



HAL
open science

Modèles physiques, images et sons : processus de synchronisation et de simulation

Jean-Loup Florens, Annie Luciani, Claude Cadoz, Benoît Chanclou, Arash Habibi

► **To cite this version:**

Jean-Loup Florens, Annie Luciani, Claude Cadoz, Benoît Chanclou, Arash Habibi. Modèles physiques, images et sons : processus de synchronisation et de simulation. 4ème Journées de l'Association Française d'Informatique Graphique, Nov 1996, Dijon, France. pp.80-93. hal-00491706

HAL Id: hal-00491706

<https://hal.science/hal-00491706>

Submitted on 8 Apr 2014

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.

MODELES PHYSIQUES IMAGES ET SONS

Processus de synchronisation et de simulation

Jean Loup FLORENS, Annie LUCIANI, Claude CADOZ
Benoit CHANCLOU, Arash HABIBI

ACROE et CLIPS - INPG - Grenoble
florens@imag.fr

RESUME

L'article présente des méthodes de synthèse d'images et de sons ainsi que leur synchronisation qui permettront de l'ordinateur de se situer comme outil à part entière dans la création audio-visuelle. Sur le plan théorique, il énonce les problèmes fondamentaux soulevés par la question de la simulation d'objets multisensoriels et sur le plan technique, il envisage diverses méthodes de mise en correspondance temporelle

MOTS-CLES

Modèles physiques, synthèse du mouvement, synthèse de l'image, synthèse de son, multimédia, multisensoriel, synchronisation

I. Exposé du problème

Certains objets sont essentiellement producteurs de sons. Ainsi des objets quasi rigides, à déformations visibles faibles et à déformations à bande passante acoustique. D'autres objets sont essentiellement producteurs de déformations visibles et non audibles, tels par exemple un coussin ou un bloc de pâte à modeler. D'autres enfin combinent les deux aspects : un objet rigide qui tombe et qui se brise par exemple.

Il existe en effet dans la nature un nombre non négligeable d'objets qui sont simultanément producteurs de sons et d'images, c'est à dire qui sont à la fois audibles et visibles. Il s'agit en fait de la majorité des objets de notre environnement. C'est pourquoi, à côté de l'illustration sonore d'images (bande son rajoutée sur des images, ou une séquence d'images) ou de l'illustration visuelle de son, le cinéma a développé tout le domaine de la prise de son en même temps que celui de la prise d'image et que le moindre de nos enregistreurs est audio-visuel (figure 1).

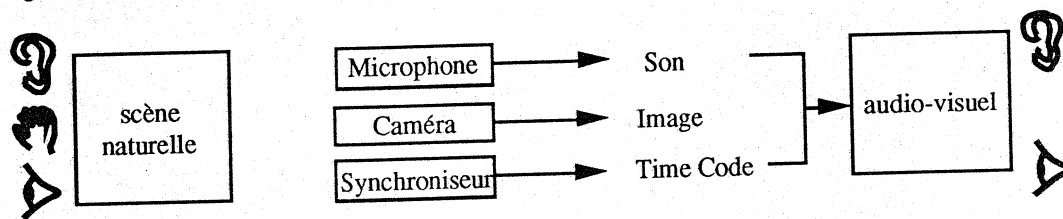


Figure 1. La situation de post-production et de post-synchronisation en audio - visuel traditionnel

En Informatique, le domaine qui jouxte celui de l'audio-visuel, est aujourd'hui celui du multimédia. En ce qui concerne la création et la production audio-visuelles (figure 2), il s'agit là essentiellement de l'activité appelée de post-production ou de post-synchronisation. On suppose que les images et les sons existent et il s'agit de les faire fonctionner ensemble.

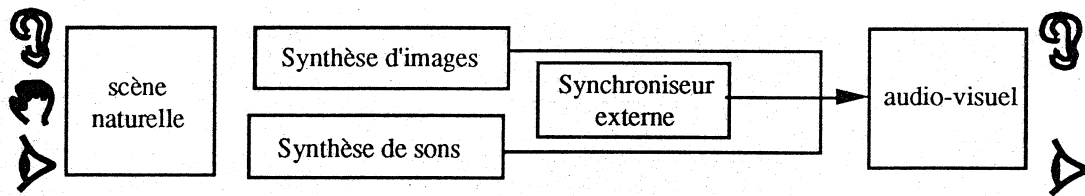


Figure 2. La situation de la post-production et de la post synchronisation en synthèse de sons et d'images

La question que nous souhaitons poser est une question amont de celles de la post-production ou de la post-synchronisation.

L'enregistrement d'images a été remplacé avec l'ordinateur par la fabrication intégrale de l'image. On peut aller aujourd'hui jusqu'à la création d'acteurs humains synthétiques qui marchent, dansent, manipulent des objets de leur environnement virtuel et même communiquent entre eux par des postures et des gestes. L'enregistrement des sons a également été remplacé par de la synthèse de sons, qu'il s'agisse de synthèse de la parole ou de synthèse de sons d'objets naturels, de bruits divers, voire même de sons d'animaux.

Au delà du multi - média, la question est donc celle de la synthèse d'objets complets sonores et visuels, c'est à dire non pas seulement d'une informatique graphique, ou d'une informatique sonore (musicale, parlée ...), mais d'une informatique "audio-visuelle".

La question a plusieurs facettes. Les deux facettes les plus fondamentales dans lesquelles il n' y a encore pratiquement pas de travaux sont:

- Celle qui prendrait en compte la synthèse d'un objet réellement multi-sensoriel (figure 3).

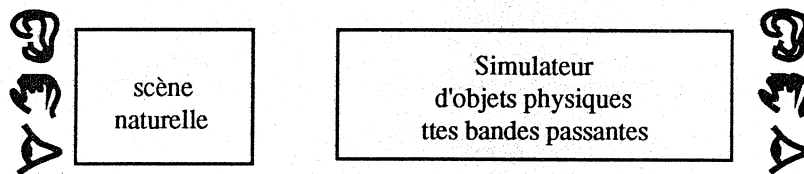


Figure 3. Simulation d'objets multi-sensoriels

- Celle qui tenterait d'avoir accès, en amont de l'image ou du son, à des entités plus abstraites que les signaux sonores ou visuels et qui représenteraient ce qui les lie. Je prendrai un exemple : lorsque l'on émet un signal vocal, on sait aujourd'hui que la vision du visage parlant est d'une grande aide pour la compréhension de la parole [1], que des paradoxes peuvent naître d'une corrélation arbitraire entre l'audition et la vision tel l'effet McGurk [2], et même au delà de cela, que la vision du parleur porte des propriétés d'anticipation du message. Autrement dit, de manière plus générale, l'identification de la parole se fait par deux modalités corrélées : la vue et l'ouïe. Mais, lors de la production de la parole, la juxtaposition entre les deux ne se fait évidemment pas au niveau de la parole émise et de l'image émise, mais au niveau du "programme moteur" qui contrôle simultanément les muscles commandant les cordes vocales, les muscles commandant la déformation des différentes cavités résonnantes et les muscles faciaux qui leur correspondent (figure 4). Il y a donc une "information de commande" très amont de la parole elle-même ou de l'image elle-même (figure 5). Cette information de commande est - elle extractible de l'image, de la parole, ou bien des deux? Si oui, on peut se poser la question de sa synthèse pour piloter par exemple deux systèmes de production séparés de parole et d'image.

