HAL
open science

Interfaçage visuel stéréoscopique : diminution de la fatigue visuelle et caractérisation de la perception des formes

Laure Leroy

To cite this version:

Laure Leroy. Interfaçage visuel stéréoscopique : diminution de la fatigue visuelle et caractérisation de la perception des formes. domain_other. École Nationale Supérieure des Mines de Paris, 2009. Français. NNT : 2009ENMP1673 . pastel-00005882

HAL Id: pastel-00005882

https://pastel.hal.science/pastel-00005882v1

Submitted on 16 Mar 2010

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.

ED n°431 : ICMS – Information, communication, modélisation et simulation

T H E S E

pour obtenir le grade de

DOCTEUR DE L'ECOLE NATIONALE SUPERIEURE DES MINES DE PARIS

Spécialité "Informatique temps réel, robotique et automatique"

présentée et soutenue publiquement par

Laure LEROY

le 15 décembre 2009

Interfaçage visuel stéréoscopique : diminution de la fatigue visuelle et caractérisation de la perception des formes

Directeurs de thèse : Philippe Fuchs, Guillaume Moreau

Jury

M. Jacques DROULEZ	Président
Mme. Corinne ROUMES	Rapporteur
M. Patrick LE CALLET	Rapporteur
M. Jacques DROULEZ	Examinateur
M. Jean-Pierre JESSEL	Examinateur
M. Philippe FUCHS	Examinateur
M. Guillaume MOREAU	Examinateur
M. Frédéric ACHARD	Examinateur invité

Remerciements

Je tiens à remercier

- Philippe Fuchs qui a dirigé cette thèse. Son œil critique et sa rigueur m'ont été très précieux pour structurer le travail et pour améliorer la qualité des différentes sections. Sa confiance et sa gentillesse m'ont énormément touchée.
- Guillaume Moreau qui a codirigé ma thèse. Il a toujours répondu présent quand j'en avais besoin, malgré un emploi du temps très chargé. Ses connaissances en informatique m'ont été d'un grand soutien.
- Jacques Droulez d'avoir accepté de présider le jury de ma thèse. Les longues conversations que nous avons eues m'ont beaucoup appris et impressionnée.
- Corinne Roumes qui a rapporté ma thèse. Sa rigueur scientifique et sa maitrise du sujet m'ont réellement fait avancer à chaque fois que j'ai eu la chance de la rencontrer tout au long de ces trois ans.
- Patrick LeCallet qui a également rapporté ma thèse. Sa gentillesse m'a beaucoup touchée. La justesse de ses remarques m'a encouragée à être plus précise dans mes futures recherches.
- Jean-Pierre Jessel et Frédéric Achard pour leur participation à mon jury de thèse et leur intérêt pour mes travaux. Leurs questions et remarques très judicieuse ont amenée des discussions très intéressantes.
- Claude Laurgeau et Arnaud de la Fortelle de m'avoir accueillie si chaleureusement dans leur laboratoire.
- Alexis Palijc qui m'a si souvent aidée au long de ma thèse, tant sur le plan humain que scientifique et technique. Il a toujours été là quand j'en avais besoin.
- Vincent Meyrueis qui a donné naissance à la forme inconnue des tests de perception et qui m'a toujours aidée à trouver les bugs de mes différents programmes.
- Pascaline Neveu qui m'a énormément aidée dans les méthodes de mesure de fatigue visuelle, et qui m'a beaucoup fait rire à chaque fois que nous nous sommes croisées.
- Véronique Chastres qui a pris le temps de m'aider pour toute la partie statistique.
- Bruno Leroy, Claire Desbant, Emmanuelle Goumarre et Patricia Azame pour les nombreuses heures passées à traquer les fautes d'orthographe de ce manuscrit
- Safwan Chendeb qui a toujours été présent dans les pires moments de doutes comme dans les plus grandes joies.
- Tous mes collègues et amis du CAOR, de la cafétéria et de l'école des Mines en général qui ont réellement ensoleillé ces trois ans.

Table des matières

INTRODUCTION **9**

CHAPITRE 1 : DE LA VISION HUMAINE AUX INTERFACES VISUELLES **11**

1.1. LA VISION 11
1.1.1. Le système visuel humain *11*
1.1.2. L'œil *12*
1.1.3. Accommodation et convergence *13*
1.1.4. La rétine *14*
1.1.5. Les fréquences spatiales dans la vision *15*
1.2. PERCEPTION DU RELIEF 17
1.2.1. Perception monoculaire du relief *17*
1.2.2. Perception binoculaire du relief *21*
1.2.3. Restitution virtuelle d'un environnement 3D *25*
1.3. CONCLUSION 32

PARTIE N°1 : PERCEPTION DES FORMES EN IMMERSION **33**

CHAPITRE 2 : ETAT DE L'ART DE LA PERCEPTION DES FORMES EN RELIEF **35**

2.1. PERCEPTION DES FORMES EN RELIEF DANS LA VISION NATURELLE 35
2.1.1. Perception des inclinaisons et obliquités *35*
2.1.2. Perception des courbures *39*
2.1.3. Perception des objets *45*
2.2. PERCEPTION DES FORMES EN ENVIRONNEMENT IMMERSIF 48
2.2.1. Influence du fond d'écran *48*
2.2.2. Influence de la distance inter-oculaire *49*
2.2.3. Influence de l'accommodation et le flou dans la vision périphérique *49*
2.3. CONCLUSION 50

CHAPITRE 3 : EXPERIMENTATIONS ET RESULTATS SUR LA PERCEPTION DE FORMES EN IMMERSION **51**

3.1. PSYCHOPHYSIQUE ET STATISTIQUE 51
3.1.1. Fondement *51*
3.1.2. Détection de seuil *52*
3.1.3. Fonction psychophysique *54*
3.1.4. Comparaison statistique de populations *55*
3.2. PROTOCOLE DE TEST 58
3.2.1. Question de recherche *58*
3.2.2. Hypothèses relatives aux questions *58*
3.2.3. Etude empirique *59*
3.2.4. Dispositif physique *65*
3.3. ANALYSES DES DONNEES 66
3.3.1. Collecte et prétraitement des données *66*
3.3.2. Hypothèse n° 1 : Il est préférable d'avoir le point de vue traqué en vision monoscopique *70*
3.3.3. Hypothèse n° 2 : Il est préférable d'avoir le point de vue traqué en vision ortho-stéréoscopique *80*
3.3.4. Hypothèse n° 3 : Il est préférable d'avoir le point de vue traqué en vision non-ortho-stéréoscopique *86*
3.3.5. Hypothèse n° 4 : Il est préférable d'avoir la vision ortho-stéréoscopique que la vision non-ortho-stéréoscopique avec le point de vue traqué *92*
3.3.6. Hypothèse n° 5 : Il est préférable d'avoir la vision ortho-stéréoscopique que la vision non-ortho-stéréoscopique sans le point de vue traqué *97*

3.3.7. Hypothèse n° 6 : Il est préférable d'avoir la vision monoscopique avec le point de vue traqué que la vision ortho-stéréoscopique sans le point de vue traqué *103*
3.4. CONCLUSION 109

CHAPITRE 4 : DISCUSSION DES RESULTATS SUR LA PERCEPTION DES FORMES EN RELIEF EN IMMERSION 111

4.1. INTERPRETATION DES RESULTATS 111
4.1.1. Interprétation générale *111*
4.1.2. Comparaison avec d'autres expériences *115*
4.2. LES LIMITES DE L'EXPERIENCE 117
4.2.1. Ordre d'exécution des séquences *117*
4.2.2. Interface visuelle *117*

PARTIE 2 : DIMINUTION DE LA FATIGUE VISUELLE EN VISION STEREOSCOPIQUE ... 119

CHAPITRE 5 : ETAT DE L'ART DE LA FATIGUE VISUELLE EN VISION STEREOSCOPIQUE 121

5.1. CAUSES DE LA FATIGUE VISUELLE EN STEREOSCOPIE 122
5.1.1. Problèmes inhérents au principe de la vision stéréoscopique *122*
5.1.2. Les problèmes secondaires de technique *125*
5.2. LES PROBLEMES DE HAUTES FREQUENCES SPATIALES 128
5.2.1. Limite en fusion *128*
5.2.2. Le confort et les hautes fréquences. *129*
5.3. CONCLUSION 130

CHAPITRE 6 : TRAITEMENT ADAPTATIF POUR LA DIMINUTION DE LA FATIGUE VISUELLE 131

6.1. EXPLICATION DE CHAQUE ALGORITHME 131
6.1.1. Les ondelettes *131*
6.1.2. BOX FILTER *146*
6.2. COMPARAISON DES ALGORITHMES 152
6.2.1. Comparaison ondelettes de Haar/BOX FILTER *152*
6.2.2. Comparaison BOX FILTER/moyenne sur carte graphique *155*
6.3. UTILISATION DE MOYENNE GLISSANTE ET AUTRES « FLOUS » 156
6.3.1. Explication *156*
6.3.2. Comparaison BOX FILTER glissant / moyenne glissante *158*
6.3.3. Autres « flous » *161*
6.4. COMPARAISON GLOBALE 165
6.5. CONCLUSION 166

CHAPITRE 7 : EXPERIMENTATIONS ET RESULTATS BRUTS SUR LA DIMINUTION DE LA FATIGUE VISUELLE EN VISION STEREOSCOPIQUE 167

7.1. INTRODUCTION 167
7.2. MOYENS DE MESURER LA FATIGUE VISUELLE 167
7.2.1. Les tests plus objectifs *167*
7.2.2. Les tests subjectifs *174*
7.3. LE PROTOCOLE 174
7.3.1. La tâche *174*
7.3.2. Les mesures de fatigues effectuées *176*
7.3.3. Déroulement *177*
7.3.4. Les sujets *177*
7.3.5. Calcul et Algorithme choisi *178*
7.3.6. Dispositif physique *180*
7.4. ANALYSE DES RESULTATS 180
7.4.1. Fonction proximum d'accommodation *180*
7.4.2. Facilité d'accommodation *181*
7.4.3. Acuité stéréoscopique *182*

7.4.4. Efficacité lors de l'exécution de la tâche *183*
7.4.5. Mesures subjectives *183*
7.5. CONCLUSION 185

CHAPITRE 8 : DISCUSSION DES RESULTATS ET DE LA METHODE 187

8.1. DISCUSSION DES RESULTATS 187
8.2. TRAITER LES MATERIAUX OU TRAITER L'IMAGE ? 187
8.2.1. Traiter les matériaux *188*
8.2.2. Traiter l'image *190*
8.3. CONCLUSION 191

CONCLUSION ET PERSPECTIVES 193

CONCLUSION 193
IMPLICATIONS INDUSTRIELLES 194
Perception des formes *194*
Fatigue visuelle en vision stéréoscopique *197*
PERSPECTIVES 198

BIBLIOGRAPHIE 199

TABLE DES ILLUSTRATIONS 205

ANNEXES 213

ANNEXE A : GRAPHIQUES SUPPLEMENTAIRES POUR LA PERCEPTION DES FORMES EN IMMERSION 215

ANNEXE B : TABLEAU RECAPITULATIF POUR LA PERCEPTION DES FORMES EN IMMERSION 236

ANNEXE C : CALCUL DE SIGNIFICATIVITE POUR L'AMPLITUDE ACCOMMODATIVE 248

ANNEXE D : COMPARAISON DE PLUSIEURS DISTRIBUTIONS 250

ANNEXE E : LIEN ENTRE FLOU ET HAUTES FREQUENCES 254

Introduction

Les dispositifs d'immersion stéréoscopique sont très répandus et tendent à l'être de plus en plus (travail en immersion, film projeté en vision stéréoscopique, écrans auto-stéréoscopiques de plus en plus nombreux). Dans certains cas ces immersions peuvent être longues (plusieurs heures) et fatigantes pour les yeux parce qu'il persiste quelques différences entre la vision naturelle et artificielle. Il faut alors agir sur la présentation de l'information pour diminuer la fatigue visuelle pour que l'observateur ait le moins de gêne possible. Nous présentons dans ce manuscrit une technique pour diminuer cette fatigue visuelle. Certaines applications demandent une perception des formes très performante (revue de projet en immersion par exemple). Nous cherchons les paramètres influant cette perception des formes et nous quantifions ces influences.

