HAL
open science

Influence de la stéréoscopie sur le mixage des ambiances "surround" au cinéma

Etienne Hendrickx, Mathieu Paquier, Vincent Koehl

To cite this version:

Etienne Hendrickx, Mathieu Paquier, Vincent Koehl. Influence de la stéréoscopie sur le mixage des ambiances "surround" au cinéma. 12e Congrès Français d'Acoustique, Apr 2014, Poitiers, France. pp.1771-1777. hal-00986896

HAL Id: hal-00986896

https://hal.univ-brest.fr/hal-00986896v1

Submitted on 6 May 2014

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

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

Influence de la stéréoscopie sur le mixage des ambiances 'surround' au cinéma

E. Hendrickx, M. Paquier et V. Koehl
UBO, Lab-STICC UMR CNRS 6285, 6 avenue Victor Le Gorgeu, CS 93837, 29238 Brest Cedex 3, France
etienne.hendrickx@univ-brest.fr

Peu d'études ont été menées sur l'influence de la stéréoscopie sur la perception d'un mixage audio au cinéma. Les témoignages de mixeurs ou les articles scientifiques montrent pourtant une grande diversité d'opinions à ce sujet. Certains estiment que cette influence est négligeable, d'autres affirment qu'il faut totalement revoir notre conception de la bande-son, aussi bien au niveau du mixage que de la diffusion.
Cette étude se concentre sur la perception des sons d'ambiance dans les enceintes « surround ». 8 séquences, extraites de films ou de documentaires 3D, ont été diffusées dans un cinéma à 44 sujets, dans leur version 2D (non-stéréoscopique) et 3D (stéréoscopique), avec plusieurs mixages différents. Pour chaque présentation, les sujets devaient indiquer si le mixage proposé leur paraissait trop frontal ou au contraire trop « surround », le but étant de mettre en évidence une éventuelle influence de la stéréoscopie sur la perception de la balance avant/arrière d'un mixage audio. L'influence du placement du sujet dans la salle a également été prise en compte. Les résultats obtenus rejoignent ceux d'une expérience préliminaire menée dans un auditorium de mixage, où les sujets se trouvaient en situation de mixeur et devaient eux-mêmes régler la balance avant/arrière : l'influence de la stéréoscopie est faible et dépend du contenu de la séquence.

1 Introduction

Certains mixeurs affirment qu'un film projeté en salle à la fois en 2D et en 3D doit être mixé différemment pour chacune des deux versions. Michael Semanick, par exemple, affirme avoir produit deux mixages très différents pour les mixages stéréoscopiques et non-stéréoscopiques du film « Alice au Pays des Merveilles » de Tim Burton, avec beaucoup plus d'effets, de musiques, d'ambiance et de réverbération dans les enceintes « surround » [1]. Les ingénieurs du son d'« Avatar » de James Cameron, ou encore ceux d'« Hugo Cabret » de Martin Scorsese, affirment par contre n'avoir que rarement senti le besoin de modifier leur mixage pour la version stéréoscopique [2, 3].

Les rares études qui ont été menées jusqu'à maintenant sur le son lié à l'image stéréoscopique se sont principalement focalisées sur la notion de réalisme physique. Ces études reposent sur l'hypothèse qu'une reproduction sonore plus proche d'un véritable champ sonore (avec des techniques telles que la WFS, Ambisonic, ou encore la restitution binaurale) devrait accroître la sensation d'immersion des spectateurs, de la même façon que l'image stéréoscopique accroit la sensation de présence, comme l'ont montré Ijsselsteign *et al.* [4].

