



**HAL**  
open science

# PRÉMISSES POUR UNE ÉTUDE SUR L'OBJET LOGICIEL AUDIOVISUEL, CONSIDÉRATIONS ET PREMIERS RÉSULTATS

Alessio Santini

► **To cite this version:**

Alessio Santini. PRÉMISSES POUR UNE ÉTUDE SUR L'OBJET LOGICIEL AUDIOVISUEL, CONSIDÉRATIONS ET PREMIERS RÉSULTATS. Journées d'Informatique Musicale, Mar 2008, Albi, France. hal-03120828

**HAL Id: hal-03120828**

**<https://hal.science/hal-03120828>**

Submitted on 25 Jan 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# PRÉMISSES POUR UNE ÉTUDE SUR L'OBJET LOGICIEL AUDIOVISUEL, CONSIDÉRATIONS ET PREMIERS RÉSULTATS

Alessio Santini  
Université Paris 8  
CICM - MSH Paris Nord  
alessio.santini@gmail.com

## RÉSUMÉ

*Dans cet article, l'auteur propose ses réflexions à propos de la composition audiovisuelle. Après quelques rappels sur les rapports image-son dans le domaine des arts numériques, les questions liées à l'approche de l'auteur seront présentées : le matériau utilisé, les problématiques liées au matériau-même et à l'écriture, les questions techniques et l'aspect logiciel. Une première production artistique sera présentée.*

## 1. INTRODUCTION

Cet article présente les prémisses de notre recherche doctorale, ayant pour sujet « l'objet logiciel audiovisuel ».

Cette recherche a lieu au sein de l'école doctorale "Esthétique, Sciences et Technologies des Arts" de l'Université de Paris VIII, dont le champ de recherche, au sein des sciences humaines, permet aux artistes chercheurs de développer des travaux de réflexion théorique sur leur propre pratique artistique en relation avec l'usage des technologies, pour une pensée artistique se développant en interaction avec le développement des outils numériques.

Nous avons dans ce cadre pour objectif de développer des stratégies de programmation pour composer des média audiovisuels. Notre intérêt est centré surtout sur l'élaboration d'une théorie et d'une pratique qui permettent de composer le son et l'image simultanément. Ce que nous visons au niveau artistique, ce sont des productions audiovisuelles expérimentales, renouvelant la relation image-son, notamment dans la « vidéo d'art ».

## 2. CONVERGENCES IMAGE-SON DANS LES ARTS NUMÉRIQUES

Il ne s'agit pas ici de dresser une liste exhaustive des expériences notables dans les domaines de l'interaction image-son dans les oeuvres audiovisuelles expérimentales, mais d'en exposer quelques-unes qui ont influencé notre réflexion personnelle.

### 2.1. L'apport du groupe de travail "Visualisation du son" de L'AFIM<sup>1</sup>.

De 2004 à 2006, le groupe de travail "Visualisation du son" de L'AFIM<sup>2</sup> [11] a approfondi des recherches basées sur l'appropriation par les artistes des perspectives de la convergence entre image et son, convergence elle-même liée à la convergence des outils de programmation informatique, en mutation permanente. Les divers développements artistiques avec les moyens de Jitter, en tant que bibliothèque vidéo numérique de l'environnement de programmation Max et de sa bibliothèque audionumérique MSP<sup>3</sup>, constituent un exemple évident d'appropriation par les musiciens du potentiel de la convergence audiovisuelle d'une part, au sens de la relation image-son dans la composition, et de la convergence numérique d'autre part. Les moyens de Pure Data associés à Gem pourraient constituer le même type d'exemple<sup>4</sup>. SuperCollider<sup>5</sup> serait également une excellente plateforme de travail pour développer des formes de convergence de l'image et du son pour une composition audiovisuel, quoique pour notre part notre préférence en tant qu'utilisateur va vers des plates formes telles que Max/Pure Data.