Le projet IHS10 (Interface Homme Système 2010), inclus dans le pôle de compétitivité Systematic, doit permettre aux concepteurs de voiture de visualiser une voiture entièrement en virtuel qui n'est pas encore construite. Un prototype réel coûte cher, aussi bien en argent qu'en temps et en matière première. L'idéal serait de pouvoir visualiser le prototype pour déceler les problèmes sans avoir à le construire. Ainsi, ce projet vise à présenter un prototype dans une salle immersive plutôt que de le construire réellement, ou en tout cas, le construire le plus tard possible. Pour ce faire, il est extrêmement important que les designers et les ingénieurs puissent percevoir ce prototype au plus près de celui qu'ils auraient perçu en réel. Il faut donc qu'une bonne perception des formes en immersion soit assurée et, si ce n'est pas le cas, il faut pouvoir quantifier ces limites. De même, les ingénieurs et designers doivent pouvoir travailler plusieurs heures sur le prototype, pour déterminer les modifications à faire sur la voiture le plus rapidement possible. De ce fait, il est important que le dispositif d'immersion ne provoque pas de trop grande fatigue visuelle. En effet, il n'est évidemment pas souhaitable de provoquer des migraines et douleurs oculaires. Nos résultats seront dont utilisés dans ce projet. L'étude sur la perception des formes permettra aux industriels de savoir la précision visuelle qu'ils peuvent attendre de leur système d'immersion pour la perception des formes, du tableau de bord par exemple.

Ce mémoire de thèse est divisé en deux parties principales. La première est consacrée à la perception des formes en immersion. Elle est composée des chapitres 2, 3 et 4. La seconde partie traite la fatigue visuelle en vision stéréoscopique. En particulier celle générée par les hautes fréquences spatiales aux endroits de fortes disparités. Cette partie est composée des chapitres 5, 6, 7 et 8.

Le premier chapitre est un chapitre général dans lequel nous aborderons le système visuel humain. Nous nous intéresserons à son anatomie, ses mécanismes et son fonctionnement global. Nous nous pencherons ensuite sur la perception du relief. Nous verrons qu'il y a plusieurs types d'indices : les indices monoculaires statiques et dynamiques, les indices proprioceptifs (accommodation et vergence) et enfin les indices stéréoscopiques. Nous expliquerons ensuite les principes de la reproduction artificielle de la vision binoculaire. Nous verrons les avantages qu'elle offre mais également les problèmes qu'elle pose. Nous présenterons quelques solutions à ces problèmes.

Le deuxième chapitre est le premier de la première partie. Il constitue l'état de l'art de la perception des formes. Nous commencerons par la perception des plans inclinés et obliques basée sur la dérivée première des disparités horizontales, sur la dérivée des parallaxes de mouvements ainsi que sur les gradients de textures. Nous poursuivrons ensuite avec la perception des courbures grâce à la dérivée seconde des disparités horizontales et des parallaxes de mouvement, ainsi que les

changements de textures et les gradients de lumière. Nous terminerons par une introduction à la perception des objets.

Le troisième chapitre est le chapitre clé de la première partie. Il débute par une introduction à la psychophysique ainsi qu'au traitement statistique des données. Nous poursuivrons par notre protocole de test et nous y exposerons nos hypothèses de travail. La moitié de ce chapitre est consacré à l'exposition des résultats de nos expériences. Nous décortiquerons hypothèse par hypothèse et configuration par configuration.

Le quatrième chapitre discute des résultats obtenus dans le chapitre trois. Nous commencerons par les résumer et par les analyser. Ensuite nous les comparerons à d'autres expérimentations plus anciennes et évoquées dans le chapitre 2. Ce chapitre clôt la première partie.

Le cinquième chapitre commence la seconde partie. Il expose un état de l'art de la fatigue en vision stéréoscopique. Elle est composée d'une partie inhérente à la vision stéréoscopique, l'autre partie est due à des problèmes de technologie, la troisième est celle que nous traiterons plus spécifiquement : les hautes fréquences aux endroits de grandes disparités.

Le sixième chapitre est plus algorithmique. Il présente les trois algorithmes que nous avons étudiés pour supprimer les hautes fréquences aux endroits des grandes disparités. Nous commencerons par expliquer les fonctionnements de la transformée en ondelettes, des Box Filter ainsi que du flou par moyenne et moyenne glissante. Une comparaison mathématique et technique entre ces algorithmes suit ces explications. Ensuite nous choisirons un de ces algorithmes pour faire nos tests de fatigue visuelle et, bien sûr, nous justifierons ce choix.

Le septième chapitre expose l'expérimentation sur la fatigue visuelle en elle-même. Il commencera par un petit état de l'art de la mesure de la fatigue visuelle. Ensuite nous exposerons notre protocole. Ce chapitre se terminera bien entendu par nos résultats sur l'influence de notre traitement sur la fatigue visuelle, sur l'efficacité de la tâche et sur le ressenti des observateurs.

Le chapitre huit clôt la seconde partie. C'est une discussion des résultats obtenus dans le chapitre précédent. Nous commencerons par les résumer et par les analyser. Ensuite nous discuterons de deux méthodes différentes de suppression des hautes fréquences.

Le chapitre neuf termine ce manuscrit. Nous y concluons notre exposé et nous expliquons les perspectives de notre travail.

Chapitre 1 : De la vision humaine aux interfaces visuelles

Dans ce chapitre, nous présentons la vision humaine et plus particulièrement la vision binoculaire. A partir des caractéristiques de la vision humaine, nous envisageons les solutions et les problèmes posés par la restitution visuelle d'un environnement virtuel. Nous explicitons plus en détails la perception du relief.

1.1. La vision

Une interface visuelle de qualité est plus délicate à concevoir qu'il n'y paraît, car la vision humaine est un sens complexe et souvent très méconnu des non spécialistes. Mais celle-ci est si naturelle qu'on en oublie sa complexité et sa puissance informative. Il est important de retenir que l'étude de ce sens, comme des autres, ne peut être parfaitement comprise qu'à partir de ses fonctionnements physiques, physiologiques et cognitifs.

Le fonctionnement du système visuel se base sur une analyse statique des images qui sont projetées sur la rétine, mais il ne se limite pas à cette analyse. Celui-ci construit dynamiquement un monde visuel. A la réception sur la rétine de la lumière, celle-ci est transformée en signaux chimiques puis électriques et ensuite des traitements spécialisés sont effectués en parallèle pour décomposer l'information. Les images formées sur les deux rétines sont décomposées selon des critères de profondeur, de forme, de couleur et de mouvement. Mais avant l'extraction des formes ou de la profondeur, des informations de bas niveau sont traitées et exploitées : des canaux fréquentiels travaillant en parallèle sont chargés d'établir des représentations de l'information visuelle à différentes échelles de l'espace observé.

Après avoir brièvement décrit l'organisation du système visuel, nous nous concentrerons sur le fonctionnement de ces canaux et présenterons une modélisation mathématique liée à la notion de fréquence dans le but d'analyser le fonctionnement du système visuel. Nous utiliserons en pratique cette analyse pour l'amélioration de la vision stéréoscopique. Mais pour cela, après l'étude des traitements monoculaires de l'information, nous aborderons l'étude de la perception de la profondeur par vision monoculaire et par vision binoculaire.

1.1.1. Le système visuel humain

Les deux yeux, récepteurs sensoriels du système visuel, participent à l'observation de l'espace. Les nerfs optiques sortant des yeux se regroupent au niveau du chiasma optique où les fibres temporales gauches rejoignent les fibres nasales droites et inversement. Les fibres du nerf optique sont alors dirigées vers le corps genouillé latéral, sauf une petite partie des fibres qui est dirigée vers le colliculus supérieur. Un des rôles supposés du corps genouillé latéral est de séparer en différentes couches cellulaires les neurones recevant les afférences rétiniennes. Du corps genouillé latéral, l'information est envoyée au cortex occipital, encore appelé cortex strié de par sa structure lamellaire. La Figure 1 montre le trajet de l'information visuelle depuis l'œil jusqu'au cortex visuel.

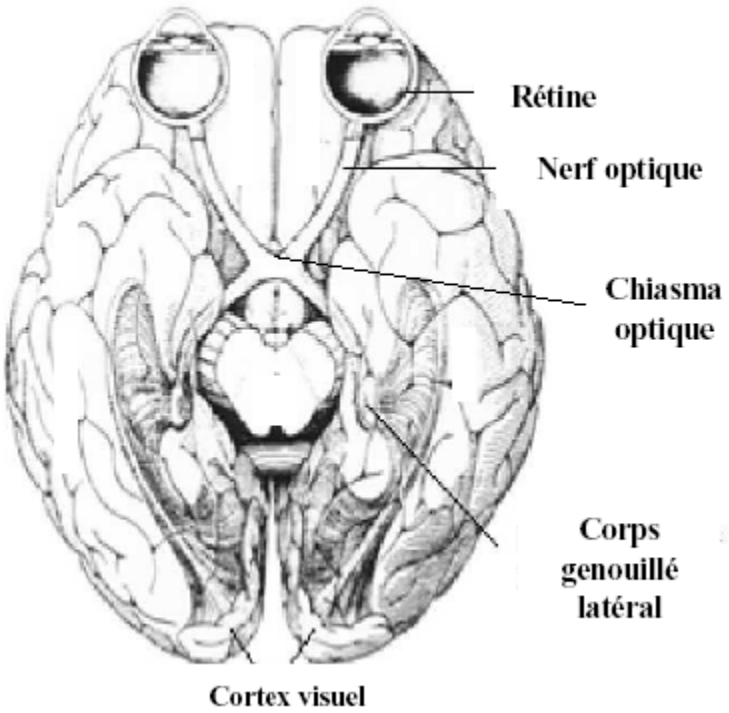


Figure 1 : projections rétiniennes vers le corps genouillé latéral, puis le cortex strié (Fuchs, 2006)

Le cortex strié, ou cortex visuel primaire ou aire V1, est organisé de manière à produire une représentation de l'information incidente sur la rétine. Comme dans le corps genouillé latéral, toutes les régions de la rétine y sont projetées et la représentation spatiale y conserve ses contiguïtés (DeValois, et al., 1988) : l'organisation du cortex strié est dite rétinotopique. La topologie est toutefois modifiée, la rétine centrale se projetant avec un facteur de dilatation, plus important que celui de la rétine périphérique, conséquence notamment de la densité de l'échantillonnage rétinien. La représentation de l'information dans le cortex strié est cependant beaucoup plus complexe que la simple organisation géométrique. Le traitement cortical de l'information ne se limite pas au cortex strié. Au-delà, une trentaine d'aires corticales a été décrite. Ces aires corticales diffèrent notamment par les propriétés de sélectivité de leurs neurones.

1.1.2. L'œil

Les principales parties de l'œil sont rappelées sur la Figure 2.

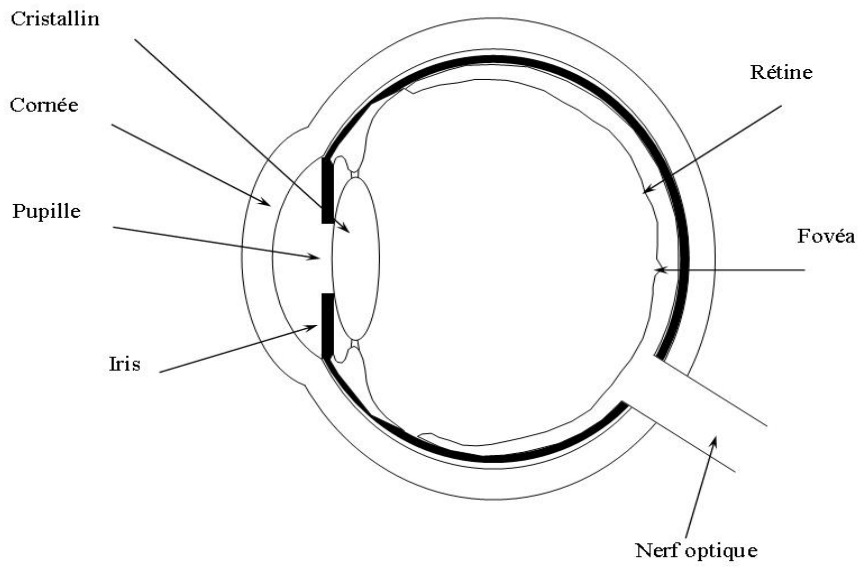


Figure 2 : les principales composantes de l'œil (Fuchs, 2006)

La fonction des yeux est de canaliser la lumière, pour des longueurs d'ondes comprises entre 400 et 700nm, émise ou réfléchie par un objet pour créer une image nette qui s'imprime sur la partie de l'œil couverte de récepteurs sensoriels : la rétine. L'œil est constitué d'une succession de milieux transparents qui jouent le rôle d'une lentille convergente, dont la focale globale peut varier par modification de la courbure du cristallin. Les rayons lumineux convergent sur le centre de la rétine, la fovéa. Celle-ci a une densité de récepteurs sensoriels nettement plus élevée que les autres zones de la rétine.