Il n'a cependant pas été encore prouvé que la stéréoscopie favorisait l'immersion car elle permettait de se rapprocher d'un champ visuel humain, et plusieurs études suggèrent d'ailleurs qu'une reproduction sonore « physiquement réaliste » n'est pas forcément appropriée à un film : quand André *et al.* [5] ont diffusé à leurs sujets un film en relief avec différents systèmes de diffusion, la WFS a été classée dernière en terme de sensation d'immersion, alors qu'il s'agissait du système le plus réaliste physiquement. Le test ne fut cependant réalisé que sur une seule séquence d'animation, avec une reproduction frontale et non à 360° autour du sujet. D'autre études sont donc nécessaires pour vérifier si une restitution du son plus réaliste physiquement peut accroitre la sensation d'immersion au cinéma [6]. Une étude récente a montré qu'une trop grande quantité d'informations spatiales pouvait réduire l'attention du spectateur à la narration [7] ce qui pourrait être une piste pour expliquer les résultats d'André *et al.*.

Pour Chion [8], les spectateurs sont tellement accoutumés aux conventions du cinéma que ces dernières sont devenues les nouvelles références du réel. Rumsey [9] suggère également que la reproduction du son a acquis ses propres standards de réalisme qui diffèrent de l'écoute naturelle.

L' « effet ventriloque » est un bon exemple pour supporter cette hypothèse : au cinéma, les dialogues sont en général reproduits sur l'enceinte centrale. Pourtant, les voix semblent véritablement provenir de la bouche des comédiens, même quand ces derniers sont sur les côtés de l'écran, voire hors-champs. Briser cette « règle » et spatialiser les voix peut même parfois être perçu négativement [10].

Il peut cependant être intéressant de s'inspirer de la réalité et d'adapter certains aspects physiques d'un champ sonore réel aux contraintes d'un système classique comme le 5.1. Chion suggère par exemple que le « surround » apporte plus de vraisemblance aux sons d'ambiance [8] car il simule notre faculté à percevoir le son à 360° et est donc sensoriellement plus convaincant.

Le but de cette étude est d'étudier les différences de perception auditive qui peuvent apparaître selon que l'on regarde un film avec l'image en 2D ou en 3D. La présente étude se focalise plus précisément sur l'influence que pourrait avoir la stéréoscopie sur la perception des sons d'ambiance au cinéma.

Si les sujets souhaitent avoir un mixage moins frontal lorsque l'image est en 3D, cela peut signifier que l'image stéréoscopique, étant plus proche de la perception humaine, incite le sujet à avoir également une reproduction sonore plus proche de la réalité, sans discrimination entre l'avant et l'arrière par exemple (à l'opposé du son « traditionnel » proposé au cinéma, où la plupart de l'énergie vient des enceintes frontales [11]). Dans tous les cas, de tels résultats suggèreraient que de nouvelles conventions, mieux adaptées à l'image stéréoscopique, sont nécessaires, et remettraient possiblement en question la pertinence des systèmes actuels (5.1, 7.1, etc.) pour le cinéma 3D.

2 Expérience préliminaire

Une expérience préliminaire [12] a été menée à l'auditorium de mixage du département « Image et Son » de l'Université de Brest, sur une configuration 5.1 traditionnelle (3 enceintes avant et deux enceintes « surround »). 11 séquences, dont la plupart filmées spécialement pour l'expérience, ont été montrées à un panel de 11 sujets (tous étudiants de la formation « Image et Son » de l'Université de Brest). Les séquences étaient montrées dans leur version stéréoscopique mais aussi non-stéréoscopique, soit un total de 22 stimuli, présentés dans un ordre aléatoire et différent pour chaque sujet.

Pour chaque stimulus, le sujet devait ajuster la balance entre les enceintes avant et arrière, à l'aide d'un bouton rotatif infini et sans gradations, afin que le sujet n'ait

aucune information visuelle ou tactile qui puisse influencer son réglage. L'expérience était répétée une seconde fois avec une pause de 15 minutes entre les deux sessions de test.

L'étude statistique des mixages a montré une influence plutôt faible de la stéréoscopie sur le mixage des séquences :

- Pour la première session du test, une seule séquence présentait une différence de balance moyenne significative entre le mixage stéréoscopique et le mixage non-stéréoscopique. Cette différence n'était cependant plus significative dans la deuxième session du test ;
- Pour la deuxième session, trois autres séquences présentaient des mixages stéréoscopiques et non-stéréoscopiques significativement différents.

