

# Projet *PHASE*

## Jouer de la musique avec un bras haptique

Jean-Philippe Lambert <lambert@ircam.fr>

### 1 Présentation du projet

*PHASE* était un projet soutenu par le réseau *RIAM* regroupant quatre partenaires associés dans la réalisation d'un dispositif d'interactions multimodales en vue de manipuler du son et de la musique : le CEA-List a effectué la recherche haptique, Haption a réalisé l'interface haptique, Ondim s'est occupé de l'intégration et de la réalisation visuelle et l'Ircam a effectué la recherche concernant les manipulations sonores et musicales et a défini les métaphores d'interaction. Le CNAM-CEDRIC a été associé à la scénarisation du jeu et à son ergonomie [GLPNV02].

Le projet *PHASE* s'est déroulé d'octobre 2003 à mi-décembre 2004. Durant les trois derniers mois, un démonstrateur a été présenté au grand public : celui-ci était installé dans la galerie des enfants du centre Pompidou, à l'occasion de l'exposition « Écoute », sous la forme d'une installation interactive offrant au grand public un jeu musical. À la différence d'un jeu vidéo, le but n'était pas d'animer des pixels sur un écran mais bien de jouer avec de la musique et ainsi susciter un éveil musical. De ce point de vue, les enfants étaient visés par le dispositif, celui-ci pouvant avoir une utilisation pédagogique [Vei01].

Le démonstrateur n'était pas l'unique but de *PHASE*, puisqu'une part importante du projet a concerné l'expérimentation et la recherche à propos de l'utilisation du geste avec des retours haptique, visuel et sonore, pour jouer de la musique. La réalisation de celui-ci a cependant permis de valider certains résultats sur un large public.

En tenant compte des différents métiers impliqués et vu les difficultés techniques, le travail de recherche et de réalisation a été décomposé en différents modules, chacun devant tenir compte de deux préoccupations majeures : quel contrôle est accessible à l'utilisateur par son geste ? Quel jeu musical est rendu possible ?

Différentes approches mettant en relation les différentes modalités ont donné lieu à différentes réalisations. Des nombreux exemples interactifs sont disponibles, sous la forme de maquettes démonstratives, plus ou moins élaborées.

## 2 Geste et retour haptique

Le geste étant le point d'entrée du système multimodal envisagé, il a été étudié particulièrement, en tenant compte du contexte spécifique du contrôle musical [CW00] avec un retour haptique [CLF90], pour lesquels des champs d'études existent. Dans notre cas, le but n'était pas nécessairement de réaliser un instrument ni d'utiliser uniquement la synthèse sonore basée sur un modèle physique, même si celui-ci reste nécessaire pour calculer le retour d'effort [cf. 2.2].

### 2.1 Geste

En raison de la complexité du dispositif haptique, surtout au début du projet, il s'est avéré important de pouvoir tester des comportements musicaux ou sonores *de manière interactive et gestuelle* sans nécessairement utiliser le dispositif haptique, pour qui les réalisations peuvent être lourdes techniquement. Un module d'abstraction du contrôle gestuel a donc été réalisé. Nous avons par exemple utilisé une tablette graphique ou une simple souris pour des réalisations rapides. Il faut évidemment le dispositif complet pour les ajustements fins et il est alors intéressant d'avoir déjà validé certains concepts et réalisations grâce à la simulation (laquelle doit évidemment tenir compte du fait qu'il manque le retour d'effort).

Nous avons pu d'autre part enregistrer des gestes, pour les rejouer (afin d'expérimenter, régler et de valider des comportements sonores dynamiques) ou les analyser (hors de leur temps d'exécution), avec ou sans le bras haptique, grâce au concours de Frédéric Bevilacqua. Une étude comparée a été effectuée entre une tablette graphique (Wacom Intuos A4 et aérographe) et un bras à retour d'effort (Virtuose 6D d'Haption). Le but était d'identifier les paramètres significatifs et leur obtention, dans notre contexte de génération sonore et musicale en temps-réel. D'autre part, cette étude devait servir à évaluer la limitation de la tablette graphique comme outil de simulation du bras haptique. *GestureRec et Matlab*

Un même ensemble de courts extraits musicaux (des phrases jouées au marimba) a été donné à écouter à plusieurs cobayes. Chacun devait, pour chaque extrait et après l'avoir préalablement écouté trois fois, tracer six fois une représentation de la musique au moment de l'écoute ; trois fois à l'aide de la tablette graphique et trois fois à l'aide du bras haptique (seul un plan horizontal était accessible, rappelant la configuration de la tablette, mais l'utilisateur n'étant pas obligé de le toucher). La représentation était librement choisie par l'utilisateur mais elle devait être cohérente au cours des six reproductions. Les exécutions ont été enregistrées puis analysées.

Comme on pouvait s'y attendre (en pensant notamment à l'utilité d'une règle pour tirer un trait droit), la position absolue et dans une moindre mesure la direction des gestes nécessitent un guide haptique. Les normes des vitesses et les accélérations relatives se sont révélées relativement robustes (après un faible filtrage passe-bas), entre les deux dispositifs, pour un même utilisateur et une

même phrase. L'inertie du bras haptique que nous utilisions au début du projet n'était cependant pas négligeable et le modèle final s'est avéré beaucoup plus performant sur ce point.