Dans ce groupe de travail, le thème de la visualisation du son a été un angle d'attaque pour

---

<sup>1</sup> Association Française d'Informatique Musicale, <http://www.afim-asso.org>

<sup>2</sup> Les développements du groupe de travail sont présentés sur le site <http://gtv.mshparisnord.org>

<sup>3</sup> Max, MSP et Jitter sont distribués par Cycling '74, <http://www.cycling74.com>

<sup>4</sup> Pure Data, de Miller Puckette, et Gem sont disponibles à l'adresse <http://puredata.org/>

<sup>5</sup> SuperCollider, développé par James McCartney, est disponible à l'adresse <http://supercollider.sourceforge.net/>

approfondir au niveau théorique ces questions de la relation image-son dans des démarches de composition musicale, visant à intégrer le visuel avec le musical, pour au fond étendre les moyens de l'expressivité sonore jusqu'au domaine visuel. On a, entre autres, abordé dans ce contexte la notion d'émulation entre ce qui de l'ordre du sonore et ce qui est de l'ordre du visuel, l'importance de la notion de rythme sonore ou/et visuel pour construction musicale composite, étendue au visuel, la question de l'intersensorialité, la perception intersensorielle du temps.

## 2.2. Opérations de sonification.

La sonification de l'image, processus inverse de la visualisation du son, constitue un autre "classique" des relations image-son, produit lui aussi de la convergence liée aux opérations de traitement offert par l'informatique (dont par exemple la sonification d'un texte, ou d'une équation). Elle est essentiellement basée sur le traitement des matrices. Cette approche a été traitée entre autres par Woon Seung Yeo et Jonathan Berger, dont nous connaissons la mise au point des techniques de sonification de l'image par *scanning* et *probing* [14]. De façon générale, les différents types d'analyse des valeurs d'une matrice vidéo, produisent différentes possibilités d'exploitation des données extraites pour produire ou traiter du son.

## 2.3. Objet "audio-visuel"

Le troisième monde possible, dans l'interaction image-son, est celui de « l'objet audiovisuel ». De notre point de vue, cet objet présente deux composants différents : le premier composant, qui correspond à sa programmation, à son code, représente l'élément productif; il est un objet ou une brique informatique qui gère les paramètres de contrôle des deux média simultanément. Le deuxième est l'objet résultant : le produit, expression d'une écriture hybride, composé intégrant des éléments sonores et des éléments visuels.

Les thématiques de l'objet logiciel audiovisuel ont déjà été traitées entre autres par Cyrille Henry autour de son projet *CHDH*<sup>6</sup> [8]. Dans sa performance, Henry, emploie des algorithmes de modélisations physiques pour la production en temps réel d'images et de sons de synthèse.

Sur l'objet audiovisuel on peut citer encore le duo *Klippav*, formé par Nick Collins et Fredrik Olofsson<sup>7</sup> [3]. Ils développent des performances où le flux audio et des captures vidéo sont traités de manière synchrone. Les fichiers sont découpés en plusieurs morceaux de différentes longueurs puis reproduits et répétés par des algorithmes contrôlés par les exécutants en temps réel, dans un contexte de performance de tendance « breakbeat ».

Nous revenons plus loin sur les aspects de la notion d'objet telle que nous l'employons, pour le moment, dans ce contexte. Avant cela, une présentation de notre propre pratique des techniques de composition granulaires s'impose.

## 3. DE LA GRANULATION DU SON À LA MICRO-FRAGMENTATION TEMPORELLE AUDIOVISUELLE

Notre parcours de pensée s'est développé en relation avec des procédés de granulation du son, que nous avons développés durant ces deux dernières années. On peut « bouleverser » le sens originel d'un objet sonore, qui se déroule sur l'axe temporel, en agissant sur les plus petites unités du flux audio. Le potentiel expressif de cette technique, largement employé de nos jours nous a suggéré d'essayer un processus similaire pour composer un clip vidéo [9].