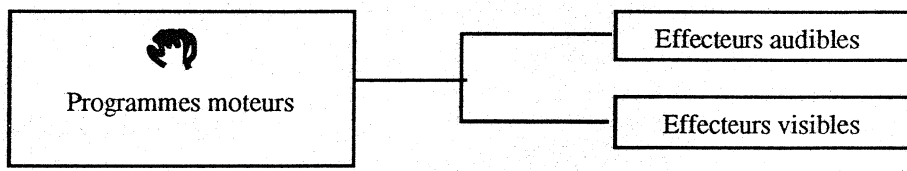


Figure 4. La notion de programme moteur

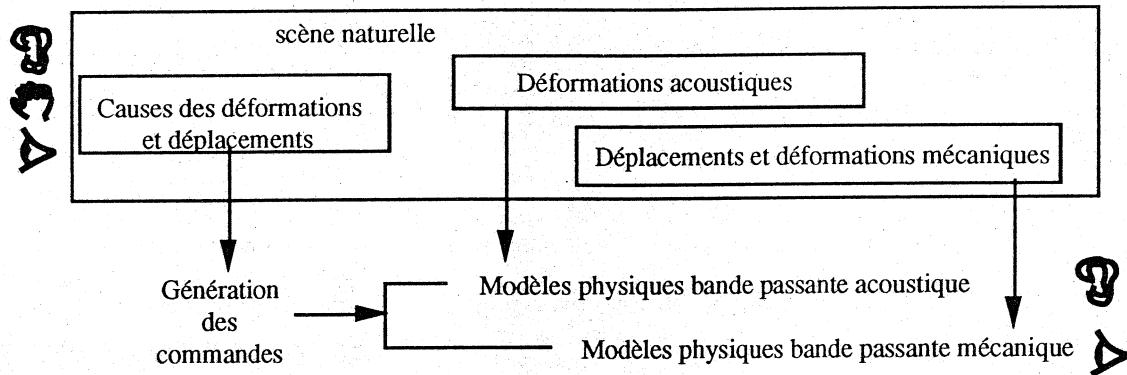


Figure 5. Structure d'un système en causes et effets

Entre les deux, toute une panoplie de modalités d'extraction d'information de l'un ou l'autre médium pour la mise en correspondance temporelle de l'un par l'autre peut être déclinée.

Dans ce travail, nous passerons en revue quelques uns de nos travaux initiaux sur cette question dans le contexte de la synthèse par ordinateur d'images et de sons synthétiques. Notre démarche est dans un premier temps une démarche extensive, en ce sens que nous avons développé un exemplaire de chaque cas possible, sans le creuser trop profondément de manière à voir l'appartenance des problèmes par catégorie et à avoir une analyse globale. Par la suite, chacune des branches pourra être travaillée pour elle-même dans une démarche plus intensive.

II. Les différentes modalités de mise en correspondance temporelle

II.1. Communication par "fichiers et modèles physiques séparés"

II.1.1. Les conditions de cette communication

L'objet naturel, qui s'adresse naturellement à nos gestes, nos yeux et nos oreilles, est un objet physique. Celui-ci, en tant que tel, est le lieu de circulation de forces et de déformations. Celles-ci peuvent se situer dans différentes zones de phénomènes : phénomènes de déformations mécaniques tactiles, mécaniques visibles, mécaniques audibles, optiques visibles...etc. Nous ne considérerons ici que des déformations mécaniques tactiles, visibles ou audibles, sans considérer les déformations optiques visibles.

Ces phénomènes mécaniques se distinguent par des zones différentes dans l'espace des variables d'état du système mécanique. Ainsi, les phénomènes mécaniques sonores sont des déformations de petites amplitudes, plutôt unidimensionnelles, plutôt centrées autour de zéro et à bande passante de l'ordre de 40 KHz. Les phénomènes mécaniques tactiles et visibles sont à bande passante de l'ordre de 1 KHz, plutôt tri-dimensionnelle, les premiers de faible amplitude de déplacement, les seconds à amplitude de déplacement plus importante. Dans les phénomènes tactiles, il faut encore distinguer les phénomènes mettant en jeu des forces faibles (tactiles) et des forces plus importantes (proprio-kinesthésique). Le propos dans ce papier n'est pas de dresser la liste exhaustive des caractéristiques de chacune des zones de déformations mécaniques des objets (ces informations se trouvent dans la littérature spécialisée sur chacun des domaines (acoustique, mécanique...)) mais d'une part de rappeler dans certaines conditions et que dans certaines conditions elles sont séparables et plutôt disjointes.

Ainsi, en première approximation, si l'on peut dire que les déformations audibles et tactiles sont à grand déplacements et à bande passante de l'ordre de 1KHz et les déformations audibles à petits déplacements et à bande passante de l'ordre de 40 KHz, un système mécanique vu sous l'angle de ces déformations mécaniques aura deux zones de fonctionnement, audibles d'une part et visibles et tactile de l'autre. On dira qu'il possède deux échelles différentes. Une analyse détaillée de la modélisation multi-échelles des systèmes dynamiques est effectuée par Habibi dans [3].

Remarque : Cette structuration des signaux produits par un objet est souhaitable pour sa modélisation car un phénomène représenté par un signal 3D contiendrait quelques Gigaoctets d'information. Et le processus générateur qui le générerait artificiellement devrait traiter de l'ordre du Teraoctet d'information par seconde.

Dans ce cas, dans quelles conditions, ces deux phénomènes disjoints seront-ils modélisables par un couplage de deux modèles différents, le premier spécialisé dans la restitution d'un type de phénomène et le second dans la restitution de l'autre type de phénomènes (figure 6)?

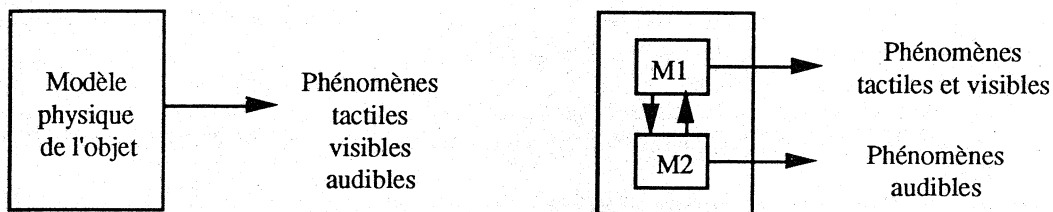


Figure 6 Dissociation du modèle physique de l'objet en deux sous-parties couplées

Nous n'entrerons pas dans le détail de cette discussion, qui est difficile et qui est la question essentielle de la validité d'une modélisation physique d'un système mécanique. Par contre, supposons ici que cela soit possible, alors dans quelle mesure pouvons-nous passer du modèle décrit dans la partie a de la figure 7 au modèle décrit dans la partie b de cette figure?

