

**“SMARTPAN” : UN NOUVEL OUTIL DE
SPATIALISATION DU SON DANS UN ESPACE
3 DIMENSIONS**

Lucas THIRON

Mémoire de Master 2 - Spécialité Son

Directeur.rice de mémoire interne : Sylvain LAMBINET

Directeur.rice de mémoire externe : Cyril HOLTZ

Responsable universitaire : Corsin VOGEL

Rapporteur.e.s : Franck JOUANNY

Novembre 2021

Résumé

Depuis le début des années 2010, le cinéma a connu une évolution technique qui a eu un impact important sur la chaîne de fabrication sonore des films : l'utilisation d'un canal de diffusion zénithal et donc de la spatialisation 3D, principalement à travers le Dolby Atmos. Tandis que cette évolution conquiert petit à petit les salles de cinéma, les outils permettant de spatialiser le son en 3 dimensions n'ont quant à eux pas encore évolué de manière efficace et significative.

En effet, on retrouve toujours aujourd'hui les classiques joysticks et potentiomètres sur les consoles de mixage, et ceux-ci sont inadaptés à un mixage 3 dimensions. Bien que ce domaine évolue et que le nombre de solutions et de prototypes augmente, les concepteurs de console n'ont majoritairement pas encore passé le cap de les intégrer à leur matériel.

Ce mémoire a pour but d'explorer l'inadéquation des outils utilisés aujourd'hui pour spatialiser des sons en 3 dimensions lors de mixages cinématographiques, ainsi que de parcourir les outils novateurs qui existent, mais également les raisons qui font qu'ils ne sont toujours pas utilisés. L'objectif sera ensuite d'essayer de proposer un nouvel outil ergonomique, simple, utilisable à une seule main, et avec le coût le plus réduit possible.

Mots clefs: Spatialisation; 3 dimensions; Smartphone; Gyroscope

Abstract

Since the beginning of the 2010s, the cinema has undergone a technical evolution which has had a significant impact on the sound production chain of films: the use of a zenithal diffusion channel and therefore of 3D spatialization, mainly through the Dolby Atmos. While this development is gradually conquering cinemas, the tools for spatializing sound in 3 dimensions have not yet evolved in an effective and significant way.

In fact, the classic joysticks and potentiometers are still found today on mixing consoles, and these are unsuitable for a 3-dimensional mixing. Although this field is evolving and the number of solutions and prototypes is increasing, most console designers have not yet passed the step of integrating them into their hardware.

This thesis aims to explore the inadequacy of the tools used today to spatialize sounds in 3 dimensions during cinematographic mixes, as well as to explore the innovative tools that exist but also the reasons why they are still not used. The objective will then be to try to offer a new tool that is ergonomic, simple, usable with one hand, and at the lowest possible cost.

Key words: Spatialization; 3 dimensions; Smartphone; Gyroscope

Remerciements

Avant de présenter mon travail je tiens à remercier tous ceux qui ont permis à ce mémoire de voir le jour.

Cyril HOLTZ, pour son accompagnement tout au long de ce projet et pour le partage de son expérience et qui a toujours su prendre du temps pour se téléphoner malgré son emploi du temps chargé.

Sylvain LAMBINET, pour son accompagnement depuis les premières idées de ce mémoire et pour ses solutions lors de blocages dans MAX.

Antoine MARTIN, pour les réflexions sur le mémoire et sur la manière de le réaliser qu'il m'a apportées.

Corsin VOGEL, pour son encadrement sérieux et engagé et pour son positivisme tout au long de l'apprentissage de cet exercice.

Les VVD Hugo et Loryne, pour leur soutien indéfectible et leur amitié sincère, même à distance.

Matéo ROUSSON, pour son ouverture d'esprit et ses discussions toujours plus intéressantes.

Aloïs DARLES, pour m'avoir supporté dans moult situations.

Merci à Florent FAJOLE pour ses conseils de lectures avisées, merci à l'équipe d'enseignement du Master SON pour leurs enseignements durant ces trois années. Merci à la promotion Son 2021. Merci à ma famille pour m'avoir toujours soutenu dans mes choix et avoir toujours eu confiance en moi.

Et merci à Emma BELY, pour la confiance qu'elle m'accorde et pour son amour et sa joie de vivre qui me permettent toujours d'avancer.

Table des matières

Introduction	7
La spatialisation dans le cinéma	9
De la monophonie au multicanal	9
De nouveaux systèmes : Le Dolby Atmos, l'Auro 3D, la WFS et le DTS:X	18
La spatialisation 2 dimensions	25
La spatialisation 3 dimensions	27
Les bases du mixage orienté objet	28
L'utilité d'un panoramique en élévation	31
Outils 1D/2D / Outils utilisés aujourd'hui	33
Potentiomètre	33
Joystick	35
Ecran Tactile	36
Problèmes	37
Outils 3D existants	39
Gant P5 (P5 glove)	39
SpaceMouse	40
Leap Motion	41
DearVR Spatial Connect	42
“SmartPan”	45
Gyroscope	45
Prototype sur MaxMSP	46
Essais	54
Conclusion	56
Références Cinématographiques	58
Références Bibliographiques	60
Webographie	61
Annexes	62

Introduction

Aujourd'hui, pour mixer du son dans un espace en 3 dimensions (selon donc 3 variables x , y et z qui sont les coordonnées plan et d'élévation), les mixeurs ont à leur disposition principalement des potentiomètres, des joysticks, un écran tactile et une souris. Ces outils sont des outils à une dimension ou deux dimensions. De plus, il est donc obligatoire d'utiliser deux outils à la fois (deux joysticks ou un joystick et un potentiomètre par exemple), ce qui n'est absolument pas ergonomique.

Il est donc nécessaire aujourd'hui de développer un nouvel outil permettant de placer une source sonore dans un espace 3 dimensions en utilisant uniquement une main et de manière intuitive.

Cette question m'était déjà venue par le passé et j'ai donc voulu réfléchir plus profondément à cela et essayer de trouver une solution pour pallier à ce problème.

J'ai donc premièrement réfléchi à l'inadéquation des outils actuellement utilisés et ai vite compris que lors d'une volonté de déplacement d'une source sonore traversant l'espace 3D (dans le plan et l'élévation) par exemple, il est très complexe de faire un mouvement fluide en une seule passe d'automatisation avec les deux mains. En effet, cette action comprise comme unifiée par le cerveau se voit décomposée en 2 actions à effectuer avec 2 mains différentes. Il est donc bien plus complexe en termes de cohérence cognitive et d'ergonomie d'effectuer cela simplement. Les mixeurs sont donc obligés de faire une première passe diagonale dans le plan puis une deuxième concernant l'élévation. Or, cela est coûteux en temps et en énergie.

J'ai donc ensuite essayé d'imaginer des systèmes mécaniques permettant un déplacement dans un espace 3 dimensions en un seul mouvement avec une seule main.

Il m'est venu 3 idées :

- Un joystick possédant également une possibilité de mouvement vertical qui pourrait donner une dimension d'élévation.
- Un bras articulé en 3 axes qui permettrait un mouvement fluide directement dans 3 dimensions.
- Une sorte de stylet permettant de pointer le placement de la source sonore dans un espace parallélépipédique.

Mais je me suis confronté très vite à de grandes complications à la fois de faisabilité, de coût, de temps de réalisation et tout simplement d'efficacité du système. J'ai également découvert plusieurs systèmes de contrôleurs audio 3 dimensions (P5 Glove, SpaceNav, Leap Motion, DearVR) dont nous développerons les bons et mauvais côtés quant à leur utilisation et leur fabrication. Mais j'ai pris la décision d'essayer de trouver un nouvel outil encore inutilisé pour le mixage et impliquant le moins de coût d'achat possible afin de le rendre accessible à tous. C'est ainsi que nous sommes parvenus à l'idée d'utiliser un smartphone.

En effet, cette solution permet premièrement un gain de temps et d'argent considérable car il n'est alors question que de développement logiciel et non plus matériel. De plus, l'outil est déjà existant, présent dans la poche de la majorité des personnes et déjà pris en main par ces personnes. Il n'y a donc que très peu d'apprentissage à effectuer et très peu d'achats également. Le smartphone pourrait permettre de remplir tous les paramètres dont nous avons besoin : Un seul outil, utilisable à une main, ergonomique et intuitif, sans besoin de construction matérielle, et peu coûteux.

L'idée serait donc d'utiliser le gyroscope intégré aux smartphones couplé à une application. Cela permettrait d'obtenir 3 coordonnées x, y et z qui seraient ensuite en communication avec un plug in de spatialisation qui est à développer également car le panner de ProTools est malheureusement très fermé et cela permettrait également de l'utiliser avec d'autres DAW que ProTools.

Je vais donc développer une application mobile permettant d'extraire des coordonnées x, y, z des mouvements de l'utilisateur, ainsi qu'un plug-in permettant de récupérer ces données et d'en faire un panning efficace permettant de placer le son dans un espace 3D.

Le but de ce mémoire est de réfléchir à l'inadéquation des outils utilisés aujourd'hui lors d'un mixage 3 dimensions, ainsi qu'analyser les nouveaux outils existants pour enfin proposer une nouvelle solution à ceux-ci.

1. La spatialisation dans le cinéma

- De la monophonie au multicanal

Depuis les prémices du cinéma, il existe une volonté de représenter le réel le plus fidèlement possible. Premièrement, grâce à l'image puis, par l'ajout du son direct lors de la sortie de l'iconique *The Jazz Singer* d'Alan Crosland (CROSLAND, 1927) en 1927, adopté comme repère historique de l'apparition du son au cinéma. Alors diffusé en monophonie, il existait déjà une volonté de diffusion multicanal afin de s'approcher encore plus d'une représentation fidèle de la réalité.

Durant les années 30, les supports optiques d'enregistrements étant limités à un canal audio, la diffusion sur divers haut-parleurs se faisait par un opérateur qui basculait le son monophonique sur les divers canaux lui étant mis à disposition. Cela permettait, entre autres, de diffuser le son des dialogues derrière l'écran et le son musical dans la fosse, afin de recréer la sensation connue des salles de cinéma de cette époque. Nous pourrions noter que, bien que complètement dans une autre optique, les premiers pas du mixage orienté objet¹ étaient posés, quoique un peu tôt. Cette tâche fastidieuse fut abandonnée rapidement au profit d'une diffusion des dialogues et de la musique dans le même haut-parleur.

Le premier film comprenant une diffusion multicanal fut *Napoléon* d'Abel Gance projeté le 11 mai 1935 dans la salle du Paramount à Paris avec l'aide d'André Debie (GANCE, 1935). Ce projet ambitieux et coûteux à l'opposé de l'unique enceinte installée alors dans les cinémas, était bien loin de la diffusion multicanal

¹ Le mixage orienté objet est une forme récente de mixage qui permet de séparer les natures de sources sonores de leurs informations de position, nous expliciterons ce principe plus tard dans ce mémoire

connue de nos jours. En effet, l'image composée d'un triptyque (en Polyvision) était accompagnée d'un haut parleur derrière chaque écran. Mais cela ne s'arrête pas là car dans la salle étaient disséminés 32 phonographes asynchrones diffusant des effets sonores selon une partition préalablement écrite.

« Cette sonorisation se révéla bouleversante lorsque le film fut présenté, pendant son exclusivité parisienne, avec cette fameuse « perspective sonore » enfin dévoilée au public. Celui-ci, littéralement enveloppé de sons, se trouvait projeté au cœur de l'action et devenait acteur à son tour [...] la rue en révolution prenait possession de la salle où les clameurs retentissaient de partout ; la tempête rugissait et déferlait avec rage sur elle et chaque spectateur se trouvait emporté, comme Bonaparte, dans le tourbillon des éléments déchaînés. ». (ICART, 1983)



*Figure 1 - Projection de "Napoléon" d'Abel Gance , avec ses trois écrans.
(Paramount, 1927)*

Ce système ne put se pérenniser, premièrement à cause des coûts humains (il fallait un opérateur par phonographe connaissant extrêmement bien le film), puis des coûts matériels (les supports de stockage étaient onéreux et de faible capacité) et enfin, cela n'allait aucunement dans la logique d'exploitation cinématographique de l'époque.

C'est en 1940 que le multicanal fait réellement son apparition dans les salles de cinéma avec deux systèmes.

Le premier, le *Vitasound* est une amélioration du principe de commutation existant sauf que la commutation ne se fait plus par un opérateur mais par une piste optique de contrôle qui commute la piste audio entre l'enceinte frontale et les enceintes arrières.

Le second, le *Fantasound*, inventé par William Garity, alors technicien dans la compagnie Walt Disney, a été développé à l'occasion de la sortie du film *Fantasia* réalisé par Leopold Stokowski. Dans son article de 1941 dans le Journal of the Society of Motion Picture Engineers, William GARITY nous explique qu'il comprenait 3 canaux destinés à alimenter les hauts parleurs placés derrière l'écran ainsi que 2 canaux alimentant d'autres hauts-parleurs placées de part et d'autre de la salle (notons la ressemblance avec l'actuel 5.1). Mais ce dispositif nécessitant un grand nombre d'enceintes (de trente à quatre-vingt en fonction des salles), il ne put être projeté que dans six salles aux Etats-Unis et fut diffusé dans le standard monophonique de l'époque pour les autres salles ce qui entraîna son échec.

Au début des années 50 apparaît la bande magnétique, plus facile à utiliser et de meilleure qualité sonore que la bande optique. Avec la meilleure qualité des enceintes, cela permettra de nouvelles inventions de systèmes sonores souvent associés à un système de prise de vue, et c'est en 1952 que nous avons pu voir apparaître une nouvelle avancée dans les systèmes de diffusion multicanal : le *Cinérama*, présenté à l'occasion de la sortie du film *This is Cinerama* (COOPER et al., 1955). Le Cinérama est un procédé à quatre projecteurs (trois pour l'image et un pour le son) et trois grands écrans incurvés (146° d'ouverture). Cinq enceintes de façade et deux enceintes de surround forment sept canaux dont, pour la première fois, des canaux Inter-Gauche et Inter-Droit². Ce système étant toujours très coûteux pour les salles de cinéma, il n'a pu se pérenniser et seulement une dizaine de films utilisant ce système ont vu le jour jusqu'aux années 60.

L'URSS propose en 1957 un système très proche : le Kinopanorama. Le procédé est également basé sur 3 projecteurs mais possède 9 canaux audio au lieu

² C'est-à-dire entre l'enceinte gauche et l'enceinte centrale et entre l'enceinte droite et l'enceinte centrale

de 7 en ajoutant un canal central-arrière et un canal zénithal au-dessus des spectateurs.

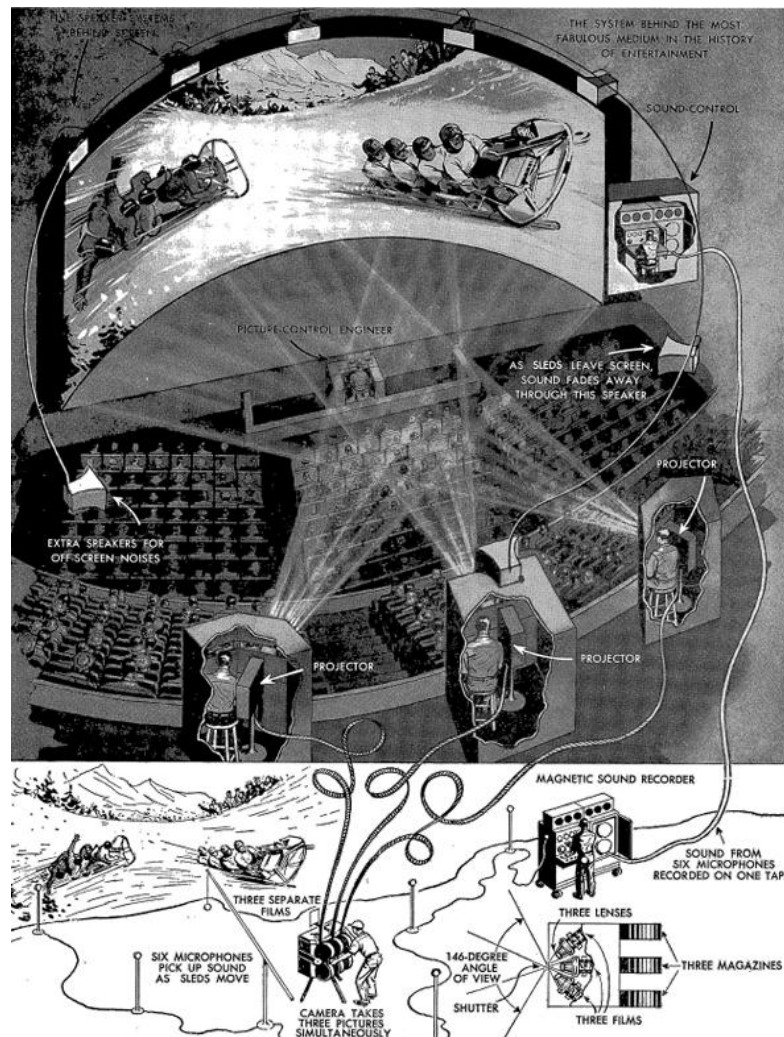


Figure 2 - Présentation du système Cinerama (Waller, 1952)

Entre 1950 et 1970, plusieurs formats de diffusion voient le jour. Nous pouvons retenir par exemple le Todd-AO (nom provenant du créateur du système Michael Todd et d'American Optical, la société d'optique qui l'exploite) ou le SuperPanavision 70, tous les deux utilisant des pellicules 70mm. Ils emploient un dispositif ressemblant à celui du Cinerama pour le son : il est basé sur six pistes magnétiques dont cinq alimentent cinq canaux frontaux et la dernière, un canal arrière. Pour les systèmes de ces années, dont la caractéristique commune est un élargissement du champ de vision du spectateur, il paraissait nécessaire aux ingénieurs de coordonner point visuel et point sonore, ce qui entraîna souvent le

placement de cinq enceintes derrière l'écran afin de couvrir toute la largeur de celui-ci. Mais ces procédés demandant encore beaucoup de changement et de coûts (la norme de l'époque est le 35mm) pour les exploitants de salle de cinéma, un autre système s'est alors imposé : le Cinemascope.