Nous avons commencé par synchroniser l'unité minimale du son avec l'unité minimale de la vidéo. Pour ce faire, nous avons utilisé le patch *K.grainSAMPLE* que nous avons programmé dans Max/MSP pour la granulation du son en temps réel<sup>8</sup>. Les paramètres dont ils disposent sont standards [9] : densité des grains, sélection du point de lecture de l'échantillon, sélection aléatoire, durée des grains, transposition, filtre, panning, et amplitude. Un objet *metro* envoie une commande de déclenchement dans les *instances* d'un objet *poly~* qui génèrent chacune un grain.

Pour synchroniser les unités minimales des deux média, nous avons d'abord envoyé le même *bang* qui déclenche les grains vers des paramètres d'un objet *jit.qt.movie* : ainsi, chaque *bang* déclenche une valeur aléatoire qui détermine l'image à lire. Si la valeur aléatoire est définie par le paramètre du granulateur qui indique la sélection aléatoire, on obtient un échantillon

---

<sup>6</sup> CHDH est une performance live, où la génération d'événements audio et vidéo est pilotée par les mêmes paramètres de contrôle. Des nombreux exemples vidéos sont disponibles à l'adresse <http://www.chdh.free.fr>

<sup>7</sup> Des nombreux exemples vidéos sont disponibles à l'adresse <http://www.klippav.org>

<sup>8</sup> Ce patch est disponible à l'adresse <http://www.cycling74.com/twiki/bin/view/Share/AlessioSantini>

audio et un clip vidéo qui trouvent bouleversée leur séquence temporelle originelle de manière synchrone, avec la perte du sens initialement porté par chacun.

Cette technique présupposait de faire correspondre chaque grain à une image. Après quelques essais il est apparu que ce n'était pas suffisant : en présence de séquences de grains très courts et de densité très haute, la recherche de la synchronie était peu efficace pour le rendu expressif de la vidéo.

Pour répondre à ce type de problèmes, plusieurs stratégies sont possibles. Par exemple, Benoît Courribet propose, afin de garder le rapport de synchronie, d'exploiter l'espace bidimensionnel de l'écran, en projetant les images en différentes dimensions et en différents endroits<sup>9</sup> [4]. Quand à nous, nous avons préféré postuler pour une autre stratégie qui sera exposée plus bas.

## 4. LE CHOIX DU MATÉRIAU SONORE ET VISUEL

Les techniques de composition granulaire du son musical ont constitué notre première approche pour une composition audiovisuelle [9].

### 4.1. La pensée de l'objet

Il est important de s'arrêter un instant sur notre pensée de l'« objet ».

En tant que musicien, notre approche de l'objet est sans aucun doute à mettre en relation avec la pensée Schaefferienne [10].

Le matériau sonore à la base de notre pratique musicale est toujours un objet sonore, presque objet trouvé, matériau préexistant sous la forme d'un phonogramme, un fichier de son, un échantillon<sup>10</sup>.

D'une certaine façon il s'agit toujours d'un matériau composé, du seul fait même de l'opération de son enregistrement.

### 4.2. Objet ou unité, définitions en cours

On pourrait substituer à la notion d'objet sonore ou musical, autrement dit, objet composé, la notion d'« unité ».

A ce point, on empruntera la notion d'unité sonore composée, telle qu'elle est actuellement développée par le musicologue Didier Guigue [5] [6].

Nous allons étendre cette notion à celle d'unité audiovisuelle composée. En effet, quant au matériau visuel, de la même façon que pour l'objet sonore, il s'agit d'un objet trouvé, d'un matériau préexistant sous la forme d'une image, qu'elle soit d'ailleurs fixe ou en mouvement (photo ou vidéo). Dans un environnement de programmation tel que Max/Msp/Jitter, faisant converger l'image et le son dans la manière d'opérer, on pourrait penser que les opérations de même nature que l'on applique à ces deux média construisent de fait une relation cohérente entre l'image et le son. Mais, pour ce qui nous concerne, et c'est là un choix expressif personnel, l'utilisation de l'image, nous paraît contrainte par le contenu sémiotique qu'elle porte<sup>11</sup> [7].

Selon nous, composer un objet audio visuel revient à manipuler au niveau élémentaire des unités audiovisuelles composées.