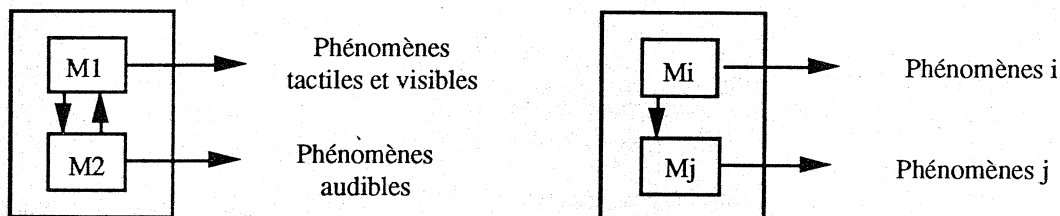


Figure 7. Découplage de sous-parties dans un modèle physique

Il suffit pour cela que l'énergie transmise de la partie 2 du modèle vers la partie 1 soit négligeable, autrement dit que l'influence de la partie 2 sur la partie 1 soit négligeable. Par exemple, que l'énergie en jeu dans le système à audible à haute fréquence se transmette peu à la partie à basse fréquence. Un raisonnement sur l'évaluation de cette condition ainsi qu'une étude détaillée des conditions mathématiques pour que cette condition soit vérifiée est effectuée par Habibi [3] à l'aide du formalisme des fonctions de transferts d'un système. Cette condition n'est pas toujours satisfaite. Il s'agit là d'une approximation dont il faudra toujours vérifier la validité. Par exemple, ce n'est évidemment pas le cas lorsqu'une vitre se brise sous l'influence d'un son. Ce n'est pas non plus le cas lorsque l'on fabrique des arrangements compacts de matériaux granulaires (cailloux, farines, poudres) par mise en vibration du système (fluidisation du système granulaire).

Si cette condition est vérifiée, alors il est possible de calculer les deux parties du modèle indépendamment l'une de l'autre, l'une servant d'entrée à l'autre (figure 8). Dans ce cas, la communication par fichier de variables d'état est correcte.

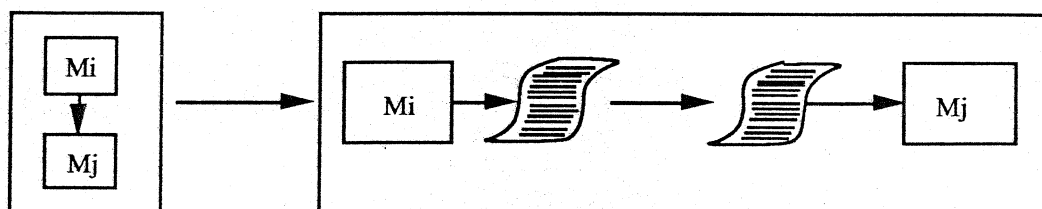


Figure 8 Condition de la communication par fichiers entre deux sous-parties découplées d'un modèle physique

Lorsque les conditions ci-dessus sont vérifiées, nous pouvons alors avoir trois cas :

- (1) contrôle du son à partir de mouvements dans l'image
- (2) contrôle de l'image à partir de mouvements dans le son
- (3) contrôle du son et de l'image à partir d'un processus physique amont ni visuel, ni auditif

A noter qu'il ne s'agit pas de post-synchronisation mais d'un système de commande dans lequel l'une des parties du modèle complet de l'objet sert de commande à d'autres.

II.1.2. Capteurs d'évènements / capteurs d'échantillons

Les fichiers de commande sont donc des sorties d'un autre modèle physique et des entrées d'un modèle physique. Il ne peut donc s'agir que de signaux temporels du type "variable d'état (fonction du temps)".

Il faut donc définir des entités que l'on pourrait appeler "capteurs" ou "détecteurs" ou observateurs" qui placés dans un modèle en cours de simulation (ou des observateurs), sans l'influencer, produiront ces signaux temporels de forces, de positions, de vitesse.

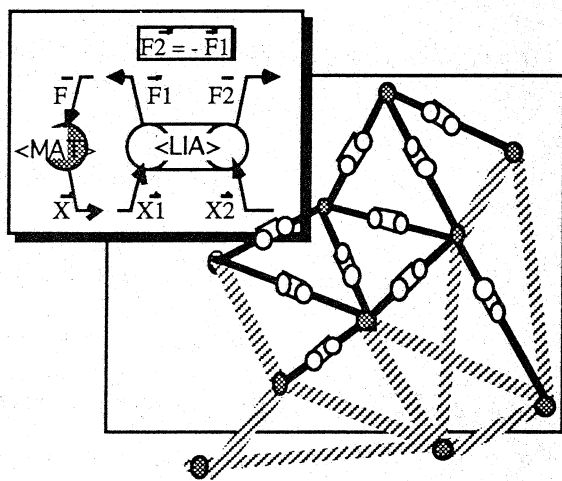
Ces détecteurs sont de deux types :

- * détecteurs de signaux : ils produisent des signaux temporels échantillonnés
- * détecteurs d'événements datés : ils produisent des informations logiques indiquant l'apparition d'un événement dans la scène physique à un certain instant. Ces variables logiques sont ensuite utilisées pour générer des signaux temporels. Un cas particulier mais très coramment utilisé aujourd'hui est le standard MIDI .

a. Détecteurs de signaux:

Les détecteurs les plus simples donnent une image échantillonnée des variables d'état directes de la simulation. Ils peuvent être placés sur n'importe quel module Cordis-Anima.

A partir du formalisme CORDIS - ANIMA [4] rappelé à la figure 9, on peut utiliser n'importe quel élément du formalisme comme capteur.



Un réseau CORDIS (en bas à gauche)
modélise tout objet physique
par un assemblage d'éléments MAT et LIA
(en haut à droite)
dont le format est donné ci-contre.

les X sont des positions
Les F sont des forces

L'implémentation standard est celle
qui associe l'inertie au MAT
et l'interaction entre masses aux LIA

Figure 9. Les principes formels du système de modélisation d'objets physiques CORDIS - ANIMA

Nous avons créé, en CORDIS - ANIMA un ensemble de capteurs de distance (figure 10a), de forces (figure 10b et figure 11b), de positions (figure 11a), de vitesses. Les signaux qu'ils produisent sont échantillonnés à la fréquence "geste", soit 1050 Hz. Ils sont extraits soit directement lorsque celle-ci

tourne à cette fréquence soit par sous- ou sur- échantillonnage. Ces détecteurs fournissent des données homogènes à des données gestuelles qui peuvent de ce fait être utilisés comme signaux d'entrée par tout modèle Cordis - Anima.

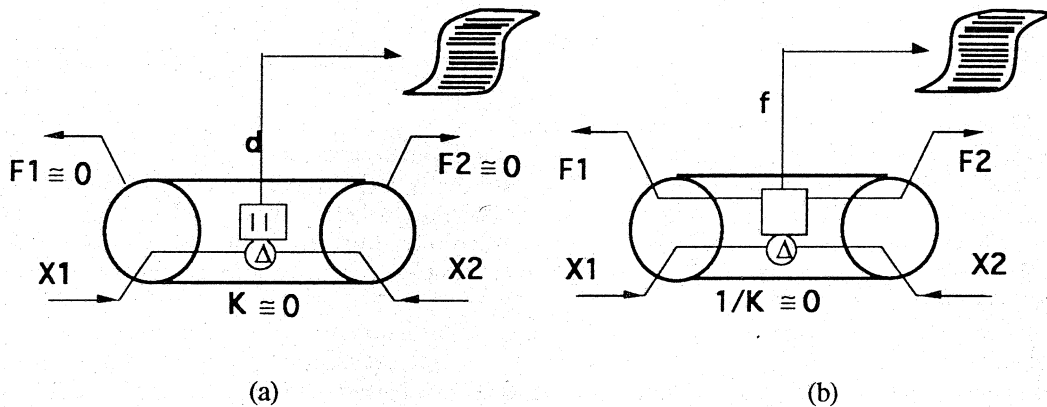


Figure 10. (a) Capteur de distance constitué à partir d'un élément de liaison élastique de raideur faible (b) capteur de force centrale constitué à partir d'un élément de liaison à raideur et viscosité fortes. (b).