Lorsque les différences étaient significatives, il s'agissait à chaque fois de mixages non-stéréoscopiques plus frontaux en moyenne que leur correspondant stéréoscopique. Ces résultats ont montré que l'effet de la stéréoscopie dépendait du contenu de la séquence. Ils ont également mis en évidence une influence forte de la session.

Une seconde expérience a été mis en place, où les sujets étaient cette fois mis dans la peau d'un spectateur plutôt que d'un mixeur, avec des séquences 3D issues de productions professionnelles, dans une véritable salle de cinéma. Plutôt que de régler eux-mêmes la balance, les sujets devaient juger l'équilibre avant/arrière de séquences dont les balances avaient déjà été fixées au préalable par les expérimentateurs.

3 Matériel et Méthode

3.1 Lieu de l'expérience

Le test s'est déroulé dans un cinéma de grande taille (22 mètres de long pour 15 mètres de large), doté d'un système 5.1 calibré selon les recommandations Dolby, avec :

- 3 enceintes à deux voies passives KCS S-2500 pour les canaux avant : Gauche, Centre et Droit ;
- 12 enceintes à 2 voies passives KCS SR-15 pour les enceintes arrière gauches et droites (couramment appelées enceintes « Surround ») ;
- 1 caisson de basse KCS C-118-A.

Les enceintes étaient alimentées par un processeur numérique Dolby CP750, dont le signal était amplifié par des amplificateurs QSC USA900, tandis que l'image était projetée par un projecteur numérique Christie CP2000-ZX, synchronisé avec des lunettes 3D actives Eyes3Shut Purple Two.

3.2 Séquences

L'expérience préliminaire a montré que l'influence de la stéréoscopie dépendait du contenu de la séquence. Il fallait donc garder pour cette seconde expérience le plus de séquences possible.

Pour que la durée du test reste raisonnable, le nombre de séquences a dû cependant être réduit à 8 (Voir Tableau 1).

5 séquences extraites de productions professionnelles ont été utilisées, ainsi que trois séquences issues de l'expérience préliminaire :

- Un dialogue dans un café, où il était apparu une différence significative entre les mixages stéréoscopiques et non-stéréoscopiques ;
- Deux marins effectuant une manœuvre à la proue d'un bateau, où il était également apparu une différence significative entre les mixages stéréoscopiques et non-stéréoscopiques ;
- Une foule dans la rue, qui n'avait montré aucune différence significative lors de l'expérience préliminaire.

3.3 Echelle de jugement

Des questionnaires papier ont été distribués aux sujets. Pour chaque stimulus leur était proposée une échelle graduée, dont les extrémités étaient labellisées « beaucoup trop frontal » et « beaucoup trop « surround » » (voir Figure 1). Le sujet devait alors choisir une des gradations en l'entourant. Il n'y avait aucun label intermédiaire prédéfini, pour éviter l'introduction de biais indésirable [13].

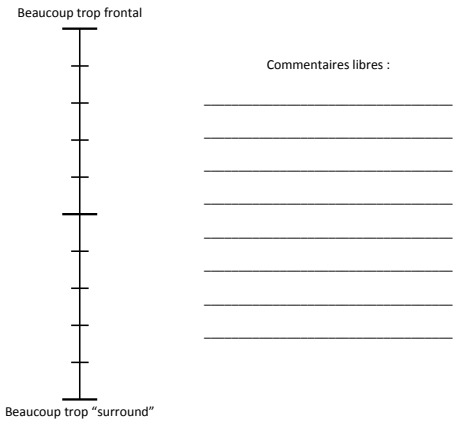


Fig. 1 Echelle de jugement pour chaque stimulus.

L'échelle était divisée en 11 gradations séparées par des intervalles égaux, ce qui signifie que l'échelle peut être considérée comme continue [14] et les données peuvent donc être analysées à l'aide de procédures paramétriques, tant que certaines hypothèses ne sont pas violées trop sévèrement [15].

