU 69] Souriau E., *La correspondance des arts*, Paris, Editions Flammarion, 1969

Musique :

[CAG 76] Cage, J., *Pour les oiseaux*, Paris, Editions Belfond, 1976

[DEL 01] Delalande F., (Sous la direction de), *Le son des musiques*, Paris, Editions Ina/Grm, 2001

[VIN 04] Vinet H., 'Explorer la musique à l'âge numérique', Les cahiers de médiologie N° 18 "Révolutions industrielles de la musique", Paris, Editions Fayard, 2004

Relation musique / peinture :

[BOS 98] Bosseur J. Y., *Musique et arts plastiques*, Paris, Editions Minerve, 1998

[KOU 03] Kouprovskaja-Bruggeman, E., *Paul Klee et son influence sur la création musicale*, Thèse Paris IV, 2003

[PAR N.C.] Parrat J., *Des relations entre la peinture et la musique dans l'art contemporain*, Grenoble, Editions Z'éditions, date n.c.

Son et cinéma :

[BEA 86] Beauvais Y., (sous la direction de), *Musique Film*, Paris, Editions La cinémathèque française, 1986

[CHI 03] Chion M., *Un art sonore, le cinéma*, Paris, Editions Cahiers du cinéma, 2003

[GHA 95] Ghali N., *L'avant garde cinématographique en France dans les années 20*, Paris, Editions Paris expérimental, 1995

[NOG 99] Noguez D., *Eloge du cinéma expérimental*, Paris, Editions Paris expérimental, 1999

[PRA 03] Pratiques N°14, *Expérimentations cinématographiques : une vision de l'art élargie*, Rennes, Presses Universitaires de Rennes, 2003

[SIT 02] Sitney A., *Le cinéma visionnaire, l'avant garde américaine 1943/2000*, Paris, Editions Paris expérimental, 1999

Hypermédia :

[BAL 00] Balpe J. P., *Les contextes de l'art numérique*, Paris, Editions Hermes sciences, 2000

[BOU 02] Bootz P., *Vers un multimédia contraint et a-média*, Paris, <http://transitoireobs.free.fr>, 2002

[CAH 00] Cahen R., *Générativité et interactivité en musique et en art électroacoustique*, Paris, <http://perso.wanadoo.fr/roland.cahen>, 2000

[GOL 94] Levin G., *Painterly Interfaces for Audiovisual Performance*, Boston, B.S. Art and Design M. I. T., 1994

[MAN 01] Manovich L., *The language of new media*, Boston, Editions MIT Press, 2001

[REN 94] Renaud-Alain A., *L'image sans gravité, la forme image aux risques de l'information* in "Revue d'esthétique", Paris, 1994

[ROU 03] [BON 03] Rousseaux F. et Bonardi A., *Comprendre des pratiques qui provoquent la musicologie*, Paris, Ichim 2003

[SCH 02] Schmitt A., *Le travail du temps : programmer "un mode d'être"* (Entretien avec Samuel Bianchini) Paris, <http://www.gratin.org/as/>, 2002

[WEI 99] Weissberg J. L., *Présence à distance*, Paris, Editions L'harmattan, 1999