On peut trouver une cohérence entre différents tracés pour un même utilisateur et également entre les utilisateurs dans le cas de phrases courtes : en vérifiant notamment si loi de Fitts [Fit54] et ses dérivées [AZ97] pouvaient s'appliquer à notre cas, l'idée est venue d'utiliser indépendamment les différentes caractéristiques liées par la loi (vitesse et rayon de courbure pour simplifier), l'utilisateur s'exprimant alors en choisissant de contrarier la loi (en forçant pour tourner rapidement suivant la vitesse de sa main). Afin de vérifier ce principe, des maquettes interactives agissant sur la musique entendue ont été réalisées et se sont avérées très intéressantes. *Esquisse musicale par exemple.*

## 2.2 Haptique

Par retour haptique, on entend retour tactile (sens tactilo-kinesthésique) et retour d'effort (sens proprioceptif). Ce retour haptique est très important pour contrôler le geste. Mais il reste à savoir quel retour proposer. On aborde alors le problème de la représentation puisqu'il faut définir quelle sensation rechercher, pour quel geste attendu et finalement pour quel accès à la musique. Ce point concerne aussi les retours visuels et sonores, par lesquels l'utilisateur situe également son action. Différentes pistes ont été explorées.

Pour que l'utilisateur puisse avoir une certaine expressivité par son geste, une idée est d'utiliser l'haptique comme guide, afin de réduire les contraintes inhérentes à un instrument qui demande une pratique pour obtenir sa maîtrise. L'apprentissage est également facilité par une scénarisation du déroulement du jeu dans l'installation publique.

Une réalisation préalable a été effectuée avec un haut-parleur ayant une membrane en plastique. L'utilisateur pouvait toucher la membrane pour sentir le son. Grâce à un microphone de contact collé à la membrane, il était possible de connaître l'influence du geste de l'utilisateur et ainsi modifier le son produit par le haut-parleur. Une boucle de rétroaction était ainsi créée. Bien entendu, les signaux touchés sur la membrane étaient adaptés à la perception tactile et un autre haut-parleur était nécessaire pour écouter le son produit. Des résultats furent déjà obtenus simplement en comprimant le spectre audio vers des fréquences ne dépassant pas 200 Herz (certaines fréquences restant encore audibles, il faudrait percer la membrane pour les éliminer). Une corrélation forte existe entre la puissance instantanée et le son entendu, une plus faible pour la fréquence fondamentale et une minime pour les caractéristiques spectrales.

Outre le fait que ce dispositif produise du son et nécessite une puissance importante pour que la membrane exerce une force sur les doigts de l'utilisateur, un haut-parleur se loge difficilement dans une poignée tenue dans la main. C'est pour cela que le CEA-List a conçu et réalisé un actionneur tactile intégré à la poignée.

Un bras haptique Virtuose 3D de Haption a été adapté pour satisfaire les

besoins du projet *PHASE*. Il fallait que le bras soit sûr, précis, léger, robuste au grand public et silencieux. Celui-ci a été intégré dans un support permettant à l'utilisateur de reposer son bras et de visualiser l'espace de travail à sa disposition, ce qui limite sa fatigue et augmente la précision des ses gestes.

### 3 Jouer de la musique

*Roland Cahen, le compositeur impliqué dans le projet a eu un rôle central dans cette recherche musicale.*

Sachant que notre principal mode d'interaction est le geste, le jeu instrumental est une possibilité, parmi d'autres, pour jouer de la musique. En effet, le mode de jeu n'est pas imposé *a priori* par le système à retour d'effort, pas plus que sa sonorité d'ailleurs [dL97]. Il faut donc les déterminer.

Plus généralement, nous avons cherché quelles libertés pouvaient être laissées au joueur pour s'exprimer dans un espace musical donné. On peut par exemple avoir accès au rythme, au timbre, à la tonalité, parmi les plus courants. Différents modes de jeu, gestuel et musical, ont été recensés.

#### 3.1 Modes de jeu

On peut se *positionner* dans un espace musical, la musique entendue n'étant pas le résultat d'un déplacement dynamique, celle-ci conservant une certaine autonomie : par exemple, on peut écouter un concert en étant placé proche d'un piano ou de la contrebasse. On peut également penser à l'ajustement de la fréquence sur un poste de réception radiophonique. *exemple du mixage par zones*

On peut également tenir compte de la *dynamique du parcours* de l'utilisateur, qui peut bien sûr se déplacer de point en point, un point d'arrêt n'étant alors qu'une singularité. De plus, les différentes musiques accessibles n'ont que peu d'autonomie, le parcours est prépondérant.

L'utilisateur peut également *conduire* la musique, c'est-à-dire que son mouvement influe sur la musique qui est entendue, par exemple comme un chef d'orchestre, qui peut modifier l'allure générale, la présence de certains instruments, etc. *Réalisation aboutie : esquisse musicale*

Enfin, le *jeu instrumental* [Cad94] est une possibilité, intéressante notamment car les sources sonores sont directement accessibles au joueur par les objets physiques les représentant. Pour véritablement jouer de la musique par contre, cela implique de véritablement travailler une gestuelle spécifique, comme pour tout instrument. On peut toutefois réduire la complexité (et donc les possibilités) de l'instrument afin de permettre à un public non exercé de le faire sonner malgré tout ou à un élève de l'apprendre.

