

# Immersion dans un monde visuel et sonore en 3D

Venceslas Biri, Patrice Bouvier, François de Sorbier de Pougnadoresse, Pascal Chaudeyrac, Benoit Piranda

Equipe SISAR, Université de Marne la Vallée  
{biri, bouvier, fdesorbi, pchaudey, piranda}@univ-mlv.fr

---

## Abstract

*Cet article présente la démarche suivie pour la création de notre salle de réalité virtuelle, aussi bien du point de vue matériel que logiciel. L'idée directrice étant de disposer d'une salle de réalité virtuelle peu onéreuse et transportable, nous présentons nos choix de conception basés sur la connaissance des mécanismes de la perception humaine.*

*Comme tout système de réalité virtuelle, notre salle vise à immerger l'utilisateur dans un espace où il pourra ressentir non seulement la présence des objets virtuels mais aussi le sentiment de présence de soi dans l'environnement virtuel. Pour recréer ce double sentiment de présence nous avons choisi d'utiliser des interfaces sensorielles les moins intrusives possibles et des interfaces motrices les plus naturelles et intuitives. Un moyen pour tendre vers l'oubli des interfaces est de détourner l'attention de l'utilisateur grâce à des applications présentant un intérêt marqué. Nous vous présentons des applications dans différents domaines : art, jeux et archéologie.*

*This article presents the approach we followed for the creation of our virtual reality room, both from the hardware point of view and the software one.*

*Our main goal to build a virtual reality room as cheap as possible and easily transportable was constrained by our knowledge of the mechanisms of human perception. Like any virtual reality system, our room aims to immerse the user in a place where he will be able to feel not only the presence of the virtual objects but also the feeling of self-presence in the virtual environment. To recreate this dual feeling of presence we used the less intrusive as possible sensor interfaces and the most natural and intuitive motor interfaces.*

*We achieve this goal by diverting the user's attention to the application itself. We present applications in various fields : art, games and archaeology.*

**Keywords:** réalité virtuelle, immersion, présence, spatialisation sonore

---

## 1. Introduction

Un système de réalité virtuelle vise à immerger un ou plusieurs utilisateurs au cœur d'un environnement artificiel dans lequel il pourra ressentir et interagir en temps réel au moyen d'interfaces sensorielles et motrices. L'expérience devra être suffisamment crédible pour leurrer les sens de l'utilisateur afin de créer, but ultime, un sentiment de présence des objets virtuels mais aussi de présence de soi au sein de l'environnement virtuel. Ce sera à ce prix que l'utilisateur acceptera de se laisser prendre au jeu.

Parmi les salles de réalité virtuelle classique rappelons l'existence du CAVE [CNSF\*92] et du Workbench [KBF\*95]. La "CAVE", Cave Automatic Virtual Environ-

ment, est composé d'une petite pièce cubique de cinq mètres environ, dont chaque mur ainsi que le sol sont des écrans où sont projetées des images synchronisées. Le Workbench est un plan de travail virtuel qui se compose de deux écrans perpendiculaires, d'un dispositif d'affichage en relief à l'aide de lunettes à cristaux liquides, d'un dispositif de capture des mouvements de la tête et d'un stylo à 6 degrés de libertés pour manipuler les objets. Enfin une autre solution consiste à faire porter à l'utilisateur un casque dans lequel est placé un certain nombre d'écrans à cristaux liquides.

Ces solutions ont en commun d'être très coûteuses et/ou très difficilement transportables. Se fixant des objectifs différents, l'équipe Sisar de l'Université de Marne la Vallée,

a monté une salle de réalité virtuelle avec du matériel grand public. Notre volonté est de disposer d'un système à moindre coût et facilement transportable. Il nous permet de tester différentes interfaces homme-machine, les moins intrusives possibles ainsi que différents algorithmes de rendu visuel et sonore afin d'immerger au mieux l'utilisateur.

Après un bref aperçu des méthodes mises en oeuvre pour l'immersion visuelle puis sonore nous détaillerons nos interfaces motrices nécessaires à l'interaction. Enfin, avant de conclure, nous présenterons quelques unes de nos réalisations. Il est important de préciser que proposer des applications présentant un intérêt marqué, ludiques, articulées autour d'un scénario fort, ou porteuses d'émotions, favorisent l'oubli des interfaces et donc facilite le sentiment de présence.

## 2. Immersion

Dans le cadre de notre salle de réalité virtuelle, l'immersion de l'utilisateur repose sur un rendu visuel en 3D basé sur la stéréovision et un environnement sonore 3D restitué à l'aide d'un dispositif 8.0. Après un bref rappel théorique nous décrirons ces dispositifs.