Les sujets étaient priés de garder leurs lunettes 3D pendant la totalité de l'expérience, même pour les séquences 2D (durant lesquelles la même image était envoyée à l'œil gauche et à l'œil droit), afin d'éviter une éventuelle influence de la perte de luminosité (qui peut aller de 40 à 70% avec des lunettes actives).

3.4 Balances sonores

Chaque séquence a été présentée aux sujets avec deux mixages différents, l'un étant plus frontal que l'autre. Si cette décision présente le désavantage de multiplier par 2 la durée du test, elle permet néanmoins de :

- réduire l'influence des expérimentateurs, qui fixent les balances avant/arrière subjectivement ;
- éviter un effet de compression des données : si un seul mixage était proposé aux sujets et que ce

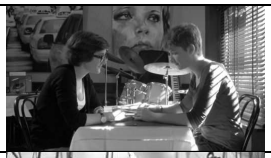
	Contenu	Valeurs de plan	Dynamique	Système d'enregistrement	Capture d'image
1	Dialogue dans un café, avec les voix diffusées sur l'enceinte centrale (Séquence tirée de l'expérience préliminaire).	Plan serré	Plan fixe	Stéréo Décorrélée	
2	Proue d'un bateau, avec deux marins effectuant des manœuvres (Séquence tirée de l'expérience préliminaire).	Plan moyen	Plan fixe	Double M/S	
3	Foule dans une rue (Séquence tirée de l'expérience préliminaire).	Plan large	Plan fixe	Double ORTF	
4	Dialogue dans un arbre (Extrait de *Souviens-moi*, de Joséphine Derobe).	Plan moyen	Caméra sur grue, tournant lentement autour des acteurs.	Stéréo Décorrélée	
5	Un homme regardant deux enfants jouer sous une grande nappe (Extrait de *Souviens-moi*, de Joséphine Derobe).	Plans serrés	Champ/ Contre champ	Stéréo Décorrélée	
6	Trois jeunes filles dansant dans une chapelle (Extrait de *Reine du Sabbat*, de Pablo Agüero).	Plan serré	Travelling arrière, lent.	Stéréo Décorrélée	
7	Une fille marchant silencieusement dans une chapelle (Extrait de *Reine du Sabbat*, de Pablo Agüero).	Plan large	Travelling avant, lent.	Stéréo Décorrélée	
8	Match de Rugby (Extrait de *One & All*, de Henry Davies).	Plans larges, moyens et serrés.	Montage rapide.	Double-M/S	

Tableau 1 : Les 8 séquences, avec leur contenu, leurs valeur et dynamique de plan, leur système d'enregistrement et une capture d'image.

dernier avait été dans l'absolu mixé trop frontalement par les expérimentateurs, les sujets n'exploiteraient probablement que le haut de l'échelle de jugement, risquant donc de minimiser des différences significatives dues à la stéréoscopie ;

- étudier d'éventuelles interactions entre le Mode Visuel et la Balance avant/arrière proposée.

Pour chaque séquence, le niveau global d'écoute a été fixé et égalisé subjectivement par les expérimentateurs [16], afin que les mixages plutôt « frontaux » soient perçus au même niveau que les mixages plutôt « surround ».

Chaque séquence a donc été présentée sous 4 modalités différentes aux sujets :

- Avec l'image stéréoscopique et le mixage « frontal » ;
- Avec l'image non-stéréoscopique et le mixage « frontal » ;
- Avec l'image stéréoscopique et le mixage « surround » ;
- Avec l'image non-stéréoscopique et le mixage « surround.

3.5 Sujets

44 personnes, non rémunérées (10 femmes et 34 hommes, âgés de 20 à 25 ans) ont pris part à l'expérience. Il s'agissait d'étudiants en formation aux métiers du son (Master Image & Son de l'Université de Bretagne Occidentale). Bien que n'ayant pas d'expérience des tests perceptifs, ces étudiants, formés à l'écoute analytique, peuvent être considérés comme "experts" [16].