1.1.3. Accommodation et convergence

En regardant un objet à moins de 65 mètres environ (Fuchs, 2006), l'image réfléchie deviendrait floue sur la rétine si l'œil ne possédait le pouvoir d'accommoder automatiquement. Sous l'action des muscles ciliaires, la puissance du cristallin varie et permet de mettre au point sur la rétine pour voir des objets proches ou éloignés (Figure 3) : c'est le phénomène d'accommodation. Le cristallin est le seul dioptre de la chaîne visuelle dont la puissance soit variable. Tous les rayons lumineux sont déviés vers la fovéa qui est le point central de la rétine. Le niveau d'accommodation est réglé pour obtenir une image nette.

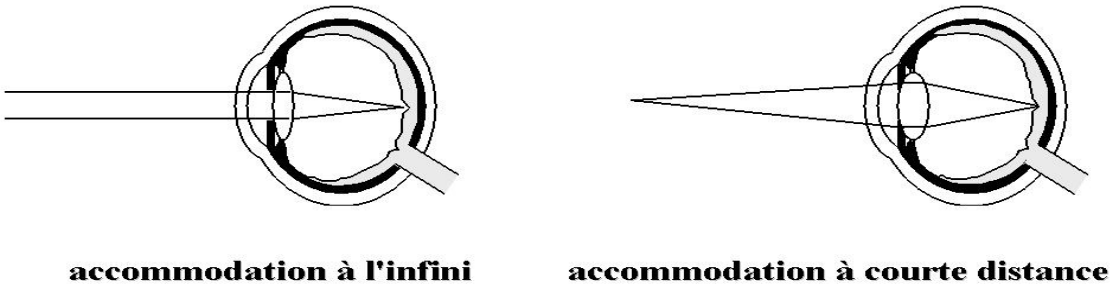


Figure 3 : l'accommodation (Fuchs, 2006)

Les muscles des globes orbitaux permettent d'orienter les deux yeux en les faisant converger vers le point de l'espace observé. C'est le phénomène de convergence (Figure 4). Comme l'accommodation, la convergence des yeux se fait inconsciemment, sauf si la personne veut loucher. Le mouvement des yeux dans les globes orbitaux, pour changer le point de convergence ou (et) pour suivre un objet en déplacement, peut être très rapide avec une vitesse maximum de 600 degrés par seconde environ.

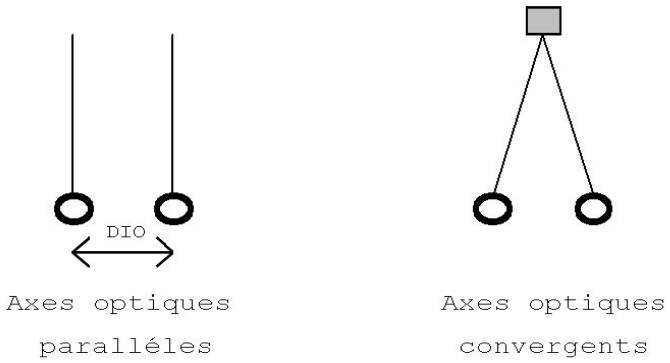


Figure 4 : la convergence

1.1.4. La rétine

La rétine est le lieu de la transformation de l'énergie lumineuse en signal électrochimique discret. Ce signal sort de l'œil par le nerf optique (Figure 5).

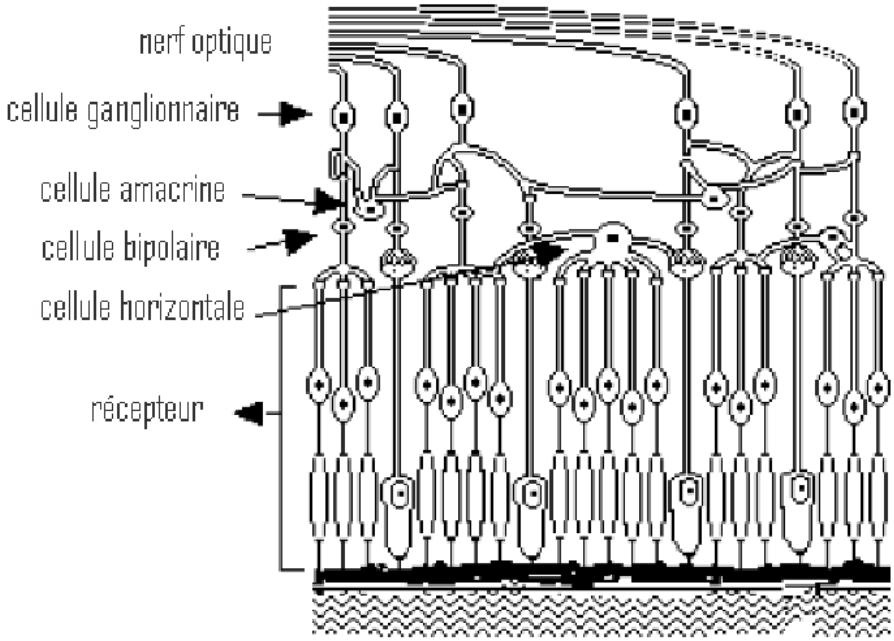


Figure 5 : la structure anatomique de la rétine (Fuchs, 2006)

Elle est composée de plusieurs couches de neurones qui recouvrent le fond de l'œil. Ces couches forment un tapis de cellules connectées horizontalement au sein d'une même couche, et verticalement d'une couche à l'autre créant une architecture neuronale fortement pyramidale. En effet, si le nombre de photorécepteurs est de l'ordre de 100 millions, le nombre des cellules ganglionnaires, dont les axones constituent le nerf optique, est de l'ordre du million. Entre les photorécepteurs et les cellules ganglionnaires, nous trouvons les cellules horizontales, les cellules bipolaires et les cellules amacrines. Paradoxalement, la première couche de cellules exposée à la lumière n'est pas celle des photorécepteurs, mais celle des cellules ganglionnaires, sauf au niveau de la fovéa. Les axones des cellules ganglionnaires sortent toutes de la rétine dans la même région dépourvue de photorécepteurs appelée papille optique. Les photorécepteurs sont reliés entre eux et aux cellules bipolaires par les cellules horizontales, qui forment la couche plexiforme externe. Les cellules bipolaires créent des liens verticaux en reliant les photorécepteurs aux cellules de la couche suivante.

1.1.5. Les fréquences spatiales dans la vision

Un son est caractérisé par sa fréquence, traduisant par là la rapidité de ses oscillations. De même, nous pouvons définir pour un stimulus visuel une mesure de la finesse de ses formes. C'est la notion de fréquence spatiale. La définition de cette fréquence est la suivante : c'est le nombre de cycles d'une oscillation périodique de luminance ou de couleur dans un degré d'angle visuel. Sa valeur est donc donnée en cycles par degré (cpd), ce que nous avons représenté schématiquement sur la Figure 6.

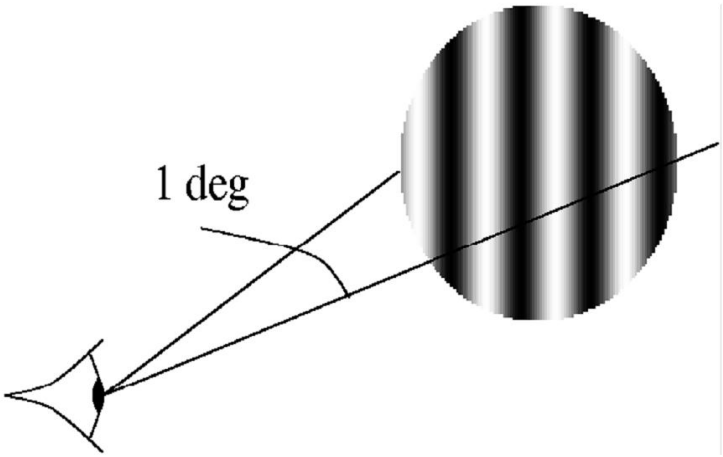


Figure 6 : définition d'une fréquence spatiale (Fuchs, 2006)

La fréquence spatiale, intuitivement décrite ici, peut être définie dans un cadre mathématique rigoureux (Perrin, 1998). Campbell et Robson ont suggéré en 1968 que le système visuel pourrait être constitué de groupes de filtres passe-bande quasi-linéaires et indépendants, chacun centré assez étroitement sur une bande de fréquences spatiales (Campbell, et al., 1968) on parle de canaux fréquentiels. La fonction de sensibilité au contraste (capacité du système visuel à détecter des différences de luminances) ne refléterait donc pas la sensibilité d'un type de cellules mais serait l'enveloppe des sensibilités de tous ces filtres. Plusieurs aspects de cette hypothèse sont extrêmement attirants et motivent un grand nombre de recherches sur le sujet :

- la similarité avec le système auditif, ajoutée à l'idée que le système nerveux traite probablement des problèmes analogues par des méthodes semblables ;
- la compression de l'information : en effet, l'information visuelle possède un grand nombre de périodicités (au moins localement) qui doivent être codées économiquement puisque le nerf optique ne possède qu'un million de fibres pour cent vingt six millions de récepteurs.

Un modèle simple de canaux fréquentiels a été décrit par Wilson et Bergen, composé de quatre filtres de tailles différentes pour l'analyse des images rétiniennes (Wilson, et al., 1979). Ces canaux ont pour largeurs centrales 3,1 min arc, 6,2 min arc, 11,7 min arc et 21 min arc au niveau de la fovéa, leur largeur croissant avec l'excentricité (Marr, 1982). Tous les auteurs ne sont pas d'accord sur le nombre de canaux fréquentiels et encore moins sur les fréquences centrales. Ainsi, Watson identifie en chaque emplacement de la fovéa sept détecteurs élémentaires qui couvrent les fréquences depuis 0,25 cpd jusqu'à 32 cpd (Watson, et al., 1983). Wilson et al. ont mis en évidence une structure avec des largeurs de bande relatives décroissant de 2,5 octaves (à 0,5 cpd) à 1,5 octaves (à 4 cpd) pour rester constantes ensuite (Wilson, et al., 1983). Les largeurs de bande sont proportionnelles aux fréquences centrales. Cela suggère une représentation mathématique espace-échelle plus qu'espace-fréquence de la scène visuelle et explique l'intérêt d'utiliser les analyses par ondelettes dans le domaine de la vision (Perrin, 1998) (voir paragraphe 6.1.1 pour l'expression mathématique des ondelettes).

L'analyse en ondelettes offre une bonne modélisation du traitement fréquentiel opéré par les cellules ganglionnaires. En conclusion, l'étude détaillée des mécanismes du traitement de l'information par le système visuel peut sembler, pour le lecteur, éloigné des préoccupations du domaine de la réalité virtuelle. Nous verrons qu'il n'en est rien. Quel intérêt y aurait-il à vouloir faire une immersion par des images en relief si après quelques instants l'utilisateur a du mal à les observer ?

1.2. Perception du relief

1.2.1. Perception monoculaire du relief

Nous évoluons dans un monde en trois dimensions et nous le percevons comme tel, que ce soit par la vision, le toucher, l'ouïe, etc. En effet, nous sommes capables de sentir, en le touchant, qu'un objet est en relief, si sa granularité est profonde, s'il a des arrêtes tranchantes... Nous pouvons également entendre lorsqu'un son provient de la gauche, de la droite, de devant ou de derrière... De même, la vision humaine est tout à fait adaptée pour évaluer le relief. En effet, grâce à nos yeux, il nous est possible de percevoir si un élément du décor est plus proche ou plus éloigné qu'un autre et ceci en fonction des ombres, de sa taille apparente, de son mouvement, de la façon dont nos deux yeux le perçoivent différemment (vision stéréoscopique), etc. Nous allons préciser en détail les différents indices intervenant dans la perception du relief, de la troisième dimension de l'espace.

Dans ce chapitre, nous ne considérons pas la perception de la profondeur par le toucher et l'ouïe pour se concentrer sur la perception visuelle des trois dimensions du monde qui nous entoure. Elle se base sur trois types d'indices et il convient, si possible, de ne pas avoir d'indices contradictoires lorsque nous restituons visuellement un environnement virtuel.

1.2.1.1. Indices statiques

Lorsque nous regardons une photo ou un dessin, nous pouvons affirmer intuitivement qu'un objet est situé devant ou derrière un autre. Pourtant, nous savons bien qu'une photo n'est pas, en général, vue stéréoscopiquement. Alors pourquoi voyons-nous du relief alors que manifestement, il n'y a pas deux points de vue différents pour chaque œil (vision binoculaire) ? Parce que nous nous servons d'indices monoculaires de profondeur. Nous nous en servons tout le temps dans notre vie quotidienne.