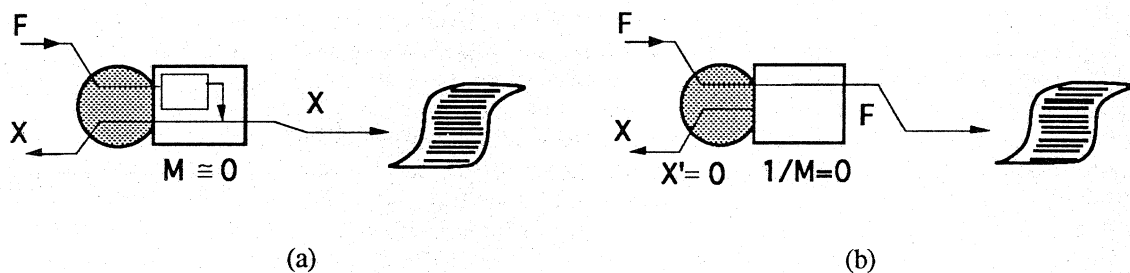


Figure 11. (a) Capteur de position absolue constitué à partir d'un élément masse de faible inertie (b) capteur de force absolue constitué à partir d'un élément masse d'inertie infinie (b).

D'autres détecteurs produisent des variables dérivées des variables directes. Nous avons ainsi également implémenté un "détecteur de pression dans une zone" par sommation des forces s'exerçant sur les particules lorsqu'elles sont dans cette zone.

D'autres capteurs de fonctionnalités plus complexes, permettant de faire de l'analyse spectrale ou modale de structure ont été implémentés, toujours dans la philosophie du modèle physique [5 et 6].

Ces détecteurs sont alors un cas particulier de générateurs de gestes pour le modèle aval, qu'il s'agisse de modèles "son" ou de modèles "image". La philosophie du modèle physique avec laquelle ils sont implémentés est fondée sur des copies de véritables instruments matériels dont on peut contrôler le comportement intrinsèque. Par exemple, lorsque l'on utilise ces capteurs pour capter le son produit dans la structure, on peut modéliser dans le capteur, les caractéristiques physiques d'un microphone par exemple.

b. Détecteurs d'événements

Les détecteurs d'événements sont des détecteurs de variables d'état complétés par un ou plusieurs seuils, les seuils étant des variables accessibles à l'opérateur. Placés sur un module Cordis, ils détectent des changements d'état. Ainsi par exemple, placés sur un module d'interaction, ils détectent s'il y a couplage ou non entre deux masses. Cette détection dépend du (ou des seuils). Par exemple, dans le cas d'une butée visco-élastique et d'un seuil sur la distance propre à ce module, si le seuil est très petit, le détecteur détectera les micro-collisions lors du "contact". L'on peut aussi détecter les collisions sur une plage de distance. Ceci permet de régler le grain de l'événement. Les détecteurs d'événements génèrent

automatiquement un fichier appelé "nom_simulation.DETECT". Dans ce fichier, se trouvent regrouper les événements datés en provenance de un ou plusieurs détecteurs placés sur la simulation. La fréquence permettant le datage est la plus basse fréquence du processus, c'est à dire c'est la fréquence image. Les événements sont donc datés en "n° image" à 25 (ou 50) images/seconde.

Des fichiers d'événements de variables dérivées des variables directes (forces, positions, vitesses) ont également été réalisés, en particulier des détecteurs seuillés de nombre de particules traversant une zone dans une tranche de temps de 1 image (1/25 seconde).

II.1.3. Quelques exemples

a. Modèle dit de la "Harpe Eolienne"

Dans l'exemple dit de la "harpe éolienne", deux modèles physiques séparés ont été créés:

- le modèle pour l'image : C'est un modèle CORDIS - ANIMA de fluides turbulents 2D [7].
- le modèle pour le son : c'est un modèle CORDIS - ANIMA dit de la "poutre frottée" [8, 9]. Une structure vibrante mécanique linéique est frottée par une interaction de frottement sec dépendant simultanément de la vitesse relative de l'organe frottant (l'archet) et de l'organe frottée (la poutre) et de la pression du premier sur le second.

Dans le modèle pour l'image, ont été placés des compteurs de particules sur une zone en certains endroits de la scène spatiale 2D. Durant la simulation du modèle image, ces capteurs fournissent des signaux donnant une représentation approchée d'une pression discrète en ces points à chaque instant (à 1050 Hz).

Dans un deuxième temps, ces signaux sont utilisés pour remplacer l'archet virtuel.

La difficulté consiste à arriver à obtenir des signaux détectés dans la première scène compatible en forme et en intensité à ceux nécessaires pour faire vibrer comme on le veut la poutre frottée. Ce travail de calibrage est d'autant plus complexe que le phénomène commandé est complexe. Ce qui est le cas ici avec le frottement de la poutre.

Les résultats (audibles sur bande vidéo) sont cependant très intéressants. Nous sommes arrivés à "faire chanter la poutre" de manière synchrone avec les turbulences du vent.

Ceci n'est en fait qu'un premier essai sur un exemple de mise au point très difficile.

Des travaux ultérieurs ont pu obtenir beaucoup plus facilement une situation identique en remplaçant le frottement d'archet directement par des micro percussions multiples, sans chercher à recréer le signal de pression.

b. Modèle dit de la "Pluie"

Le modèle de la pluie relève du troisième cas décrit dans le paragraphe II.1.1 : "contrôle du son et de l'image par un processus physique amont, ni visuel, ni auditif". Nous avons donc trois modèles différents, l'un appelé "Modèle dynamique pour la structure événementielle" sera le modèle amont aux deux autres modèles indépendants "Modèle pour le son" et "Modèle pour l'image". Les trois modèles se calculent indépendamment, le premier nécessairement avant les deux autres.

Modèle dynamique pour la structure événementielle

Un premier modèle CORDIS - ANIMA représentant un ensemble de particules tombant en chute libre sur une ligne 1D visco-élastique entre des parois mobiles 1D va générer un fichier d'impacts datés (figure 12). Nous avons ainsi généré une organisation physique temporelle des impacts (Physique puisque la séquence temporelle est produite par un mouvement de chute avec rebonds sur des obstacles physiques). La densité des impacts est un critère important dans ce type de phénomènes. Elle est contrôlée par un modèle physique de raréfaction et de condensation. Pour cela, les particules tombent dans un goulet plus ou moins étroit, composé de parois mobiles et du fait de leur mouvement brownien émergent dû à leurs interactions de collisions, des particules peuvent être éjectées de la scène. La position des obstacles du goulet, la position et la vitesse initiales des particules ainsi que leur nombre au départ contrôlent l'organisation temporelle des impacts ainsi que l'évolution macroscopique telles qu'une raréfaction par exemple.