En 1953, la 20th Century Fox présente *La Tunique* d'Henri Koster (KOSTER, 1953) en Cinémascope. Ce procédé d'anamorphose permettant d'obtenir un ratio d'image de 2,35:1 à partir d'une pellicule 35mm désanamorphosée, fut inventé en 1926 par Henri Chrétien. La bande magnétique contient alors 4 pistes composant un LCRS³, et le procédé d'anamorphose permet aux exploitants de salles de seulement acheter une lentille plutôt qu'un projecteur ce qui en fait en partie son succès. Plus de 10000 cinémas sont équipés pour diffuser du Cinémascope à la fin des années 50.

Jusqu'aux années 70, le cinémascope reste le système le plus utilisé avant l'apparition de nouveaux procédés.

Le Dolby Stéréo (ou Dolby Surround) viendra, en 1976, remplacer ces anciens procédés et uniformiser ainsi que standardiser la diffusion sonore au cinéma. La firme américaine Dolby est alors connue pour ses réducteurs de bruit. Mais elle propose alors un système matriciel qui permet d'obtenir 4 canaux à partir de seulement 2 pistes optiques appelées Lt et Rt, et offrant donc la possibilité de diffuser un mixage LCRS à moindre coût pour les studios. Les cinémas étant déjà équipés pour du Cinémascope, il ne leur reste qu'à obtenir un décodeur Dolby Stéréo. Ajoutons à cela que les enceintes Inter-Gauche et Inter-Droite sont ré-utilisées comme renfort de basse dans la version 70 mm et que ce système est entièrement compatible avec les salles encore équipées en Mono et l'on obtient un format proposant pour la première fois peu de changements pour les salles de cinéma, qui sera vite adopté par la majorité. La sortie en 1977 de *Star Wars* de George Lucas (LUCAS, 1977) concrétise l'utilisation du Dolby Stéréo.

³ Left Centre Right Surround

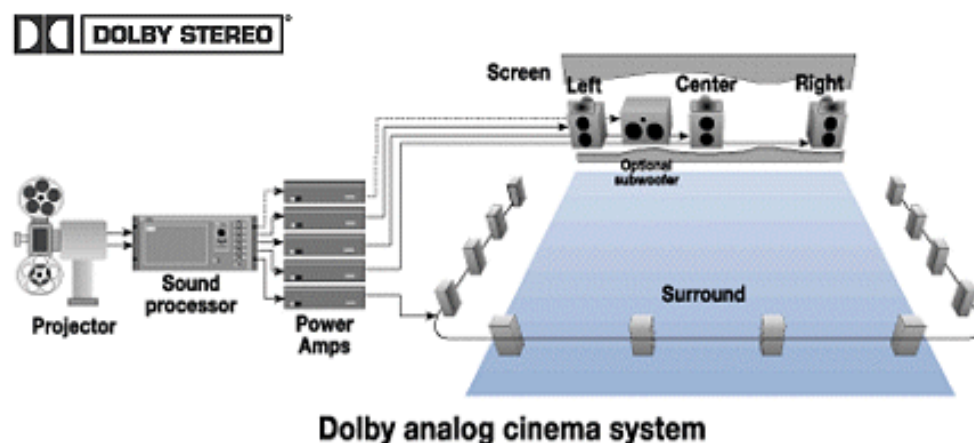


Figure 3 - Configuration pour une salle au format Dolby Stéréo (Dolby, 2006)

En 1987, la SMPTE (Society of Motion Picture and Television Engineers) travaille sur le développement d'un nouveau standard pour le son numérique au cinéma. Accompagnant la démocratisation du Compact Disc, ce standard a deux objectifs principaux :

- L'utilisation de canaux discrets uniquement, sans matricage.
- Une qualité sonore au moins égale à celle du Compact Disc.

“Tomlinson Holman⁴ fait donc une proposition de normalisation du système 5.1 qui, sans le savoir, va permettre de standardiser l'écoute cinéma jusqu'à aujourd'hui. Ce système présente trois canaux frontaux et deux canaux arrières (tous les cinq étant pleine bande), mais utilise également un canal supplémentaire réservé aux fréquences basses, le LFE (Low Frequency Effect).” (DENIZOT, 2016)

Trois principaux systèmes vont cohabiter par la suite pendant plus de 10 ans : Le Dolby SR-D⁵ créé en 1992 pour *Batman Returns* de Tim Burton (BURTON, 1992), le DTS (Digital Theater Systems) et le SDDS (Sony Dynamic Digital Sound) créés tous deux en 1993 pour les sorties respectives de *Jurassic Park* de Steven Spielberg (SPIELBERG, 1993) et de *Last Action Hero* de John McTiernan (MCTIERNAN, 1993). Ces trois systèmes proposent une piste optique Dolby Stereo pour assurer une rétrocompatibilité.

⁴ Professeur de techniques cinématographiques à l'University of South California et inventeur de la charte THX

⁵ Dolby Spectral Reduction-Digital qui deviendra plus tard le Dolby Digital

Le Dolby Digital ainsi que le DTS offrent un mixage 5.1 mais de manières différentes. Dolby utilise directement la pellicule pour y inscrire l'information sonore avec une compression en AC-3 (compression 1/12) tandis que DTS compresse le son en $\frac{1}{3}$ et obtient donc une meilleure qualité sonore en plaçant l'information sonore sur un CD dont le lecteur externe est synchronisé à l'image par une piste de TimeCode inscrite sur la pellicule. Le système DTS coûte également moins cher à l'installation.

Le SDDS, de son côté, inscrit également l'information sonore sur la pellicule mais propose un mixage 7.1 avec des canaux Inter-Gauche et Inter-Droit. Sony mise alors sur la compatibilité avec les anciennes salles équipées pour les projections 70mm, bien que ce système coûte plus cher.

Les trois concurrents sont soutenus par des studios concurrents également : Columbia soutient le SDDS, Universal soutient le DTS tandis que la Paramount et Disney soutiennent Dolby. Dans tous les cas, les trois systèmes utilisent une partie différente de la pellicule donc il est possible d'imprimer les trois sur une seule copie pour permettre aux exploitants d'être sûrs de pouvoir projeter le film.

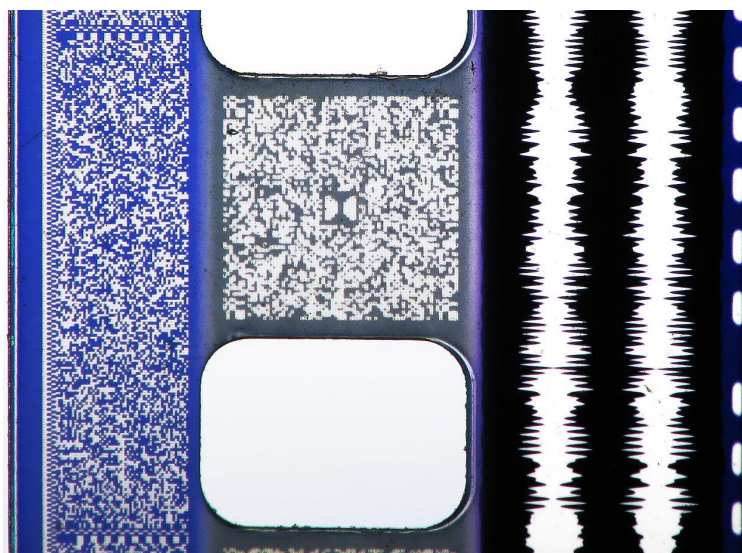


Figure 4 - Pellicule 35 mm comportant de gauche à droite les pistes audio : SDDS, Dolby Digital, Dolby Stéréo classique et le TimeCode du DTS sur le bord droit.

(Creative Commons Licensed, 1993)

Les années 1990 sont une période de multiplication des copies possibles de film et du nombre de cinémas (qui ouvrent alors équipés pour le son numérique). Tout cela va permettre au 5.1 de se démocratiser à grande échelle.

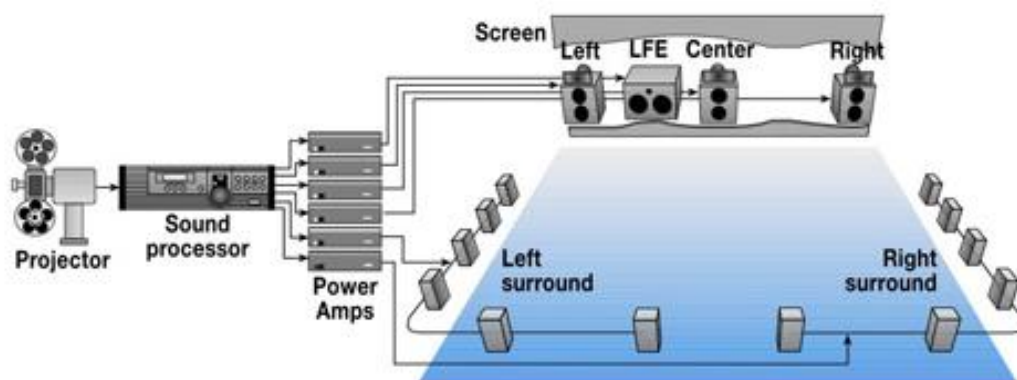


Figure 5 - Configuration pour une salle au format Dolby Digital (Dolby, 2006)

Entre 1993 et 2006, DTS et Dolby font évoluer régulièrement leurs procédés 5.1 en y ajoutant les uns après les autres, des canaux, toujours selon leurs procédés matriciels et d'encodage propres. Bien que Dolby et DTS soient à peu près égaux en nombre de salles équipées de leurs systèmes et de films à leurs formats (Dolby étant tout de même légèrement en avance), SDDS est en retard. La firme arrête la vente de son système aux exploitants au début des années 2000, après avoir perdu petit à petit du terrain face à ses concurrents. Pour Dolby, les années 2000 sont marquées par une réédition de leur système Pro Logic en Dolby Surround, Pro Logic II, Pro Logic IIx, Pro Logic IIz. Ces procédés permettent d'étendre le nombre de canaux d'un programme stéréophonique en les reconstituant par matricage.

A partir de 2006, le cinéma numérique se déploie dans le monde et les avancées technologiques connexes permettent désormais d'intégrer 16 canaux audio dans un seul DCP (Digital Cinema Package), et ce, sans compression de données. Dolby profite de ce contexte pour présenter son nouveau système Dolby Surround 7.1 à l'occasion de la sortie du film *Toy Story 3* de Lee Unkrich en 2010 (UNKRICH, 2010). Différemment du SDDS, il est question ici de séparer les deux

canaux surround en quatre canaux Lss et Rss (Left/Right Side Surround) et Lrs et Rrs (Left/Right Rear Surround).

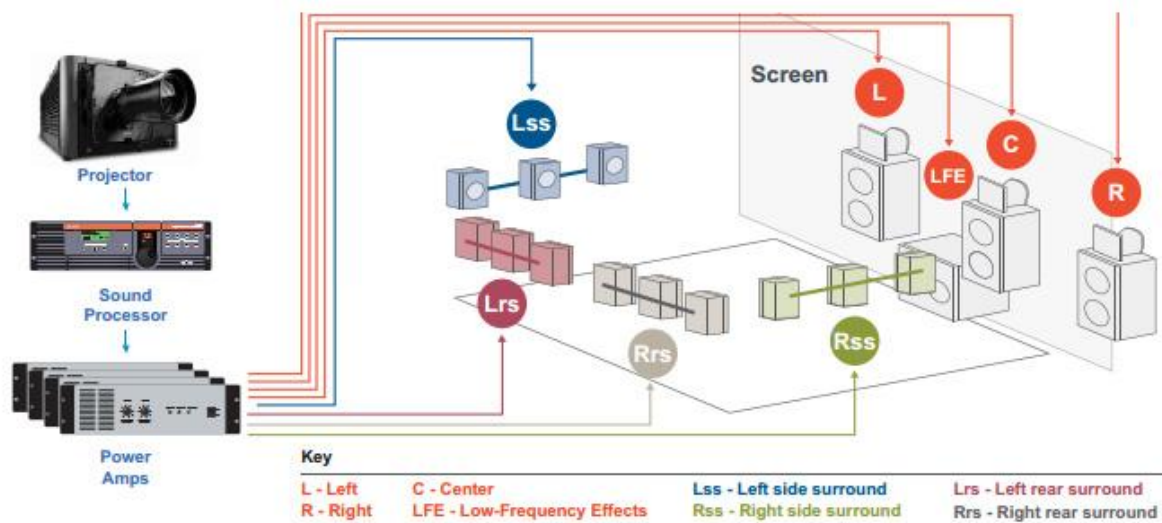


Figure 6 - Configuration pour une salle au format Dolby Surround 7.1

Bien que le format 5.1 reste aujourd'hui le standard SMPTE⁶, le nombre de salles équipées d'une diffusion 7.1 a explosé. Mais ce n'est pas tout car, de la même manière que le cinéma évolue sur les formats d'images avec la 3D, ou encore l'IMAX, de nouveaux systèmes sonores sont apparus et offrent de nouvelles perspectives pour les bandes sonores de films.

⁶ Society of Motion Picture and Television Engineers

- De nouveaux systèmes : Le Dolby Atmos, l'Auro 3D, la WFS et le DTS:X

- L'Auro-3D

L'Auro-3D est un format de diffusion développé en 2005 par la société belge Auro Technologies et appartient maintenant à la société Barco. Le principe de base est d'ajouter au format classique 5.1 une dimension verticale tout en restant dans un principe *channel-based* (ce modèle est plus précisément expliqué par la suite).

Dans sa version classique de cinéma (11.1) l'Auro-3D est composé :

- d'un niveau "bas" : un 5.1 classique
- d'un niveau "haut" : un 5.0 situé environ 30° au dessus du niveau bas
- d'un canal "top" ou "voice of god" situé au plafond au-dessus des spectateurs.

Il existe une version 13.1 également où sont ajoutés les canaux latéraux du 7.1 de Dolby, ce qui permet une adaptation aux salles déjà équipées de la sorte.

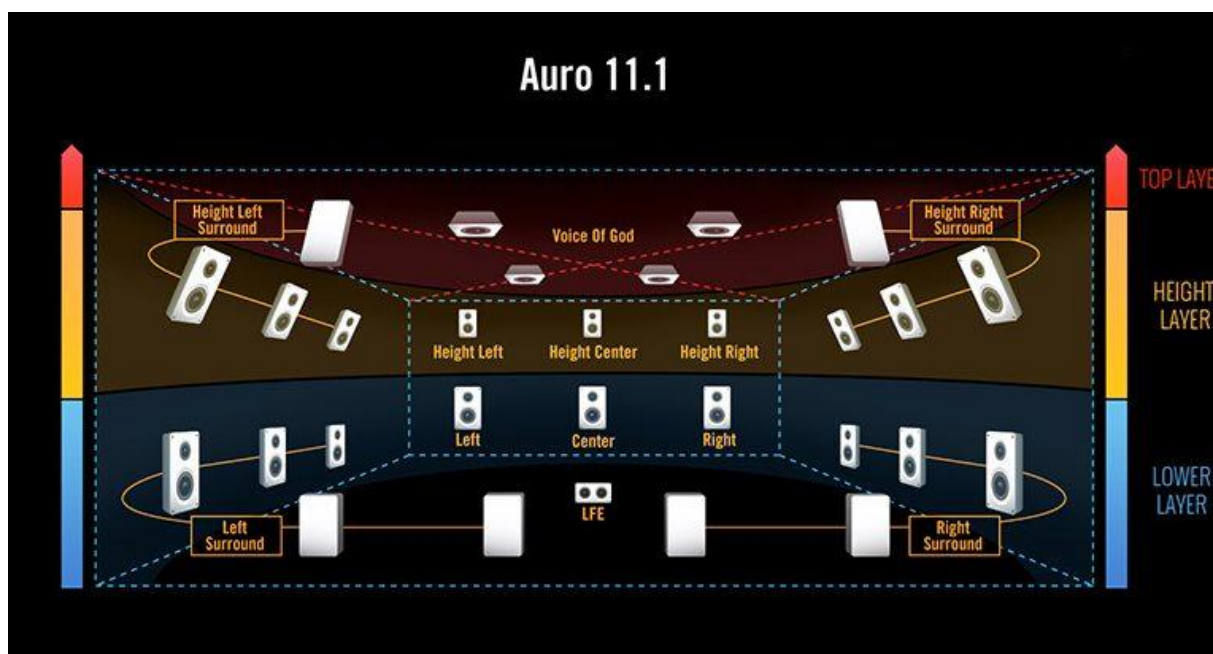


Figure 7 - Configuration Auro 3D 11.1 (Auro-3D, 2006)

Aujourd'hui l'Auro-3D est présent dans plusieurs centaines de salles dans le monde et cette technologie a été utilisée pour plusieurs films tels que *Red Tails* de Anthony Hemingway (HEMINGWAY, 2012), *Rise of the Guardians* de Peter Ramsey (RAMSEY, 2012), ou encore *The Croods* de Kirk DeMicco et Chris Sanders (DEMICCO et SANDERS, 2013).