La perspective

La perspective est utilisée par les peintres et les dessinateurs depuis bien longtemps. Cette technique est la plus connue pour donner une impression de relief. Il y a plusieurs perspectives (Fuchs, et al., 2003). La perspective « cavalière » est utilisée en dessin industriel car elle est assez simple à dessiner, les lignes non parallèles au plan sont tracées comme des « fuyantes parallèles » (Figure 7). La perspective photographique est le résultat de la projection des objets sur le capteur plan photographique et elle est utilisée pour les dessins « réalistes », les lignes perpendiculaires au plan du papier se rejoignent au « point de fuite » (Figure 8).

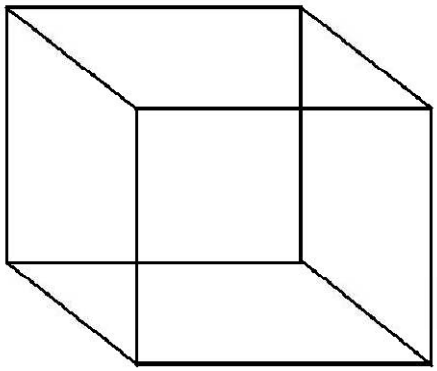


Figure 7 : perspective industrielle ou cavalière

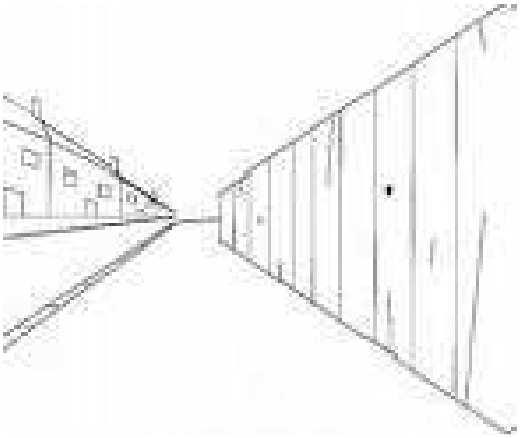


Figure 8 : perspective artistique ou photographique

Différence d'intensité lumineuse

Selon la variation de la luminosité sur l'objet, il est souvent possible de déterminer quelle est l'orientation de l'objet par rapport à la surface du papier ou de l'écran (Figure 9).

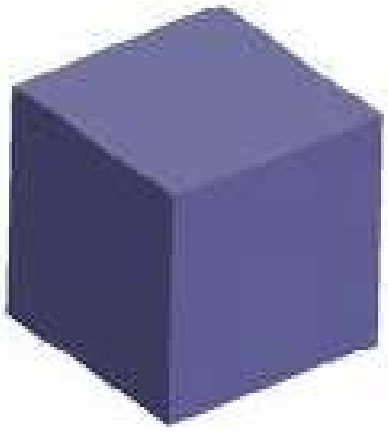


Figure 9 : le relief est perçu grâce à la variation de luminosité

Dimensions relatives

Intuitivement, si nous avons deux objets qui normalement ont la même dimension mais qui sur l'image présentent deux tailles différentes, nous aurons tendance à supposer que celui qui est le plus grand doit être plus proche (Figure 10).

Figure 10 : intuitivement, les fleurs sont perçues comme étant de plus en plus éloignées

Occultation

Lorsqu'un objet en cache un autre, nous en déduisons a priori que l'objet caché est plus loin que l'autre. Notons que, sur la Figure 11, nous supposons que c'est un rectangle qui est caché par l'ellipse, alors qu'il n'est pas dessiné entièrement. Notre cerveau complète donc les informations manquantes d'une image par ce qui lui semble le plus évident.

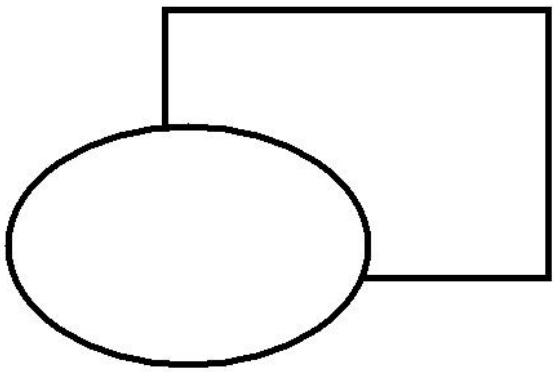


Figure 11 : l'ellipse est perçue devant le rectangle

Variation des fréquences spatiale de la texture

Une texture proche est plus précise, plus nette, qu'une texture éloignée. Notre œil ne perçoit que les basses fréquences d'un objet lointain (Olivia, et al., 2006). Lorsque les détails disparaissent progressivement sur une image, on a l'impression que l'objet s'éloigne, surtout si cet effet est lié à un effet de taille (Figure 12).

Figure 12 : les détails de la texture sont moins visibles au loin

Variation de visibilité

Si la photo a été prise à l'extérieur, l'atmosphère atténue les détails. Plus le paysage est flou, plus il est lointain (Figure 13). Cet indice peut donc compléter le précédent : le gradient de texture.

Figure 13 : au loin, le paysage devient flou

1.2.1.2. Indices cinétiques

Les mouvements des images sur la rétine constituent un moyen très puissant de perception du relief. Certaines études montrent que la perception du relief grâce au mouvement est indépendante des autres indices de relief (Palmer, 1999). La perception du relief grâce à cet indice est très rapide, elle est même plus rapide que la stéréoscopie, surtout si les images sont difficiles à fusionner.

Les parallaxes de mouvements : Lors de nos déplacements, les images des objets proches se déplacent plus rapidement sur la rétine que celles des objets lointains.

Par exemple, vous êtes dans un bois, vous ne regardez que les troncs d'arbres en fermant un œil. Il est difficile de dire avec précision quel arbre est le plus proche. Sur la rétine, ils n'ont certes pas le même diamètre, mais il est impossible de savoir si le plus gros est réellement plus proche, ou s'il est simplement plus vieux et donc plus gros ! Par contre, si vous vous déplacez, vous allez voir des arbres qui semblent se déplacer plus rapidement que les autres, ce sont les arbres les plus proches.

Lors d'un mouvement, certains objets vont passer devant d'autres. L'objet couvert ou découvert se trouve logiquement derrière un autre. C'est exactement le même phénomène que celui expliqué au paragraphe 1.2.1.1, qu'une composante dynamique vient améliorer.

Mouvement de l'observateur ou de l'objet : Le mouvement de l'observateur ou de l'objet ne produira pas la même perception. Si l'environnement (ou un objet) se déplace, il y a moins d'informations. En effet, lorsque nous changeons de position, le système vestibulaire de notre oreille interne nous transmet certaines informations supplémentaires, comme notre accélération. La perception en profondeur est donc plus performante lorsque l'observateur se meut dans l'espace.

1.2.1.3. Indices proprioceptifs : l'accommodation et la convergence

Les indices proprioceptifs, accommodation et convergence, sont ajustés par le système visuel. C'est par la valeur de commande des muscles orbitaux et ciliaires que le cerveau a une perception proprioceptive de la profondeur de l'objet observé. Rappelons que la vision naturelle humaine est une vision active. Les yeux sont presque toujours en mouvement pour observer l'environnement et les cristallins se déforment pour permettre l'accommodation. Par contre, un appareil photographique observe en vision passive l'environnement réel. La profondeur de champ de toute image photographique est un indice de profondeur qui correspond partiellement au phénomène d'accommodation. On peut donc utiliser des effets de profondeur de champ sur des images de synthèse pour donner plus de « relief » à l'espace observé.

1.2.2. Perception binoculaire du relief

Nous percevons également la profondeur en nous servant de nos deux yeux. Ils ne sont pas situés au même endroit, donc ils ne reçoivent pas tout à fait la même image. Ces différences dans l'information sont utilisées par le cerveau pour en déduire la profondeur des objets observés. Contrairement à la perception du relief grâce aux indices monoculaires, c'est un phénomène neurobiologique et non cognitif (Figure 17).

Cette capacité physiologique peut être exploitée en vision artificielle. Il suffit de prendre deux images au même instant mais dont les points de vue sont séparés d'une certaine distance horizontalement (idéalement la même qu'entre les deux yeux) et de projeter chaque image sur l'œil approprié.

1.2.2.1. La vision binoculaire

Nous définissons la DIO comme la distance interoculaire (ou inter-pupillaire), la distance séparant les deux yeux, ou plus précisément les centres des deux pupilles (en moyenne 6,5cm) (Fuchs, 2006).

La disparité horizontale

La disparité rétinienne est la différence des angles formés sur les deux rétines par les rayons lumineux provenant des deux points A et F, F étant le point observé par le sujet. Si nous nous référerons à la Figure 14, la disparité vaut |α - β|.

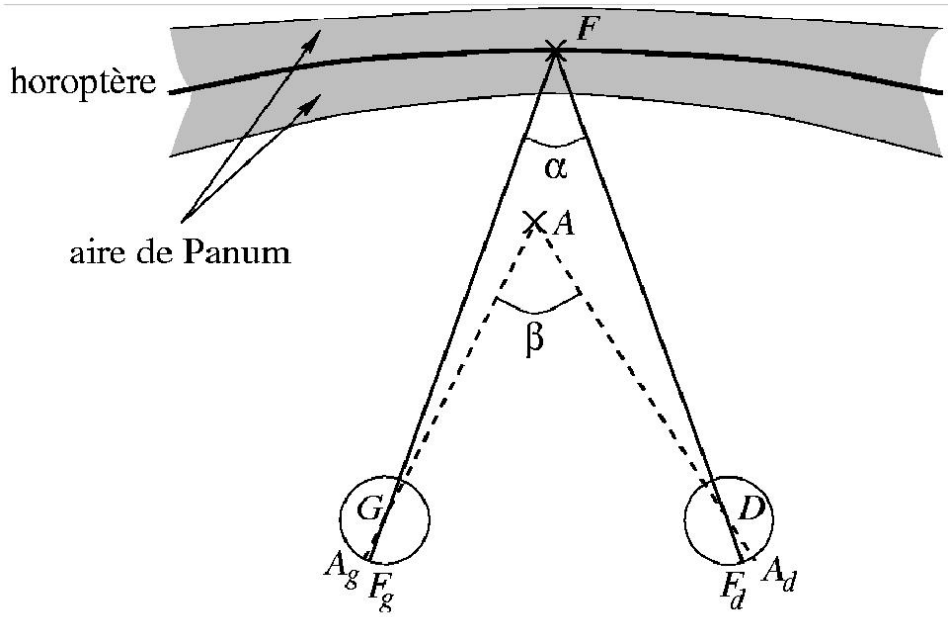


Figure 14 : aire de Panum (Fuchs, 2006)

Lors d'une convergence en un point de l'espace (le point F sur la Figure 14), il existe certains points pour lesquels la disparité est nulle. Ces points sont situés sur l'horoptère. Il existe une différence entre l'horoptère empirique et l'horoptère théorique, qui est le cercle passant par les deux points nodaux (Howard, et al., 1995). Cette différence est appelée la déviation de Hering-Hillebrand (Bergholm, et al., 1996). Il n'y a qu'un seul horoptère par point de convergence : sur la Figure 15, les yeux convergent, mais si leur convergence change, l'horoptère changera lui aussi.

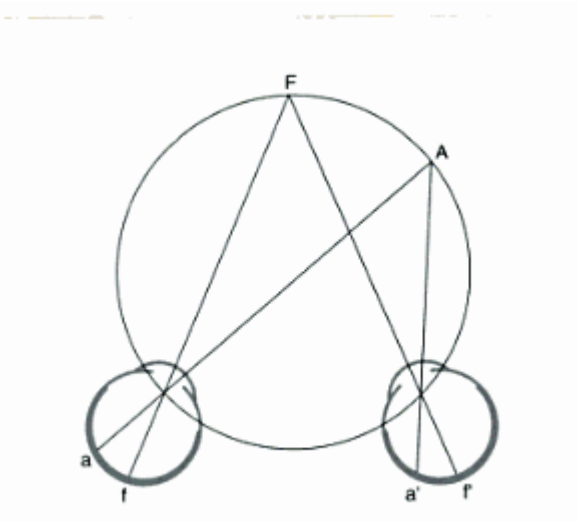


Figure 15 : horoptère théorique dit géométrique (Espinasse-Berrod, et al., 2008)

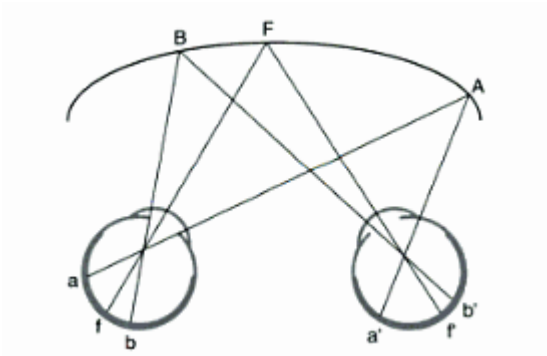


Figure 16 : horoptère empirique, dit longitudinal Hering-Hillebrand (Espinasse-Berrod, et al., 2008)(Bergholm, et al., 1996)

Les points situés sur l'horoptère sont fusionnables : le cerveau perçoit les images de chaque œil comme provenant du même point (Figure 16). Il existe une zone appelée "aire de Panum" où tous les points situés dans cette zone peuvent être fusionnés même s'ils ne sont pas exactement sur l'horoptère (Figure 14), c'est le phénomène de stéréopsie. Les points situés hors de l'aire de Panum sont vus en double.