### 2.1. Restitution panoramique par stéréoscopie

blabla

#### 2.1.1. Rendu de la scène

Un objectif de JIM 3D est de reproduire le principe de restitution de l'environnement virtuel qui est utilisé dans le jeu vidéo MYST. Ce choix est basé sur le fait que le contenu des scènes apparaît être plus réaliste comparé aux méthodes de rendu traditionnelles que nous avions pu expérimenter jusqu'ici dans nos applications destinées à notre salle de réalité virtuelle. Le rendu d'une scène se fait en deux étapes : création du contenu puis affichage des mosaïques générées. D'une manière théorique nous cherchons à créer une image simulant une vision à 360 degrés depuis une position donnée et, dans le jeu, à l'afficher sur un cylindre. En pratique, nous utilisons un logiciel de modélisation type 3DSMax ou Maya qui nous permet d'obtenir le rendu réaliste que nous désirions.

A l'aide d'un script et à partir de la scène 3D, nous produisons un ensemble d'images qui mises bout à bout forment la mosaïque de la vision à 360 degrés. Cette dernière est finalement appliquée en tant que texture sur un cylindre créé via l'API OpenGL.

La principale contrainte de cette technique est de conserver la résolution de la mosaïque d'images lors de la phase de rendu. Les solutions appropriées sont, entre autres, l'utilisation d'une mosaïque de résolution importante, l'utilisation de paramètres optimales pour le cylindre (rayon, hauteur, discréétisation, etc) ou un nombre important d'images composant la mosaïque.

#### 2.1.2. Utilisation stéréoscopique des panoramiques

Pour intégrer la salle de réalité virtuelle et se détacher du mode de visionnage de MYST, JIM 3D devait pouvoir diffuser des images stéréoscopiques ajoutant ainsi une sensation de profondeur à l'utilisateur.

Le principe d'une image stéréoscopique est de fournir à chaque œil une vision légèrement décalée d'un même point de vue d'une scène. De cette façon il est possible de simuler le fonctionnement de la vision humaine. Il existe divers dispositifs pour faire cette association œil-image. Nous pouvons citer par exemple les visiocasques ou les lunettes filtrantes (rouge-cyan, à obturation électronique, etc).

Nous ne pouvons pas appliquer la stéréoscopie directement sur le cylindre en utilisant une simple rotation des deux points de vue car des problèmes surviendraient au niveau de l'axe inter-oculaire. Pour appliquer le principe stéréoscopique à des images panoramiques, plusieurs solutions ont déjà été proposées [Stereo Panorama with a Single Camera, Parallel-Perspective Stereo Mosaics] mais ces méthodes ont le désavantage de ne pas être temps réel. Ce critère étant à la base du fonctionnement de JIM 3D, nous avons opté pour une technique mettant en œuvre non plus un mais trois cylindres. Ils sont positionnés sur les sommets d'un triangle équilatéral : trois mosaïques d'images sont donc produite à partir du logiciel de modélisation. Pour donner la sensation de relief dans l'application JIM 3D, il suffit alors de sélectionner le bon cylindre pour chaque œil en fonction de l'orientation du point de vue. Les résultats obtenus permettent d'apprécier une profondeur dans notre jeu et le saut d'image lié au changement de cylindre pour un œil ne se ressent quasiment pas.

#### 2.1.3. Intégration dans notre système immersif

A l'heure actuelle la quasi totalité des systèmes de réalité virtuelle utilise un dispositif visant à stimuler le sens visuel de ses utilisateurs. Trois critères fondamentaux doivent être respectés pour que ces derniers aient le sentiment d'être immergés dans l'environnement virtuel.

Tout d'abord le champ de vision de l'utilisateur doit être principalement occupé par l'image virtuel pour éviter que des éléments du réel ne viennent trop interférer et réduire l'effet immersif. Les solutions existantes sont diverses comme l'emploi d'un visiocasque ou alors d'un écran de projection ayant une taille suffisamment importante (5 mètres sur 3.50 mètres dans le contexte d'intégration de JIM 3D).

Le second critère vise à utiliser des images stéréoscopiques de sorte que l'utilisateur puisse percevoir la profondeur (cf. section précédente). Deux images, destinées à chacun des deux yeux de l'utilisateur, sont projetées sur l'écran au moyen de deux vidéo-projecteurs. Pour que les yeux de l'utilisateur puisse les différencier, nous utilisons

dans notre salle de réalité virtuelle un ensemble de filtres polarisants linéaires. Dès lors, l'unique contrainte pour l'utilisateur est le port de lunettes dédiées à l'utilisation de ce type de filtrage et n'a donc pas à subir le poids d'un équipement plus important.