Comme DTS et Dolby auparavant, l'Auro 3D possède son système "home-cinéma" avec des versions plus légères de ses procédés en 9.1 ou 10.1.

- Le Dolby Atmos

Parmi les nouveaux systèmes, le plus implanté dans le cinéma mondial est le Dolby Atmos. Présenté en 2012 par Dolby, le premier film diffusé dans ce format est *Rebelle* des studios Disney en 2012 (ANDREWS et CHAPMAN, 2012). Mais c'est *Gravity* d'Alfonso Cuarón en 2013 (CUARÓN, 2013) qui assoit la popularité de ce système. Aujourd'hui, avec plus de 2000 salles équipées dans le monde dont plus de 75 en France, Dolby possède une avance considérable face à ses concurrents, et commence à s'imposer comme futur monopole du son 3D dans les salles de cinéma.

Comme ses concurrents principaux l'Auro-3D et le DTS:X, le Dolby Atmos a pour vocation d'ajouter une dimension verticale dans la diffusion. Le placement des enceintes diffère de l'Auro-3D : Horizontalement il n'y a toujours qu'une rangée d'enceintes mais faisant un tour complet de la salle (ajout de surround entre l'écran et les surrounds déjà présents), il y a également un ajout d'enceintes inter-gauche et inter-droite ainsi que 2 rangées d'enceintes au plafond (ces enceintes sont séparées de la même manière que les enceintes latérales : entre 2 et 3 m) et 2 subwoofers vers l'arrière de la salle.

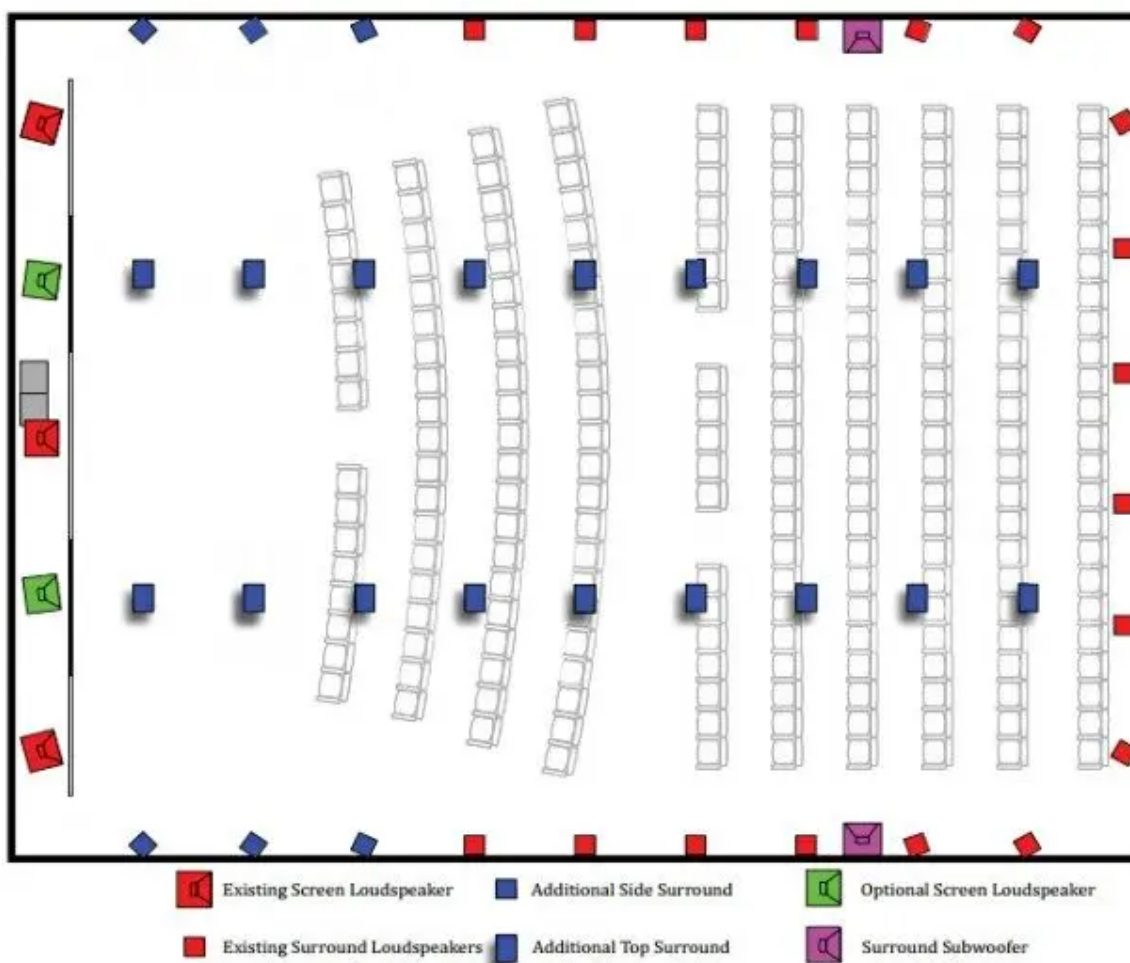


Figure 8 - Implémentation d'un système Dolby Atmos dans une salle de cinéma. On observe en rouge les enceintes préexistantes et en bleu, rose, et vert les enceintes ajoutées. (Dolby, 2012)

Mais la plus grande différence se situe dans le mixage qui ici est un mixage orienté objet, ce qui signifie que les informations audio sont séparées des informations spatiales qui lui sont associées.

Le Dolby Atmos propose de combiner un flux classique mixé en 9.1 appelé "bed", principalement utilisé pour des ambiances ou des musiques, avec un ensemble "d'objets" sonores qui peuvent être positionnés indépendamment dans l'espace selon les informations de spatialisation associées. Chaque objet sonore aura donc son information de spatialisation associée qui, par l'algorithme de calcul de Dolby, sera rediffusé sur une ou plusieurs enceintes. Il n'y a donc plus de contrainte de diffusion par canal, chaque enceinte est indépendante. Un flux Atmos

entier est composé de 128 pistes comportant minimum 1 bed de 10 pistes (9.1) et donc au maximum 118 objets sonores indépendants.

Lors du mixage, le mixeur aura le choix de placer les sons dans un bed (et donc les mixer traditionnellement en 9.1) ou de les considérer comme objets. Un grand avantage de ce système est sa transportabilité : du fait de la séparation du contenu audio de ses informations de position, le rendu ne dépend pas de la salle de diffusion. C'est le processeur de rendu propriétaire de Dolby, le RMU (Render Master unit), qui adapte la restitution à la configuration d'enceintes de la salle concernée. Il n'est donc pas nécessaire que chaque salle soit équipée exactement de la même manière, le RMU calculera en fonction de la position des enceintes comment alimenter celles-ci.

Malheureusement, la technologie Dolby Atmos conserve encore aujourd'hui beaucoup de secrets. Leur politique ne nous permet pas d'avoir accès à l'ensemble des processus et langages utilisés lors des calculs de son algorithmique.

- DTS:X

Le DTS:X (auparavant DTS MDA ou Multi-Dimensionnal Audio) est un système de diffusion 3 dimensions basé sur le modèle orienté objet. Comme le Dolby Atmos ou l'Auro-3D, son but est d'ajouter une dimension de hauteur aux systèmes de diffusion classiques 5.1 ou 7.1. Le système DTS:X est plus adaptable que le système Dolby Atmos : en effet, l'ajout d'enceintes au plafond n'est pas nécessaire, bien que recommandé, tout comme l'ajout d'enceintes sur le mur de l'écran. En fait, le DTS:X a comme fonctionnement d'adapter le processeur d'écoute à la salle. Les placements d'enceintes sont effectués en fonction des possibilités de la salle (sa taille, sa construction, son architecture), puis en sort un fichier de configuration qui sera traité par le processeur audio pour rendre les données audio et de spatialisation issues du mixage le plus fidèlement possible lors de la diffusion.

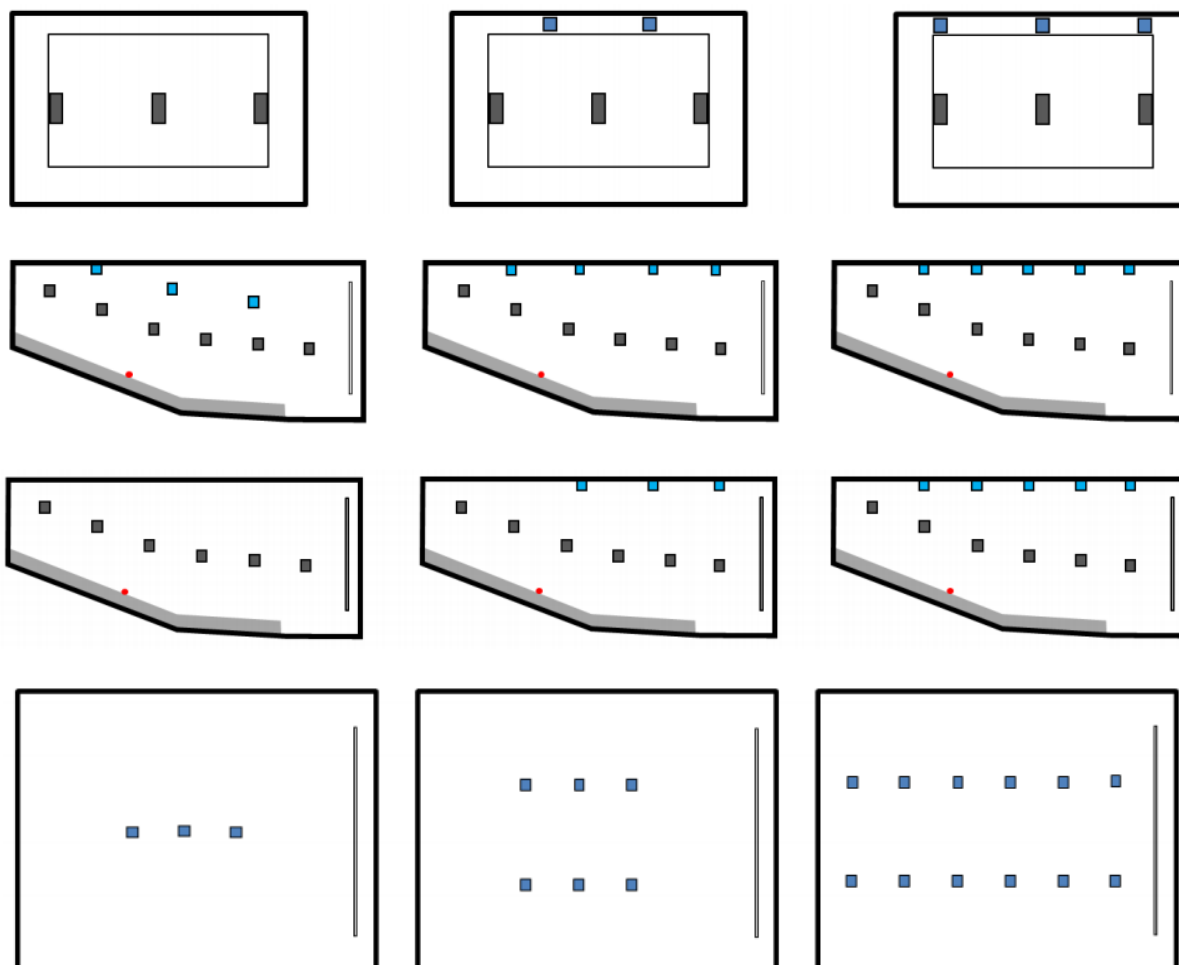


Figure 9 - Schémas représentant les possibilités d'ajout d'enceintes pour un système DTS:X sur respectivement le mur de l'écran, les murs de côté (pour salles sans enceintes au plafond puis pour salles avec enceintes au plafond) et enfin sur le plafond. (DTS, 2019)

Bien que ce système soit relativement proche du Dolby Atmos, il est tout de même plus adaptable à la salle et possède des enceintes en hauteur derrière l'écran, ce que le Dolby Atmos ne possède pas. Mais DTS a proposé ce système quelques temps après l'Atmos et c'est donc pour l'instant le géant Dolby qui continue de conquérir le marché actuel des salles de cinéma.

- La WFS

“La WFS (Wave Field Synthesis) a été développée à la fin des années 1980. Elle se base sur le principe de Huygens, qui dit que chaque point d’un front d’onde généré par une source primaire peut être considéré comme une source secondaire émettant des ondes secondaires.”(TIJOU, 2016). Le principe de la WFS, tel que proposé par Berkhout en 1988, est de créer une approximation d’un front d’onde à partir de hauts-parleurs servant alors de sources secondaires. Comme l’on peut le voir sur la Figure 10, en superposant alors les ondes ponctuelles de ces hauts parleurs, on retrouve une onde primaire synthétisée proche d’une restitution physique naturelle. La particularité de la WFS est de tenter de reproduire un champ acoustique naturel tandis que les autres systèmes utilisent des illusions auditives, des images fantômes et simulations de latéralisation pour simuler un champ sonore. Naturellement, ,comme le Dolby Atmos, la WFS a une approche du mixage orientée objet.

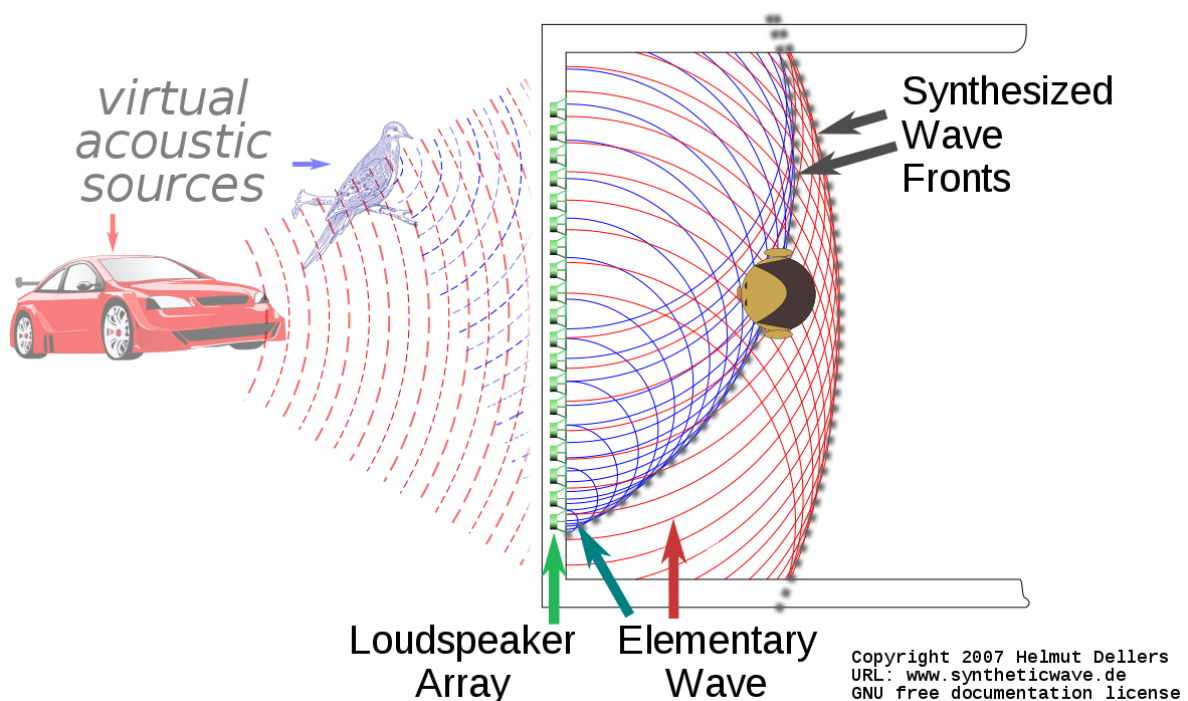


Figure 10 - Illustration du principe de Huygens dans le cadre d’une restitution en WFS (Helmut Dellers, 2007)

Dès les années 2000, la WFS est envisagée comme système de diffusion au cinéma. La première salle de cinéma équipée en WFS (avec 192 hauts-parleurs) est à Ilmenau en Allemagne. S'en suivront quelques installations aux Etats-Unis. Ce projet est surtout supporté par l'entreprise allemande IOSONO, faisant partie du groupe BARCO depuis 2014 (également distributeur de l'Auro-3D). Le premier film utilisant ce procédé est *Les Immortels* de Tarsem Singh (SINGH, 2011).

La WFS possède plusieurs avantages non négligeables pour le cinéma : une meilleure restitution de l'espace et de la profondeur, et une précision de localisation presque au niveau de celle d'une source réelle. Mais son principal atout est l'absence théorique de "Sweet Spot". Effectivement, lors du mixage, le mixeur est positionné au "sweet spot", une zone plus ou moins centrale définie par sa distance avec les haut-parleurs gauche et droits. Or, les systèmes de diffusion existants ne peuvent reproduire un champ sonore cohérent qu'en cette zone centrale correspondant à un nombre de places limité dans la salle. La WFS quant à elle, restitue un champ acoustique proche du réel ce qui augmente considérablement la taille de cette zone d'écoute. C'est-à-dire que tous les spectateurs présents dans celle-ci percevront la source d'un son provenant de la même direction, comme si elle était réelle. En théorie, le principe de la WFS donnerait une zone d'écoute infinie mais la technologie limitant le nombre de hauts-parleurs, la WFS n'est aujourd'hui qu'une approximation, et cette zone reste limitée, bien que très large.