Ce processus génère un fichier d'impacts datés.

Deux versions ont été réalisées :

- (v1) fichier d'événements logiques: instant d'impact, identité de la particule percutante, identité de la particule percutée (point du sol),
- (v2) fichier d'événements physiques : instant d'impact, identité de la particule percutante, identité de la particule percutée (point du sol), + force (et /ou vitesse) au moment de l'impact

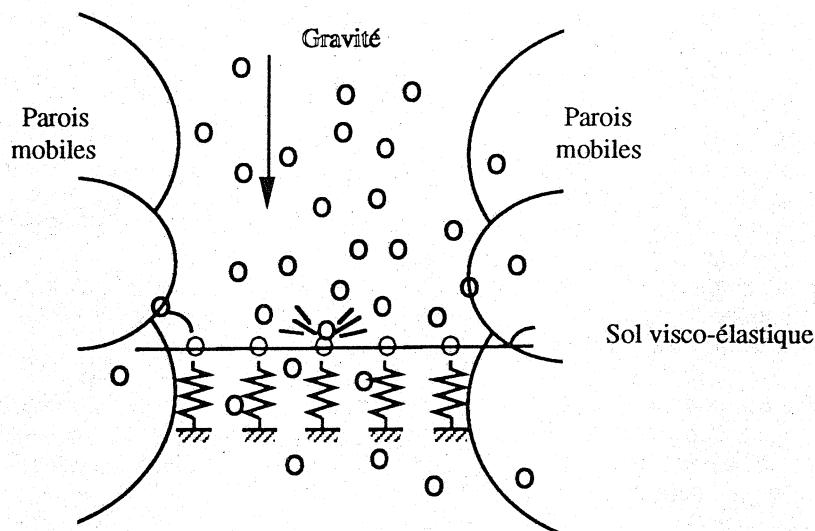


Figure 12. Modèle physique d'un générateur d'impacts physique

Modèle pour le son

Les impacts commandent dans un deuxième temps l'excitateur d'une structure vibrante. La structure vibrante est composée d'un ensemble d'oscillateurs élémentaires 1D masse - ressort - frottement représentant des modes d'une structure vibrante dont on maîtrise donc très bien les propriétés acoustiques. L'excitateur est :

- soit un ensemble de masses identifiées lâchées d'une certaine hauteur initiale constante (Version v1)
- soit un ensemble de masses identifiées lâchées soit avec une hauteur ou une vitesse initiale variable en fonction de la force d'impact captée. (version v2)
- soit directement par un signal de tupe "impulsion de force" à l'instant de l'impact (version v2)

Modèle pour l'image

Le modèle pour l'image est une surface déformable sur laquelle les particules tombent sous le contrôle du fichier d'impacts. En terme de contrôle, il est donc similaire au modèle pour le son. Ces deux modèles diffèrent par les points suivants :

- les oscillateurs pour le son sont des oscillateurs 1D et la surface déformable pour l'image est une surface à 2 degrés de liberté de propagation et à 1 degré de liberté de déformation
- le nombre de points de la struture sonore est égale aux nombre de modes de déformations acoustique sde celle - ci. Il est de l'ordre de la dizaine au maximum. Au contraire, le nombre de points de l'objet déformable pour l'image est égal au nombre de pixels voulus soit pour une excellente image de propagation de "ronds dans l'eau" 1240 * 1080 masses ponctuelles toutes reliées par des visco-élasticités selon un schéma de grille 2D1/2.

On modélise ainsi le comportement dynamique d'une surface d'eau. Si l'on affecte à chaque point de cette surface une loi d'éclairage, on obtient simultanément une méthode de visualisation de cette surface à la résolution du pixel. Les lois d'éclairage peuvent être de plusieurs type, par exemple :

Intensité du Pixel = a * enfoncement de la masse correspondante

Intensité du pixel = k * Dz pixels voisins

Ainsi, la particule *i* du générateur d'impact produit à la fois un son et à la fois une onde visible sur la surface de l'eau. L'effet est d'un grand réalisme. Il permet aussi d'aller vers "hyper-réalisme". Par exemple, la particule *i* du générateur d'impact est associé à un oscillateur sonore toujours le même et qu'elle tombe toujours au même endroit de la surface liquide visuelle, il produit le fait que un son particulier est associé à une petite zone de l'espace. la surface liquide est homogène visuellement et acoustiquement découpée en petis instruments sonores.

II.2. Modèle dit de la "Granule" : modèles couplés geste - image - son

Nous allons maintenant nous intéresser aux cas pour lesquels la communication par fichier entre deux parties du modèle de l'objet multisensoriel n'est pas possible. Il y en a deux catégories :

- la première catégorie est celle pour laquelle on ne peut découper le modèle de l'objet en deux (resp. plusieurs) sous-parties couplées, l'une produisant un type de phénomène, l'autre un autre type de phénomènes. Dans le paragraphe II.1.1., nous avons pris comme exemple de ce cas la fluidisation de matériaux granulaires par exemple ou la propagation d'une fracture dans un matériau qui sont deux cas de problèmes physiques très complets et très complexes.
- la deuxième catégorie est celle pour laquelle il est possible de structurer le modèle physique de l'objet en plusieurs parties couplées, les unes produisant le son, les autres le mouvement visuel, les autres les mouvements tactiles par exemple.

La première catégorie ne pose des problèmes de modélisation et de connaissance physiques mais elle ne pose pas de problème de structuration du modèle. Le modèle doit être écrit comme un tout, être calculé aux bandes passantes temporelles et spatiales les plus élevées présentes dans l'objet. Elle pose des problèmes de puissance de calcul.

La seconde catégorie permet de diminuer la quantité de calcul. Elle est d'autant plus efficace que les sous-parties de l'objet présentent chacune des simplifications liées à la spécificité du phénomène qu'elles doivent produire. Par exemple, si la partie devant fonctionner à fréquence acoustique peut être considérée comme linéaire et à 1 degré de liberté. Elle se pourra alors se représenter par synthèse modale comme un ensemble d'oscillateurs élémentaires indépendants du second ordre.

Nous nous intéresserons ici à cette deuxième catégorie et pour cela nous allons prendre un exemple significatif de ce cas. Lors d'une percussion d'un objet par un autre manipulé à la main, la dynamique du choc est contrôlé par l'opérateur et dépend étroitement de la perception que celui a de la collision : avant et après le choc du poids de l'objet qu'il tient et pendant le choc de la matière des deux objets et de la nature de leur couplage, après le choc des nouvelles conditions initiales imposées à l'objet tenu par l'objet percuté. Dans ce cas, il est clair que le système ne peut pas se représenter par deux sous parties indécouplées : d'une part la perception tactile par l'opérateur de l'objet percuté est essentielle et d'autre part, la dynamique de l'action de l'opérateur est également essentielle pour le son produit par l'objet percuté.