### 3.2 Quelle musique jouer ? Comment ?

Il existe différentes musiques à jouer ou rejouer : les musiques connues sont identifiables et il est aisé de s'y repérer, mais elles sont par ailleurs peu malléables car enregistrées sur un support audio. On peut cependant y naviguer [Cah02], principalement en modifiant la temporalité de la lecture.

Un travail sur la matière sonore fixée a donc été réalisé en utilisant notamment la synthèse granulaire, celle-ci offrant un accès privilégié au temps, indépendamment de l'effet de hauteur, en contrepartie d'un son qui peut être assez typé. Il est ainsi possible de *rejouer* une musique de différentes manières [cf. 4.3].

On peut également réaliser des musiques adaptées aux manipulations attendues (comme le mixage de pistes ou la modification de tonalité). Celles-ci sont alors disponibles soit sur un support audio (ce qui nous ramène au cas précédent), soit selon une structure musicale, par exemple des notes MIDI. Cette structure musicale peut aussi être utilisée pour des musiques existantes, notamment sur des extraits choisis dont une forme de *partition* est disponible. De nombreuses manipulations musicales ont ainsi été réalisées sur plusieurs structures musicales, dont le rythme (quantification, durées, tempo), l'intensité (quantification, sélection), la hauteur (transposition, changement d'échelle, changement de mode tonal) et l'orchestration de voie (en hauteur, en temps et en timbre).  
*exemple interactif MiColl*

On peut enfin concevoir des générateurs [Cah00] permettant d'interagir sur des points musicaux particuliers, adaptés à certaines formes de jeu. Le générateur, qui dispose d'une certaine autonomie vis-à-vis du joueur, assure que les contraintes musicales choisies soient respectées pendant le jeu avec les différents contrôles dynamiques proposés. Afin de disposer de plusieurs situations de jeu musical, un système de préréglages est alors indispensable. Enfin, pour construire un parcours dynamique à moyen et long terme – une musique –, il est utile de pouvoir interpoler entre différentes situations au cours du jeu. Comme les paramètres produisant la musique sont explicitement connus dans le cas de la musique générée, il est alors possible de définir leur interpolation. C'est dans ce contexte qu'un générateur de flux staccato à trois voies a été programmé.  
*réalisation de Staga*

L'utilisation de structures musicales proches de notes nécessite une génération sonore, comme un échantillonneur MIDI, avec les possibilités et les limitations de ce type de synthèse sonore : un échantillonneur logiciel a été réalisé *ad hoc* et l'intégration fluidsynth (de Peter Hanappe) avec les fontes sonores dans le système s'est révélée très efficace.

Quels que soient les systèmes de génération ou de transformation utilisés, il est toujours possible de profiter des différentes particularités, notamment en les détournant, mais il devient alors difficile de généraliser la méthode. Le (vaste) problème reste le suivant : comment combiner les problématiques d'écriture et de restitution sonore, suivant le jeu musical que l'on souhaite rendre possible ?

## 4 La métaphore, lien entre le geste et la musique

Le geste de l'utilisateur s'effectue dans l'univers spatial réel (contenant sa main) afin de manipuler de la musique. Les différents retours (haptique, sonore et visuel) influent sur son geste. Il y a une boucle de rétroaction directe de l'haptique, puisque le geste et le retour d'effort sont localisés à la main de l'utilisateur. La cohérence entre les différentes modalités est garantie par la *métaphore* qui fait le lien logique entre celles-ci et donc entre le monde réel où se trouve la main et le monde virtuel musical. La métaphore est donc particulièrement importante [HWP02] et détermine la puissance d'expression du système [RWDD97], à commencer par le son que celui-ci peut produire, ce qui est le but de notre système multimodal. À ce titre, le son et la musique peuvent également influencer les retours haptique et visuel perçus par l'utilisateur.

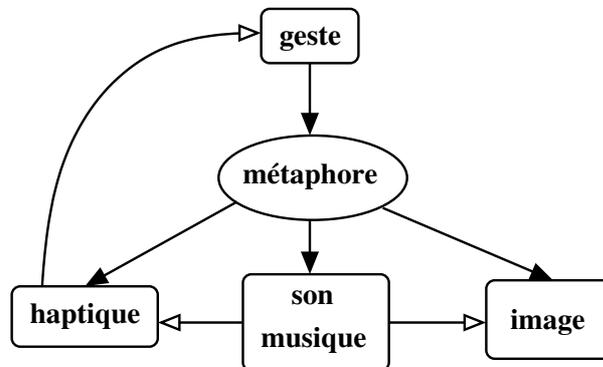


FIG. 1 – métaphore d'interaction multimodale avec entrée gestuelle

Notre système multimodal intégrant l'haptique présente différentes contraintes : le geste s'effectue dans un univers spatial avec une certaine dynamique temporelle ; le bras haptique utilise un moteur de physique pour calculer le retour d'effort (les objets sont déterminés en forme, masse et propriétés de contact ; ils interagissent via des forces) ; l'ensemble des modalités doit être perceptivement cohérent pour l'utilisateur. En tenant compte de ces contraintes et suivant les relations souhaitées entre les modalités, *le but étant de jouer de la musique*, il est possible de définir de nombreuses métaphores et donc d'aboutir à différentes réalisations, plus ou moins proches de la réalité. Plusieurs types de métaphores ont été expérimentées et sont présentées ensuite. Elles peuvent bien entendu se combiner afin d'enrichir l'univers d'interaction proposé à l'utilisateur.