Figure 11 - La salle du Mann's Chinese Theater à Los Angeles, équipée en WFS.

Cependant, de par sa configuration impliquant un nombre considérable de haut-parleurs, et par son manque de dimension verticale réelle, la WFS n'est pour l'instant pas réellement implantée dans le paysage cinématographique actuel.

- La spatialisation 2 dimensions

Depuis plusieurs dizaines d'années, comme expliqué auparavant, le format de diffusion de référence est le 5.1. C'est donc un format de diffusion horizontal, donc à deux dimensions, utilisant un modèle channel-based. La spatialisation lors de ce système de diffusion se fait par l'envoi des sons dans les cinq canaux d'enceintes qui entourent le spectateur, ainsi que le canal LFE (Low Frequency Effect). Il n'est donc nécessaire en théorie que de peu d'informations de spatialisation. En effet, le placement des sources sonores se fait par le biais de potentiomètres ou de joysticks. Les potentiomètres correspondent au placement gauche-droite (Left-Right) et avant-arrière (Front-Rear) à l'intérieur d'un carré admettant le canal gauche comme son coin supérieur gauche et le canal arrière droit comme son coin inférieur droit. On peut voir sur la figure 12 les cinq canaux délimitant le carré correspondant à la schématisation de la salle de diffusion, ainsi que les paramètres de position.

On peut donc observer les divers paramètres classiques d'une spatialisation horizontale :

- "front": correspond à l'avant de la ligne tracée sur laquelle va se déplacer le son
- "rear": correspond à l'arrière de cette même ligne
- "f/r" (Front/Rear) : correspond au placement avant-arrière du son le long de cette ligne
- "center%" : correspond au pourcentage d'envoi du son dans le canal central

Les valeurs affichées correspondent au pourcentage de différence par rapport au centre du carré. Il est possible et courant de lier "front" et "rear" ce qui correspond alors au placement "gauche/droite".



Figure 12 - Capture d'écran de la fenêtre de panning de ProTools lors d'un envoi dans un bus 5.1

- La spatialisation 3 dimensions



Figure 13 - Capture d'écran de la fenêtre de panning de 3DAW de Fairlight lors d'une spatialisation Dolby Atmos

Lors de spatialisation en 3 dimensions, d'autres paramètres entrent en jeu dans les plug-ins de spatialisation :

- "Haut/Bas" : qui ici représente la composante d'élévation du son
- "Rotation" : qui correspond, pour les objets multicanaux directionnels, à leur rotation horizontale dans l'espace
- "Pan V" : qui correspond à la rotation verticale des objets multicanaux directionnels
- "Taille" : qui va correspondre à la taille de l'objet et donc la place qu'il prendra dans la restitution sonore

- Les bases du mixage orienté objet

Channel-based :

Aujourd'hui, il existe actuellement trois approches principales de la spatialisation dans la diffusion cinématographique. En premier, il y a le modèle orienté canal (ou speaker-feed en anglais). Ce modèle consiste en l'alimentation (feed) de chaque enceinte (speaker) individuellement, c'est-à-dire, que le signal sonore est pensé pour chaque enceinte à un emplacement précis de la salle de cinéma et chaque canal correspond à une enceinte. La configuration du système de diffusion découle donc du nombre de canaux utilisés. Ce modèle est le plus répandu de par sa simplicité et sa transparence de fonctionnement. Lorsque le système de diffusion est connu (Stéréo, 5.1, 7.1..), tout ce système fonctionne très correctement. Mais il arrive qu'un mixage soit conçu pour un système de diffusion et finisse diffusé dans une autre configuration (par exemple un film mixé en 5.1 et diffusé en 7.1). Dès lors on peut observer une distorsion de l'espace sonore entre l'auditorium de mixage et la salle de cinéma.

Scene-based :

Le second modèle existant est celui de l'ambisonie, principalement développé par Michael Gerzon. Celui-ci, expliqué par exemple lors de la 2eme Conférence Internationale AES : *The Art and Technology of Recording* de Mai 1984 par Gerzon lui-même, se base sur la description et la décomposition du champ acoustique en harmoniques. Ce principe s'appuie sur des harmoniques sphériques allant fréquemment de 1 à 7 (64 canaux) pour décomposer le champ acoustique. La différence avec le premier modèle est que l'on ne ramène plus le son à une source acoustique mais à un champ complexe en 2D ou 3D. Ce procédé est également appelé HOA pour High Order Ambisonic. Le travail et la diffusion d'un tel signal requiert simplement une matrice qui calcule le signal selon la configuration d'enceintes employée. Cette technique n'alimente pas directement les enceintes mais possède l'information directionnelle d'un champ sonore grâce à un encodage

du son en composantes spatiales d'un repère cartésien (donc selon les axes x, y et z).

Object-based :

Le modèle orienté objet a été proposé par TUD (Technology University of Delft) et amené en Europe par le projet Carrouso (BRIX et al.).⁷ Le projet Carrouso a, entre autres, instauré une norme qui permet de représenter les métadonnées audio aujourd'hui : le MPEG-4 BIFS. Ces métadonnées sont utilisées dans le modèle orienté objet pour attribuer des informations spatiales aux informations audio brutes. Le format orienté objet propose donc une description du champ spatial comme des événements sonores diffusés à des positions spécifiques de la salle. Il est donc différent des premiers modèles dans lesquels le son est spatialisé par rapport au spectateur (système égocentrique) et prend la direction d'une spatialisation du son par rapport à la salle de diffusion (système allocentrique).

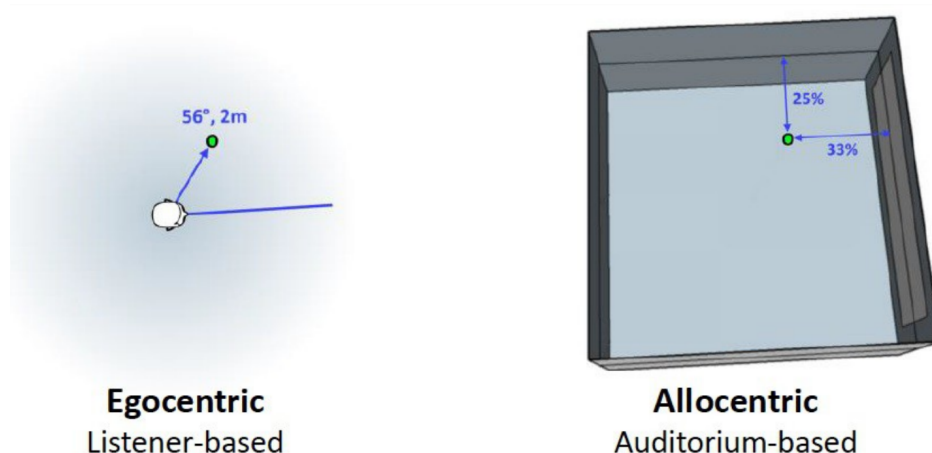


Figure 14 - Illustration de la différence entre le modèle égocentrique et le modèle allocentrique (SMPTE)

⁷ Sandra BRIX, Thomas SPORER et Jan PLOGSTIES, « CARROUSO-An European approach to 3D-audio », Convention Paper 5314, Presented at the 110th AES Convention, Amsterdam, 2001.

Le format orienté objet se base donc sur trois caractéristiques : (DENIZOT, 2016)

- Position de la source : coordonnées spatiales (cartésiennes ou polaires)
- Caractéristiques de la source : taille, directivité
- Paramètres de la salle de diffusion : taille, réverbération...

C'est donc un modèle très adaptable qui ne prend aucunement en compte le placement des hauts parleurs, il peut donc s'adapter aux diverses configurations de systèmes de diffusion basées sur ce modèle (Dolby Atmos, WFS...). Le système choisi tentera par des algorithmes de restitution de rendre en temps réel les informations audio et spatiales en concordance avec celles qui lui sont communiquées.

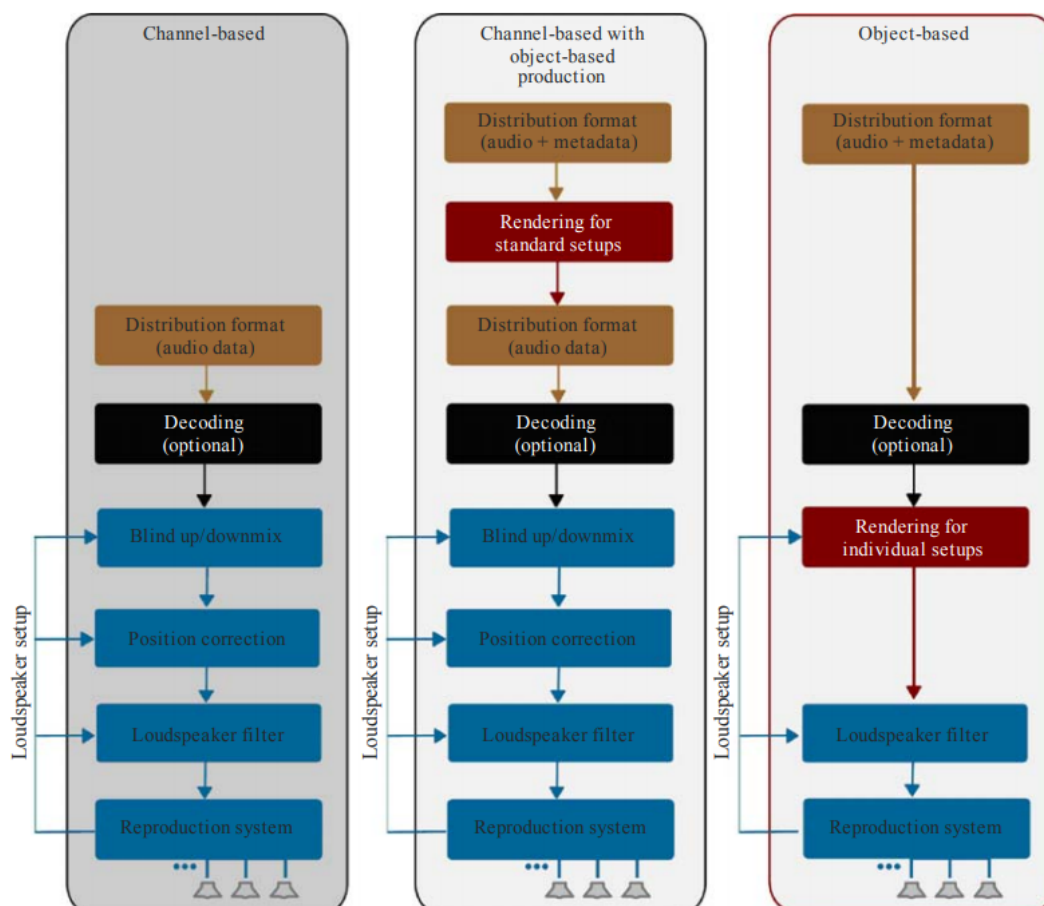


Figure 15 - Schéma comparatif des différentes possibilités de flux audio numériques depuis l'arrivée des formats orientés objet. (ITU-R BS.2159-4)

- L'utilité d'un panoramique en élévation

Depuis le développement du Dolby Atmos et du DTS:X (pour ne citer que les plus présents dans la diffusion cinématographique aujourd'hui), cette apparition de canaux zénithaux fait beaucoup débat. En effet, pour certains, il ne s'agit que d'un gadget et d'un argument de vente sans intérêt artistique, tandis que pour d'autres, c'est un vrai apport dans la spatialisation sonore d'un film et un réel impact dans sa narration et son émotion.

Dans une diffusion 5.1 ou 7.1, les canaux surround (arrières) ne sont classiquement utilisés quasiment que pour diffuser des ambiances ou des parties musicales, mais très peu d'effets y sont diffusés, ainsi que presque jamais de voix. En effet, d'expérience personnelle lors de mixage, j'ai pu observer l'effet d'une voix diffusée dans une enceinte surround : les spectateurs se retournent, ou du moins se déconcentrent de l'écran, car ils croient percevoir quelqu'un leur parler derrière eux. Or le cerveau va naturellement essayer de regarder la personne qui s'adresse à nous.

Lorsqu'il est question d'effets, ils sont utilisés avec parcimonie dans les canaux surround car la hauteur des enceintes fait que l'on a pas une sensation de hauteur humaine, mais quelque chose d'un peu plus élevé, ce qui peut donner des situations compliquées : si l'on imagine un plan en plongée où un hélicoptère dépasse la caméra par le bas de l'écran, il pourrait être une idée d'effet de faire parvenir le son de l'hélicoptère depuis les surrounds, or il sera perçu comme au-dessus du spectateur alors que celui-ci est sensé être en-dessous.

Dans une diffusion horizontale, il est donc rare d'entendre ces choses dans les surrounds qui sont donc présents plutôt pour une question d'immersion que de localisation et c'est pour cela que l'apparition des canaux surrounds et leur utilité a longtemps été discutée. Or, les ambiances diffusées à l'arrière provoquent une sensation d'enveloppement et donc d'immersion dans le film. L'image étant "plate", il est facile de se sentir séparé du film et de perdre cette sensation d'immersion.

Et c'est cet enveloppement qui est principalement recherché aujourd'hui avec les systèmes de diffusion 3D. Le Dolby Atmos a conquis les salles de cinéma depuis bientôt 10 années et l'expérience des mixeurs, avec qui j'ai pu discuter, qui en ressort est assez unanime : la spatialisation 3D permet une homogénéisation, ainsi

qu'une sensation aérée de l'espace. Les placements d'effets peuvent être désormais très précis et par exemple placés au centre de la salle, ce qui était impossible auparavant. Néanmoins, le cinéma français de nos jours use moins de ces éléments précis en mouvement que son équivalent américain. Effectivement, les choix cinématographiques en France vont plutôt axer l'utilisation du Dolby Atmos sur l'enveloppement et l'ambiance que sur les mouvements précis d'objets, bien qu'ils soient tout de même présents. Finalement, nous retrouvons bien une plus grande immersion avec ces nouveaux formats. La demande des spectateurs étant dans divers domaines une immersion plus grande (Réalité Virtuelle, Image 3D...), il est sûr que ces formats sonores 3D sont également un argument de vente, surtout lorsqu'ils peuvent être portés par une firme telle que Dolby. Mais il est nécessaire de faire la part des choses : toute avancée technologique découle d'un besoin ou d'une demande, et cette avancée technologique devient donc souvent un commerce, mais c'est également une étape supplémentaire vers de nouveaux horizons et de nouvelles ouvertures. Chaque progression, avec le temps de son expérimentation et de son apprentissage, offre de nouvelles possibilités dont chacun utilise ce qu'il juge nécessaire à son art ou à son but.

En seconde partie, ce mémoire visera à comprendre les outils de spatialisation utilisés aujourd'hui et pourquoi ils ne sont pas adaptés à un mixage 3 dimensions.

2. Outils 1D/2D / Outils utilisés aujourd'hui

- Potentiomètre

Le potentiomètre est un outil utilisé depuis des dizaines d'années dans des domaines plus que variés. Le principe de base d'un potentiomètre électronique est simple : il s'agit d'une résistance variable à 3 bornes. Celle du centre est un curseur dont on peut faire varier la position. Ce système permet de recueillir, entre la borne reliée au curseur et les deux autres, une tension qui dépend de la position du curseur et la tension appliquée évidemment.

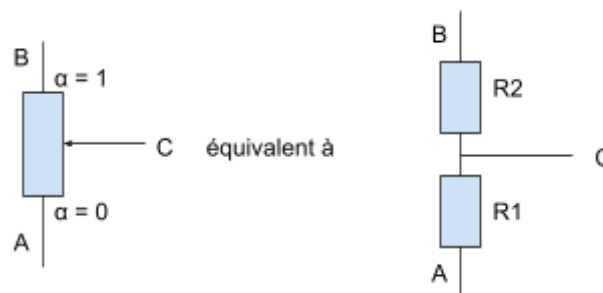


Figure 16 - Schéma équivalent d'un potentiomètre

Il existe aujourd'hui des potentiomètres numériques qui sont très répandus sur les consoles de mixage. Visuellement, ils ressemblent aux potentiomètres classiques mais n'envoient que des informations numériques bornées par 2 valeurs limites selon leur but. Ils sont donc pour certains à rotation infinie afin d'être adaptables à toute sorte de but (ils peuvent tout autant faire varier la fréquence d'un filtre qu'un envoi auxiliaire). Il existe aussi des potentiomètres linéaires (ou fader) qui sont plus généralement assignés à la gestion d'un volume ou d'un envoi, ils fonctionnent sur le même principe de base, mais sous une autre forme.

Dans la spatialisation, les potentiomètres rotatifs sont régulièrement utilisés. En effet, dans une spatialisation stéréophonique, un seul potentiomètre suffit à placer le son de gauche à droite, c'est d'ailleurs une des premières raisons de son existence dans le domaine du son. Lorsqu'il est question de spatialisation 2

dimensions, il se décompose en 2 variables (gauche-droite et avant-arrière). Les mixeurs ont appris de nos jours à instinctivement utiliser les 2 potentiomètres conjointement afin d'écrire une automation de mouvement en une seule fois, bien qu'il faille de la pratique pour faire cela facilement et avec réussite. Lors de spatialisation 3 dimensions, la tâche se complique : il devient quasiment impossible d'effectuer un mouvement avec 3 potentiomètres en une seule fois (le 3eme gère l'élévation), les mixeurs sont donc forcés d'effectuer cette automation en 2 fois au minimum. Les potentiomètres servent également à définir tous les autres paramètres de spatialisation (divergence et envoi LFE par exemple).