Nous allons étudier sur un exemple de ce type dit "modèle de la granule" les différents problèmes liés à l'obtention d'un modèle optimal structuré en sous-parties couplées spécialisées.

II.2. 1. "Granule" : modèle de référence.

L'objet est constitué d'une cavité mobile dans un espace naturel à deux dimensions, et manipulable par un transducteur gestuel à retour d'effort [10]. Cette cavité contient des grains agglomérés et possède des propriétés vibratoires capables de produire du son sous l'effet des impacts des grains. La cavité est constituée d'un certain nombre de particules inertes liées par des éléments élastiques faiblement amortis. L'objet solide mobile à l'intérieur de cette cavité est lui aussi constitué de particules inertes liées par des éléments élastiques mais plus fortement amortis. La délimitation d'une zone fermée et sa forme et donc l'aptitude de la cavité à contenir l'objet solide et l'empêcher de s'en échapper sont obtenues par un réglage adéquat des éléments de liaison liant les particules de la cavité à celle des grains qu'elle doit contenir. Ces éléments de liaison réalisent l'interaction d'exclusion ou "butée". Nous avons choisi la configuration "intérieure" (ou bulle) qui permet de former plus facilement une cavité même à partir d'un très faible nombre de particules. Nous pouvons utiliser par exemple 3 particules comme cela est représenté sur la figure 13. Jusque là le modèle ne diffère pas de ceux déjà réalisés dans le cadre de la synthèse en temps réel d'images en mouvements avec contrôle instrumental [11, 12] 13].

La production du son est obtenue à partir des vibrations de la cavité produites par les impacts. Il est donc nécessaire que les constantes fréquentielles de ces modes d'oscillation se situent dans un spectre acoustique intéressant ce qui nécessite que la fréquence de calcul de l'ensemble du modèle soit de l'ordre d'une dizaine de kHz. Mais cette condition n'est pas suffisante pour une représentation correcte des forces de percussion. Durant la phase de contact, le système se comporte comme une nouvelle cavité dont

les fréquences propres doivent être sensiblement plus élevées que celles de la cavité en oscillation libre, de l'ordre de quelques dizaines de kHz ce qui conduit à des fréquences de calcul du même ordre.

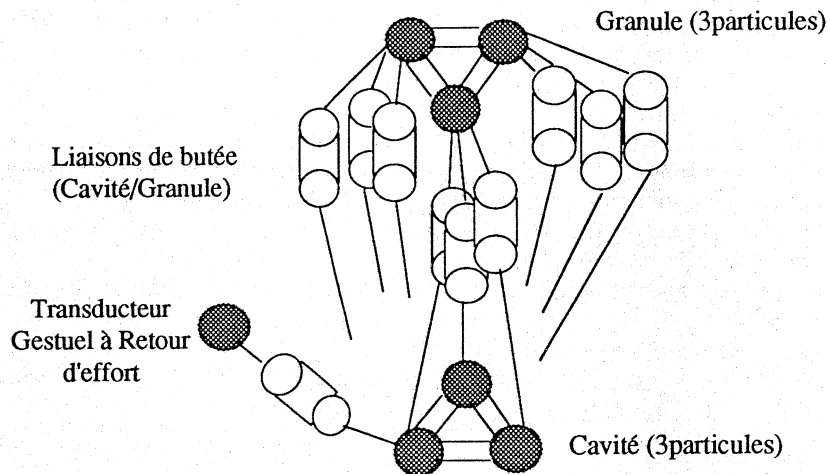


Figure 13 Le modèle CORDIS-ANIMA de référence pour la granule (a).

Les deux objets "cavité" et "granule" sont constitués ici de 3 particules liées par des éléments de liaisons élastiques et amortis (REF) et par des éléments de liaison dont la caractéristique élastique est non linéaire (b) et réalise la fonction de butée intérieure (BUT).

Compte tenu de ceci certaines difficultés apparaissent alors :

- La première concerne la précision des calculs. Le mouvement des particules de la cavité est la superposition du mouvement moyen de celle-ci et des mouvements vibratoires. Or les rapports d'amplitudes entre ces deux mouvements peuvent atteindre l'ordre de 10^9 pour des configurations de paramètres physiques usuelles et pertinentes quant au rendu acoustique. La représentation en format flottant à 28 bits de mantisse dont nous disposons, est alors insuffisante.

- Une autre difficulté est liée à la capture des mouvements vibratoires pour la production du signal sonore. Dans les modèles d'objets sonores plus classiques [14,15,16], dans lesquels la structure vibrante n'est pas soumise à des mouvements lents de grande amplitude, on utilise un module d'observation qui capte le mouvement d'un ou plusieurs points du modèle et transmet le signal au périphérique de sortie sonore après conversion en format fixe de 16bits. Cette dernière contrainte rend le module inutilisable dans le cas présent, toujours en raison de la composante lente de grande amplitude.

- Au delà de cette question de calcul, subsiste un problème d'indétermination du modèle quant à la relation exacte liant les mouvements vibratoires, bidimensionnels, de la cavité et le signal acoustique unidimensionnel qui en résulte. La projection cartésienne, consistant à n'utiliser qu'une des directions de vibration a été employée dans des simulations en temps différé ([17] et n'est pas toujours satisfaisante puisque certaines directions d'impact sont "muettes" et cette difficulté ne se résout pas par combinaison linéaire des composantes de la vibration.

- Enfin, bien que le modèle présenté soit très simple, sa réalisation à 40kHz nécessite une puissance de calcul très importante. Nous l'avons évaluée à environ 300 Mflops pour une cavité à 15 masses et une granule à 3 masses. La complexité est essentiellement due aux éléments de liaison en nombre important et qui comportent des calculs géométriques (distance, direction) de coût élevé sur les machines de structure classique actuelles.

Ces différentes considérations nous ont conduits à transformer profondément ce premier modèle et à aboutir à un modèle constitué de deux parties couplées :

- la partie visuelle comporte des objets à mouvements bidimensionnels (espace naturel) et est simulée à fréquence faible de 1kHz.
- la partie acoustique est à mouvements unidimensionnels et simulée à fréquence élevée de 44kHz.

Nous présentons d'abord la transformation conduisant à la réduction à une dimension de la partie acoustique et qui suppose une même fréquence de simulation. Nous présentons ensuite les transformations supplémentaires nécessaires pour la réalisation de la simulation à deux fréquences.

II.2.2. Transformation du modèle : séparation en partie visuelle et sonore. Modules de projection radiale.

Dans ce nouveau modèle, les éléments de liaison établis entre les masses de la cavité et celles de la granule réalisent, comme dans le cas précédent une interaction d'occupation exclusive de l'espace, mais la sphère d'exclusion peut maintenant être le siège d'oscillations radiales (sphère pulsante) par couplage à une structure vibrante passive et à mouvements unidimensionnels. Ce sont, bien sûr, ces mouvements qui confèrent au modèle ses propriétés acoustiques.

Les relations ci-dessous expriment ce couplage. Nous les comparons à celle de l'interaction classique de butée employée dans le premier modèle.