### 4.1 Interaction directe

L'interaction directe s'apparente à une forme de bruitage ou de sonification : pour certaines interactions entre les objets physiques (manipulés par l'utilisateur ou non) correspondent des interactions sonores. Lorsque celles-ci sont bien choisies, elles sont rapidement comprises par l'utilisateur et participent à son immersion dans le monde virtuel, car les modalités sont proches de celles perçues couramment dans la réalité (même si le résultat sonore proprement dit peut

s'en écarter significativement). Les objets physiques sont des sources sonores et le jeu avec ces objets peut s'apparenter au jeu instrumental.

Chaque objet défini dans le moteur de physique dispose de caractéristiques fixes (forme, taille, masse, matériau, etc.) et dynamiques, ces dernières pouvant être événementielles (collision avec un autre objet, perte de collision) ou continues (position, orientation, vitesse, force exercée par un autre objet...).

Pour chaque objet physique, des comportements sont déterminés, afin d'y associer des objet sonore et visuel qui les partagent. Il s'agit d'une stratégie de mise en correspondance (*mapping*) par couches [HWK99] : à partir des objets physiques sont extraits des comportements qui sont appliqués à des objets sonores, comme le montre partiellement (il manque notamment les caractéristiques fixes de objets physiques et le visuel) le schéma de la figure 2.

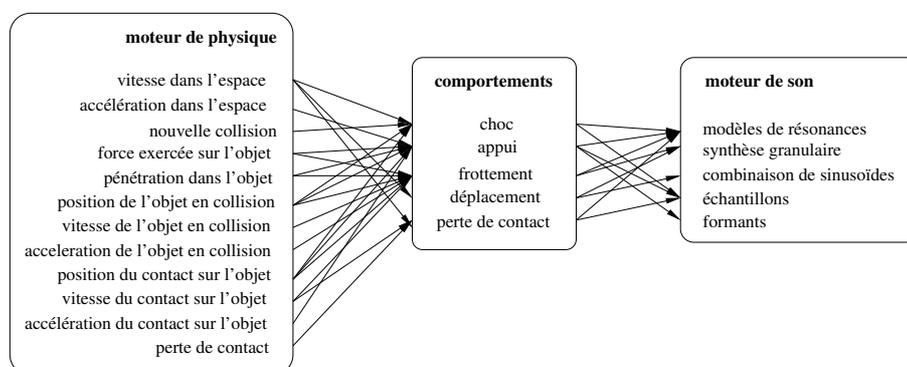


FIG. 2 – schéma de mise en correspondance à travers des comportements

Les comportements sont observés pour un couple d'objets, chacun étant simultanément la référence des comportements sonores et visuels :

**choc** Du point de vue de la physique, l'événement de contact entre deux objets n'est associé qu'à une position dans l'espace à un instant donné : aucune force n'a encore été exercée. Cependant, l'énergie cinétique relative des deux objets au point de contact (intégrant la translation et la rotation de chaque objet) peut nous permettre de prévoir quelle sera l'énergie mise en jeu lors du choc, quitte à remettre ensuite en cause cette première estimation. Un événement de choc est défini par celle-ci ainsi que par la position sur l'objet. Pour un couple d'objet ayant déjà un point de contact, tout nouveau contact est ignoré. Cette simplification peut s'avérer limitative mais réduit grandement les événements.

**appui** Deux objets en contact peuvent s'appuyer plus ou moins l'un sur l'autre. On peut mesurer les forces appliquées ou encore la distance de pénétration dans le cas d'un objet mou. Il est également intéressant d'observer la variation de l'énergie cinétique relative (les accélérations linéaires et de rotation pondérées par les masses), selon la normale au contact entre les deux objets considérés : par exemple, lorsqu'un cube roule sur un plan sans décoller, on peut avoir accès, de manière continue, aux appuis successifs des faces du cube sur le plan.

**frottement** Deux objets en contact peuvent également frotter l'un contre l'autre.

On peut là encore mesurer les forces appliquées ou observer la variation de l'énergie cinétique relative entre les deux objets, mais cette fois dans le plan tangent. Il est pratique d'utiliser un repère lié à l'objet observé et le type celui-ci dépend donc de la forme de la surface de l'objet.

**perte de contact** Un événement de perte de contact a lieu lorsque deux objets n'ont plus aucun point de contact en commun. La seule information disponible, outre le moment d'occurrence est alors l'énergie cinétique relative des deux objets.

Un comportement propre a également été défini pour chaque objet : le **déplacement** de chaque objet selon le point d'observation de l'utilisateur : il s'agit simplement des vitesses linéaires et de rotation, permettant notamment d'avoir des effets de vent relatif.