Figure 17 - Capture d'écran du panner 7.1 de ProTools (Avid Technology)

- Joystick

Le joystick est un outil datant des années 60, qui lui aussi fut utilisé dans plusieurs domaines, bien qu'il fut popularisé par le jeu vidéo. Il a comme principe de base d'associer deux potentiomètres dont les curseurs sont dans des directions perpendiculaires. Ainsi, les valeurs de résistance des potentiomètres varient indépendamment en fonction de la position du joystick et cela permet d'obtenir 2 données (X et Y). Dans le domaine du mixage, le joystick apparut lors de la numérisation du son afin d'indiquer des données de spatialisation 2 dimensions (classiquement pour du 5.1). Le joystick motorisé, quant à lui, apparaît sur une console Harrison en 1999 et permet une plus grande aisance à la relecture d'automation. Les différents joysticks présents sur les consoles de mixage présentent une sensibilité plus ou moins grande en fonction de leur fabrication ainsi qu'un retour haptique ou non.

Les joysticks sont donc utilisés à l'origine pour de la spatialisation 2 dimensions (gauche-droite et avant-arrière). Ils s'utilisent assez instinctivement pour placer un son autour de l'auditeur. Dès lors que l'on est dans un espace 3 dimensions, un deuxième joystick sert donc au paramètre d'élévation mais il est très ardu en un seul mouvement des deux mains d'effectuer aisément un mouvement qui comprend les 3 dimensions.



Figure 18 - Module de spatialisation avec 2 joysticks de l'Avid S6 (Avid, 2020)

- Ecran Tactile

Une surface tactile est un outil qui combine l'affichage et le pointage. Le premier écran tactile a été créé en 1971 par Sam Hurst, chercheur au Laboratoire national d'Oak Ridge aux Etats-Unis. Jusqu'en 2007, où Apple présente l'iPhone équipé d'un écran tactile capacitif, les écrans tactiles étaient résistifs. C'est-à-dire qu'ils étaient composés de 2 plaques conductrices très proches séparées par un isolant et lorsqu'une pression est effectuée sur une plaque, elle se déforme et vient en contact avec l'autre plaque, ce qui crée une information électrique ensuite utilisée par le logiciel pour définir sa position.

Les écrans tactiles capacitifs, eux, sont composés d'une plaque de verre qui accumule les charges. Et lorsqu'un doigt (ou une surface conductrice) est en contact, des charges s'enfuient par celui-ci, ce qui crée un déficit sur la plaque et dont la différence avec tous les bords de l'écran est calculée afin d'obtenir sa position.

Il existe des surfaces tactiles qui peuvent servir à la spatialisation. Par exemple, sur les dernières consoles d'Avid, il est possible d'ajouter un Ipad. Il peut servir à visualiser toutes sortes de choses, mais également à spatialiser des sons. On peut y voir donc le panner de ProTools et, de la même manière qu'avec un joystick, on peut placer un son dans un espace 2 dimensions. On y retrouvera également tous les paramètres du panner ProTools à disposition, bien que l'utilisation principale reste une alternative au joystick.



Figure 19 - Utilisation du panoramique avec un Ipad (Avid, 2020)

- Problèmes

Je n'ai pas personnellement d'expérience dans le mixage 3 dimensions. C'est pourquoi j'ai décidé d'interviewer plusieurs mixeurs afin de mieux comprendre les problématiques qui s'échappent des outils actuellement utilisés pour le mixage 3 dimensions.

J'ai donc pu discuter avec 2 mixeurs et un recorder, et j'ai pu comprendre premièrement les habitudes en place dans le mixage 3 dimensions.

Lors de mixages 3D, les mixeurs ne peuvent pas effectuer de mouvement de source sonore dans l'espace 3 dimensions en une seule fois. Comme nous l'avons vu précédemment, les outils présents devant eux sont :

- potentiomètres
- joysticks
- écran tactile
- souris

Ce sont donc deux outils 1D et deux outils 2D. Il est donc obligatoire de combiner 2 outils pour obtenir un mouvement 3 dimensions. Selon les mixeurs avec qui j'ai pu échanger à propos de leurs techniques de travail ou de celles de leurs collègues, les habitudes de combinaison peuvent différer : Un ingénieur du son décrit les 3 composantes spatiales uniquement avec les potentiomètres (il lui faut donc a minima 2 mouvements successifs, un premier pour les composantes X et Y, puis un second pour la composante Z d'élévation), un autre utilise les joysticks (le premier pour les composantes horizontales, le second pour la composante verticale) mais est également forcé de l'effectuer en 2 étapes successives, car il est très compliqué pour le cerveau de dissocier ces 2 mouvements effectués sur le même outil mais dont les conséquences sonores sont différentes. Le dernier utilise l'écran tactile pour faire le mouvement horizontal, puis effectue l'élévation au potentiomètre ou au joystick. Mais l'outil qui revient au final très régulièrement est la souris. La console Avid S6, largement répandue dans les auditoriums de mixages actuels, ne possède pas de joysticks motorisés, ce qui en complique fortement l'utilisation. De plus, la configuration des potentiomètres est compliquée lors de volontés légèrement différentes de la norme (par exemple lors d'égalisation dynamique sur un plug-in

FabFilter) et ils ne sont pas extrêmement précis. Le software d'Avid lié à la S6 proposé sur l'Ipad n'étant pas optimal non plus, il est souvent plus rapide de faire les étapes de spatialisation tout simplement à la souris, afin d'optimiser le temps passé à placer un son, dans des périodes de mixages toujours plus courtes.

Il est aisé d'observer qu'aucun de ces processus n'est optimal. Premièrement, une perte de temps s'opère sur chaque décomposition d'étape d'écriture d'automatisation de mouvement. Or, les mouvements pouvant être nécessaires un grand nombre de fois lors du mixage d'un film, le temps perdu cumulé peut devenir extrêmement grand. Ensuite, il est moins agréable de ne pas avoir le contrôle facile sur les mouvements que l'on veut donner au son dans l'espace. Le mixeur veut pouvoir agir facilement sur le placement du son et surtout instinctivement, sans avoir à réfléchir à quelle main décrit quel mouvement, sans que les mouvements de son corps soient en accord avec les mouvements sonores.

Les professionnels avec qui j'ai discuté sont unanimes : il est nécessaire de mettre à disposition des mixeurs un outil capable de décrire 3 directions à la fois, en un seul mouvement, et intuitivement. Les outils pour l'instant présents ne sont que des outils dérivés d'outils encore plus anciens (le potentiomètre était à la base prévu pour de la stéréophonie) et ne sont absolument pas prévus pour une spatialisation 3 dimensions.

C'est pourquoi, il y a, depuis quelques années, plusieurs tentatives d'innovations d'outils permettant une spatialisation 3 dimensions qui ont été effectuées. Nous allons présenter les principales.

3. Outils 3D existants

- Gant P5 (P5 glove)

Le P5 Glove est un gant infrarouge développé par la société Américaine “Essential Reality”, à l’origine dédié pour le jeu vidéo. Il se compose d’un gant équipé de capteurs de flexion sur chaque doigt et d’un récepteur infra-rouge qui permet de localiser la position du gant et de chaque doigt ainsi que les rotations de la main.

Tim KREGER (BENCINA, 2010) eu l’idée de développer un driver Mac OS X (P5osc) en 2004, qui permet de récupérer les informations transmises par le récepteur et de les convertir en informations OSC. Ce procédé, décrit sur un article de blog de son ensemble musical Simulus, permet donc de placer un son ainsi que d’agir sur divers paramètres de celui-ci uniquement en bougeant la main dans l’espace et en pliant ou non des doigts.



Figure 20 - Photographie d'un P5 Glove (Nytrix, 2006)

Les tests, effectués par Christopher HUMPHRIES dans son mémoire *Developing interactive strategies for three dimensional sound spatialization*

(HUMPHRIES, 2010), de ce gant ne semblent pas particulièrement concluants. Tout d'abord, la distance de connexion entre le gant et le récepteur n'excède pas 1m, ce qui peut être rapidement handicapant. Et bien que l'effet Kinésonique⁸ soit théoriquement plutôt satisfaisant, le délai introduit par la connexion infra-rouge entre le mouvement du corps et celui du son rend la pratique réelle compliquée. A cela s'ajoutent les douleurs que peuvent provoquer la tenue d'un bras à l'horizontale pendant des périodes répétées.

- SpaceMouse

La SpaceMouse est un contrôleur 3D développé par la société "3D Connexion". Il s'agit d'un capuchon à 6 degrés de liberté qui permet de naviguer dans un espace 3 dimensions avec aisance. Premièrement développé pour la conception de projets architecturaux et de CAO, un pilote gratuit a été développé par une communauté d'internet (le SpaceNav) coordonnée par John TSIOMBIKAS (TSIOMBIKAS, 2018) afin d'obtenir plus de libertés sur l'utilisation des informations transmises par la SpaceMouse et d'en extraire des données OSC.

Les tests par Christopher HUMPHRIES de ce système ont l'air plutôt concluants : la prise en main est intuitive, les mouvements simples et n'impliquent aucune douleur ou fatigue musculaire. Les points négatifs sont la marge de mouvement faible qui implique une précision de positionnement amoindrie, ainsi qu'une variation de vitesse de déplacement du son affaiblie, bien que très correctes.



Figure 21 - Photographie d'une SpaceMouse Compact (3DConnexion, 2021)

⁸ En anglais "Kinaesonic", un terme introduit par Mark Bokowiec et Julie Wilson-Bokowiec décrivant la relation du mouvement corporel lié au mouvement du son.

- Leap Motion

Le Leap Motion Controller est un objet développé par la société américaine Ultraleap qui regroupe Leap Motion et Ultrahaptics. Il s'agit d'un émetteur et récepteur composé de LEDs et d'une caméra infrarouge (UltraLeap, 2020) qui permettent de calculer en temps réel la position de mains positionnées au-dessus de lui. Selon UltraLeap, la précision de cet appareil est actuellement sans pareille dans le domaine commercial.

La société australienne Fairlight a conçu un plug-in de spatialisation 3 dimensions (le AirPan) qui, combiné avec un Leap Motion Controller, permet avec juste une main de placer une source sonore ainsi que déterminer plusieurs paramètres de celui-ci dans un espace 3 dimensions. En bougeant la main dans un cube de 60 cm de côté au dessus du Leap Motion Controller et avec certains mouvements de doigts, il est possible "d'attraper" le son, de réduire ou augmenter sa composante stéréophonique-monophonique, ou d'agir sur sa divergence par exemple.



Figure 22 - Photographie des données reçues par le Leap Motion Controller modélisées sur un écran (Ultraleap, 2021)

En 2014, Fairlight est une des premières sociétés à intégrer un nouvel outil de spatialisation 3 dimensions dans leur DAW et leur console de mixage (les consoles Fairlight possèdent le software Fairlight intégré).

Ce système semble très approprié à l'utilisation, l'effet kinésonique est très présent, le seul problème pourrait être la fatigue et la douleur venant du fait d'avoir le bras levé tout au long du placement sonore.

- DearVR Spatial Connect

La société allemande DearReality, fondée en 2014 par Achim Fell et Christian Sander et intégrée en 2019 au groupe Sennheiser, est spécialisée dans la spatialisation sonore et plus particulièrement par le biais de la Réalité Virtuelle (VR). Ils ont conçu un spatialiseur utilisable en VR : le DearVR Spatial Connect.

Il s'agit d'un programme standalone compatible avec les casques de VR Oculus et HTC. Avec 2 outils dans chaque main il est possible de contrôler tous les paramètres sonores possibles d'une session de mixage, le contrôle principal étant le pointage qui permet de placer la source sonore dans un espace 3 dimensions dans lequel on est immergé. Ce système compatible avec les DAW Reaper et Nuendo via leur plug-in de panning 3 dimensions DearVR Pro, est principalement développé pour le jeu vidéo ou la vidéo immersive, mais peut être utilisé dans le domaine cinématographique.



Figure 23 - Montage photographique montrant l'utilisation du dearVR Spatial Connect (DearReality, 2021)

Ces systèmes plus ou moins adaptés ne sont que les plus avancés et commercialisés, bien que très récents. Il existe également d'autres systèmes plus prototypiques ou "homemade" (comme le Wheel-e-bin ou le ScreenSpat) mais il existe peu d'informations et de tests à leur propos et, étant non commercialisés, il est difficile de les inclure dans une liste de candidats sérieux à un nouvel outil de spatialisation 3 dimensions.

Les 4 systèmes d'outils présentés précédemment pourraient tous être des options envisageables pour le futur, mais n'ayant pas eu l'occasion de les essayer (les tests évoqués ont été effectués par Christopher HUMPHRIES), je ne puis expliquer réellement leurs avantages et leurs défauts, bien que la SpaceMouse me paraisse le système le plus adapté car il est le seul à posséder un système de retour haptique qui permet de placer une limite aux bordures de la zone de spatialisation et qui n'induit aucune tenue de bras dans l'air sans repos.

Mais les fabricants de consoles et de DAW sont pour l'instant sûrement frileux à l'idée de se plonger entièrement dedans (à part Fairlight), d'autant que ces outils impliquent des coûts supplémentaires possiblement élevés qui ne seraient pas forcément appréciés des utilisateurs.

C'est pourquoi l'idée d'utiliser un smartphone comme outil m'est venue. En effet, cet outil étant présent dans la plupart des poches, il ne serait nécessaire que d'une application et d'un plug-in afin de faire fonctionner le système que je vais vous présenter par la suite.

4. “SmartPan”

Le “SmartPan” est un système utilisant l’écran tactile d’un smartphone, ainsi que son gyroscope. Par plusieurs configurations possibles, il permet de placer un son dans un espace 3 dimensions aisément.

L’idée de ce système nous est venue après avoir établi certains objectifs qu’il devait remplir :

- L’outil doit être utilisable à une seule main
- Il doit permettre de placer un son dans un espace 3 dimensions, donc avec 3 coordonnées spatiales X, Y et Z
- Il doit être intuitif, qu’il ne soit pas nécessaire de regarder sa main pour effectuer un mouvement

Tout ceci en essayant que le système ait le coût le plus réduit possible.

Le Leap Motion Controller m’a donné cette idée de mouvement dans l’espace avec la main mais, celui-ci étant coûteux, nous avons pensé au smartphone. Ce dernier possède les capteurs nécessaires à remplir ces objectifs (un écran tactile et un gyroscope), il est utilisable à une main, et constitue un objet de notre quotidien, il ne nécessite pas d’apprentissage. De plus, il est déjà présent à nos côtés et donc le coût en est uniquement logiciel. Il nous fallait donc ensuite récupérer les données que son gyroscope pouvait fournir.

- Gyroscope

“Le gyroscope est un instrument basé sur le principe de la conservation du mouvement angulaire. Cet appareil donne ainsi la position angulaire sur un, deux ou trois axes par rapport à un référentiel inerte. Il fut inventé par le physicien français Léon Foucault en 1852.”(Futura-Science, 2019)

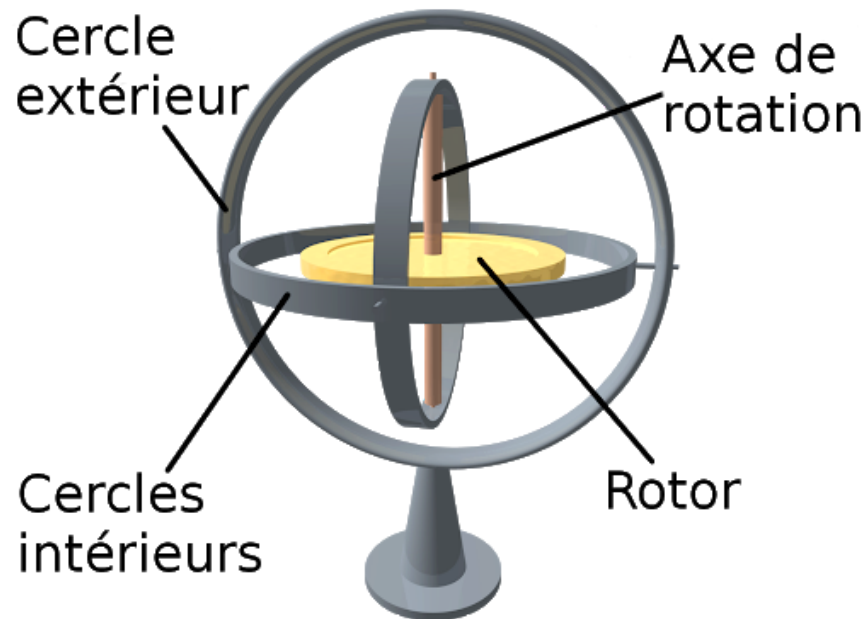


Figure 24 - Représentation d'un gyroscope (Lucas Vieira, 2006)

Ce principe est aujourd'hui miniaturisé dans ce qui est appelé gyroscope MEMS (Microelectromechanical System) et est donc contenu dans une puce adaptable à tout circuit imprimé. C'est ce système qui est utilisé dans les smartphones afin de déterminer leur orientation (ou position angulaire). On peut récupérer les données du gyroscope du smartphone qui sont donc 2 coordonnées (x et y) dont les valeurs natives sont contenues entre -10 et 10.

- Prototype sur MaxMSP

Pour commencer le système SmartPan, nous avons premièrement essayé de récupérer les données gyroscopiques de mon smartphone. Nous avons donc utilisé une application existante : TouchOSC. Cette application permet de transmettre en OSC ou en midi des informations tactiles, gyroscopiques ou d'accéléromètre via un réseau wifi.