L'aire de Panum n'est pas forcément symétrique. Certaines personnes ont une aire de Panum plus proéminente à l'arrière de l'horoptère. Pour certaines personnes, c'est exactement l'inverse (Howard, et al., 1995). L'aire varie également en fonction du temps d'exposition et de la fréquence spatiale (voir paragraphe 5.2).

La perception de profondeur stéréoscopique est réalisée par le système visuel de l'observateur, indépendamment des indices monoscopiques (Howard, et al., 1995). Pour preuve, les expériences de Julesz (1971) montrent à l'aide de stéréogrammes de points aléatoires que le relief peut être perçu sans aucun indice monoculaire (Figure 17). Par contre, cette perception est relative et non absolue. Il est plus difficile de donner la profondeur exacte d'un objet sans repère que sa profondeur par rapport à un autre (Julesz, 1971).

Les stéréogrammes de points aléatoires sont des couples « image gauche - image droite » qui séparément ne représentent, comme leur nom l'indique, que des points aléatoires (voir Figure 17). Mais lorsque ces couples sont présentés en stéréoscopie, les personnes qui se servent d'indices stéréoscopiques voient apparaître une forme en relief. En effet, l'image gauche est composée de points

réellement aléatoires, mais l'image droite est une copie de la première dont une forme est décalé vers la droite. C'est cete même forme qui est perçu en relief.

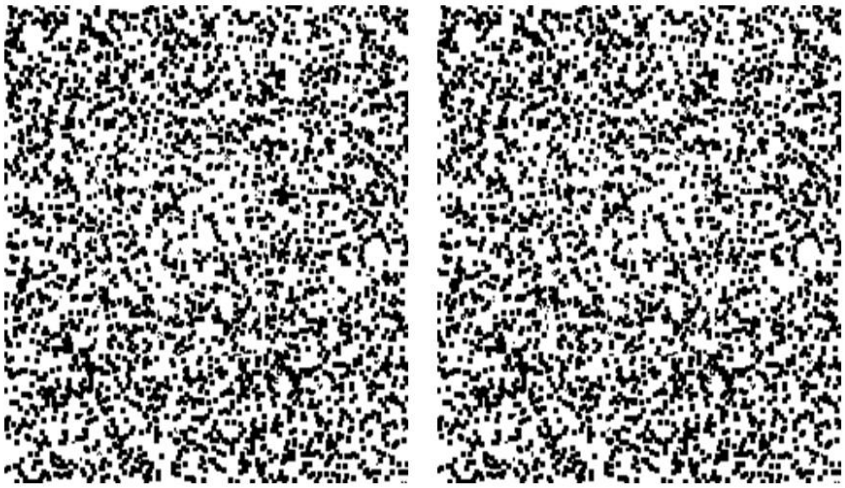


Figure 17 : stéréogramme de points aléatoires de Julesz (Julesz, 1971)

Les disparités verticales

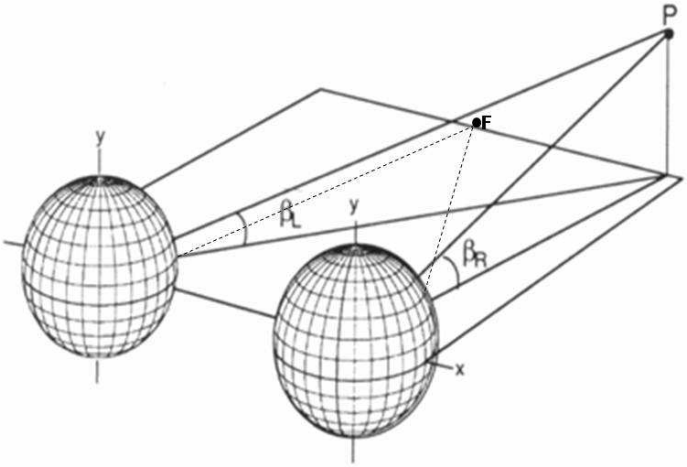


Figure 18 : les yeux observent le point F, il y a des disparités verticales au point P

La disparité verticale absolue (β) d'un point quelconque de l'espace est mesurée par la différence d'élévation de ce point dans les deux yeux : $\beta = \beta_G - \beta_D$ (Espinasse-Berrod, et al., 2008)(Figure 18). Nous traitons donc toujours des disparités verticales et horizontales, mais les disparités verticales se concentrent dans la périphérie du regard. Elles sont notamment utiles lors de la perception de formes (Devisme, 2008). La Figure 19 représente les disparités d'un plan fronto-parallèle, on distingue les composantes horizontales et verticales en chaque point.

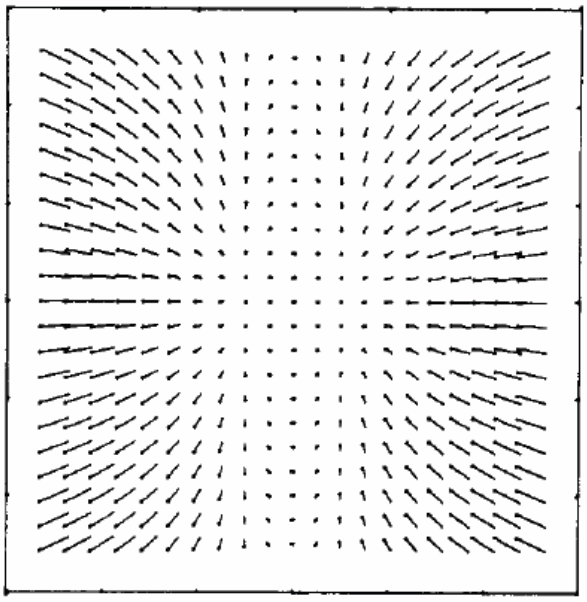


Figure 19 : représentation vectorielle des disparités d'un plan fronto-parallèle (Nakayama, 1983)

Nous avons parcouru le fonctionnement de la vision humaine en condition normale et naturelle. Dans le paragraphe suivant, nous exposons les moyens utilisés pour reproduire cette vision le plus fidèlement possible en environnement immersif 3D.

1.2.3. Restitution virtuelle d'un environnement 3D

Nous avons vu les différents indices dont se sert notre cerveau pour percevoir le relief. Nous allons à présent voir comment nous pouvons reproduire ces indices de façon artificielle. Nous passerons en revue les différentes combinaisons, des plus simples aux plus complexes techniquement.

1.2.3.1. Monoscopique statique

C'est évidemment la façon la plus simple de visualiser un espace 3D. Il faut tenir compte de tous les indices monoculaires cités dans le paragraphe 1.2.1.1. C'est ce que nous faisons à chaque fois que nous dessinons, prenons une photo, créons un monde virtuel statique projeté sur un écran monoscopique.

1.2.3.2. Monoscopique avec mouvement de l'objet ou de l'environnement

Reproduire le mouvement de l'environnement ou d'un objet est relativement aisé en virtuel. Il suffit de déplacer l'objet virtuel ou le monde virtuel par rapport à la caméra virtuelle. Ainsi, les parallaxes de mouvement peuvent apparaître. Cette technique suppose qu'aucun mouvement ne soit incompatible avec la scène.

Cette technique est utilisée dans les films. Elle est également souvent utilisée lors de conception assistée par ordinateur : l'ingénieur bouge virtuellement (généralement avec la souris) la pièce qu'il vient de créer pour mieux percevoir son relief.

1.2.3.3. Monoscopique avec point de vue traqué

Comme nous l'avons indiqué dans le paragraphe 1.2.1.2, cette perception du relief par le mouvement est plus cohérente, donc plus efficace que lorsque l'environnement bouge, le sujet res-

tant immobile. Ceci est dû aux informations données par le système vestibulaire du sujet. Il faut donc que la position du sujet soit connue par le système et que la caméra virtuelle soit positionnée en fonction de celle-ci (Figure 20). Le temps de latence du système doit être extrêmement court sous peine de provoquer une fatigue très prononcée, voire des nausées (voir le paragraphe 5.1.2.3).

Figure 20 : système pour suivre la position et l'orientation du sujet (ART-tracking)

1.2.3.4. Stéréoscopique

Le principe de la stéréoscopie consiste à afficher deux images légèrement différentes sur un même plan. Chaque image correspond à un œil, chaque œil ne voit qu'une seule image. Pour presque chaque point de l'image de gauche correspond un point homologue sur celle de droite, plus ou moins décalé. C'est précisément ce décalage qui donne la perception de relief. Pour des images en relief affichées sur un écran stéréoscopique, la parallaxe horizontale angulaire est définie par l'angle DPG, formé par les deux points homologues des images droite et gauche vues du point de vue de l'observateur (Figure 21). Notons que dans ce type de restitution, nous n'utilisons en général que les disparités horizontales (voir paragraphe 1.2.3.5). Les parallaxes horizontales créent en conséquence des disparités rétiniennes sur les yeux de l'observateur, lui procurant la perception binoculaire de la profondeur relative des points homologues observés.

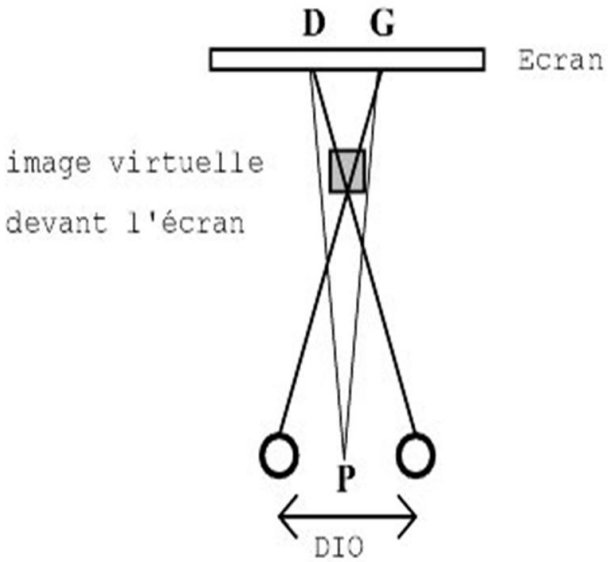


Figure 21 : définition de la parallaxe horizontale angulaire (Fuchs, 2006)

Quand les points homologues ont une parallaxe nulle et qu'ils sont donc affichés aux mêmes endroits, l'objet virtuel se situe au niveau de l'écran (Figure 22).

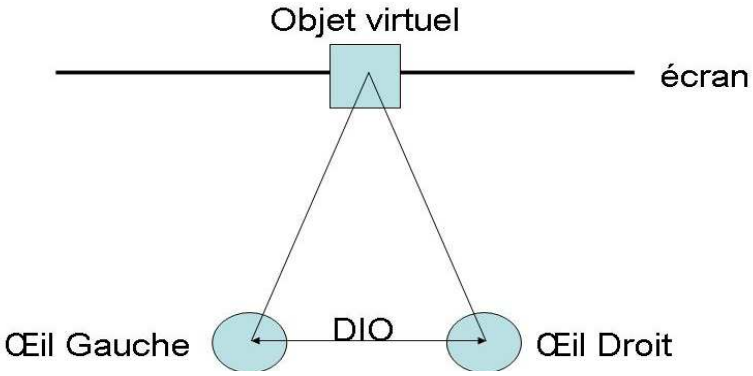


Figure 22 : vision avec parallaxe nulle

Quand la parallaxe est positive (les points homologues gauches et droits sont respectivement affichés à gauche et à droite), l'objet observé est situé virtuellement derrière l'écran. (Figure 23)

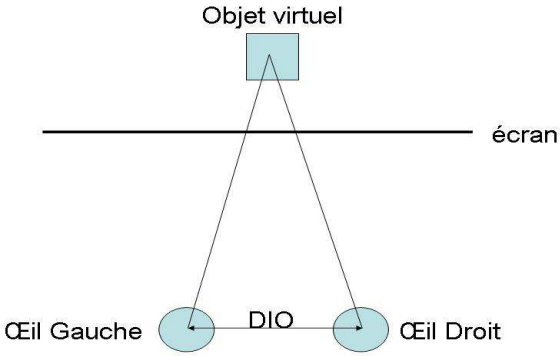


Figure 23 : vision avec parallaxe positive

Quand la parallaxe est négative (les points homologues gauches et droits sont respectivement affichés à droite et à gauche) l'objet observé est situé virtuellement devant l'écran. (Figure 24)

Notons que les disparités rétiniennes ne doivent pas être trop prononcées pour permettre à l'observateur de fusionner les deux images. En effet, elles doivent rester dans l'aire de Panum pour être fusionnables. En conséquence, les parallaxes horizontales angulaires doivent être limitées. C'est la principale condition à respecter lorsque nous utilisons des images stéréoscopiques. Des études expérimentales (Valyus, 1962) ont démontré la difficulté de fusionner deux images planes fournissant des parallaxes horizontales (et donc des disparités rétiniennes) d'angle supérieur à 1,5 degrés. Fuchs donne une limite un peu inférieure pour la limite de parallaxe horizontale : 1,2◦ (Fuchs, et al., 1995). Ces variations sont normales car elles sont variables pour chaque individu qui tolère plus ou moins les images en relief et qui force ou non son système visuel à fusionner les images. Ces variations de limites de fusion dépendent aussi d'autres critères : mouvement ou non des objets, fré-

quences spatiales… Nous les aborderons au paragraphe 5.2. A noter que 3 à 5% (peut-être plus) de la population n'arrive pas à fusionner des images en relief sur un écran.