Nous avons vu dans la partie précédente que le moteur de JIM 3D permettait de récupérer facilement une image destinée à chacun des deux yeux de l'utilisateur. Ayant respecté cette obligation, JIM 3D s'intègre aisément dans notre salle de réalité virtuelle ce qui nous permet d'apprécier encore plus notre application.



Figure 1: Notre écran de 6x3 mètres.

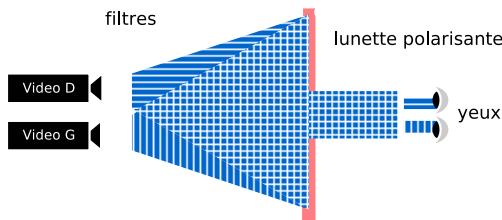


Figure 2: Principe de la stéréoscopie.

## 2.2. Son

### 2.2.1. Théorie Duplex

Avant toutes choses nous rappelons que la plage des fréquences audibles pour l'Homme s'étend de 20Hz à 20kHz. C'est en 1907 que Lord Raleigh [Ral07] définit la théorie Duplex sur la perception sonore en trois dimensions. Cette théorie approxime une tête humaine par une sphère munie de deux orifices diamétralement opposés pour représenter les oreilles. En dehors des cas où la source sonore se trouve sur le plan médian entre les oreilles, les distances de la source sonore aux oreilles ipsilatérale et controlatérale sont différentes. Cet écart, dû au temps de parcourt différent, est à l'origine de l'interaural time delay (ITD). Cependant l'ITD est exploitable seulement pour



Figure 3: Utilisation du flou pour attirer l'attention.

les fréquences inférieures à 1500Hz<sup>†</sup>. En dessous de cette fréquence les ondes ne sont pas perturbées par des obstacles de la taille d'un visage. On notera que l'ITD intervient particulièrement pour la localisation des sons dont l'attaque est rapide. Pour des sons continus ou dont l'attaque est plus lente le cerveau utilisera plutôt l'interaural phase différence (IPD) qui traduit la différence de phase au niveau des oreilles. Au dessus de 1500Hz l'interaural level delay (ILD) entre en jeu. De part leur longueur d'onde ces ondes rebondissent sur la tête et donc ne parviennent pas ou alors affaiblies (par diffraction) à l'oreille controlatérale. Selon Carlile [Car96] une atténuation de 40dB est obtenue pour une fréquence de 3kHz.

La limite de la théorie Duplex est qu'elle donne à faire croire que l'on entend en trois dimensions parce que l'on dispose de deux oreilles. Pour s'en convaincre il suffit de savoir qu'une personne sourde d'une oreille, on parlera d'audition monaurale, peut localiser en trois dimensions une source sonore [WK97]. Pour cela elle se base sur l'apprentissage de signatures spectrales non pas associées seulement à un type de son mais à un type de son et sa position spatiale relative par rapport à lui. Ces variations spectrales reposent sur des interférences causées par les épaules, la tête et le pavillon<sup>‡</sup>. Cette capacité contribue aussi dans le cas classique

<sup>†</sup> Ce qui correspond à une longueur d'onde d'environ 22cm soit la largeur de la tête.

<sup>‡</sup> La forme du pavillon fait que celui-ci influe sur les sons de fréquence supérieures à 2000Hz.

d'audition binaurale à affiner la localisation et plus particulièrement l'élévation, à faciliter la distinction avant/arrière et à placer les sources lorsqu'elles sont sur le plan médian. Ces interférences peuvent être simulées grâce aux head related transfer function (HTRF) puis être raisonnablement restituée lors de l'écoute au casque. Afin d'être complet on peut signaler que la connaissance a priori de la source sonore augmente la précision de la détection.

On peut affirmer que l'ouïe est un sens multi-modale, en effet l'homme, de façon inconsciente, effectue de légers mouvements avec la tête. En tendant l'oreille on cherche à obtenir en quelques sortes d'autres images sonores de la scène et donc affiner la précision de la détection. De surcroît on ressent les ondes sonores avec tout le corps notamment les plus graves qui sont autant ressenties qu'entendues.

Précisons qu'il existe de fortes interactions entre la perception visuelle et sonore, citons deux exemples: la ventriloquie et l'effet McGurk [MD76] où la vision de lèvres articulant une syllabe alors que le son émis est une autre syllabe fait reconnaître à l'utilisateur une troisième syllabe différente.