3.6 Groupes

La perception de la balance entre enceintes avant et arrière dépend directement du placement du sujet dans la longueur de la salle [17]. Il a donc été décidé de séparer les sujets en 2 groupes :

- Le premier groupe en proximité de l'écran (rangées n°2, 3, 4 et 5) ;
- Le second groupe au « Sweet Spot » (rangées n° 10, 11, 12 et 13),

en supposant que l'effet de la rangée est négligeable à l'intérieur d'un groupe. Avec 22 sujets par groupe, répartis sur 4 rangées, les sujets étaient assis à une distance maximale de 1,35 mètre du centre de la rangée. Cette distance étant faible par rapport à la taille de la salle (22 mètres de long pour 15 mètres de large), il est également supposé que les positions latérales ont une influence négligeable sur la perception de la balance avant/arrière [17].

3.7 Répétitions

La première expérience avait montré une influence de la session. Il a donc été décidé de répéter le test trois fois de suite.

Les stimuli duraient chacun 20 secondes. A la fin d'un stimulus, des lumières s'allumaient pour permettre aux sujets d'écrire leur réponse. Au bout de 20 secondes, les lumières s'éteignaient et le stimulus suivant était joué. Avec 96 stimuli à juger, la durée de l'expérience était d'1 heure et 4 minutes, ce qui est supérieur aux recommandations de Bech (30-40 minutes) [15] mais inférieur à celles de Schatz, qui obtint des résultats encore fiables après 90 minutes d'évaluations de qualité, malgré l'apparition chez ses sujets de signes de fatigues mesurables [18]. Les sujets ont tous passé le test en même temps, sans pause entre chacune des trois sessions du test pour éviter qu'ils ne communiquent entre eux.

Tableau 2 : Résultats de l'ANOVA.

Source	SS	DF	MS	F	p
S	743.993	6.698	111.081	21.356	0
V	33.375	1	33.375	9.566	0.004
B	691.159	1	691.159	156.687	0
R	22.596	2	11.298	3.145	0.048
G	30.004	1	30.004	4.4E+16	0
S*V	21.325	7	3.046	2.38	0.022
S*B	211.016	7	30.145	16.475	0
S*R	64.012	9.335	6.857	3.25	0.001
S*G	184.527	7	26.361	5.297	0
V*B	0.487	1	0.487	0.445	0.509
V*R	8.417	2	4.208	4.08	0.02
V*G	22.571	1	22.571	6.469	0.015
B*R	27.362	2	13.681	14.604	0
B*G	57.836	1	57.836	13.111	0.001
R*G	0.445	2	0.222	0.062	0.94
S*V*B	12.912	7	1.845	1.627	0.127
S*V*R	37.491	9.728	3.854	1.925	0.042
S*V*G	13.778	7	1.968	1.538	0.154
S*B*R	35.855	14	2.561	1.932	0.021
S*B*G	28.922	7	4.132	2.258	0.03
S*R*G	11.436	14	0.817	0.581	0.881
V*B*R	0.021	2	0.011	0.011	0.989
V*B*G	2.917	1	2.917	2.663	0.11
V*R*G	12.366	2	6.183	5.994	0.004
B*R*G	2.555	2	1.277	1.363	0.261
S*V*B*G	24.181	7	3.454	3.047	0.004
S*V*R*G	30.037	14	2.146	1.542	0.091
S*V*B*R	45.138	14	3.224	2.532	0.002
S*B*R*G	14.718	14	1.051	0.793	0.677
V*B*R*G	0.836	2	0.418	0.447	0.641
S*V*B*R*G	19.249	14	1.375	1.08	0.373

4 Résultats

Après avoir vérifié les hypothèses de normalité, d'homoscédasticité et de sphéricité, une analyse de variance mixte à 5 facteurs a été effectuée sur les données, pour étudier les effets des facteurs suivants :

Intra-sujets

- S : Séquence (8 niveaux) ;
- V : Mode Visuel (2 niveaux : non-stéréoscopique vs stéréoscopique) ;
- B : Balance Sonore (2 niveaux : mixage « plutôt frontal » vs mixage « plutôt surround ») ;
- R : Répétition (3 niveaux).