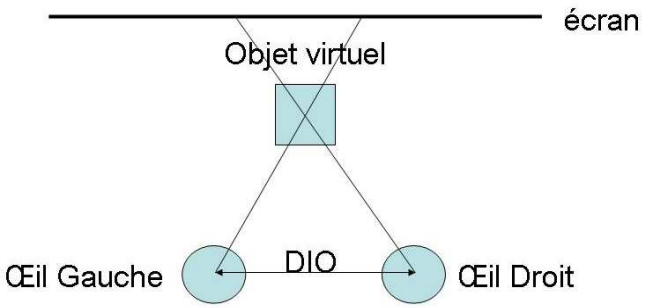


Figure 24 : vision avec parallaxe négative

Avantage de la restitution virtuelle de la vision binoculaire

Ces parallaxes horizontales reproduisent les disparités rétiniennes et permettent une perception de relief indépendante des indices monoscopiques. Il est préférable qu'elles restent cohérentes avec les indices monoculaires pour éviter des conflits entre les indices provoquant de la fatigue visuelle. Elles permettent donc d'avoir une impression de relief et d'immersion beaucoup plus efficace.

Inconvénients de la restitution virtuelle de la vision binoculaire

Le conflit accommodation/convergence

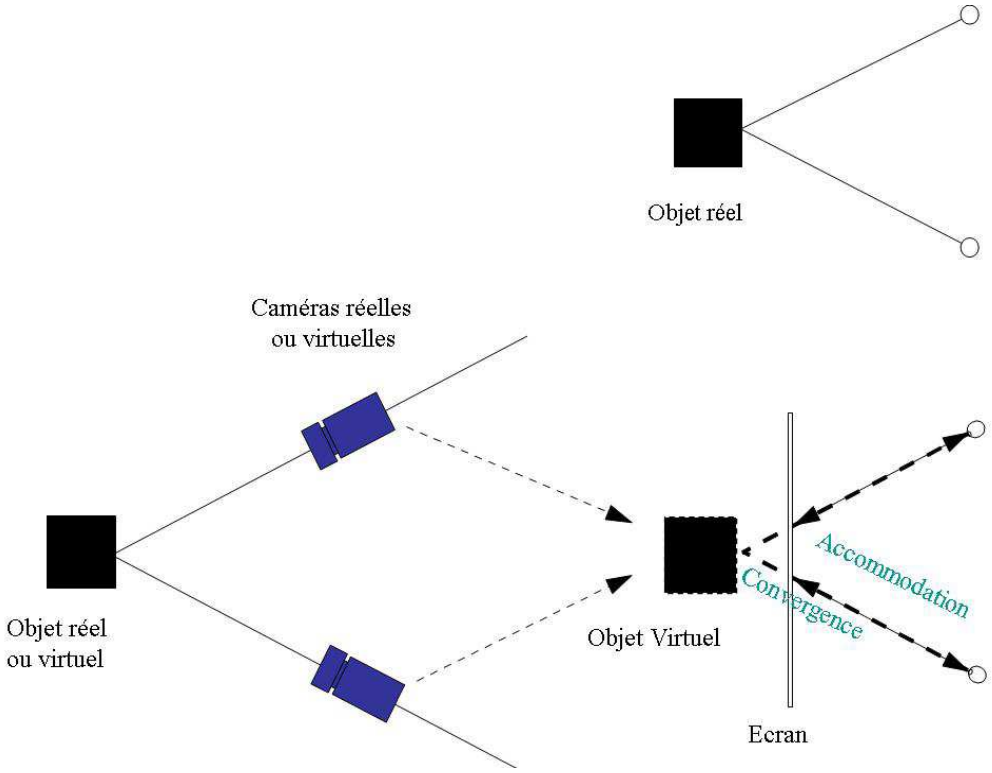


Figure 25 : différence entre vision naturelle et restitution stéréoscopique (Fuchs, 2006)

L'accommodation et la convergence des yeux étant toutes deux liées à la profondeur de l'objet que l'on observe, il en découle une relation entre celles-ci. Cette correspondance n'est pas innée chez l'homme mais est apprise expérimentalement et inconsciemment par l'enfant (Fuchs, 2006). Cette relation naturelle peut être mise en défaut artificiellement lorsque l'on regarde des images stéréoscopiques sur un écran : un objet virtuel positionné en arrière de l'écran fait converger les axes optiques sur lui tandis que les yeux accommodent sur l'écran (voir Figure 25). Cette modification de la relation convergence-accommodation est susceptible de créer des difficultés visuelles et de la fatigue pour l'utilisateur. Nous explicitons plus en détail ce problème au paragraphe 5.1.1. (Fuchs, 2006)(Perrin, 1998)(Roumes, 1995).

Ce conflit ne pourra être évité quelque soit la technique de projection utilisée. Il conviendra donc de le minimiser en limitant les écarts entre la position de l'objet virtuel (point de convergence) et la position du plan de projection (point d'accommodation).

Les contradictions entre les indices monoscopiques et stéréoscopiques

Les contradictions entre les indices stéréoscopiques et les autres indices de relief sont néfastes. Sur les petits écrans, représenter un objet "sortant" de l'écran et partiellement occulté par le bord est une erreur courante. Le cerveau n'arrive plus à savoir si l'objet est devant ou derrière le bord de l'écran. En effet, l'objet est perçu devant le bord de l'écran grâce aux indices stéréoscopiques, mais également derrière lui à cause de l'occultation qui est un indice monoscopique. Ce problème peut également se produire lorsque dans un monde de synthèse les occultations entre les objets sont mal gérées. Par exemple, un objet plus éloigné n'est pas caché par un objet situé binoculairement devant.

De même, tous les indices de lumière et de texture doivent être pris en compte correctement lors de la génération d'images de synthèse en 3D. Nous évitons ainsi les contradictions entre les indices monoscopiques et stéréoscopiques.

Les mouvements pseudoscopiques

Si l'observateur se déplace parallèlement à l'écran, il voit les objets qui se déplacent anormalement : il perçoit des mouvements appelés mouvements pseudoscopiques (Figure 26). L'espace se déforme artificiellement, créant des incohérences visuelles par rapport au déplacement de l'observateur.

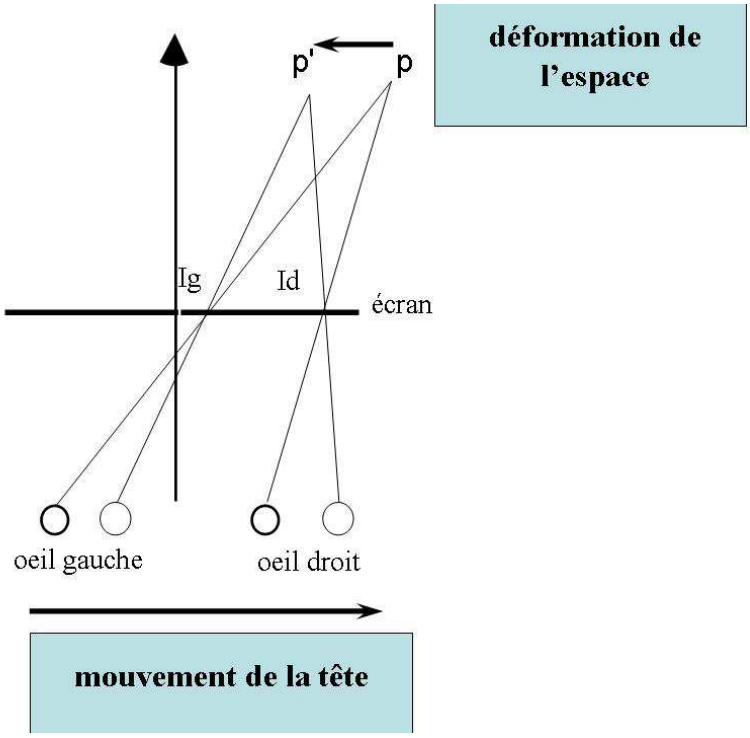


Figure 26 : mouvements pseudoscopiques parallèles à l'écran

Les yeux de l'observateur se translatant vers la droite, il voit le point P (théoriquement fixe) se déplacer vers la gauche au point P', car les points Id et Ig sont fixes. Ce phénomène est peu perceptible quand l'observateur est assis, impliquant une translation latérale de la tête trop faible pour entrevoir ces faux mouvements. Un schéma similaire permet de comprendre que si l'observateur recule ou avance face à l'écran, il y a étirement ou raccourcissement de la profondeur perçue (Figure 27).

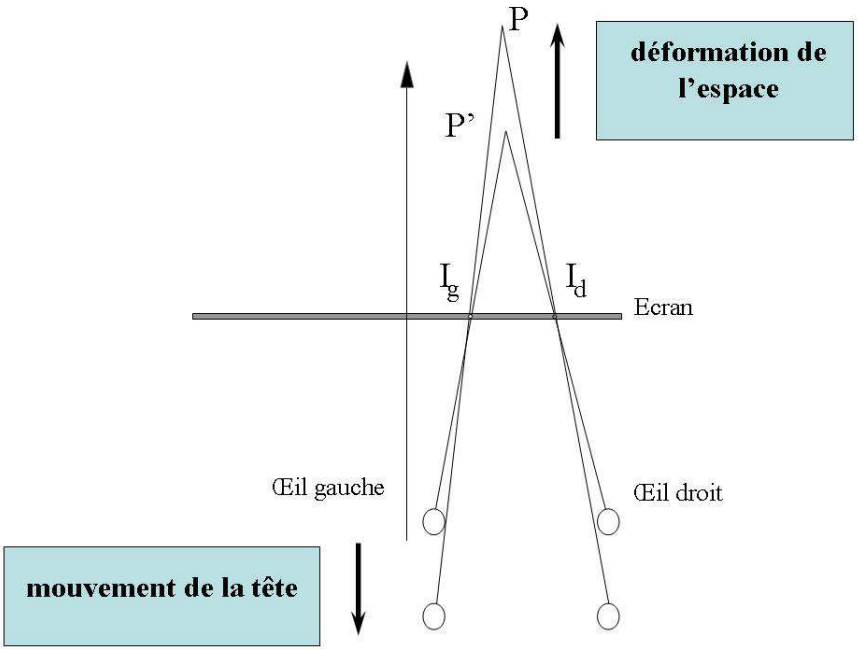


Figure 27 : mouvements pseudoscopiques perpendiculaires à l'écran

1.2.3.5. Stéréoscopique avec point de vue traqué

Si l'observateur est susceptible de bouger la tête et si on désire parfaire la vision, il est souhaitable de déterminer la position de la tête (et donc de son point de vue) grâce à un capteur de localisation spatiale. Les images stéréoscopiques sont alors modifiées en temps réel en plaçant les deux points de vue des caméras virtuelles en conséquence. Ceci est possible avec la même technique que celle exposée au paragraphe 1.2.3.3, en association avec les techniques de vision stéréoscopique.

Cette technique permet de supprimer les mouvements pseudoscopiques. Elle permet également de cumuler les indices binoculaires aux indices cinétiques. L'impression de relief est beaucoup plus rapide et l'immersion d'autant plus forte. Il reste cependant un inconvénient pour la restitution virtuelle.

Sur quels critères doit-on se baser pour positionner et orienter les deux caméras fictives ou réelles ? A priori, on peut penser que l'on doit faire converger les deux axes optiques vers la zone principale à observer. Dans ce cas, les projections sur les deux plans donnent des images avec des parallaxes horizontales mais aussi verticales (Figure 28).