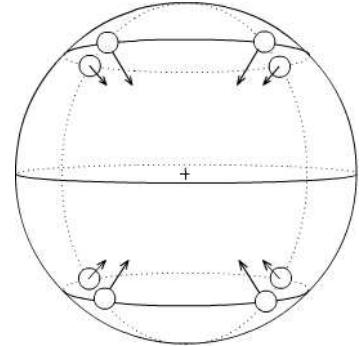
### 2.2.2. Dispositif de restitution

Comme nous l'avons signalé dans l'introduction nous recherchons les interfaces les moins intrusives, aussi nous nous sommes orienté sur la restitution sonore via des haut-parleurs plutôt qu'au casque qui peut se révéler pesant et source d'inconfort. Nous y voyons plusieurs avantages : nous n'avons pas besoin d'utiliser de HTRF génériques qui par définition ne sont pas celles propre à l'utilisateur. Celui-ci pourrait en avoir sa capacité de localisation affectée. Comme nous l'avons dit au paragraphe précédent l'utilisateur effectue des mouvements de la tête inconscients pour affiner la localisation. Ces mouvements seront effectués par notre utilisateur or sauf au prix d'un tracking très précis il est impossible d'adapter les HTRF à ses mouvements. Il en résultera pour l'utilisateur une impression de sons provenant de l'intérieur de la tête. De plus le son étant une onde il se propage aussi par l'intermédiaire du squelette. Par exemple, il est impossible de simuler les sons très graves qui sont plus ressentis que entendus. Enfin entre l'écouteur et l'oreille se crée une cavité qui a sa propre réponse impulsionale qui doit être compensée via les HTRF.

### 2.2.3. Placement des enceintes

L'audition étant un sens omnidirectionnel il est naturel de représenter l'espace sonore par une sphère dont le centre serait la tête de l'utilisateur. Aussi le but est d'approximer au mieux cette sphère. Afin de conserver le caractère transportable voulu pour notre système de réalité virtuelle nous avons décidé de n'utiliser que 8 enceintes placées aux coins d'un cube inscrit dans la sphère. Néanmoins il est clair que plus le nombre d'enceintes est grand plus le placement de la

source sonore virtuelle sera précis. Notre système nous permettra de situer la provenance du son en tout point du cube.



**Figure 4:** Position et direction des enceintes dans la sphère.

### 2.2.4. Notre système

La spatialisation du son est réalisée à l'aide d'une carte son Terratec 7.1 qui était originellement conçu pour une spatialisation planaire (avant centre, avant droit, avant gauche, surround droit, surround gauche, arrière gauche, arrière droit et caisson de basse). Cette configuration a été modifiée en un système 8.0 créant un espace 3D pour le son. Les huit haut-parleurs ont les mêmes caractéristiques et sont placés aux extrémités d'un cube inscrit dans la sphère ayant pour centre la tête de l'utilisateur. Le son restitué n'est pas de haute qualité à cause du remplacement du caisson de basse par une enceinte standard, cependant, une grande part du spectre audible est restituée.

Trois bibliothèques sonores ont été testées sous Linux, les deux premières étant d'un niveau plus élevé que la troisième, OpenAL [Ope], FMOD [FMO] et Jack-it [Dav]. Elles ont toutes un système de lecture et de positionnement dans le fichier sonore et proposent toutes des interfaces de buffers permettant la communication entre différents modules audio. OpenAL et FMOD proposent toutes deux une architecture prête à l'emploi pour jouer un son sur les systèmes de sonorisation classiques tels que la stéréo ou la spatialisation planaire 5.1 (planaire, car aucune information sur l'élévation du son n'est transmise). Avec Jack-it, il est possible de partager des buffers concurrents<sup>§</sup> entre plusieurs applications et entrées/sorties avec une faible latence. Les huit enceintes ont été définies en tant que sorties indépendantes, c'est avec cette disposition que les algorithmes de spatialisation sphérique ont été écrits. L'avantage majeur de Jack-it comparé à l'utilisation seule des pilotes Alsa ou OSS est qu'un mixage optimisé des buffers est effectué avant leur envoi vers une sortie donnée. L'inconvénient majeur est qu'aucun des éléments de haut niveau tels que la conversion de fréquences d'échantillonnages, l'effet Doppler,

<sup>§</sup> opérant simultanément

l’atténuation de certaines fréquences en fonction de la distance ou la réverbération liée à l’environnement du son n’est implémenté.

Le but de la spatialisation est de positionner une source sonore dans un environnement 3D, il n’y a par conséquent aucune raison de lire des fichiers stéréo entrelacés. Notre système se limite donc à la lecture de fichiers WAV mono.