Inter-sujets

- G : Groupe (2 niveaux : groupe 1 « Proche de l'écran » vs groupe 2 « au Sweet Spot »).

Les résultats de l'analyse de variance sont présentés dans le Tableau 2.

4.1 Modes Visuels : Stéréoscopie vs. Non-Stéréoscopie

Les résultats montrent une influence significative du Mode Visuel sur les notes des sujets (V : $F = 9.566$, $p = 0.004 < 0.01$) avec une valeur moyenne de 5.54 pour les balances stéréoscopiques et de 5.36 pour les balances non-stéréoscopiques, ce qui montre que les sujets ont globalement perçu les mixages avec images 3D comme étant plus frontaux que ceux avec images 2D.

Cependant, la significativité de l'interaction Séquence - Mode Visuel (S*V : $F = 2.38$, $p = 0.022 < 0.05$) montre que cet effet du Mode Visuel n'est pas le même en fonction de la séquence. Un post-hoc LSD de Fisher, comparant l'effet du Mode Visuel séquence par séquence (figure 2), montre même que le mode visuel n'a d'impact que pour les séquences 1 (dialogue dans un café) et 2 (marins effectuant des manœuvres sur un bateau).

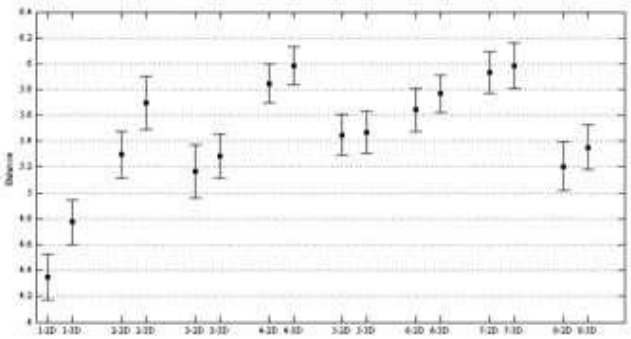


Figure 2 : Effet du Mode Visuel (2D « non-stéréoscopique » vs. 3D « stéréoscopique ») en fonction de la Séquence.

4.2 Influence de la Répétition

Les résultats indiquent une influence de la répétition (R : $F = 3.145$, $p = 0.048 < 0.05$), avec pour les 3 sessions des moyennes successives de 5,517, 5,486, et 5,348. L'interaction Mode Visuel - Répétition est également significative (V*R : $F = 4.08$, $p = 0.02 < 0.05$) et montre que les différences de perception dues au mode visuel sont de moins en moins significatives au cours des sessions (p <

0.001 pour la session 1 et p = 0.024 pour la session 2) jusqu'à ne plus être significatives du tout lors de la troisième et dernière session (p = 0.276, figure 3).

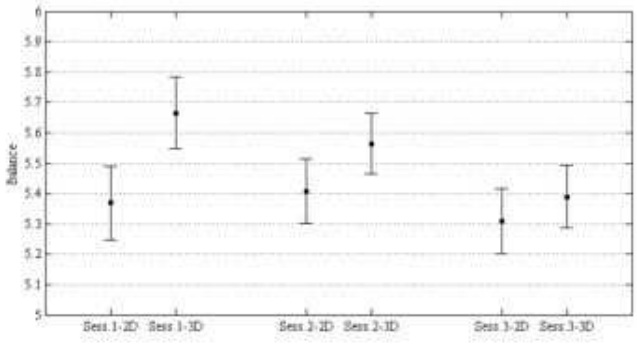


Figure 3 : Effet du Mode Visuel en fonction de la Session.