Les objets sonores exploitent, différemment selon leur classe, les différents comportements physiques, selon les multiples synthèses utilisées simultanément et éventuellement chaînées : échantillons, modèles de résonances (particulièrement adaptés pour un rendu réaliste [vdDP]), combinaison de sinusoïdes, synthèse granulaire, formants. Parmi tous les objets constituant la scène, les quatre ou cinq les plus proches du point d'écoute sont sélectionnés et effectivement calculés. Ceux-ci utilisent alors le Spat Ircam [JW95] pour le rendu de position dans l'espace, comprenant le rayonnement dans l'air et l'effet Doppler associé au déplacement.

Tout ceci est valable *a priori* pour tout objet, mais il est quelquefois intéressant d'en différencier certains, comme celui manipulé directement (appelé *avatar*). L'utilisateur peut ainsi percevoir et jouer avec différents comportements d'origine physique mais utilisés musicalement, selon une scène composée pour proposer des relations entre les différents objets.

## 4.2 Zones d'interaction spécifiques

L'espace géométrique accessible à l'utilisateur par sa main détermine l'espace musical accessible via son avatar. Il est alors possible d'utiliser l'espace géométrique ou une partie de celui-ci comme terrain de jeu musical, de différentes façons.

Une approche relativement proche de l'interaction directe, mais sans utiliser nécessairement plusieurs objets physiques et sonores ni chercher un quelconque réalisme, consiste à partir du monde physique virtuel (quitte à en construire un) et à essayer de l'écouter de différentes manières. Par exemple, il est possible d'écouter différents paysages par un simple mixage entre deux sources sonores en fonction de l'altitude. *exemple de l'écoute d'image*

Il est par ailleurs possible de définir une topologie spécifiant le comportement sonore suivant les positions et les chemins accessibles. Par exemple dans la réalisation du *mixage par zones*, les différents éléments musicaux sont répartis dans l'espace sous la forme d'ellipsoïdes pouvant s'interpénétrer. L'accès aux éléments se fait par l'avatar qui peut entrer dans les différents éléments, se

déplacer à l'intérieur et en sortir. Les zones sont ressenties grâce à l'haptique comme des objets mous. Une légère résistance s'oppose à leur pénétration ou à leur sortie, qui restent possibles : ces actions résultent d'un véritable choix de l'utilisateur. La réactivité au geste est immédiate (entrée ou sortie, mouvement à la surface ou dans une zone) mais le geste n'est pas nécessaire au déroulement de la musique, les processus sonores pouvant se dérouler de manière autonome. En particulier, le positionnement sonore n'est effectué que lorsque le joueur est en contact avec au moins une zone, pas lorsqu'il décolle. Ce décollage est important car il permet alors de découpler le geste de la musique entendue et donc de préparer des actions sans perturber les processus en cours. Ceux-ci réagissent selon des zones délimitées mais sont combinables, les différentes zones pouvant se recouvrir. Dans la réalisation qui a été effectuée, on a un son associé à chaque zone et son volume qui dépend de la position à l'intérieur de celle-ci. L'image sert à représenter les différentes zones dans l'espace. Les couleurs servent à déterminer des liens entre différentes zones et à représenter l'activité des processus sonores, ce qui est utile notamment avant d'entrer dans les zones, donc avant de les entendre. Il s'agit ici d'une navigation sonore, le joueur ayant différentes variantes à portée de main. L'intérêt de la composition réside alors dans l'exploration et le parcours de cet espace de jeu.

Le terrain de jeu peut également être dérivé de processus musicaux ou sonores et ainsi constituer une véritable cartographie de ceux-ci, permettant au joueur de s'y repérer. Ceci a notamment été réalisé pour les *montagnes de la concordance harmonique*, d'après un travail de Jean-Marc Chouvel [CD94], afin d'explorer des accords de trois tons de manière intuitive.

Sans définir précisément le terrain de jeu, il reste possible d'encourager le joueur à suivre certaines directions ou courbures privilégiées et de borner ses déplacements, ce qui est très intéressant dans des processus par ailleurs relativement libres, comme *l'esquisse musicale* ou l'interaction avec des processus de synthèse sonore, comme *les herbes*.

### 4.3 Rejouer une musique

Il s'agit ici de rejouer une musique qui a été fixée, à plus ou moins long terme, et qui a été tracée dans l'espace accessible au joueur. On peut penser à une bande magnétique ou à un sillon de disque vinyle, qui vont être lus par des têtes de lecture. En considérant la musique selon la dimension temporelle uniquement, se positionner à un point de cette trace revient à écouter un temps donné.

La trace est matérialisée par un sillon creusé dans une surface, à la manière d'un vinyle mais celle-ci n'est pas forcément linéaire : sa forme peut représenter l'évolution musicale à moyen terme. Au fond de ce sillon se trouve une texture visuelle et haptique représentant l'information sonore et musicale à court terme.