La conversion de fréquences d’échantillonnage a été réalisée à l’aide de filtres à temps variant [PM96] qui ont l’avantage d’intégrer les processus d’interpolation et de décimation dans le filtrage. Cette méthode produit un filtrage très optimisé tout à fois adapté à une application temps réel. Les conversions ont été réalisées pour les fréquences usuelles: 44.1 kHz (qualité CD), 48 kHz (qualité DAT) et 96 kHz (qualité studio). Ce module a été créé pour permettre une importation rapide de sons au sein des installations (sans avoir à effectuer de conversions préalables) et pour rester compatible avec la majorité des cartes son fonctionnant sous Linux.

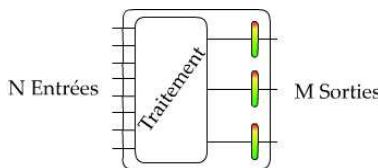


Figure 5: Représentation du module LADSPA.

Les modules de spatialisation ont été implémentés à l’aide de LADSPA [LAD]; c’est une API simple permettant de développer des modules audio tels que celui défini dans l’illustration 5. Ces modules ont N entrées qui peuvent être mutées et M sorties. Pour chaque sortie, le volume peut être contrôlé (pour afficher un vumètre par exemple). Le traitement est la partie qui détermine les modifications appliquées au son et/ou les liens effectués entre entrées et sorties. Un grand nombre de modules libres de droits ont été implémentés avec LADSPA, dont une collection signée Steve Harris [Har]. Un module de réverbération a été utilisé pour créer différentes ambiances en fonction des propriétés de la scène; une grande réverbération permet de simuler une cathédrale ou un rêve, une réverbération plus courte est utilisée pour une pièce de petite taille ou pour donner à l’utilisateur une sensation d’enfermement.

Un filtre passe bas a été utilisé pour simuler l’atténuation des hautes fréquences en fonction de la distance son/utilisateur. Le module de spatialisation est le dernier maillon de la chaîne de traitement, comme le montre l’illustration 6. Il récupère une entrée, le son provenant du module de réverbération et dispose de 8 sorties liées aux sorties de la carte son.

Les positions des haut-parleurs et du son sont converties du repère euclidien vers un repère sphérique avec la tête de

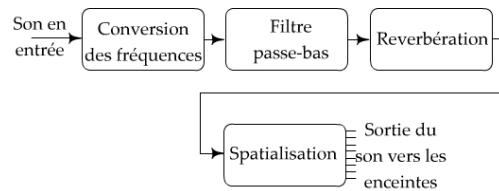


Figure 6: Traitement d’un son simple.

l’utilisateur comme origine dans les deux mondes, réel et virtuel.

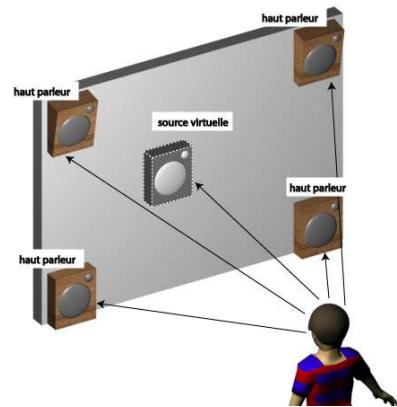


Figure 7: Modulation d’amplitude.

Pour placer les sources virtuelles, nous allons jouer sur la variation du volume sonore des enceintes. Le volume de sortie d’un haut-parleur  $i$   $h_{i,vol}$  est calculé comme suit: la position du son est définie par  $(s_p, s_\phi, s_\theta)$  avec  $s_p$  la distance,  $s_\phi$  la latitude et  $s_\theta$  l’azimut. La position des huit enceintes est définie par  $(h_{i,p}, h_{i,\phi}, h_{i,\theta})$ .

$$\left\{ \frac{-\pi}{2} \leq h_{i,\theta} - s_\theta \leq \frac{\pi}{2} \right\} \left\{ \frac{-\pi}{2} \leq h_{i,\phi} - s_\phi \leq \frac{\pi}{2} \right\}$$

L’azimut et la latitude sont comparés et si la condition précédente est vraie alors la relation suivante est appliquée pour obtenir le volume:

$$h_{i,vol} = s_{vol} \cdot |\cos(h_{i,\theta} - s_\theta)| \cdot |\cos(h_{i,\phi} - s_\phi)|$$

$s_{vol}$  étant le volume de la source. Enfin  $s_p$  permet enfin de moduler le volume sonore restitué en fonction de la distance entre la source virtuelle et l’utilisateur.