4.3 Groupes : Proximité de l'écran vs. « Sweet Spot »

Les résultats indiquent une influence significative du groupe (G : F = 4.4E+16, p < 0.001) avec une valeur moyenne de 5.535 pour le groupe 1 et une valeur moyenne de 5.366 pour le groupe 2 ce qui montre que les sujets assis en proximité de l'écran ont globalement perçu les mixages comme étant plus frontaux que les sujets assis au « Sweet Spot ». L'interaction entre le Mode Visuel et le Groupe (V*G : F = 6.469, p = 0.015 < 0.05) est également significative. Un post hoc LSD de Fisher montre qu'il n'y a finalement d'effet significatif du Mode Visuel que pour le groupe 2 (figure 4), soit le groupe assis au « Sweet Spot » (p < 0.001), tandis qu'il n'y en a pas pour le groupe 1 assis en proximité de l'écran (p = 0.700).

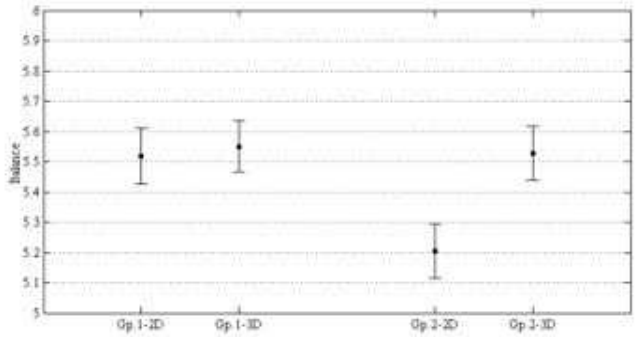


Figure 4 : Effet du Mode Visuel en fonction du Groupe.

4.4 Balance : « Plutôt Frontal » vs. « Plutôt Surround »

Les sujets ont globalement bien perçu la différence entre balances « plutôt frontales » et « plutôt surround » (B : F = 156.687, p < 0.001) avec une moyenne de 5.855 pour les balances frontales et une moyenne de 5.046 pour les balances « surround ». Cependant, l'interaction Mode Visuel - Balance (V*B : F = 0.445, p = 0.509) n'est pas significative, ce qui montre que la balance avant/arrière proposée aux sujets n'a pas d'influence sur l'éventuelle différence d'équilibre avant/arrière que vont ressentir les sujets lorsque l'image va passer de 2D à 3D (ou inversement).

5 Discussion

Les résultats montrent un effet global du Mode Visuel : lorsque l'image passe en stéréoscopie, le mixage tend à paraître plus frontal (ou autrement dit, à « manquer d'arrière »), ce qui est cohérent avec les témoignages de plusieurs ingénieurs du son [1], qui affirment produire des mixages plus « surround » pour la version 3D d'un film que pour sa version 2D.

Cependant, l'analyse des interactions montre que cet effet du Mode Visuel ne concerne globalement que 2 séquences sur les 8 proposées et n'est valable que pour certains emplacements dans la salle (ici, l'effet ne fonctionne qu'au « Sweet Spot » et disparaît en proximité de l'écran).

Cet effet a également tendance à s'atténuer au cours des sessions jusqu'à disparaître complètement lors de la dernière session. Au vu des témoignages des sujets à l'issue de l'expérience, il semblerait que la fatigue, favorisée par la diffusion d'images 3D et par la monotonie des séquences répétées chacune 12 fois, soit la principale responsable de ce phénomène. Les sujets avaient tendance à perdre progressivement leur esprit critique (ne trouvant les mixages ni trop frontaux ni trop « surround » qu'importe le mode visuel) et finissait par n'exploiter que le milieu de l'échelle de gradations, ce qui explique non seulement pourquoi les différences sont de moins en moins significatives, mais aussi pourquoi les moyennes se rapprochent sensiblement de la note 5/10 au cours des sessions. Cependant, le nombre de séquences présentant des différences significatives reste faible, même en début de test (3 sur 8, moins de la moitié).