On pourrait au passage, évoquer la proposition des Unités Temporelles Sémiotiques, non pas au niveau sonore, où contrairement à l'image, la sémiotique y est très problématique. En milieu visuel, la proposition des Unités Temporelles Sémiotiques trouverait peut-être mieux sa place qu'en milieu sonore, d'autant qu'au niveau de la macroforme, nous ne visons surtout pas une approche narrative, pas plus que des développements linéaires simples. On pourrait parler peut-être également d'unités sémiotiques visuelles. Notons que nous sommes en cours de définition des concepts, aussi nous nous réservons dans notre travail futur de faire évoluer ces définitions.

### 4.3. Anthem de Bill Viola

Dans une perspective d'interaction des significations des deux média, la maîtrise de ces caractéristiques peut offrir de puissants moyens d'expression.

À ce propos, nous trouvons un exemple admirable d'utilisation d'écriture audiovisuelle dans l'oeuvre *Anthem* de Bill Viola<sup>12</sup> [13]. Viola y utilise des signes reconnaissables dans les images qui supportent et enrichissent l'articulation du son. Souvent, en présence du même son (le cri d'une jeune fille traité de différentes manières), les différentes images qui lui sont associées, offrent un changement dans la perception que l'on a de ce son. Les fréquences, les signes et les éléments rythmiques des deux média se fondent dans un langage dont l'hybridation apparaît consistante.

<sup>9</sup> Des extraits des travaux de Benoît Courribet sont disponibles à l'adresse <http://gtv.mshparisnord.org/spip.php?article20>

<sup>10</sup> Même au sens le plus réduit du terme, le fragment temporel le plus court offert par un système de production sonore donné. Par exemple, 1s/44100

<sup>11</sup> Dans notre recherche, l'étude de la sémiotique de l'image s'impose

<sup>12</sup> L'oeuvre est disponible à l'adresse <http://www.ubu.com/film/viola.html>

## 5. LES LOGICIELS ET LA MÉTHODE DE COMPOSITION

Pour revenir à ce qui a été dit plus haut, la logique de la granulation n'est qu'un moment de notre recherche, une première tentative de délinéarisation du "discours" audiovisuel. Les exigences de composition et montage peuvent changer selon les occasions, et c'est pour cela que nous avons privilégié l'idée de développer un système modulaire, flexible et dynamique de création : plusieurs briques logicielles sous formes de sous-patches de type *Bpatcher* traitent simultanément l'image et le son avec les mêmes commandes. Des exceptions peuvent exister : nous expliquerons par la suite comment des modules qui récupèrent des données d'analyse audio ou vidéo peuvent enrichir la palette du compositeur avec des logiques de ré-injection.

Nous nous occupons d'abord de créer une base de données audiovisuelles, pour passer ensuite au montage. Jusqu'au montage, les fichiers sont traités de manière synchrone. Considérons d'abord deux types de synchronie : une « synchronie absolue » et une « synchronie relative ». La première se vérifie quand les longueurs des deux média sont mises en relation et identifiées : chaque échantillon du fichier audio est mis en relation fixe avec une image du fichier vidéo. En considérant la longueur du fichier audio (qui change dynamiquement avec la vitesse de lecture), la vitesse de lecture de l'objet *jit.qt.movie* change également. Cette modalité est efficace quand le contenu audio du *buffer~* est lu dans un objet *groove~* : hors du domaine du micro-temps, elle est très pratique pour l'application des effets et pour l'écriture, comme pour la création et la répétition de boucles audiovisuelles.

La synchronie relative se vérifie quand les micro-fragments audio et vidéo sont sélectionnés avec les mêmes réglages aléatoires, mais après étalonnage exclusif. Les réglages agissent sur le développement temporel des fichiers, mais l'algorithme est exclusif pour chaque médium.