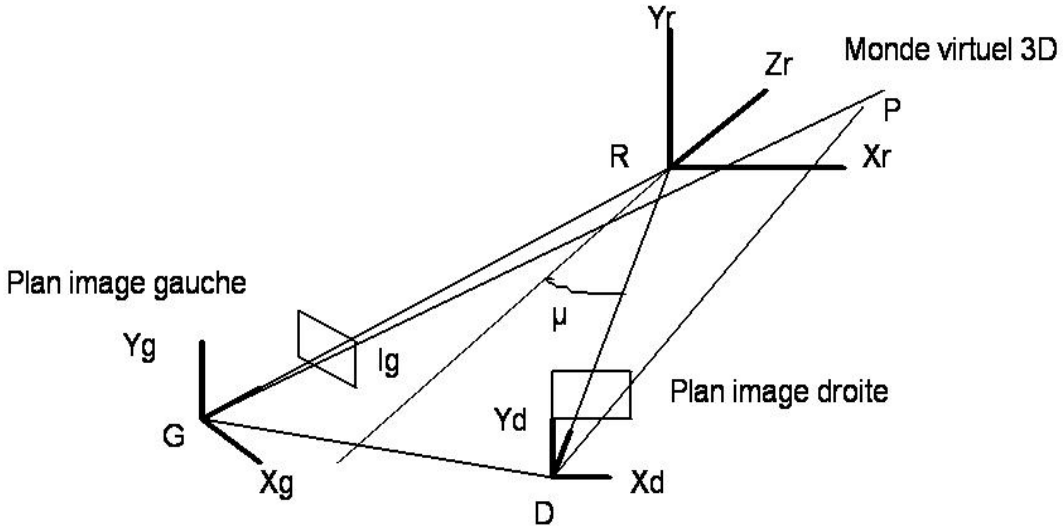


Figure 28 : positionnement des deux caméras fictives ou réelles dont les axes convergent vers un point (Fuchs, 2006)

Mais ce cas n'est guère souhaitable car cela impose des parallaxes verticales qui rendent difficile la fusion des images. Par exemple, les projections droite et gauche de la face avant d'un cube centré sur l'axe Zr sont exposées sur la Figure 29.

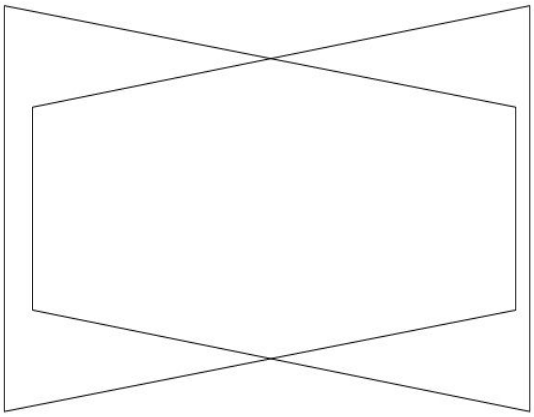


Figure 29 : les projections droite et gauche de la face avant d'un cube centré

Nous avions vu au paragraphe 1.2.2.1 que les disparités verticales sont normales en vision naturelle dans la périphérie de ce que nous regardons. Elles font partie intégrante de notre perception du relief. Cependant, lorsqu'il s'agit d'une restitution virtuelle, nous ne savons pas où regarde le sujet, sauf si nous traquons les yeux de l'observateur. Si nous faisons l'hypothèse que le sujet regarde le point central de l'écran, les disparités verticales sont réparties sur toute l'image. Si le sujet regarde effectivement le point central, cela ne pose aucun problème de fusion. Par contre, s'il regarde un autre point de l'écran, il y aura des disparités verticales qui ne devraient pas exister (voir paragraphe 5.1.1.2).

1.2.3.6. Stéréoscopique avec point de vue et yeux traqués

Cette technique permet d'utiliser les disparités verticales dans les images stéréoscopiques. Elle consiste à traquer non seulement le point de vue, mais également les yeux de façon à savoir précisément le point de l'écran regardé par l'observateur. Grâce à cette information, nous pouvons ajuster l'orientation des caméras virtuelles très précisément et notamment lier la convergence des caméras virtuelles sur la convergence des yeux de l'observateur.

Toutefois, il est important d'avoir un système extrêmement rapide, réactif et précis. En effet, les yeux bougeant très rapidement, le capteur et le traitement informatique devront aller presque aussi vite, une latence, même petite, générant une fatigue visuelle.

1.3. Conclusion

Dans ce chapitre, nous avons synthétisé le fonctionnement du système visuel humain. L'anatomie de l'œil et de ses composantes ont été décrites avant de préciser la façon dont les informations sont acheminées et analysées. Nous avons également expliqué comment nous percevons le relief en vision naturelle grâce aux indices monoscopiques statiques et dynamiques, aux indices proprioceptifs et, bien sur, grâce aux indices binoculaires. Nous avons ensuite synthétisé la façon dont nous reproduisons un environnement virtuel en utilisant ces indices de relief.

Dans la suite de ce manuscrit, nous aborderons, dans un premier temps, la perception des formes et, dans un second temps, la fatigue visuelle provoquée par la vision stéréoscopique en environnement immersif.

Partie n°1 : perception des formes en immersion

Chapitre 2 : Etat de l'art de la perception des formes en relief

Selon la définition de Koenderink, la forme d'un objet est une propriété intrinsèque de la surface de celui-ci. Elle ne dépend pas du point de vue de l'observateur (Koenderink, et al., 1992). Nous allons donc dans ce chapitre étudier la façon dont nous percevons cette propriété des objets qui nous entourent. Nous aborderons ensuite une introduction à la perception des objets eux-mêmes et nous terminerons ce chapitre sur quelques expériences menées en environnement immersif.

2.1. Perception des formes en relief dans la vision naturelle

2.1.1. Perception des inclinaisons et obliquités

2.1.1.1. Analyse des disparités binoculaires

L'obliquité est la rotation d'une surface par rapport à l'axe vertical (Figure 30). Une surface oblique crée une « rampe » de disparités horizontales appelée disparité de compression-expansion ou gradient horizontal de disparités horizontales (Devisme, 2008). La disparité horizontale étant directement liée à la profondeur et une surface oblique subissant une variation horizontale et progressive de profondeur, il est normal d'avoir un gradient horizontal de disparités horizontales.

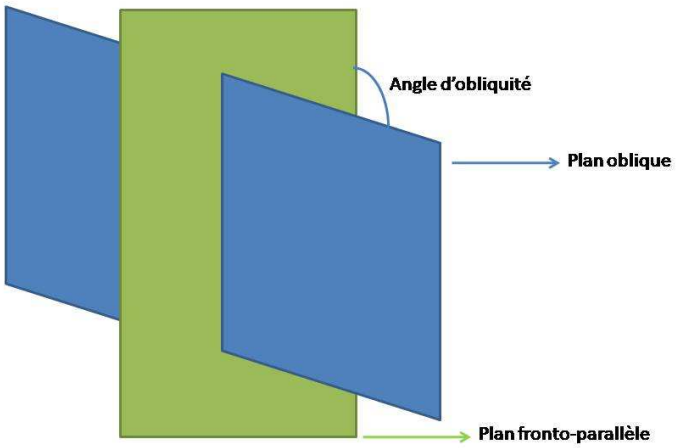


Figure 30 : un plan oblique produit un gradient horizontal de disparités horizontales

L'inclinaison d'une surface, quant à elle, est la rotation de celle-ci par rapport à un axe horizontal (Figure 31). La profondeur d'une surface oblique variant de façon verticale sur la surface, elle produit un gradient vertical de disparités horizontales, appelé également cisaillement horizontal.

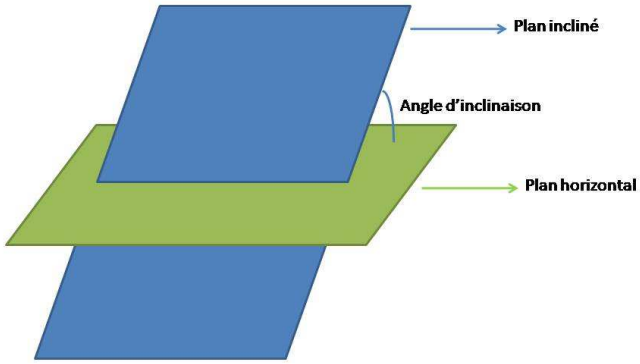


Figure 31 : un plan incliné produit un gradient vertical de disparités horizontales

Dans un stéréogramme de points aléatoires (voir paragraphe 1.2.2.1), une surface inclinée ou oblique est plus longue à percevoir que deux surfaces adjacentes tournées ou inclinées dans des directions opposées (Gillam, et al., 1988). En effet, les variations continues de la disparité appliquées sur tout le stimulus sont difficilement perçues en profondeur. Un gradient uniforme de disparités est plus difficilement détectable qu'une discontinuité dans le gradient de disparité (McKee, 1993).

La perception d'obliquité ou d'inclinaison d'une surface peut être produite par des disparités horizontales seules, ou plus exactement par un gradient horizontal ou vertical de disparités horizontales seules. Dans ce cas, nous parlons d'effet géométrique. Par contre, quand la perception de l'obliquité ou d'inclinaison d'une surface est due uniquement à des disparités verticales (et donc un gradient horizontal ou vertical de disparités verticales) nous parlons d'effet induit (Gillam, et al., 1983)(Ogle, 1950)(Figure 32).

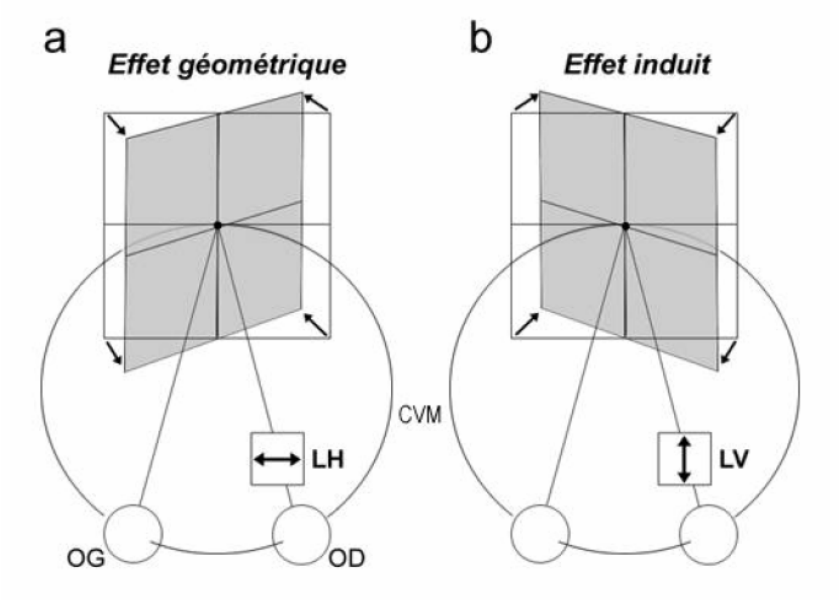


Figure 32 : effet induit et géométrique pour la perception d'une surface oblique (Devisme, 2008)

Prenons un plan incliné et un plan oblique de même angle par rapport au plan fronto-parallèle, l'angle d'inclinaison du plan incliné sera perçu comme étant plus grand que celui de l'obliquité du plan oblique. De même, l'angle seuil de détection d'inclinaison est estimé à 1,3° tandis que celui de l'obliquité est de 2,1°. Nous voyons donc bien que la perception de pente est anisotropique (Cagenello, et al., 1993).

Quelques expériences ont été menées en laboratoire pour étudier des plans inclinés et obliques. Elles ont montré que le gradient de disparités en fonction de l'inclinaison ou de l'obliquité du plan dépend également de la distance d'observation. Ces expériences étaient menées sur des plans uniques. Cependant, dans notre environnement réel, nous n'observons pas de plans uniques, mais plutôt un très grand nombre de plans ayant tous des inclinaisons et obliquités différentes. Lorsqu'il y a plusieurs plans dans la scène observée, l'information est redondante et permet de lever les ambiguïtés. Ainsi, Gillam a montré que les inclinaisons et obliquités sont vues plus rapidement et plus précisément lorsqu'il y a plusieurs plans dans la scène. C'est notamment la présence de discontinuités de gradient entre les plans qui améliore cette perception (Gillam, et al., 1983).

2.1.1.2. Analyse des parallaxes de mouvement

Lorsque la tête de l'observateur se déplace latéralement, les surfaces inclinées et obliques produisent également des gradients de parallaxes de mouvement très semblables aux gradients de disparités binoculaires. En effet, si nous comparons les parallaxes de mouvement du paragraphe 1.2.1.2 avec les disparités rétiniennes du paragraphe 1.2.2.1 nous voyons qu'il y a de fortes similitudes.