6 Conclusion

Les résultats de la présente étude rejoignent ceux obtenus lors de l'expérience préliminaire : l'influence de la stéréoscopie apparaît comme étant faible et dépend étroitement du contenu de la séquence. Cette influence semble également fragile dès lors qu'on s'écarte des positions optimales d'écoute ou que la fatigue devient trop importante.

Un prochain test subjectif tentera de déterminer si les séquences significatives correspondent aux séquences donnant aux sujets la plus grande sensation de profondeur visuelle.

Remerciements

Les auteurs tiennent à remercier Erwan le Morvan (Université de Brest), Pierre Souchar (3D Fovéa), le cinéma « Le Bretagne » de Saint-Renan, les réalisateurs Joséphine Derobe, Pablo Agüero et Henry Davies, ainsi que tous les étudiants de la formation Image & Son de Brest (ISB) qui ont participé aux tests subjectifs. Cette recherche s'inscrit dans le projet « Cross Channel Film Lab 2 », sélectionné dans le cadre du programme européen de coopération transfrontalière INTERREG IV A France (Manche) - Angleterre, cofinancé par le FEDER.

Références

[1] V. Gambier, Nouvelle approche pour la bande sonore d'un film en relief, *Mémoire de fin d'études, ENS Louis-Lumière*, page 55 (2010).

[2] M. Coleman, *The sound of Hugo*. http://soundworkscollection.com/videos/hugo.

[3] M. Coleman, *The sound of Avatar*. http://soundworkscollection.com/videos/avatar.

[4] W. Ijsselsteijn, H. de Ridder, J. Freeman, S. E. Avons, and D. Bouwhuis, Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence, *Presence-Teleop. Virt.*, vol. 10, no.3, p. 298-311 (2001).

[5] C. R. André, M. Rébillat, and B. F. G. Katz, Sound for 3D cinema and the sense of presence, *Proceedings of the 18th International Conference on Auditory Display* (2012).

[6] C. André, J. Embrechts, and J. G. Verly, Adding 3D sound to 3D cinema: Identification and evaluation of different reproduction techniques, *Audio Language and Image Processing*, pages 130-137 (2010).

[7] M. Schmidt, S. Schwartz, and J. Larsen, Interactive 3D audio: enhancing awareness of details in immersive soundscapes? *133ème Convention AES*, Preprint 8780 (2012).

[8] M. Chion, *L'audiovision*, Armand Colin, pages 93-128 (2005).

[9] F. Rumsey, Spatial quality evaluation for reproduced sound : terminology, meaning, and a scene-based paradigm, *J. Audio Eng. Soc.*, Vol. 50, No. 9 (2002).

[10] M. Chion, *Un art sonore, le cinéma*, Les Cahiers du Cinéma, page 122 (2003).

[11] L. Jullier, *Le son au cinéma*, Les Cahiers du Cinéma, page 17 (2006).

[12] E. Hendrickx, M. Paquier, V. Koehl, Should a movie have two different soundtracks for its stereoscopic and non-stereoscopic versions? A case study on the front/rear balance, *Proceedings of the International Conference on 3D Imaging* (2013).

[13] E.C. Poulton, *Bias in quantifying judgments*, Lawrence Erlbaum Associates, Hillsdale, United States of America, 1992.

[14] J.C. Nunnally, I.H. Bernstein, *Psychometric Theory*, McGraw-Hill, 1994.

[15] S. Bech, N. Zacharov, *Perceptual Audio Evaluation*, Wiley-Blackwell, pages 162-169, page 302 (2006).

[16] AES recommended practice for professional audio – subjective evaluation of loudspeakers, *J. Audio Eng. Soc.* (1996, reaffirmed 2007).

[17] F. Toole, *Sound reproduction*, Focal Press, pages 328-330 (2008).

[18] R. Schatz, S. Egger, K. Masuch, The impact of test duration on user fatigue and reliability of subjective quality ratings, *J. Audio Eng. Soc.*, Vol. 60, No. 1/2, (2012).