La tête de lecture (*TL*) est matérialisée par une sphère que manipule le joueur. Il peut la positionner dans la trace, qu'il voit et sent par les retours visuel et haptique. Lorsque la *TL* est à l'intérieur du sillon, elle prend la couleur

de la trace et émet le son correspondant à l'instant associé. Si le joueur parcourt la trace, il rejoue la musique qui y est enregistrée. À la différence d'un système utilisant un diamant et un vinyle, la vitesse de parcours n'influe pas sur la hauteur, grâce à l'utilisation de la synthèse granulaire. (*voir l'exemple interactif SogsBrowser avec l'objet sogs~ de Norbert Schnell*)

Grâce à la forme du sillon et au retour haptique lorsque la *TL* appuie dessus, il est facile au joueur de suivre la trace, celle-ci servant de guide. Le joueur peut toutefois *choisir* de s'en écarter. Toute position en dehors de la trace est associée au point de la trace le plus proche, ce qui est valide pour tout point l'espace géométrique mais pertinent seulement lorsque la *TL* est proche de la trace. Aussi, deux types d'écart à la trace ont été définis : l'écart latéral, selon la tangente à la surface gravée par le sillon et l'écart en altitude, selon la normale à la surface gravée. Ces deux écarts ont pour origine la trace. Lorsque l'écart latéral augmente, la musique reproduite par la *TL* est de moins en moins forte jusqu'à disparaître à l'extérieur du sillon. Lorsque l'écart en altitude augmente, le moment de relecture est de plus en plus étendu (l'utilisation de la synthèse granulaire permet d'avoir de nombreux points de lecture autour du moment correspondant à l'aplomb de la *TL*).

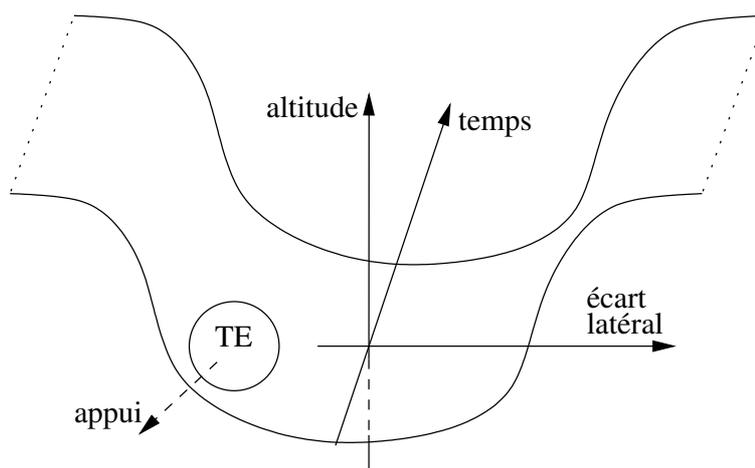


FIG. 3 – tête de lecture (TE) dans le repère relatif au moment de lecture sur la trace

Lorsque la trace et donc la musique enregistrée défilent dans l'espace, comme quand le disque vinyle sur lequel est gravé le sillon tourne, on ne contrôle plus un temps absolu mais une vitesse relative, créant ainsi des accélérations ou des ralentissements par rapport à la vitesse de référence (qui est celle du disque). On peut enfin contrôler le défilement du temps dans l'espace en contrôlant la vitesse du disque et ainsi conduire la musique qui y est enregistrée. Ceci est réalisé en appliquant un frein à la vitesse nominale du sillon selon les écarts latéral et en altitude et selon l'appui de la *TL* sur le sillon. De cette manière, tous les contrôles sont coordonnés par le geste de l'utilisateur.

Afin d'assurer une musicalité à la relecture, nous nous sommes aperçus qu'il était nécessaire de limiter les variations rapides de vitesse relative lorsque la

vitesse du disque devenait importante. (Pour assurer une bonne réactivité du système, il faut cependant les conserver quand le disque a une vitesse faible ou est arrêté.) Il est possible de ne pas prendre en compte toutes les variations de position de la *TL* Mais puisque la position temporelle entendue ne correspond alors plus forcément à celle visualisée, il convient alors de trouver un compromis acceptable pour les différentes modalités mises en œuvre. Il est plus intéressant d'utiliser l'haptique afin de varier l'inertie de la *TL* en fonction de la vitesse du disque pour véritablement limiter ses mouvements, ce qui est perceptible par l'utilisateur.

## 5 L'installation pour le grand public



FIG. 4 – photographie du démonstrateur : le bras haptique et l'écran vidéo

La métaphore permettant de rejouer la musique, avec les contrôles de vitesse, a été particulièrement développée pour être au cœur du jeu proposé au grand public. En plus de la tête de lecture (*TL*), une tête d'écriture (*TE*) a été ajoutée. On entend et on voit la *TE* qui génère une musique suivant les actions du joueur. Cette musique est déposée dans l'espace sous la forme d'une trace visuelle et haptique que la *TE* laisse derrière elle quand elle se déplace. Le joueur manipule la *TL*. Lorsqu'il se place sur la trace, il sent et entend la musique qu'il est en

train de rejouer. Il poursuit la *TE*. C'est la trace dans le sillon qui fait le lien entre la *TE* et la *TL*. Elle est inscrite par la *TE* et se met en mouvement lorsque la *TL* est dedans. Ainsi, la *TL* a un déplacement relatif par rapport à la trace et le joueur contrôle indirectement la vitesse de la musique qu'il rejoue. Pour suivre ou rattraper la *TE*, le joueur doit bien suivre la musique grâce à la trace.