Le premier pas dans la création du matériau de base consiste à importer un fichier audio et un fichier vidéo (ils peuvent bien sûr être aussi les deux composants d'une même source) dans le patch *Max/MSP/Jitter* programmé pour la micro-fragmentation temporelle en temps réel. Les fichiers sont traités selon un processus de bouleversement sur l'axe temporel : le flux audio selon la technique de granulation, la matrice vidéo selon un système de sélection des images à lire. Après la découverte des limites dont nous avons parlé plus haut,

qui émergeaient de la synchronie grain / image, nous avons introduit un système de réglages pour éliminer la synchronie dans le cas de certaines valeurs de densités des grains trop hautes. Cette absence de synchronie peut être substituée par plusieurs *presets* d'effets vidéo qui sont donc déclenchés à l'occasion, ou aussi avec un changement d'état (0. ou 1.) du *rate*<sup>13</sup> de l'objet *jit.qt.video*. Si le *rate* est 0. chaque grain déclenche une image, et l'absence de synchronie conduit à une image fixe ; si le *rate* est 1. le fichier est lu normalement et chaque grain provoque un saut dans la lecture. Dans ce cas, l'absence de synchronie conduit à une lecture normale du fichier. Normalement, en utilisant des fichiers vidéo qui n'ont pas un développement temporel originel, comme dans le cas de l'enchaînement de plusieurs photos, il est plus efficace d'utiliser la fonction *rate = 0.*

Le résultat de cette première opération est ce que l'on peut appeler une « unité audiovisuelle composée ». La question de la méthode de l'écriture de l'unité reste ouverte à différentes solutions : pour l'instant nous avons essayé avec un séquenceur réalisé dans *Max/MSP* et avec un gant muni de capteurs (dans le cadre d'une expérimentation pour l'usage sur la scène). Quelle que soit leur origine, les données sont stockées dans un objet *coll* : une phase d'écriture et une phase de lecture sont possibles. La phase de lecture permet au compositeur d'évaluer et de corriger la phrase après l'avoir écrite ou performée.

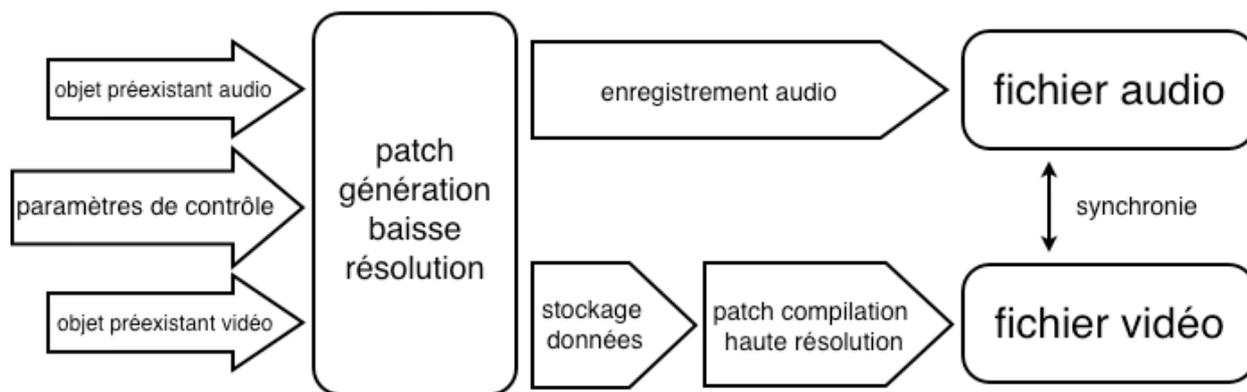
Il est possible de transformer l'unité audiovisuelle obtenue en appliquant des effets audio et vidéo, jusqu'à l'itération. Nous avons développé une série des patches qui appliquent des effets aux deux média de manière synchronisée. Pour l'instant il s'agit de patches indépendants mais nous considérons l'idée de développer une suite de petits patch à enchaîner de manière modulaire selon les besoins du moment.

Dans chacune de ces deux phases (génération, et application d'effets) nous avons ajouté un passage pour la compilation en temps différé du fichier vidéo afin d'obtenir un résultat de qualité qui, dans le cadre du temps réel, aurait été compromis, surtout au niveau de la fréquence d'images qui baisse de manière importante si on demande à la machine à la fois d'exécuter les commandes et de compiler le résultat dans un nouveau fichier.