L'OSC (ou Open Sound Control) est un format de transmission de données entre machines (ordinateurs, synthétiseurs, ou autres) développé pour du contrôle en temps réel. Il utilise les protocoles UDP (User Datagram Protocol) ou TCP

(Transmission Control Protocol) et apporte des améliorations de rapidité et de flexibilité par rapport à la norme MIDI. Nous utiliserons l'UDP pour sa simplicité.

Le protocole UDP qui permet la transmission, sans connexion préalable, de données de manière simple entre 2 entités définies chacune par une adresse IP et un numéro de port.

Max est un logiciel créé par Miller Puckette, lorsqu'il travaillait à l'IRCAM, dans les années 1980. Il le décrit comme "a way of combining pre-designed building blocks into configurations useful for real-time computer music performance"⁹ (Puckette, 2002). "Concrètement, c'est un logiciel qui permet de faire en direct de la synthèse sonore, de l'analyse, de l'enregistrement, ainsi que du contrôle d'instruments MIDI."(MARTIN, 2016) Appelé Patcher, puis Max/MSP, c'est Max 8 la dernière version en date. De nos jours c'est Cycling'74, une société américaine fondée en 1997, qui s'occupe du développement de Max. Miller Puckette a également créé Pure Data, une version open source de Max, lorsqu'il quitta l'IRCAM autour de 1990.

Max est un environnement de programmation visuellement très simple et qui ne requiert aucune compétence en langage informatique. Basé sur la notion de "patcher", c'est-à-dire des boîtes qui communiquent entre elles par des fils. Les différents types de boîtes sont : Objet, Message, Toggle, Button, Number et Slider. Les boîtes Objet sont les plus utilisées car ce sont elles qui permettent d'y insérer toutes sortes de fonctions, avec chacune leur nombre d'entrées et de sorties spécifiques. Il existe des fonctions d'opérations mathématiques, de traitement de signal audio et vidéo, et de gestion d'informations MIDI. Chaque boîte objet assortie à sa fonction, lors de la réception d'un message, réalisera sa fonction puis répondra éventuellement par un ou plusieurs autres messages aux boîtes auxquelles elle est reliée.

L'objet de base de mon patcher Max est l'objet "udpreceive" accompagné d'un numéro de canal, qui permet de récupérer les messages OSC transmis sur ce canal.

⁹ Un moyen de combiner des blocs pré-construits dans des combinaisons utiles pour la performance musicale informatique en temps réel

Grâce à l'objet "udpreceive" dans Max, nous avons réussi à récupérer les données gyroscopiques et à les visualiser en temps réel (celles-ci se présentent sous la forme de messages /accxyz suivi des 2 coordonnées de rotation du gyroscope ainsi que celle de l'accéléromètre).

L'accéléromètre mesure la vitesse verticale du téléphone mais pas la position. Nous avons premièrement pensé à utiliser cette troisième coordonnée pour l'élévation, mais comme il ne mesure que la vitesse et non la position, il est nécessaire de faire des mouvements assez brusques pour obtenir une donnée, ce qui ne serait par exemple pas adapté à un mouvement lent lors d'un mixage.

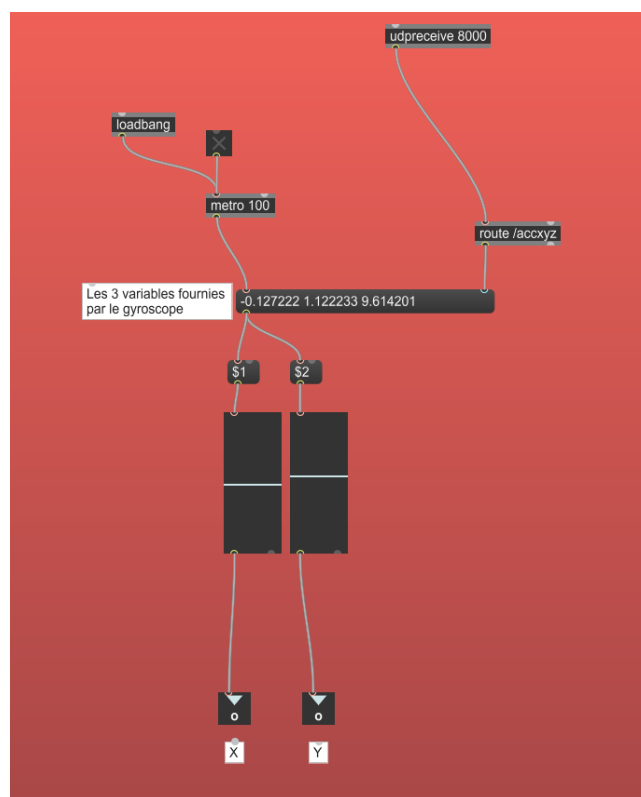


Figure 25 - Capture d'écran du sous-patch permettant de récupérer les données du gyroscope

De ceci, nous avons créé premièrement un panner stéréo (donc 2 canaux) basé sur des multiplications d'un son mono par les données du X du gyroscope afin de l'envoyer dans des canaux L et R.

Nous avons ensuite développé ce panner pour en faire un panner 5 canaux. En séparant les données X et Y fournies par le gyroscope et grâce à un ensemble de calculs entre les signaux visant à obtenir des données positives et comprises

entre 0 et 1 pour multiplier l'amplitude des signaux, et à l'outil "pictslider", nous avons pu faire varier le placement d'un son mono dans 5 canaux, le tout contrôlé par le gyroscope du smartphone.

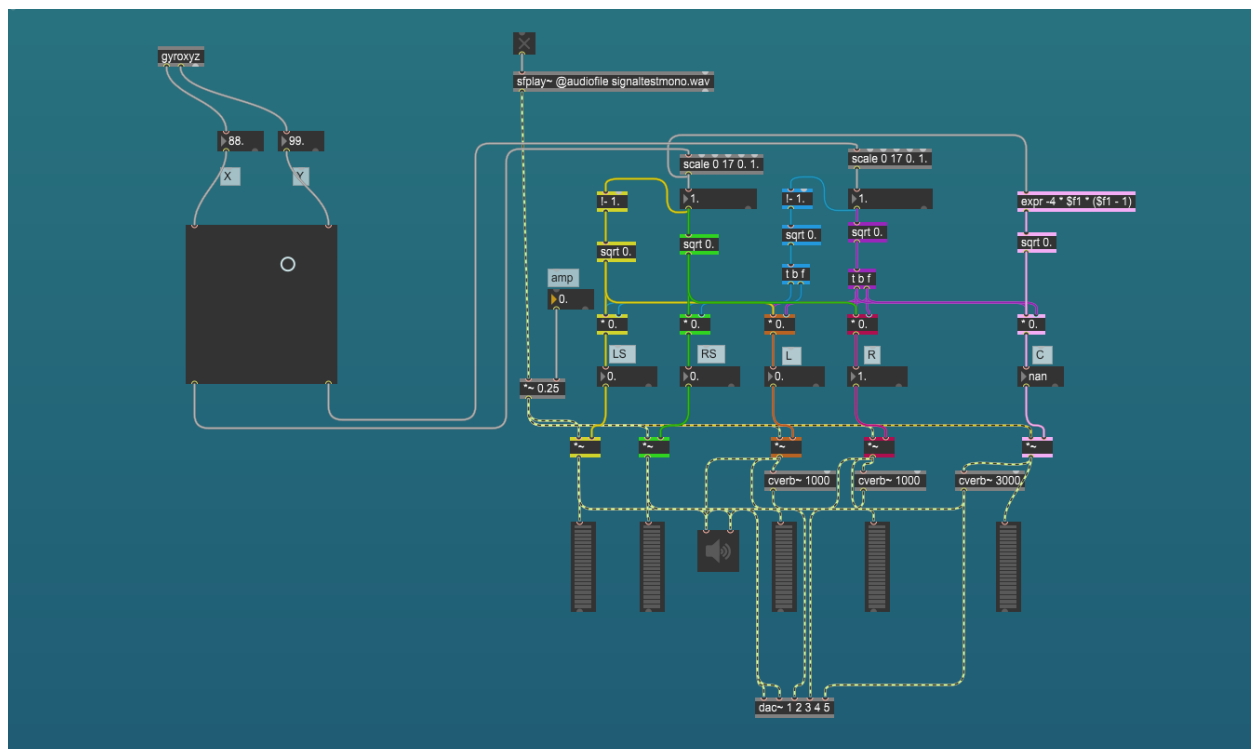


Figure 26 - Capture d'écran du panner 5.1

Cela étant fait, nous nous sommes rendus compte que le placement horizontal (XY) dirigé par le gyroscope était assez approximatif. Or, ce sont les mouvements les plus fréquents et donc qui se doivent d'être les plus précis possibles. Nous avons donc décidé d'inverser notre vision et d'utiliser le gyroscope uniquement pour l'élévation et le tactile pour le placement horizontal.

Nous avons donc créé un layout de TouchOSC afin d'avoir un pad XY sur l'écran sur lequel on peut agir et qui envoie des messages osc avec la forme /1/xy1 (ce qui correspond au pad xy numéro 1 de la page numéro 1 du layout) suivi des 2 coordonnées. En séparant les informations du gyroscope et celles du tactile dans notre sous-patch nous avons une base solide pour continuer.

Il s'agissait ensuite d'agir sur les paramètres d'un plug-in de spatialisation avec ces données. Nous avons choisi pour l'expérimentation du prototype d'utiliser Reaper car c'est un logiciel gratuit et très ouvert sur les types de communications

possibles. En effet, on peut configurer un contrôleur OSC dans les paramètres en y entrant l'adresse IP de réception ainsi que le canal.

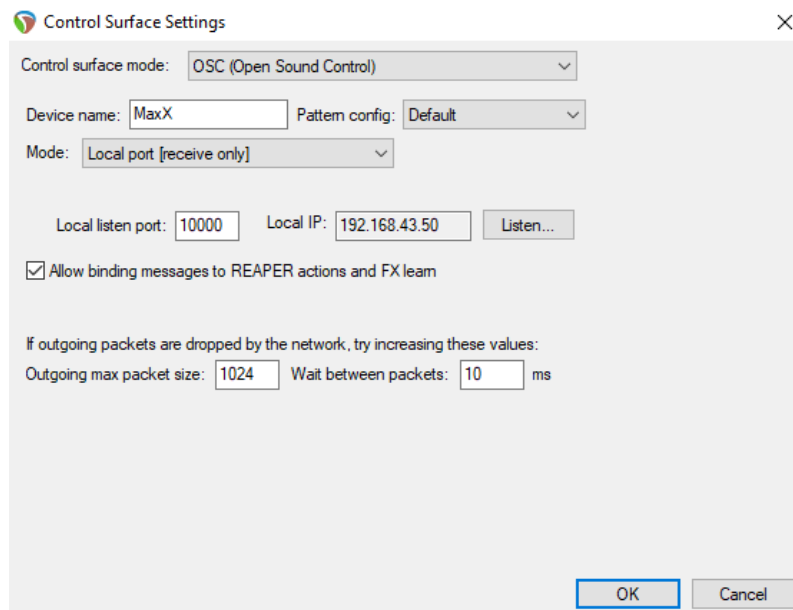


Figure 27 - Capture d'écran de la fenêtre de paramétrage de contrôleur OSC dans Reaper

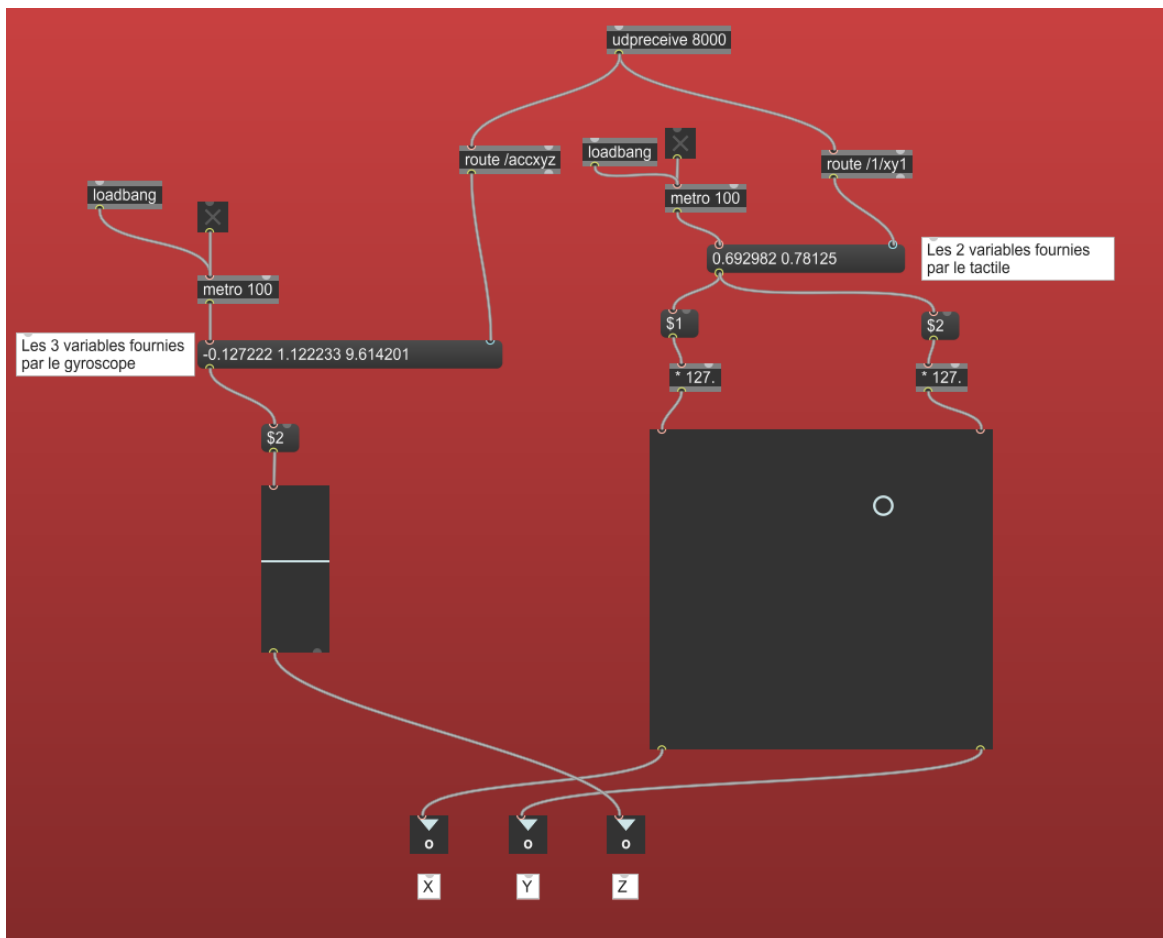


Figure 28 - Capture d'écran du sous-patch récupérant les données du gyroscope et du tactile

Après réflexion, nous avons pensé qu'il fallait avoir le choix du paramètre que le gyroscope ou l'écran tactile pourrait gérer. Nous avons donc créé un nouveau layout TouchOSC comportant un pad xy, ainsi qu'un fader pour la donnée d'élévation.

Nous avons donc implémenté différents circuits, lissé les données envoyées, ajouté des possibilités d'éteindre les sources, ainsi que de choisir de quelle source proviendrait chaque composante spatiale. Ci-dessous le patcher max obtenu :

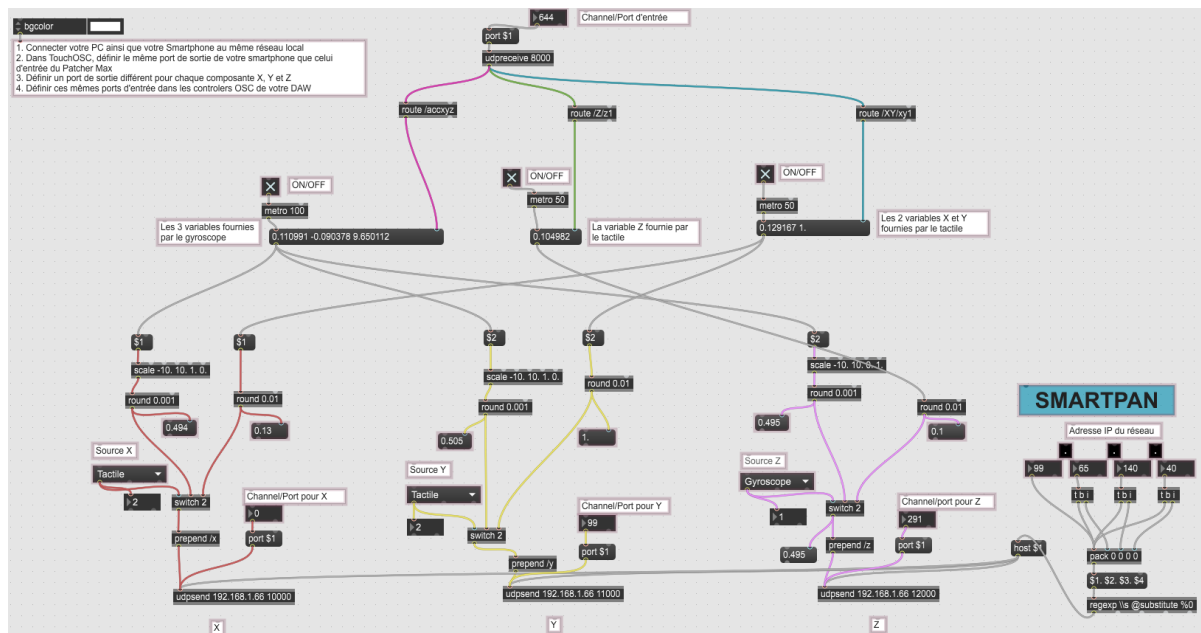


Figure 29 - Capture d'écran du patcher dans sa version étendue

Ce patcher comporte donc l'entrée "udpreceive" permettant de recevoir les données OSC de TouchOSC qui se séparent en /accxyz (les données du gyroscope), /XY (les données du pad xy) et /Z (les données du fader). S'en suivent des objets "scale" qui permet d'échelonner les données contenues entre -10 et 10 entre 0 et 1, et "round" qui permet d'arrondir les données à moins de décimales que nativement. Puis, des objets "switch" permettent de choisir les sources (gyroscope ou tactile) pour chaque composante, puis encore les objets "prepend" qui permettent d'ajouter un préfixe à la donnée afin d'avoir un message OSC défini (/x, /y ou /z), et enfin les objets "udpsend" qui permettent le renvoi de ces données OSC dans des ports définis récupérables dans le plug-in de spatialisation de notre DAW.