Le choix de cette métaphore a été motivé par plusieurs raisons. L'analogie avec le monde réel est appréciable car on peut penser à un disque vinyle dont le sillon serait gravé (par la *TE*) et parcouru par un diamant que l'on tiendrait en main (la *TL*). La course-poursuite, ici entre la *TE* et la *TL*, est un ressort de jeu courant. De plus, le principe musical de rejeu est proche du canon ou de l'improvisation rythmique en imitation à deux. Enfin, cette métaphore a l'avantage de faciliter le geste, la trace déposée au fond d'un sillon servant de guide pour rejouer la musique inscrite dans l'espace.

Le joueur devait écouter la *TE* et jouer avec. Cette situation musicale est courante, mis à part qu'il n'y a pas de quantification tonale ou rythmique. On a une variation continue de la vitesse relative des deux protagonistes. Il peut alors devenir difficile de distinguer les deux acteurs musicaux, ce qui devient une problématique musicale en soi : pour le dispositif, car on veut favoriser le jeu, mais également pour le joueur qui doit chercher à écouter l'autre et à s'écouter pour réussir à se situer dans l'ensemble.

Le mode de jeu compétitif avec la *TE* était entraînant, mais n'était toutefois pas imposé. À tout moment, il était possible pour le joueur de s'arrêter, sortir du sillon et jouer avec les objets sonores ou zones d'interactions, qui d'ailleurs n'étaient pas inactif pendant le défilement du sillon. Ceux-ci participent en permanence à l'immersion et à l'éveil sonore et musical du joueur, par la richesse de sonorités et des interactions proposées. Lorsque le joueur était arrêté, la *TE* attendait alors non loin en silence pour repartir dès que le joueur reprenait de la vitesse. Il n'y avait pas à proprement parler de score, mais la progression du joueur était évaluée afin d'ajuster le comportement du système : il n'y avait pas de différence entre un joueur qui ayant de la difficulté à avancer et un autre plus contemplatif, mais dans les deux cas le système ralentissait, facilitant ainsi la prise en main du système par le grand public.

L'installation *PHASE* fut très fréquentée : en ne comptant que les personnes ayant joué, il y eut environ 50 personnes par jours, 6 jours par semaine pendant 3 mois, soient plus de 3 000 joueurs. Ce fut l'occasion de confronter le prototype au grand public (enfants et adultes, musiciens ou non, joueurs de jeu vidéo, aveugles, enseignants, etc.). Cette expérience fut très positive car l'accueil du public fut enthousiaste d'une part et nous pûmes d'autre part étudier et consolider le système, sur des points matériel et logiciel.

## 6 La réalisation du prototype

Le dispositif réalisé dans *PHASE* permet d'utiliser plusieurs types de métaphores multimodales dans une même scène de jeu interactif. Le système est un prototype complexe intégrant quatre ordinateurs.

Le premier (RTAI-Linux) gère les capteurs de position et les moteurs du bras haptique. Il intègre notamment différentes sécurités, ce qui est indispensable pour toute utilisation, grand public ou non.

Le second (Windows XP) intègre le moteur d'interactions physiques (Vortex de CM-Labs) interagissant avec le bras haptique ainsi que le moteur de scénario et d'affichage (Virtools). Deux projecteurs vidéo associés à des lunettes polarisées permettent aux utilisateurs d'avoir une vision stéréoscopique de la scène, ce qui favorise l'immersion et facilite le geste.

Le troisième (MacOS X) reçoit les informations, en fonction des informations fournies par le scénario et le moteur de physique via OSC, le protocole standard pour l'audio [WF97], et s'occupe de la génération sonore et musicale (MAX/MSP), de la tête d'écriture (*TE*), de la tête de lecture (*TL*), du sillon et des autres objets sonores de la scène, selon les métaphores décrites précédemment. Il a été fait usage de nombreux développements de l'équipe « applications temps-réel » de l'Ircam, notamment l'environnement FTM.

Un quatrième ordinateur (MacOS X) spatialise le son sur le dispositif de restitution (8 haut-parleurs tout autour du joueur et de l'écran) selon les positions des objets, transmises également par OSC depuis le moteur de physique. Le rendu des zones est également confié à cette machine car leur rendu spatial, non nécessairement réaliste, peut nécessiter un traitement particulier.

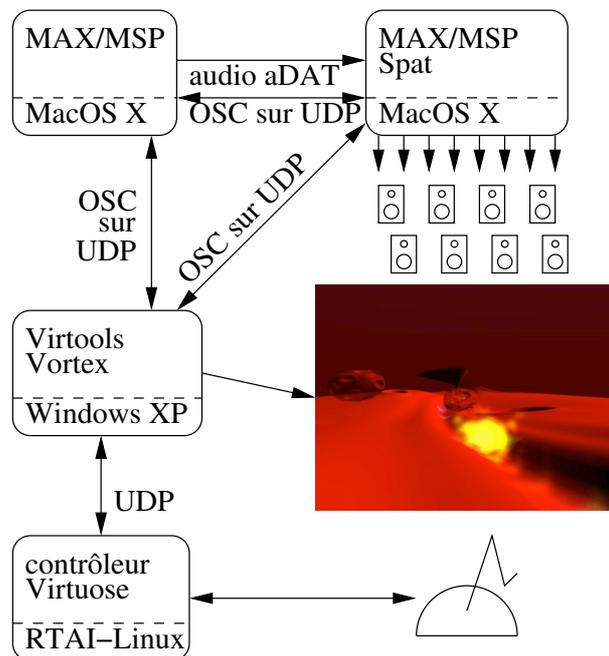


FIG. 5 – schéma de l'architecture du système

Parmi les différents développements, la gestion du temps fut cruciale, notamment pour les têtes de lecture et d'écriture. La base de temps était répartie entre les moteurs de physique et de son, sur deux ordinateurs différents. Lors