Interaction de butée élastique:

$$r = \|X_c - X_g\|, \quad u = \frac{X_c - X_g}{r}$$

$$F_g = -F_c = u \cdot \Phi(r)$$

Interaction de butée avec couplage radial

$$r = \|X_c - X_g\|, \quad u = \frac{X_c - X_g}{r}$$

$$F_g = -F_c = u \cdot \Phi(r - x_v)$$

$$f_v = \Phi(r - x_v)$$

X_c, X_g, F_c, F_g représentent respectivement les positions (2D) et les forces d'interaction d'une masse de la cavité et d'une masse de la granule.
 x_v et f_v sont la position et la force au point de couplage à la structure vibrante.

Ceci est une représentation modale de la cavité qui se trouve ainsi décomposée en un mode de mobilité équivalent à une cavité évoluant dans l'espace 2D mais dépourvue de propriétés vibratoires, et en modes vibratoires équivalents à des résonateurs attachés au sol. On constate qu'à la sommation des déplacements correspond une diffusion de la même force sur les deux systèmes (figure 14), ce qui est cohérent avec cette décomposition et confère au modèle sa cohérence énergétique.

Cette nouvelle représentation n'est pas équivalente au modèle initial mais autorise une grande liberté quant à la constitution des propriétés acoustiques de l'objet. Le modèle de la partie vibratoire peut être de topologie quelconque : la question de correspondance entre les propriétés acoustiques des deux modèles n'est pas abordée ici, mais elle ne pose aucun problème de principe. Les difficultés liées à la superposition des mouvements acoustique et visuel n'existent plus dans la simulation de ce nouveau modèle.

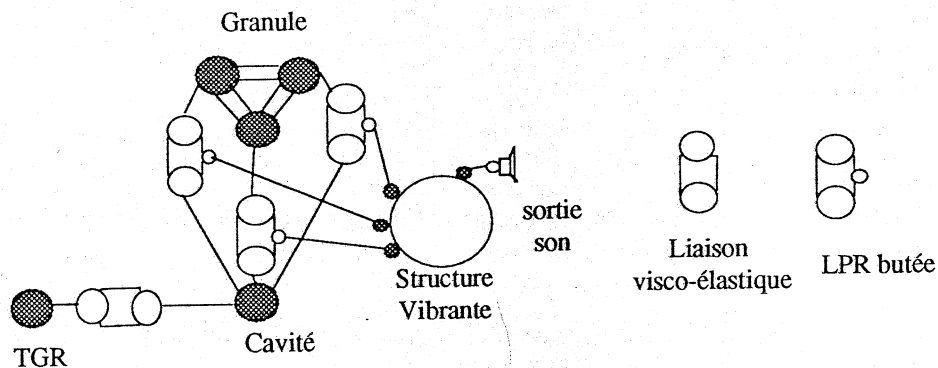


Figure 14. Modèle de la granule à une fréquence avec séparation des parties acoustique. Ce modèle utilise un module à projection radiale de type butée.

Nous avons introduit le couplage radial à partir du module de butée. Cette opération est bien sûr possible à partir de tout type d'élément de liaison du formalisme CORDIS et permet d'engendrer une classe de modules "à projection radiale" qui permettent de coupler deux systèmes définis sur des espaces de mouvements différents. Les relations introduites ci-dessus à propos de la butée se généralisent aisément.

Les propriétés de ces modules sont :

- ils vérifient le principe d'action - réaction et d'invariance par translation et rotation sur chacun de leur bipoint. Le point non re
- ils possèdent les mêmes propriétés énergétiques que le module <LIA> à partir duquel ils sont construits. Leur structure générale est représentée à la figure 15.

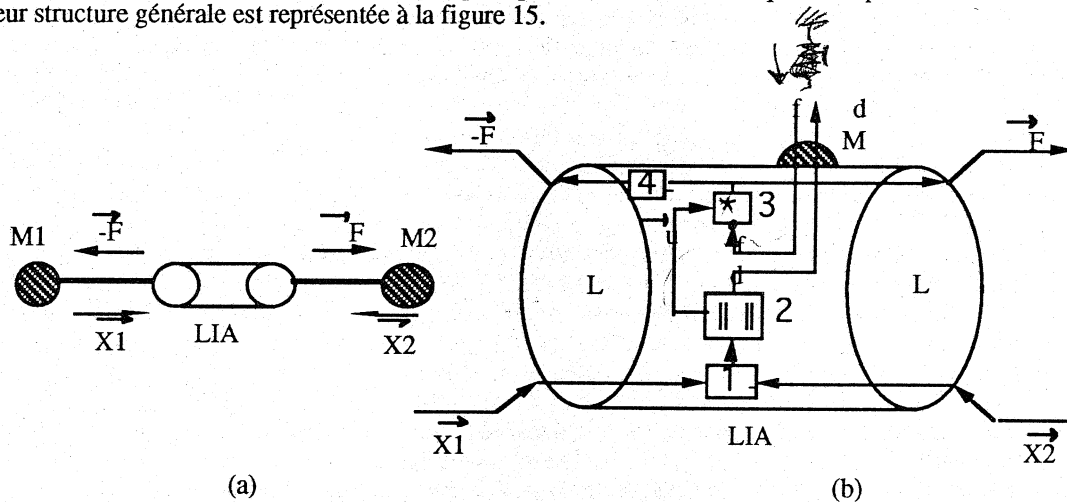


Figure 15. Module de liaison général (a) type <LIA> et module à projection radiale quadripolaire (b). La boîte oblongue appelée LIA représente dans les deux cas un opérateur produisant un signal de force radiale f à partir d'un signal de distance. Cet opérateur peut être linéaire (visco-élasticité) ou non (butée)
Opérateur 1 : soustraction - Opérateur 2 : distance - Opérateur 3 : projection - Opérateur 4 : inversion
(Simulation projection)

Bien que ce nouveau modèle doive, comme le premier, être calculé à la plus haute fréquence de Shanon présente dans les sous-parties de l'objet, c'est à dire celle correspondant à la bande passante acoustique, ce nouveau modèle permet un gain substantiel en puissance de calcul au fur et à mesure que l'on complexifie la partie acoustique, indépendamment de la partie visuelle. Le coût de cette complexification est peu pénalisant puisqu'elle n'affecte que la partie linéaire et unidimensionnelle du modèle.

II.2.3. Modèle à temps discret à 2 fréquences

Une deuxième optimisation du modèle consiste à réduire la fréquence de simulation de tous les éléments à déplacements bidimensionnels c'est à dire le mode de mobilité de cavité et les grains percuteurs, à la valeur de la fréquence d'échantillonnage des signaux gestuels. Le gain, sur la puissance de calcul nécessaire, de cette dernière transformation est important puisque tous les calculs de distance sont maintenant effectués à basse fréquence. La partie visuelle peut alors être traitée dans un environnement similaire à celui déjà existant pour la synthèse d'image en temps réel, c'est à dire le modèleur simulateur ANIMA [4].

La technique de séparation des parties hautes et basses fréquences est semblable à celle employée en synthèse de son par simulation d'instruments à percussion élémentaires que nous décrivons brièvement pour introduire le problème.

Modèle bifréquence d'un instrument à percussion élémentaire.