### 2.3. Synchronisation des évènements visuels et sonores

Il est primordial pour la réussite d’une application de réalité virtuelle, que d’une part les différents évènements sensoriels soient synchrones et d’autre part de maintenir le lien

de causalité entre les actions de l'utilisateur et les réactions du système. Pour ce faire notre système utilise deux threads, le premier gère les entrées, par exemple la capture de la position des Leds et la partie rendu graphique. Lors du déplacement de l'utilisateur ou d'objets dans la scène, un message est envoyé au second thread, le thread audio. Ce dernier gère, en fonction de la position relative de l'utilisateur par rapport aux sources sonores, la restitution du son ou non. Dans le cas où il y a restitution, le processus décrit au paragraphe précédent est exécuté (voir figure 6).

### 3. Interaction

Outre le fait que l'interactivité en temps réel soit une des conditions liée à la définition d'un système de réalité virtuelle, elle permet aussi d'accroître la sensation d'être immergé dans le monde virtuel. L'interactivité induit une connaissance de la position d'un utilisateur et des mouvements des parties de son corps par rapport au système. Plusieurs interfaces ont déjà été mise au point dans cette optique dont des traqueurs magnétiques, des gants de données ou des osquelettes.

Dans le contexte de notre salle de réalité virtuelle, nous avons fait le choix d'employer des interfaces visuelles sous la forme de diodes électroluminescentes sur lesquelles nous appliquons un tracking. Ces dernières ont en effet l'avantage d'être légères ce qui évite à l'utilisateur de se munir d'un appareillage trop contraignant (poids, résistance,...) et qui diminuerait la sensation d'immersion.

Notre système est composé de deux caméra placées de manière à ce qu'elles couvrent la plus grande zone possible devant l'écran et que l'on puisse suivre les déplacements de l'utilisateur. Les caméras sont préalablement calibrées au moyen d'outils disponibles dans la bibliothèque OpenCV. A partir des images transmises par ces dernières nous faisons une mise en correspondance des sources lumineuses en nous servant des données chromatiques ou géométriques (diodes positionnées de manière bien spécifique). La dernière étape consiste alors à croiser ces informations et les paramètres des caméras pour en déduire la position dans l'espace de chacune de ces sources lumineuses.

L'utilisateur a à sa disposition un certain nombre d'outils lui permettant d'interagir avec l'environnement virtuel dans lequel il évolue. Nous avons équipé une paire de lunettes polarisantes de diodes pour, dans un premier temps, évaluer la position de l'utilisateur dans notre système de réalité virtuelle et dans un second temps de connaître l'orientation de sa tête pour, par exemple, estimer la direction de son regard. Un autre exemple est l'utilisation d'un stylo modifié par l'ajout de deux diodes qui permet à un utilisateur d'effectuer, dans le cadre d'une de nos applications, un picking de points de contrôle pour modifier la géométrie d'un objet. Il devient simple d'imaginer que n'importe quel objet (pistolet en plastique, pinceau...) peut devenir une nouvelle interface de notre système à partir du moment où nous

l'équipons de diodes. De quelque manière que ce soit, ces appareillages sont conçus dans l'esprit de les rendre le plus naturel possible pour l'utilisateur.

Dans le cadre de JIM 3D nous avons créé une nouvelle interface nous permettant d'agir sur l'univers virtuel. Un seul outil, un gant sur lequel des diodes ont été appliquées, est utilisé pour effectuer une rotation du point de vue et lancer des événements associés à des objets spécifiques présents dans la scène. Les diodes permettent de calculer la position et la direction d'un des doigts de la main (l'index par exemple).

La jouabilité de JIM 3D est similaire à celui du très célèbre jeu vidéo d'aventure "MYST 3 : Exile", c'est à dire que notre application offre un système de vision à 360° depuis une position fixe. Pour modifier le point de vue et conserver l'effet immersif, l'utilisateur doit effectuer un mouvement horizontal de sa main et quasiment parallèle à l'écran. Ce geste vers effectué la droite provoque une rotation vers la droite et vice versa. De cette manière l'utilisateur a l'impression de maîtriser naturellement le changement de point de vue.

Dans les différentes scènes que l'on peut créer, il est possible d'interagir sur des objets dans le but, par exemple, de les modifier ou de les utiliser. Pour que l'utilisateur puisse lancer ce genre d'événements, il doit pointer l'objet du doigt et s'il maintient cette pose au moins une demi-seconde alors l'action est déclenchée. Pour l'aider dans cette phase, un pointeur est affiché à l'écran ce qui permet à l'utilisateur de connaître la zone visée. Dans un scénario utilisé par JIM 3D, l'utilisateur a la possibilité de presser les touches d'un piano, choisir de se déplacer dans un autre lieu ou manipuler divers objets.