Normalement on joue l'improvisation ou les indications du séquenceur dans un patch en basse résolution vidéo et on stocke les données des paramètres vidéo dans une nouvelle liste *coll*. Cette liste est ensuite ouverte dans un nouveau patch en haute résolution qui

---

<sup>13</sup> La valeur du *rate* contrôle la vitesse de lecture. Si cette valeur est 1. le fichier est lu avec la vitesse originelle



**Figure 1.** Schéma de la production d'une unité audiovisuelle : l'audio est enregistré tout de suite, la vidéo demande un passage supplémentaire de stockage des données et une compilation image par image en temps différé.

compile le fichier vidéo image par image, en temps différé. Le composant audio, qui n'entraîne pas ce type de problèmes, est enregistré dans un fichier dans le patch en basse résolution.

Quand on a compilé une quantité satisfaisante des unités audiovisuelles composées, on peut passer à la phase du montage. Jusqu'ici nous avons utilisé un logiciel de montage vidéo, comme on utilise normalement un logiciel de montage audio : les phrases audiovisuelles sont disposées sur la timeline, coupées, déplacées, répétées, etc. Le montage est nécessairement un processus différent selon le média : pour les composants sonores, on emploiera le mixage audio, et pour les composants image, les techniques d'incrustation vidéo.

La dernière phase de montage consiste dans l'application de quelques effets pour amalgamer certains éléments de manière arbitraire. De plus, l'utilisateur peut choisir de ne pas utiliser tous les composants vidéo, ou d'ajouter des sons : dans une logique de montage de plusieurs unités, il n'est pas nécessaire qu'à chaque nouvel événement audio corresponde un nouvel événement vidéo. Le compositeur peut utiliser le son détaché de son clip de différentes manières : on peut par exemple modifier un paramètre général du master vidéo « en suivant » le son<sup>14</sup>.

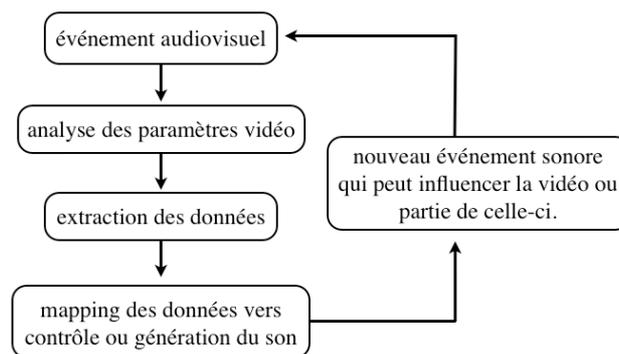
### 5.1. Des modules de base réalisés

- delay → ancrage de l'image : ce module permet de traiter une unité audiovisuelle. La même commande envoie des valeurs soit au temps de délai audio du module, soit à l'ancrage de l'image sur l'axe  $x$ ,  $y$ , ou les deux.

- amplitude audio → *jit.op* : ce module gère la vidéo en rapport aux résultats de certaines opérations mathématiques entre le niveau d'amplitude du son et les valeurs des cellules de la matrice vidéo en entrée. L'utilisateur peut sélectionner la fonction opératoire de l'objet *jit.op* pour changer d'opération mathématique.

- VA.gen

Le module *VA.gen* analyse la matrice vidéo et extrait les données concernant les niveaux RGB (*red, green, blue*) et HSL (*hue, saturation, luminance*). Pour chaque niveau on peut analyser sa valeur numérique (donnée continue) et la vitesse avec laquelle cette valeur change (donnée *bang*). Si la vitesse de changement est plus grande que la valeur-seuil sélectionnée par l'utilisateur, un *bang* est



**Figure 2.** Schéma de réinjection des données avec le module *VA.gen*

<sup>14</sup> Exemple : un son qui consiste en quatre événements percussifs qui se répètent peut correspondre à une perturbation dans la vidéo compilée

déclenché ; le fonctionnement est presque le même que le relevé des transitoires d'attaque dans un flux audio. L'utilisation de ce module est très utile pour créer un effet de « réinjection » des valeurs vidéo vers le contrôle des paramètres du son, et donc, si on veut, de la vidéo encore.