Le modèle représente grossièrement un élément d'instrument à clavier de type Dulcitone ou Celesta [18]. Il s'agit d'un instrument à percussion dans lequel le percuteur (marteau) est entraîné par une touche de clavier à travers une simple liaison élastique. Une condition importante pour que la percussion fonctionne est que durant la phase de contact entre le percuteur et la structure vibrante, les forces exercées sur le

percuteur par le dispositif d'entraînement, ici la touche et la liaison élastique, soient faibles vis à vis des forces de réaction de la structure vibrante. Cette condition est dans le cas présent obtenue par une valeur adéquate (suffisamment faible) de la raideur de la liaison d'entraînement, les constantes de temps du système "percuteur / liaison d'entraînement" sont alors importantes vis à vis de celles des modes de la structure constituée en phase transitoire de contact par le percuteur et la structure vibrante.

Lorsque cette condition est réalisée, on peut réaliser la simulation en double fréquence dans des conditions acceptables, la partie HF étant constituée du percuteur, de la liaison de butée percuteur/structure vibrante et de la structure vibrante ; la partie BF est constituée de la liaison d'entraînement et évidemment du module TGR (transducteur Gestuel à Retour d'effort).

La transposition directe de ceci au cas de la granule est immédiate mais apporte peu : Les masses de la cavité, de la granule ainsi que les liaisons de butée doivent toujours être calculées à haute fréquence. Il faut alors introduire un nouvel artifice dans le modèle si l'on désire atteindre le but recherché, et calculer toute la partie 2D en basse fréquence.

Le percuteur doit alors être décomposé en deux éléments.

Modèle bifréquence à double percuteur.

La frontière entre les deux zones à haute et basse fréquence se situe alors au voisinage immédiat des structures vibrantes représentant les modes vibratoires. Aux transformations de sur-échantillonnage et sous échantillonnage subies par les signaux au franchissement de cette frontière, il est théoriquement nécessaire d'adjoindre des filtres reconstituteurs et anti-repliement. De tels filtres, outre un coût en calcul élevé, peuvent produire d'importants retards de phase et une déstabilisation de la simulation. Pour résoudre ce problème, nous avons alors introduit un artifice fonctionnel qui s'apparente à un mécanisme à échappement. Cet artifice ainsi que toutes les transformations précédentes respectent les propriétés énergétiques du modèle initial. En particulier les propriétés dynamiques des contacts grains-cavité tels qu'elles apparaissent visuellement dépendent étroitement des paramètres d'élasticité et de viscosité des modes vibratoires. Ceci garantit la haute cohérence des effets sonores et visuels ainsi produits.

III. Conclusion

Nous avons voulu décrire et en partie résoudre les problèmes inhérents à la modélisation et à la simulation physique multisensorielle. Cette modélisation nous semble évidemment un point d'avenir pour les développements du domaine de la représentation des objets par l'ordinateur, plus couramment appelé le domaine des objets virtuels.

D'une manière plus générale, les méthodes abordées ici ne s'appliquent pas seulement à des phénomènes dynamiques séparables en phénomènes acoustiques, tactiles ou visuels. Elles peuvent servir de base à toute optimisation du modèle, en amont de l'optimisation des procédés calculatoires dès lors qu'un modèle valable de l'objet peut être un modèle dit "multi - échelle".

En effet, en terme de modélisation physique, les optimisations conduisant à des gains les plus spectaculaires en terme de gain de calcul tout en restant à bonne résolution spatiale et/ou temporelle, sont les optimisations portant sur le modèle lui même avant les optimisations calculatoires. La question reste alors celle de la préservation de la validité du modèle dans la zone de fonctionnement souhaitée.

IV. Bibliographie

- [1] ADJOUDANI A., GUIARD - MARIGNY T., Le GOFF B., BENOIT C., Un modèle 3D de lèvres parlantes, XX ième Journées d'Etude de la Parole, Trégastel, Juin 94
- [2] [McGURK H. & MacDONALD J., 1976, Hearing lips and seeing voices, Nature, 264, 746-748
- 3] HABIBI A., Visualisation d'objets très déformables : relation entre le mouvement, la forme et l'image, Thèse de doctorat 3ième cycle, INPG, à paraître, Décembre 1996
- [4] CADOZ (C), LUCIANI (A) & FLORENS (JL), "CORDIS-ANIMA : a Modeling and Simulation System for Sound and Image Synthesis- The General Formalism", Computer Music Journal, Vol 17-1 spring 1993, MIT Press
- [5] INCERTI (E), CADOZ(C) - "Time-Frequency Decomposition by means of Physical Model", ISMA (Dourdan, France), July 2-6 1995, pp 458 - 464.

- [6] INCERTI (E), "Synthesis and Analysis Tools with Physical Modelling : an Environment for Musical Sound Production" - Proc. of ICMC'96, Hong-Kong, August 1996, à paraître
- [7] LUCIANI (A), HABIBI (A), VAPILLON (A), DUROC (Y) - "A Physical Model of Turbulent Fluids", 6th Eurographics Workshop on Animation and Simulation- Maastricht 1995 - Springer Verlag Ed. "Computer Science Series - pp16-29 - 17 Sept Aarhus, Denmark. ICMC 94 Proceedings - 1994 pp. 11-18
- [8] FLORENS (JL), "Excitateurs complexes", Rapport interne ACROE
- [9] FLORENS (JL) & CADOZ (C), "Modèles et simulations en temps réel de cordes frottées", 1er Congrès Français d'Acoustique - SFA - Editions de Physique - Lyon Avril 1990.
- [10] LUCIANI (A), CADOZ (C), FLORENS (JL), "The CRM device : a force feedback gestural transducer to real-time computer animation", Displays , Vol. 15 Number 3 - 1994 - Butterworth-Heinemann, Oxford OX2 8DP UK, pp. 149-155.
- [11] LUCIANI (A), JIMENEZ (S), FLORENS (JL), CADOZ (C) & RAOULT (O), "Computational physics : a modeler simulator for animated physical objects", Proceedings of the European Computer Graphics Conference and Exhibition. Vienna, Austria, Septembre 91, Editeur Elsevier
- [12] CHANCLOU (B), LUCIANI (A) - "Physical models and dynamic simulation of planetary motor vehicles with a great number of degrees of freedom"- Proc of IAS 4 Conf. - IOS Press - 1995 - pp 465-472
- [13] JIMENEZ (S), "Modélisation et simulation physique d'objets volumiques déformables complexes", Thèse de l'INPG, Spécialité Informatique, INPG, Grenoble Nov. 1993.
- [14] DJOHARIAN (P), "Generating Models for Modal Synthesis", CMJ Vol 17-1 spring 1993
- [15] FOURCADE (P), CADOZ (C), "Sound Synthesis by Physical Modelling : an elementary Striker", Proc. of Forum Acusticum, Antwerpen, 1996
- [16] INCERTI (E), CADOZ (C) - "Topology, Geometry, Matter of Vibrating Structures Simulated with CORDIS-ANIMA Sound Synthesis Methods"- Proc. of ICMC - Banff (Canada), MIT Press, 1995, pp 96-103
- [17] LUCIANI (A), "Modèles physiques et animation", Rapport de recherche "Lifia 111-Acroe 90-16" - Grenoble 1990, (Exposé Imagina 90)
- [18] BOUASSE H., Verges et plaques, Cloches et carillons, Librairie Albert Blanchard, 1987