**Figure 8:** Diodes fixées sur les lunettes pour détecter l'inclinaison de la tête.

### 4. Applications

L'un des premiers essais réalisés dans la salle de réalité virtuelle, est la restitution en relief de tableau grâce à une carte de profondeur. Cette incursion dans les arts a donné lieu à un outil de peinture virtuelle en quatre dimensions, où les couches de peinture pouvaient se mélanger lorsqu'un



**Figure 9:** Exemple d'outils modifiés.

pinceau virtuel repassait sur une couche précédente si celle-ci n'avait pas séché (intégration du facteur temps). Toujours dans le domaine de l'art, l'Amorphous Ubiquity [AP04] est une installation interactive en temps réel où l'utilisateur est immergé dans un labyrinthe en trois dimensions en perpétuel mouvement. Ce dernier est le résultat de l'influence d'un dialogue gestuel, sonore et corporel de l'utilisateur avec le système.

Un environnement virtuel mérovingien a aussi vu le jour dans le cadre d'une collaboration entre l'équipe de recherche et la mairie de la commune hébergeant notre laboratoire. Celle-ci ayant mis à jour l'un des sites mérovingiens les plus importants d'Europe au cours de fouilles archéologiques, nous a demandé d'en faire une reconstitution méticuleuse. Cette espace ayant pour but à la fois de mettre en valeur auprès des habitants l'histoire de leur commune mais aussi de fournir aux archéologues un outil pour évaluer ou proposer à leurs confrères des hypothèses de travail.

La salle permet aussi de diffuser des films en 3D. Ces derniers, filmés avec deux caméras correctement écartées et orientées sont visionnés de façon synchronisés via les deux vidéo projecteurs.

Le pantin est une application permettant de piloter un avatar à partir d'une extraction de la silhouette de l'utilisateur provenant de l'une de nos deux caméras. La personne filmée sur un fond blanc est détournée en temps réel et son squelette de contrôle est extrait et sert à piloter l'avatar.

L'une des dernières applications développées pour la salle de réalité virtuelle est un jeu interactif semblable à "Myst" [MYS] (voir figure 10). L'intérêt de ce développement fut d'expérimenter la restitution des scènes, non pas via un ensemble d'objets en 3D mais via une série de photographies. Si ce processus était déjà connu sous le terme de mosaïque, il n'avait jamais été expérimenté dans le cadre d'une restitution panoramique en stéréo. Dans la pratique les scènes sont réalisées avec des logiciels type 3DSMax puis un script construit un cylindre d'images à partir de plusieurs centaines de rendu. Dans le jeu, lors du déplacement de l'utilisateur, les deux photographies les plus pertinentes sont utilisées pour former une image stéréoscopique. On peut ainsi obtenir en

temps réel (plus de 25 images par seconde) des images dont la qualité aurait nécessité un temps de calcul plus important.



**Figure 10:** Une application ludique.

## 5. Conclusion

Nous avons présenté dans cet article notre salle de réalité virtuelle qui présente les qualités d'être peu coûteuse, facilement transportable et évolutive car, excepté l'écran, elle n'est composé que de matériel grand public. De plus, du point de vue logiciel, elle n'utilise que des développements opensource disponibles sous Linux.

L'intérêt de cette salle est d'offrir la possibilité d'immerger visuellement et de façon sonore un utilisateur et de lui proposer d'interagir avec des interfaces intuitives. Ainsi cette salle s'inscrit parfaitement dans le monde du jeu car elle tend à rendre le gameplay intuitif et amplifie les sensations du joueur.

Enfin la réalité virtuelle ne doit pas se limiter à stimuler de façon disjointe les sens de l'utilisateur mais doit proposer une perception multisensorielle. La finalité serait de proposer de nouvelles sensations, nées de l'intersensorialité, quasi synesthésiques.

## References

- [AP04] ALMIRON M., PIRANDA B.: Ubicuidad amorfa. *CiberArt Bilbao 2004* (avril 2004).
- [Car96] CARLILE S.: Virtual auditory space: Generation and applications. *New York: RG Landes* (1996).
- [CNSF\*92] CRUZ-NEIRA C., SANDIN D. J., FANTI T. A. D., KENYON R. V., HART J. C.: The cave: Audio visual experience automatic virtual environment. *Commun. ACM Siggraph 35(6)* (1992), 64–72.
- [Dav] DAVIS P.: Jack audio connection kit. <http://jackaudio.org/>.
- [FMO] Fmod. <http://www.fmod.org/>.
- [Har] HARRIS S.: Swh plug-ins. <http://plugin.org.uk/>.
- [KBF\*95] KRUEGER W., BOHN C., FROHLICH B., SCHEUTH H., STRAUSS W., WESCHE G.: The responsive workbench. *IEEE Computer 28, 7* (1995).