## 6. PRODUCTION ARTISTIQUE

### 6.1. *Persistenza del Ferro*

*Persistenza del Ferro* est une pièce audiovisuelle de 4 minutes composée en utilisant les programmes et les techniques décrites dans cet article<sup>15</sup>. La base de données est constituée par des échantillons audio enregistrés en différentes occasions, et par des clip vidéo et des photos qui témoignent de l'état présent et passé des mines de fer d'un village sur une île de Toscane, en Italie.

Le but était de traiter la thématique des conditions de travail difficiles des mineurs de la première moitié du siècle passé, de manière à exprimer la fatigue physique et une forme de violence. On peut considérer le résultat final comme un court-métrage non-narratif où des éléments intelligibles émergent sporadiquement depuis un continuum audiovisuel en grande partie abstrait.

La composition a été beaucoup influencée par la technique de traitement et montage de la musique granulaire [1] [2] [12] elle est perceptible dans le résultat de la composante audio et dans la structure générale de la pièce.

### 6.2. Un premier bilan

L'analyse des qualités et limites de la pièce audiovisuelle que nous avons réalisée nous offre l'occasion de réfléchir plus avant sur la technique de création et écriture, et sur la gestion des événements dans le flux du temps.

Ayant mené essentiellement des études musicales, notre travail est encore beaucoup conditionné par celles-ci : notre objet audiovisuel appartient encore d'une certaine manière au domaine de la visualisation du son, parce que sa structure est encore fortement dépendante des logiques constructives musicales. Si, dans le patch utilisé pour générer les unités audiovisuelles, il est vrai que les paramètres sont communs et que l'analyse des caractéristiques formelles des média est seulement un enrichissement accessoire, ces paramètres communs sont aussi presque identiques à ceux d'un granulateur audio.

Une des limites expressives de la pièce que nous avons produite est l'utilisation des couleurs : la très haute saturation, par exemple, a été en partie nécessaire pour pallier à la basse qualité des sources vidéo aussi

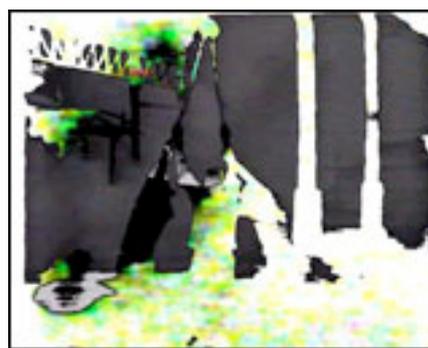


Figure 3. Images extraites de *Persistenza del Ferro*

bien qu'à notre inexpérience.

De l'analyse de *Persistenza del Ferro*, nous avons trouvé des éléments de réflexion supplémentaires, parmi lesquels la gestion du silence (ou l'absence du son) dans un régime de synchronie audiovisuelle. On peut bien illustrer ce concept avec un exemple précis : on peut dans la pièce distinguer clairement, au milieu de plusieurs incrustations, la figure d'un mineur battant le fer d'une longue masse. Le spectateur suit alors le développement d'un mouvement qu'il reconnaît et qui lui suggère l'arrivée d'un « coup sonore ». Cette attente donnée par une expérience partagée (le coup de marteau produit un coup sonore) offre un grand potentiel expressif. Dans notre pièce le mouvement du marteau se répète deux fois, l'une après l'autre ; la première est synchronisée avec un son percussif granulaire, la deuxième se produit en l'absence du coup correspondant. Ce silence, l'absence d'un son qu'on attendait « visuellement », est tel un événement musical, et doit donc être inséré dans la palette du compositeur.