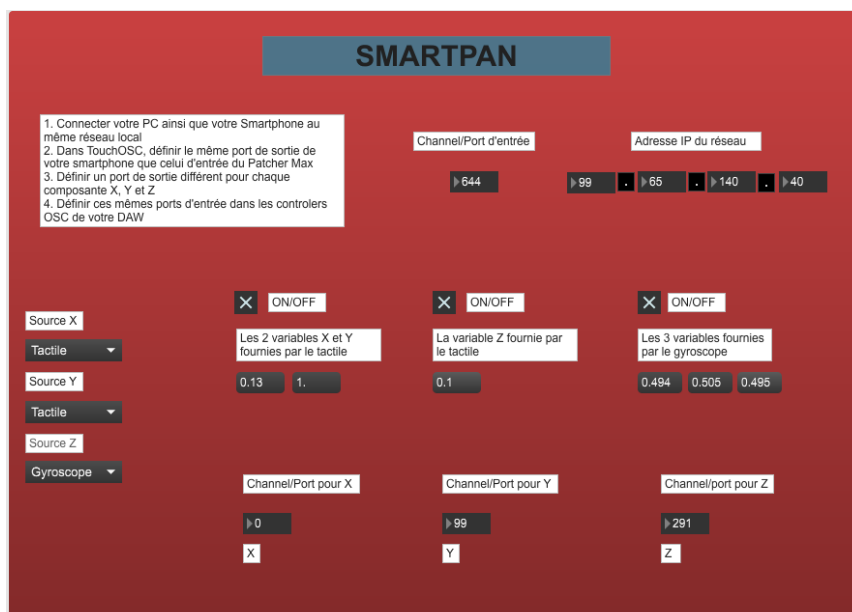


Figure 30 - Capture d'écran du patcher sous sa version présentation

Une dernière étape à consisté à ajouter le paramètre de divergence. Lors des mixages, l'affectation des joysticks est généralement la suivante : le premier joystick pour les coordonnées X et Y, le second pour la coordonnée Z, décrite verticalement, et le paramètre de divergence, horizontalement. Nous avons donc voulu reproduire une similitude (toujours avec un choix de la source de cette donnée, par le gyroscope ou le tactile) et il est possible d'avoir donc par le gyroscope l'élévation, ainsi que la divergence. Nous obtenons ceci :

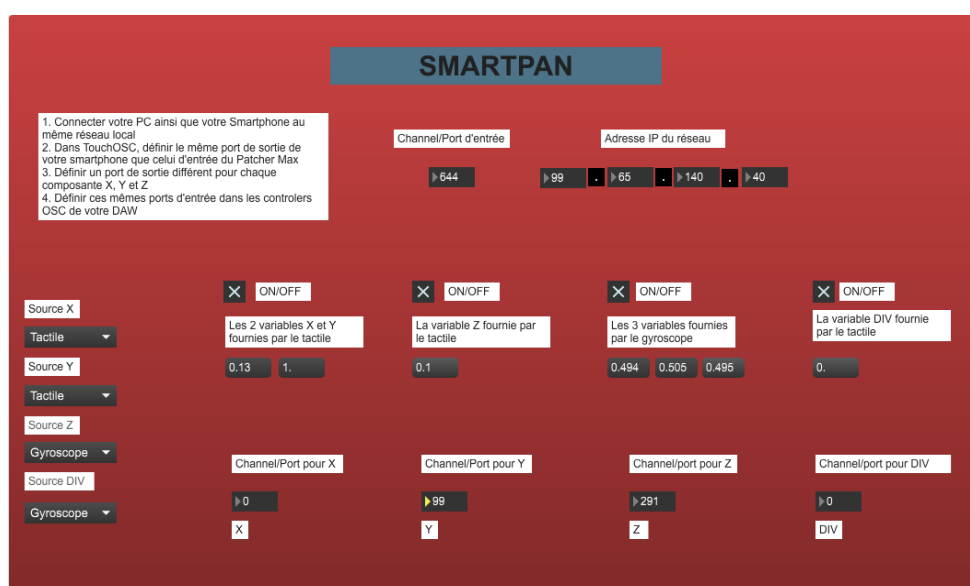


Figure 31 - Capture d'écran du patcher sous sa version présentation

Et les différentes pages de layout de TouchOSC créés se retrouvent comme ceci :

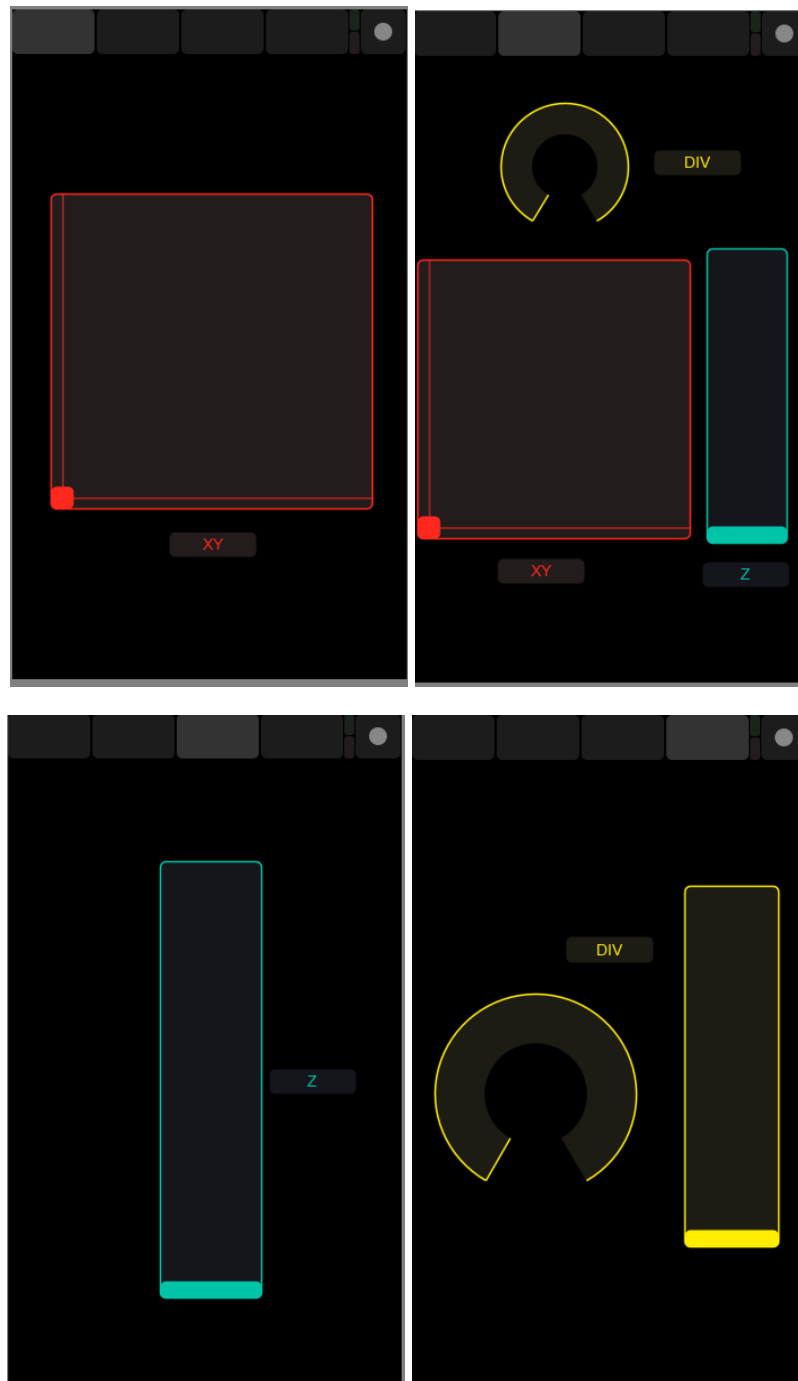


Figure 32 - Captures d'écran des fenêtres disponibles sur le layout de TouchOSC créé

- Essais

Dans un premier temps, j'ai réalisé quelques essais de mon côté. Ceux-ci s'avèrent plutôt concluants. J'arrive à contrôler les différents paramètres de spatialisation dans le plug-in de spatialisation de Reaper sans trop de difficultés. Les principaux problèmes que j'ai pu observer sont les suivants :

- Pour Reaper, lors de l'assignement de chaque contrôleur OSC aux paramètres du plug-in, les données envoyées en simultanément se chevauchent et il faut donc activer un par un les paramètres dans le patcher Max pour chaque assignation. Ce n'est pas un grand problème, mis à part l'ajout de fastidiosité lors de la mise en place du système.
- La précision du gyroscope du smartphone n'est pas toujours optimale. Surtout lors des valeurs limites (lorsque le téléphone est par exemple dans une position verticale), les valeurs transmises varient beaucoup sans raisons. Néanmoins, cela diffère selon les smartphones : le smartphone que j'utilise pour la majorité des tests est un Oppo mais il m'est arrivé de tester ce système sur un Iphone récent et les résultats étaient plus convaincants. Je pense également que cela pourra être ajusté si l'application est développée de bout en bout, car il pourra être possible de lisser les valeurs limites.

Lors du mois de Septembre 2021 j'ai par la suite eu l'occasion d'effectuer des tests perceptifs auprès de plusieurs personnes. J'ai effectué ces tests dans la salle multicanal de l'ENS Louis-Lumière les 9 et 10 Septembre. Le système de diffusion de la salle est alors de l'ambisonic d'ordre 3, c'est-à-dire 21 haut-parleurs répartis en 4 hauteurs : une couronne basse composée de 5 haut-parleurs équidistants, une couronne moyenne (à hauteur de tête) composée de 10 haut-parleurs équidistants, une couronne haute composée de 5 haut-parleurs équidistants, et enfin un haut-parleur situé au-dessus du point d'écoute, à la verticale.

Pour que le SmartPan fonctionne dans cette salle, il m'a fallu configurer plusieurs plug-ins de la suite de plug-ins de la marque IEM afin de convertir mes coordonnées

cartésiennes fournies par le patch Max en coordonnées polaires, plus adaptées à l'ambisonic. Il m'a fallu ensuite un encoder et un decoder ambisonic pour me permettre d'agir sur la position du son avec le plug-in de spatialisation d'ambix.

J'ai ensuite préparé 4 types de sons à faire écouter aux personnes participant au test :

- Plusieurs sons monophoniques de bruitages et effets (des sons de pas, d'hélicoptère et de vaisseau)
- Une voix monophonique
- Une ambiance de forêt stéréophonique
- Une musique stéréophonique

J'ai ensuite invité plusieurs personnes à venir faire le test qui se déroulait comme suit :

- J'introduisais le sujet de mon mémoire et leur expliquais le fonctionnement du plug-in de spatialisation.
- Dans un premier temps je leur demandais de faire bouger avec la souris les sons monophoniques autour d'eux, puis les sons stéréophoniques. De cette manière ils se familiarisaient avec le plug-in et le système de diffusion.
- Dans un second temps, je leur demandais de définir un ou plusieurs mouvements assez précis et d'ensuite essayer de reproduire ces mouvements les yeux fermés, avec les 4 types de sons.
- Ensuite, je leur expliquais l'utilisation du SmartPan et leur faisais répéter les étapes faites à la souris avec le SmartPan, yeux ouverts puis fermés, avec les 4 sons différents.
- Enfin, je leur soumettais un questionnaire (cf annexe).

Dans ce questionnaire, je questionne principalement l'appréciation du SmartPan, son ergonomie et sa précision. J'ai pu obtenir 8 résultats : 3 étudiants de 4ème année de la section Son de l'ENS Louis Lumière, 2 étudiants de 3ème année, 1 étudiant de 1ère année, ainsi qu'un enseignant. La moyenne d'âge se situe à environ 24,1 ans.

Dans l'ensemble les résultats sont homogènes : ces 8 personnes ont apprécié l'idée principale, bien qu'elle mérite encore de l'amélioration.

Présentation des résultats :

- 62,5% des personnes n'ont pas retrouvé les mêmes sensations entre l'utilisation de la souris et du SmartPan
- 85,7% des personnes ont préféré utiliser la souris les yeux ouverts et le SmartPan les yeux fermés
- 75% des personnes ont trouvé le SmartPan moyennement précis tandis que 25% l'ont trouvé très précis
- 50% ont trouvé le SmartPan précis en horizontal tandis que 37,5% l'ont trouvé précis en vertical et 12,5% ne l'ont trouvé précis dans aucun des deux axes.
- 75% des personnes ont trouvé la prise en main du SmartPan ergonomique bien que 37,5% ont trouvé le tactile peu pratique.
- 87,5% des personnes ont préféré l'utilisation du SmartPan à celle de la souris.
- 50% ont préféré pour la relation directe entre les mouvements du corps et ceux du son.
- 100% des personnes utiliseraient le SmartPan lors d'un mixage
- 100% des personnes l'utiliseraient pour des sons précis en mouvement, 37,5% pour des voix, 25% pour des ambiances et 0% pour des musiques
- 75% des personnes l'utiliseraient de temps en temps lors du mixage (pour une majorité des sons, selon le type de films), 12,5% l'utiliseraient très souvent et 12,5% l'utiliseraient très peu.

Par la suite, j'ai pu rencontrer Cyril HOLTZ et Antoine MARTIN (mixeur et programmeur à HAL audio), à qui j'ai pu présenter le SmartPan. Nous avons pu essayer les différentes manières d'utiliser ce dispositif. Je n'ai pas pu effectuer les mêmes tests qu'auparavant, tant les conditions étaient différentes, j'ai préféré avoir leur retour direct d'utilisation. Le regard expérimenté de ces personnes a pu apporter des critiques extrêmement intéressantes :

- Le lissage des données est à améliorer. On peut observer une “vibration” du point de localisation du son lorsque le téléphone est immobile.
- Les valeurs extrêmes sont compliquées à atteindre. Or, ce sont des valeurs avec lesquelles il faut travailler régulièrement lors d'un mixage (par exemple à l'extrême gauche ou droite de l'image).
- Lorsque le SmartPan est en utilisation gyroscopique du XY et tactile du Z : pour effectuer un mouvement gauche-droite le long de l'écran, on se retrouve avec le poignet tordu vers l'avant ce qui perd énormément de liberté de mouvement, de précision, et ce qui est douloureux.
- Un problème général : finalement, en combinant le tactile et le gyroscope, on en arrive à combiner deux mouvements pas forcément intuitifs, et le résultat n'en est pas si différent des outils utilisés de nos jours. De plus, instinctivement, nous allons utiliser deux mains pour combiner ces deux mouvements (gyroscope et tactile). L'objectif n'est donc pas atteint.

Mais aussi des points positifs et des idées d'améliorations :

- Il est très agréable d'avoir des grands pads et potentiomètres rotatifs faciles d'accès sur le tactile.
- Lorsque le SmartPan est utilisé avec le tactile pour le XY et le gyroscope pour le Z, son utilisation est assez facile, intuitive, et plutôt précise.
- Il pourrait être intéressant de pouvoir bloquer un axe pour ne pas avoir à le gérer dans le mouvement (par exemple bloquer les axes X et Z et n'agir que sur l'avant-arrière avec le Y)
- Il est agréable d'avoir plusieurs paramètres accessibles sur l'écran (par exemple la divergence)

Toutes ces remarques d'un point de vue d'une utilisation professionnelle par des personnes expérimentées ont beaucoup apporté à mes réflexions.

Avec Cyril HOLTZ et Antoine MARTIN nous avons pu ensuite essayer ensemble un plug-in de spatialisation très similaire au SmartPan, sorti quelques semaines plus tôt : Le SpaceController (Sound Particles, 2021).

Space Controller est un plug-in développé par Sound Particules, sorti à la fin du mois d’Août 2021. Ce plug-in a pour principe d’utiliser, comme pour SmartPan, un smartphone pour placer un son dans un espace 3 dimensions.



Figure 33 - Capture d’écran du plug-in Space Controller (Sound Particules, 2021)

La grande différence avec SmartPan est que Space Controller utilise le “pointage” du smartphone pour placer le son. En effet, dans l’utilisation du SmartPan on pouvait imaginer une bille placée sur un plateau et que l’on déplace en penchant le plateau. Dans l’utilisation du Space Controller, on imagine plutôt avoir un laser dans la main que l’on pointe pour indiquer la position de la source sonore.



Figure 34 - Capture d'écran de l'application mobile Space Controller (Sound Particules, 2021)

Nous avons pu observer que Space Controller est extrêmement précis et fluide, il est ergonomique et intuitif, et très bien développé.

En effet, sur la partie application, il est très simplifié : un gros bouton pour activer l'envoi de données de position, et le choix du canal sur lequel on agit (il est possible d'agir avec plusieurs smartphones simultanément sur différents canaux).