[LAD] Linux audio developer's simple plugin api (ladspa).  
<http://www.ladspa.org/>.

[MD76] MCGURK H., DONALD J. M.: Hearing lips and seeing voices. *Nature* (December 1976), 746–748.

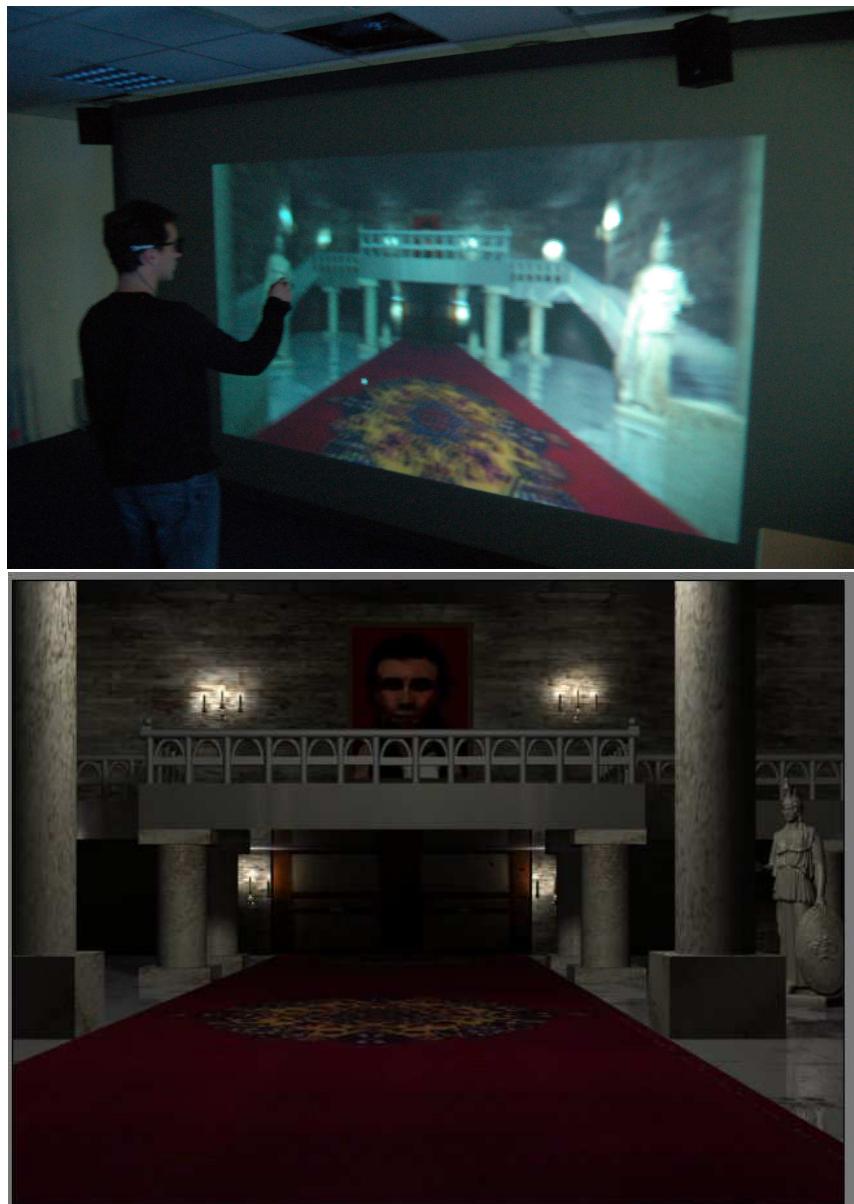
[MYS] Myst. <http://www.myst.com/>.

[Ope] Open audio library (openal).  
<http://www.openal.org/>.

[PM96] PROAKIS J. G., MANOLAKIS D. G.: *Digital Signal Processing, principles algorithms, and applications*. Prentice Hall International Editions, 1996.

[Ral07] RALEIGH L.: Our perception of sound direction. *Philosophical Magazine* (1907).

[WK97] WIGHTMAN F. L., KISTLER D. J.: Monaural sound localization revisited. *J. Acoust. Soc. Am* 101 (February 1997), 1050–1063.



**Figure 11:** *Une application ludique.*