Une autre idée intéressante a émergé de nos expérimentations : dans le cas d'un réseau interactif inter-médial audiovisuel, où l'interaction est donnée soit depuis un mapping des données, soit depuis une analyse physique, il peut arriver qu'un médium puisse renforcer l'autre. Si on produit un son qui, pour des questions

<sup>15</sup> Il est possible de voir la pièce à l'adresse : <http://www.alessiosantini.eu/pags/media.htm>

physiques n'est pas évident à l'oreille (tel qu'un mouvement ondulatoire de très basse fréquence), on peut en avoir, par interaction, une représentation vidéo évidente. Cette représentation visuelle met souvent en évidence à l'oreille ce qui ne l'était pas avant. Tout se passe comme si la vidéo augmentait le son, ou mieux : étendait son pouvoir expressif.

## 7. CONCLUSIONS

Dans cet article, nous avons présenté les prémisses d'un travail autour de l'objet audiovisuel. Nous avons souligné certains aspects de la notion de convergence, l'importance de l'approche granulaire pour une technique de micro-fragmentation temporelle, la question du matériau audiovisuel, les différences entre objet et unité. Nous avons présenté notre approche logique et compositionnelle, illustrée par un exemple de production artistique et son bilan.

Pour synthétiser, il apparaît que l'objet logiciel audiovisuel concerne les opérations possibles dans un environnement de programmation en perpétuelle extension, pour la production d'objets audiovisuels composés, autrement dit, la production d'unités audiovisuelles composées. Ces unités audiovisuelles minimales sont, pour ce qui nous concerne en cohérence avec la notion d'unité sémiotique visuelle. Ces unités audiovisuelles peuvent donner lieu à des traitements multiples.

À la suite de cette première étape de recherche, il apparaît indispensable d'approfondir nos recherches en direction de :

- un examen psycho-physique et cognitif des échelles perceptuelles intersensorielles du temps audiovisuel
- l'acquisition d'une connaissance approfondie des caractéristiques techniques et formelles de la vidéo
- les développements de plusieurs briques logicielles à intégrer dans notre environnement modulaire en construction
- la composition de pièces audiovisuelles à proposer au public et à analyser
- la mise à point d'un patch qui puisse nous permettre de composer du matériau audiovisuel dans un contexte de performance.

## RÉFÉRENCES

[1] Budón, O., "Composing with Objects, Networks, and Time Scales: An Interview with Horacio Vaggione", *Computer Music Journal*, Vol 24, N 3, The MIT Press, 2000

[2] Caires, C., "Vers une écriture des processus de micromontage", *JIM* 2003

[3] Collins N., Oloffson F., "Klippav : Live Algorithmic Splicing and Audiovisual Event Capture", *Computer Music Journal*, Vol 30 N 2, The MIT Press, 2006

[4] Courribet, B. "Compositing audio-visual art : the issue of time and the concept of transduction", actes des SMC 2006, Marseille

[5] Didier, G., "Esthétique de la Sonorité - quelques prémisses théoriques et quelques propositions pratiques", Conférence du 11 Janvier 2008 à l'université de Paris VIII

[6] Didier, G., *Esthétique de la sonorité*, ouvrage à paraître chez L'Harmattan, Paris

[7] Eco, U., *Trattato di semiotica generale*, Bompiani, 1975

[8] Momeni A., Henry C., "Dynamic Independent Mapping Layers for Concurrent Control of Audio and Video Synthesis", *Computer Music Journal*, Vol 30 N 1, The MIT Press, 2006

[9] Roads, C., *Microsound*, MIT Press, Camb, MA, 2001

[10] Schaeffer, P., *Traité des objets musicaux*, Seuil, Paris, 1966

[11] Sedes, A., et al., "Groupe de travail Visualisation du Son", *JIM* 2005

[12] Vaggione, H., "Notes sur Atem", *Espaces composites : essais sur la musique et la pensée musicale d'Horacio Vaggione*, L'Harmattan, Paris, 2007

[13] Viola, B., *Anthem*, videotape NTSC, 11:30 min, color, sound, 1983

[14] Yeo W.S., Berger, J., "Application of image sonification methods to music", *Proceedings of the 2005 International Music Conference. International Computer Music Association, Barcelona, Spain, 2005*