du parcours de la *TE*, une trace de temps était posée dans l'espace du moteur de physique alors que la musique générée était entendue et enregistrée dans la mémoire du moteur de son. Lors du parcours de la *TL*, le point le plus proche dans la trace était trouvé et la différence avec le temps courant de la *TE* donnait le retard utilisé pour reproduire la musique entendue précédemment. Entre chaque information temporelle du moteur de physique (environ toutes les millisecondes), le moteur de son avait une nécessaire autonomie pour calculer les échantillons audio (à une fréquence d'échantillonnage stricte de 44 100 Herz). De cette manière, les différentes modalités restaient cohérentes.

## 7 Conclusion

L'utilisation d'un dispositif haptique pour jouer de la musique soulève de nombreuses questions : techniques, scientifiques et musicales. Dans le cadre du projet *PHASE*, des outils matériels, logiciels et méthodologiques ont été développés, pour permettre l'élaboration et la réalisation de métaphores ayant un sens musical. Ceux-ci sont maintenant disponibles et utilisables via le prototype existant.

La réalisation d'une installation interactive pour le grand public sous la forme d'un jeu musical intègre différentes métaphores, montre la validité et l'intérêt d'un tel dispositif et ouvre la voie à de nombreuses manipulations musicales gestuelles.

## Références

- [AZ97] John Accot and Shumin Zhai. Beyond fitts' law : Models for trajectory-based hci tasks. In *ACM/SIGCHI : Conference on Human Factors in Computing Systems (CHI)*, 1997.
- [Cad94] Claude Cadoz. Le geste canal de communication homme-machine. La communication « instrumentale ». *Sciences Informatiques - Numéro Spécial : Interface Homme-Machine*, 13(1) :31–61, 1994.
- [Cah00] Roland Cahen. Générativité et interactivité en musique et en art électroacoustique, 2000.
- [Cah02] Roland Cahen. Navigation Sonore Située, 2002.
- [CD94] Jean-Marc Chouvel and Boris Doval. La concordance harmonique. Vers une nouvelle approche de la consonance. In *Proceedings of the International Conference on Music Perception and Cognition (ICMPC)*, 1994.
- [CLF90] Claude Cadoz, Leczek Lisowsky, and Jean-Loup Florens. Modular Feedback Keyboard. In *Proceedings of the International Computer Music Conference (ICMC)*, 1990.
- [CW00] Claude Cadoz and Marcelo M. Wanderley. Gesture - Music. In M.M. Wanderley and M. Battier, editors, *Trends in Gestural Control of Music*, pages 71–94. Ircam - Centre Pompidou, 2000.

- [dL97] Serge de Laubier. Le méta-instrument a-t-il un son ? Émergence de lois ou de constantes dans le développement d'instruments virtuels. In GMEM, editor, *Colloque les Nouveaux Gestes de la Musique*, Marseille, 1997.
- [Fit54] Paul M. Fitts. The information capacity of the human motor system in controlling the amplitude of the movement. *Journal of Experimental Psychology*, 47 :381–391, 1954.
- [GLPNV02] Viviane Gal, Cécile Le Prado, Stéphane Natkin, and Liliana Vega. Writing for video games. In *Proceedings Laval Virtual (IVRC)*, 2002.
- [HWK99] Andy Hunt, M. Marcelo Wanderley, and Ross Kirk. Toward a Model for Instrumental Mapping in Expert Musical Interaction. In *Proceedings of the International Computer Music Conference (ICMC)*, 1999.
- [HWP02] Andy Hunt, Marcelo M. Wanderley, and Matthew Paradis. The importance of parameter mapping in electronic instrument design. In *Proceedings of the Conference on New Instruments for Musical Expression (NIME)*, 2002.
- [JW95] Jean-Marc Jot and Olivier Warusfel. Spat : A Spatial Processor for Musicians and Sound Engineers. In *CIARM*, Ferrara, Italy, 1995.
- [RWDD97] Joseph Butch Rován, Marcelo M. Wanderley, Shlomo Dubnov, and Philippe Depalle. Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. *Kansei - The Technology of Emotion*, 1997.
- [vdDP] Karl van den Doel and Dinesh K. Pai. Modal Synthesis for Resonating Objects. To appear in a book.
- [Vei01] Anne Veitl. Quelles ressources technologiques pour renouveler les pédagogies de la musique ? Présentation critique d'outils. Technical report, Ministère de la Culture - DMDTS, 2001.
- [WF97] M. Wright and A. Freed. Open Sound Control : A New Protocol for Communicating with sound Synthesizers. In *Proceedings of the International Computer Music Conference (ICMC)*, 1997.