Sur la partie plug-in, il est bien plus complet : il y a au centre une belle visualisation de la position, à droite les niveaux par canal audio, et à gauche les informations de position. Mais il est également doté d'un large panel d'options : il est possible de bloquer les données d'un ou plusieurs axes, d'être en mode "miroir", de n'agir que sur la partie haute de l'élévation, d'agir en rotation d'une source, ainsi que d'agir sur une source mono, stéréo ou 5.0.

Lors de nos essais de ce système, nous avons été bluffés par la qualité de ce dernier. Cela m'a également permis d'observer ce qui aurait pu être un objectif final de ce projet et d'avoir devant moi un moyen de comparer les idées de SmartPan et de Space Controller.

La première chose qui m'a marquée est l'utilisation de la donnée de rotation "boussole". Je n'ai pas pour certitude que Sound Particules utilisent les données fournies par la boussole du smartphone pour cela, mais c'est ce qui me paraît le plus sensé. Effectivement, c'est une idée à laquelle je n'avais pas pensé et qui solutionne énormément de problèmes, comme celui d'avoir 3 données de position sans utiliser l'écran tactile. Cela apparaît immédiatement plus intuitif alors de pointer l'emplacement de la source sonore plutôt que de pencher le smartphone afin de faire "glisser" le son vers la direction voulue.

La seconde chose à laquelle je n'avais pas pensé est la multitude de possibilités d'agir sur la source sonore. Le fait de pouvoir agir sur un seul canal d'un son polyphonique, le mode "miroir", le mode rotation et le blocage d'axes sont des outils que Cyril HOLTZ trouve très importants lors d'un mixage. Je n'avais pas encore eu cette vision pratique du métier, et c'est bien ici où l'on ressent le professionnalisme du développement de Space Controller.

Mais essayer Space Controller avec Cyril HOLTZ et Antoine MARTIN nous a également permis de voir un point positif présent dans SmartPan qui n'existe pas dans Space Controller. Celui-ci concerne surtout l'application sur smartphone : il est finalement agréable d'avoir accès à plusieurs paramètres sur le téléphone, comme la divergence. Cela permet de ne pas avoir à alterner entre l'utilisation du téléphone et de la souris. Et il est pratique que l'accès à ces paramètres se fasse par l'utilisation de gros potentiomètres ou faders. C'est le point négatif de Space Controller à notre sens : l'application est un petit peu trop simple.

- Discussion

Les résultats des tests effectués à l'ENS Louis-Lumière suggèrent que la plupart des sujets ont apprécié le SmartPan, l'ont trouvé plutôt précis et ergonomique bien qu'il mérite des améliorations comme supprimer le tactile et lisser les données de localisation afin d'augmenter la précision. Mais tous ont trouvé l'idée intéressante et son utilisation serait viable dans un contexte de mixage.

Il est tout de même à mettre en perspective que l'échantillon testé est assez faible et peu varié. Aucun des sujets n'a d'expérience de longue durée de mixage cinéma, et ils sont très jeunes. De plus, aucun n'a d'expérience de mixage en 3 dimensions.

C'est pourquoi le retour de Cyril HOLTZ et d'Antoine MARTIN par la suite fût extrêmement important. Leur expérience des métiers liés au mixage cinéma et à la diffusion sonore 3 dimensions a apporté un regard critique plus proche de la réalité du métier.

Il en est tout de même ressorti que l'idée générale leur plaît et, bien qu'il y ait des améliorations à effectuer sur le SmartPan, il en demeure plutôt fonctionnel pour le moment.

Le but final du SmartPan est encore loin d'être accompli. Mais la sortie du Space Controller m'a permis d'observer en quelques sortes ce but final. En effet, il faut premièrement ne plus avoir de prototype dans Max et ne plus dépendre de touchOSC ou d'un plug-in de spatialisation déjà existant.

A terme, il faudrait idéalement créer une application de smartphone avec les mêmes fonctionnalités principales que TouchOSC mais avec plus de paramétrage possible de l'envoi de données gyroscopiques, ainsi que tous les paramètres présents dans mon patcher Max. Il faudrait également créer un plug-in de spatialisation pouvant récupérer des données OSC, et qui s'adapte aux formats Atmos, DTS X, Auro-3D, ainsi que les plus classiques 5.1, 7.1 et stéréo, avec tous les paramètres présents classiquement sur les plug-ins de spatialisation. SmartPan

nécessiterait également tous les paramètres de déplacement des sources présents dans Space Controller comme le verrouillage d'axes, le mode miroir ou le mode rotation. Le plug-in devrait également avoir une visualisation 2D et 3D du placement du son dans l'espace, ainsi que pouvoir être accepté par les diverses DAW (aux formats .vst, .aax...).

J'avais pensé l'idée d'un système émetteur-récepteur où le récepteur serait placé sous le téléphone afin d'avoir une donnée de hauteur du téléphone qui pourrait être utilisée pour l'élévation. Il serait alors possible de s'affranchir entièrement du tactile pour le placement en 3 dimensions. Mais cela impliquerait des coûts matériels supplémentaires et pourrait donc rester une option, mais pourrait fournir une autre manière d'obtenir ces 3 dimensions que le pointage.

Conclusion

Dans une recherche éternelle de plus de réalisme, les techniques utilisées dans le cinéma n'ont fait qu'évoluer vers une immersion grandissante. Succombant à la demande croissante, sont apparus les systèmes de diffusion sonore multicanaux qui n'ont, eux aussi, fait qu'accroître le nombre de leurs canaux de diffusion. C'est ainsi que nous en sommes arrivés aujourd'hui aux systèmes à haut nombre d'enceintes que peuvent être le Dolby Atmos ou le DTS:X. Malgré cette évolution technologique, les outils utilisés dans ces systèmes n'ont que peu évolué. Les mixeurs utilisent encore aujourd'hui des outils inadaptés à leur pratique ; des outils qui ne sont que dérivés d'anciens outils et ne sont pas repensés entièrement d'un point de vue ergonomique. Les joysticks, les potentiomètres et écrans tactiles ne permettent pas aux mixeurs de spatialiser un son en 3 dimensions facilement et rapidement comme il leur est demandé, et ces professionnels du son reviennent donc régulièrement à l'utilisation de la souris, qui est par excellence un outil générique.

C'est pourquoi il est nécessaire de créer de nouvelles choses, d'essayer de comprendre les problématiques inhérentes à la pratique de la spatialisation dans le mixage trois dimensions. Plusieurs systèmes ont commencé à émerger ces dernières années (comme le Leap Motion combiné à l'Airpan sur les consoles Fairlight), mais ceux-ci sont restés, soit à la marge, soit prototypiques.

Le Space Controller qui a été proposé il y a peu par la société Sound Particules constitue un réel candidat aux solutions de toutes ces problématiques. En effet, il possède toutes les qualités requises pour un travail de spatialisation 3 dimensions fiable, ergonomique, intuitif et à faible coût.

Le smartphone est un outil détenu par un grand nombre de personnes, or c'est un outil formidable possédant de grandes qualités qui ne sont pas toujours utilisées dans leur entièreté. C'est pourquoi il m'a paru intéressant d'en dériver des possibilités pour essayer de trouver une solution à ces problématiques de mixage 3 dimensions. Tous les smartphones possédant un écran tactile, ainsi qu'un gyroscope, il m'a paru possible d'en utiliser les données afin de spatialiser un son dans un espace 3 dimensions. C'est ce que j'ai essayé de faire avec le système SmartSpat présenté précédemment. Ce projet est loin d'être terminé mais, à terme,

il pourrait être similaire, après des couches d'améliorations issues de batteries de tests auprès de mixeurs, à Space Controller.

Les diverses personnes pratiquant ces métiers, avec qui j'ai eu la chance de discuter, étaient unanimes : il risque d'y avoir de plus en plus de films demandant une spatialisation 3 dimensions, mais cela implique souvent plus de temps de travail. Or, les temps de mixage et de montage son sont de plus en plus réduits par les productions, et les outils à disposition ne permettent pas de gagner du temps. Il est donc important de pouvoir gagner ce temps en pouvant faire confiance à un outil intuitif. Je n'avais pas de certitudes que le SmartPan pourrait plaire, mais la sortie du Space Controller m'a montré que l'idée d'utiliser un smartphone à des fins de spatialisation était dans l'air du temps. Je pense que cela pourrait plaire à plusieurs mixeurs d'avoir sous la main un outil plus ergonomique, permettant un gain de temps et d'énergie.

Je suis fort heureux d'avoir pu découvrir plus de choses sur les pratiques concrètes du mixage pratiqué par diverses personnes. Tout ceci m'a amené à réfléchir beaucoup, à la fois sur le côté technique des choses, et aux moyens d'améliorer la qualité du travail futur. Mais cela m'a aussi fait repenser mon idée du métier de mixeur et que, bien que les productions cinématographiques laissent de moins en moins de temps de travail et sont encore frileuses à l'idée de la 3D, les volontés artistiques cinématographiques évoluent, et je n'ai nul doute que les métiers participant à leurs concrétisations aussi.

Chaque avancée technologique aura son lot d'artistes prêts à l'explorer.

Références Bibliographiques

DENIZOT, Florent, *L'Approche orientée Objet Au Cinéma : Entre Surenchère Technologique et Outil de Création Sonore*, Mémoire (sous la direction de Mohammed Elliq), Master Son, Ecole Nationale Supérieure Louis-Lumière, 2016

DOLBY LABORATORIES INC., *Dolby Atmos Specifications : issue 3*, Documentation de spécification technique, 2015, Consultable sur www.dolby.com.

DOLBY LABORATORIES INC., *Dolby® AC-4 : Audio Delivery for Next-Generation Entertainment Services*, Document de normalisation, 2015.

DOLBY LABORATORIES INC., *Un son de prochaine génération pour le cinéma*, Documentation de présentation technique, 2012, Consultable sur www.dolby.com.

GARCIA Jérémie, FAVORY Xavier, BRESSON Jean, *Trajectoires: a Mobile Application for Controlling Sound Spatialization*, CHI EA '16: ACM Extended Abstracts on Human Factors in Computing Systems., San Jose (US), Mai 2016

GARITY, William, Fantasound, *Journal of the Society of Motion Picture Engineers*, vol 36, Janvier 1941

GERZON, Michael, Ambisonic Surround-Sound Mixing for Multitrack Studios, Convention Paper C1009, *Presented at the 2nd AES International Conference: The Art and Technology of Recording*, Mai 1984

HUMPHRIES, Christopher, *Developing interactive strategies for three dimensional sound spatialization*, Mémoire (sous la direction du Professeur Michael Clarke et Mr Mark Bokowiec), University of Huddersfield, Janvier 2010

ICART, Roger, *Abel gance ou le prométhée foudroyé*, Lausanne, L'âge D'homme, coll. *Histoire Théorie*, 1983.

IVES, David, DEMONBREUN, Sam, *Audio Console with Motorized Joystick Panning System*, Patent No.: US 6,813,530 B1, LaVergne (US), 2004

KIM, Syng N., *Analog Joystick Controller*, Patent Number 4,587,510, Hoffman Estates (US), 1986

LITKE, David, *Fractured Sounds, Fractured Meanings : A Glove-Controlled Spectral Instrument*, UBC School of Music/MAGIC Group, Vancouver, 2007

MARIE, Baptiste, *Une Interface Tangible pour le Spatialisateur de L'IRCAM*, Mémoire (sous la direction de Claude Gazeau et Thibaut Carpentier), Master Son, Ecole Nationale Supérieure Louis-Lumière, 2011

MARTIN, Antoine, *Création de la Footstep, Control interface d'échantillonnage dédiée aux bruits de pas pour le cinéma*, Memoire (sous la direction de Sylvain LAMBINET), Master Son, Ecole Nationale Supérieure Louis-Lumière, 2016

MCGEE, Ryan Michael, *Sound Element Spatializer*, Mémoire (sous la direction de Curtis ROADS), Master of Science In Media Arts & Technology, University of California Santa Barbara, Decembre 2010

PUCKETTE, Miller, "Max at Seventeen", *Computer Music Journal*, pp. 31-43, publié le 26 Avril 2002.

TIJOU, Clement, *Cohérence audiovisuelle spatiale au cinéma : Influence de la vitesse et de la nature des objets sonores*, Memoire (sous la direction de Etienne HENDRICKX), Master Son, Ecole Nationale Supérieure Louis-Lumière, 2016

WILSON-BOKOWIEC Julie et BOKOWIEC Mark, "Kinaesonics: The intertwining relationship of body and sound", *Contemporary Music Review*, Vol. 25, pp. 47-57, Février 2006.

Références Cinématographiques (par année de sortie)

CROSLAND Alan, *The Jazz Singer (Le Chanteur de Jazz)*, 88 min, États-Unis, sorti en France le 4 octobre 1929, 1927

GANCE Abel, *Napoléon*, 330 min, France, sorti le 7 avril 1927, 1935

COOPER Merian, Von FRITSCH Gunther, SCHOEDSACK Ernest B. et TODD Michael JR, *This Is Cinerama*, 115 min, États-Unis, 1955

KOSTER Henry, *The Robe (La Tunique)*, 135 min, États-Unis, sorti en France le 4 décembre 1953, 1953

LUCAS Georges, *Star Wars*, 121 min, États-Unis, sorti en France le 11 Septembre 1977, 1977

BURTON Tim, *Batman Returns*, 126 min, Etats-Unis, sorti en France le 15 Juillet 1992, 1992

SPIELBERG Steven, *Jurassic Park*, 127 min, Etats-Unis, sorti en France le 20 Octobre 1993, 1993

MCTIERNAN John, *Last Action Hero*, 130 min, Etats-Unis, sorti en France le 11 Août 1993, 1993

UNKRICH Lee, *Toy Story 3*, 103 min, Etats-Unis, sorti en France le 14 Juillet 2010, 2010

SINGH Tarsem, *Immortals (Les Immortels)*, 110 min, Etats-Unis, sorti en france le 23 Novembre 2011, 2011

HEMINGWAY Anthony, *Red Tails*, 125 min, États-Unis, sorti en France le 25 mai 2012, 2012

RAMSEY Peter, *Rise of the Guardians*, 97 min, États-Unis, sorti en France le 19 décembre 2012, 2012

ANDREWS Mark et CHAPMAN Brenda, *Brave*, 93 min, États-Unis, sorti en France le 1er août 2012, 2012

DEMICCO Kirk et SANDERS Chris, *The Croods*, 98 min, États-Unis, sorti en France le 3 Juillet 2013, 2013

CUARÓN Alfonso, *Gravity*, 91 min, États-Unis, sorti en France le 23 octobre 2013, 2013

Webographie

Sound Particles : Space Controller, publié le 28 Août 2021

<https://soundparticles.com/products/spacecontroller>

Connect Your DAW to the VR World, publié le 26 avril 2021

<https://www.dearvr.com/products/dearvr-spatial-connect>

Ultraleap Hand Tracking Overview, publié en Mars 2020

<https://docs.ultraleap.com/hand-tracking/>

BENCINA, Ross, *P5 Glove Developments*, publié le 11 Août 2010

<http://www.simulus.org/p5glove>

TSIOMBIKAS, John, *SpaceNav*, publié le 14 Avril 2018

<http://spacenav.sourceforge.net/index.html>

Gyroscope : qu'est-ce que c'est ?, publié le 19 Février 2019

<https://www.futura-sciences.com/tech/definitions/technologie-gyroscope-11121/>

Annexes

06/10/2021 20:48

Test SmartSpat

Test SmartSpat

1. Avez-vous retrouvé les mêmes sensations en utilisant les 2 systèmes ?

Une seule réponse possible.

- Oui
 Non

2. Avez-vous retrouvé les mêmes sensations en utilisant les 2 systèmes ?

Une seule réponse possible.

- Oui
 Non

3. Avez-vous trouvé plus simple le test les yeux fermés ou ouverts ?

Plusieurs réponses possibles.

- Yeux ouverts à la souris
 Yeux fermés à la souris
 Yeux ouverts avec le SmartSpat
 Yeux fermés avec le SmartSpat

4. Avez-vous trouvé le SmartSpat précis dans la localisation ?

Une seule réponse possible.

- Très précis
 Moyennement précis
 Peu précis
 Inutilisable

06/10/2021 20:48

Test SmartSpat

5. Le SmartSpat est-il plutôt

Plusieurs réponses possibles.

- Précis en horizontal
 Précis en vertical
 Aucun des deux

Autre : _____

6. Avez-vous trouvé la prise en main du SmartSpat ergonomique ?

Plusieurs réponses possibles.

- Oui
 Non
 Le tactile n'est pas pratique
 Le gyroscope n'est pas pratique

Autre : _____

7. Quel système avez-vous préféré ?

Une seule réponse possible.

- Souris
 SmartSpat (avec smartphone)

8. Pourquoi ?

9. Utiliseriez-vous le SmartSpat lors d'un mixage ?

Une seule réponse possible.

- Oui
 Non

06/10/2021 20:48

Test SmartSpat

10. Si oui pour quel type de sons ?

Plusieurs réponses possibles.

- Pour des sons précis en mouvement
- Pour des voix
- Pour des ambiances
- Pour des musiques

Autre : _____

11. Si oui, l'utiliserez-vous fréquemment ?

Plusieurs réponses possibles.

- Très souvent (pour quasiment tous les sons à spatialiser)
- De temps en temps (pour une majorité des sons, selon le type de film)
- Peu (2 à 3 fois par film maximum)

Autre : _____

12. Qu'amélioreriez-vous dans ce système ?

Ce contenu n'est ni rédigé, ni cautionné par Google.

Google Forms