Lors d'un mouvement latéral de l'observateur par rapport à une surface oblique, des variations continues des parallaxes entre une position A et une position B se produisent. Pour une vitesse de la tête de l'observateur donnée, le taux instantané d'expansion ou de compression horizontal dépend de l'obliquité de la surface, nous appelons ceci un gradient horizontal de vitesses horizontales (Horizontal gradient of horizontal velocity). Cependant, il faut noter qu'une surface oblique fixe placée devant un observateur en mouvement latéral n'est pas la seule façon de produire un gradient horizontal de vitesses horizontales. Une surface d'obliquité différente placée de façon excentrée ou

en rotation ou une surface fronto-parallèle qui se déplace en profondeur en même temps que le déplacement latéral de l'observateur produit le même gradient (Allison, et al., 2003).

Cette ambigüité peut être levée grâce aux parallaxes verticales. En effet, lorsqu'une surface oblique est excentrée par rapport à l'observateur, ou qu'une surface fronto-parallèle se déplace en profondeur, cela affecte également les disparités verticales qui s'en trouvent modifiées. Il est alors possible à l'observateur de lever l'ambigüité puisque les disparités verticales et les parallaxes verticales dépendent de l'excentricité (voir Figure 19).

De même, une surface inclinée produit une variation continue des parallaxes de mouvement nommée gradient vertical de vitesses horizontales, elle produit les mêmes ambigüités que pour un gradient horizontal.

2.1.1.3. Analyse des textures

« *Une texture correspond au « dessin » supporté par une surface (plane ou courbe). Les éléments structurants de la texture possèdent des caractéristiques associées à des propriétés de répartition spatiale (périodicité, aléatoire, irrégularité, variations lentes) communes sur l'ensemble de la région considérée.* »(Massot, 2006).

Gibson est le premier à évoquer un gradient de texture qu'il définit comme des changements de taille ou de forme de petits éléments de texture qui apparaissent dans beaucoup de surfaces de notre environnement (Gibson, 1950). Ces gradients de texture nous donnent des informations sur la profondeur des objets regardés (voir paragraphe 1.2.1.1) mais également sur l'orientation des surfaces (Figure 33, Figure 34) et, nous le verrons, sur la courbure de ces dernières (voir paragraphe 2.1.2.3).

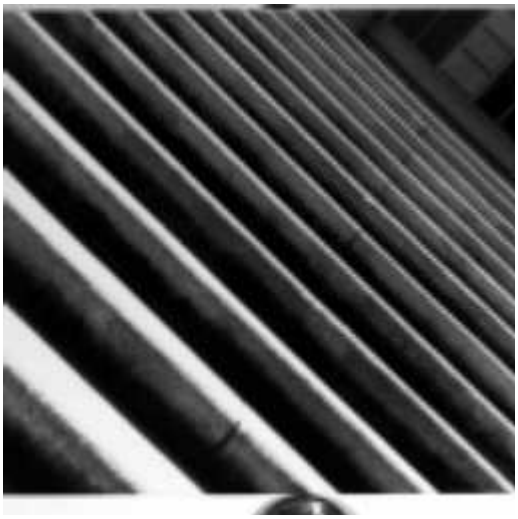


Figure 33 : le gradient de la texture nous incite à percevoir un plan incliné

Figure 34: le gradient de la texture nous incite à percevoir un plan oblique

La texture est une information importante dans la vision en relief, cependant, contrairement à certains autres indices (disparités binoculaires, parallaxes de mouvement, ombre...) il est difficile d'en établir une description mathématique (Massot, 2006). Ces indices apparaissent lors de la projection d'une surface du monde en trois dimensions (qu'il soit réel ou virtuel) sur le plan d'images en deux dimensions (que ce soit l'écran ou la rétine). Cette projection produit des déformations dans les éléments de la texture.

Stevens a démontré que les deux aspects de la texture (la taille et la forme) donnent des éléments d'information différents sur l'orientation de la surface (Stevens, 1979). La taille des éléments peut être utilisée pour estimer la distance relative des différentes parties de la surface étudiée, et donc retrouver l'orientation de la texture de l'objet. Ceci n'est vrai que si les éléments de la texture ont la même taille. C'est un exemple d'hypothèse heuristique que fait notre cerveau pour percevoir le relief. Pour lui, les éléments de la texture qu'il regarde sont tous à peu près semblables en taille et en forme. Beaucoup d'illusions d'optique sont justement basées sur des textures qui ne présentent pas cette caractéristique. La forme projetée des éléments de la texture peut également nous informer de l'orientation de la surface. Le cerveau pour en déduire ce genre d'information se base évidement sur l'hypothèse citée plus haut que tous les éléments de la texture sont similaires.

2.1.2. Perception des courbures

2.1.2.1. Analyse des disparités binoculaires

La dérivée seconde des disparités correspond à la variation du gradient de disparités, donc à une surface courbe. Elle est également appelée « disparity curvature » qui pourrait se traduire par courbure de disparité. Contrairement à la dérivée première de disparités, la seconde est normalement invariante en fonction de la distance d'observation. Elle n'est cependant pas totalement invariante :

- elle varie en fonction de l'excentricité de la surface par rapport au point de vue de l'observateur,
- la courbure de disparité varie également en fonction du gradient de la normale de la surface.

Des variations sinusoïdales de la disparité permettent d'étudier la sensibilité aux disparités en fonction de la modulation de fréquence. Or la mesure des seuils de discrimination des variations de la disparité en fonction de la fréquence de cette variation montre que la fonction de transfert est

un filtre passe-bande. Autrement dit, il existe une fréquence de variation de la disparité (donc une courbure) optimale pour chaque excentricité par rapport au point de vue de l'observateur (Prince, et al., 1998)(Figure 35).

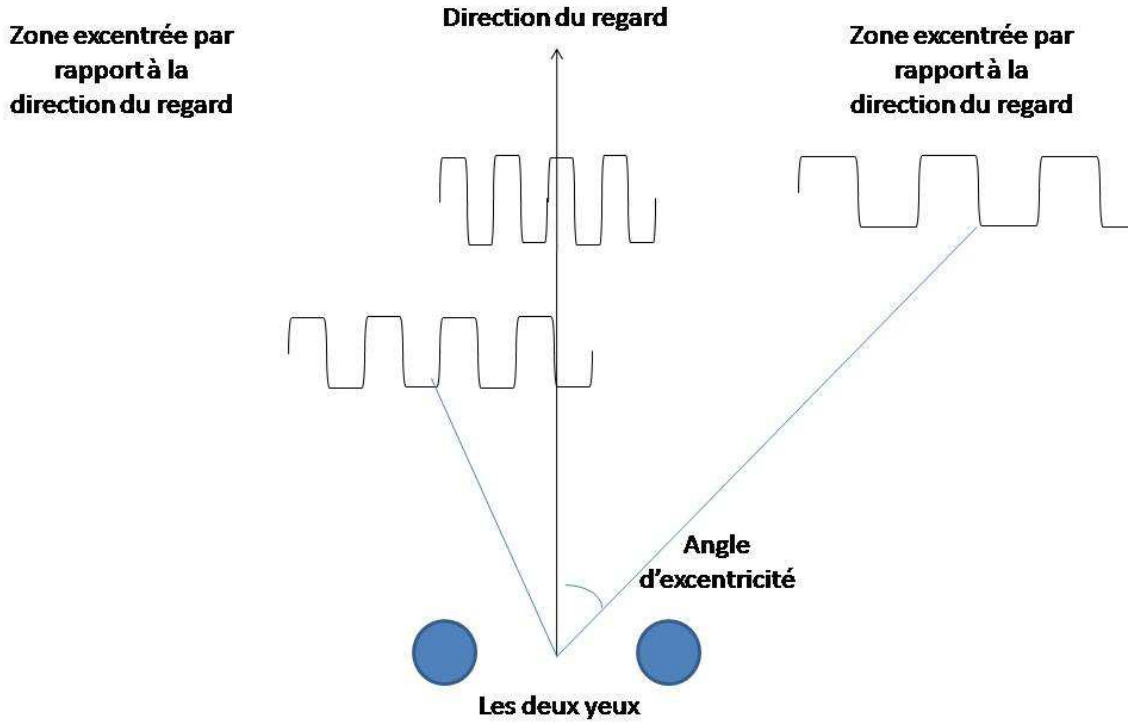


Figure 35 : la fréquence de variation des disparités optimale dépend de l'excentricité (vue du dessus)

Cependant, les variations de courbures verticales et horizontales ne sont pas traitées de la même façon par le cerveau. En effet, il semblerait qu'il y ait une certaine insensibilité aux variations verticales pour les faibles fréquences. Ceci mettrait en évidence une certaine anisotropie stéréoscopique (Devisme, 2008)(Bradshaw, et al., 1999).

Rogers et Cagenello ont montré que les surfaces courbes sont plus facilement perçues que les surfaces planes inclinées (Rogers, et al., 1989). Les seuils de discrimination d'orientation de courbure pour les surfaces paraboliques dans des stéréogrammes de points aléatoires sont très petits. Les auteurs ont déterminé que le seuil moyen de courbure de disparité détectable est de 0.02 min arc/deg^2 (pour un dispositif occupant 2,66° du champ visuel). Cependant, ce seuil dépend de la taille du dispositif de visualisation dans le champ visuel : plus celui-ci est grand, plus le seuil sera bas (et donc les courbes mieux perçues) (Rogers, et al., 1989). Ce seuil dépend également de l'orientation du cylindre parabolique. En effet, il sera 1,5 fois plus élevé pour un cylindre orienté verticalement.

Les cylindres paraboliques utilisés jusqu'ici sont très utiles pour étudier la discrimination des orientations de courbure. Cependant, Johnston suggère plutôt d'utiliser des cylindres circulaires, ceci permettrait une étude de la perception de forme plus aisée, les sujets se référant à leur standard interne (Johnston, 1991). Il place un cylindre supposé circulaire, mais il modifie le diamètre en profondeur et demande aux sujets si la base du cylindre est bien un cercle. Il refait cette expérience plusieurs fois en déplaçant le cylindre pour mesurer l'évolution de la perception des formes en fonction

de la distance entre le cylindre et l'observateur. Il en conclut que notre perception des formes est assez faible.

- un cylindre placé à 53cm de l'observateur était vu circulaire alors que sa profondeur ne dépassait pas 67% de son rayon,
- un cylindre placé à 214cm de l'observateur était vu circulaire alors que sa profondeur dépassait 175% de son rayon (Figure 36).

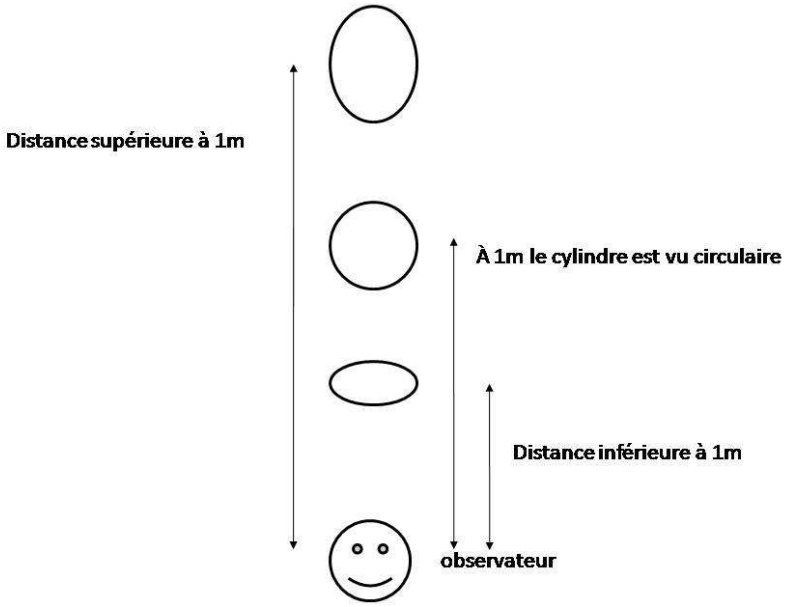


Figure 36 : le cylindre est vu déformé selon la distance à laquelle il se trouve (vue du dessus)

Il attribue cette mauvaise performance à notre difficulté à établir une distance absolue en profondeur ; avec comme conséquence qu'il n'y a que pour une distance d'environ 1m que le sujet peut apprécier correctement les formes. Il montre que lorsque la distance entre le sujet et la surface courbe est sous-estimée, la courbure est elle-même sous-estimée et à l'inverse, lorsque cette distance est surestimée, la courbure est également surestimée (Scarfe, et al., 2006). De plus cette expérience se déroulait dans le noir, ce qui rendait impossible aux sujets de s'aider d'autres indices pour avoir une bonne vision des distances absolues. En effet, Glennerster a réalisé sensiblement la même expérience avec plus d'indices de profondeur, et les résultats étaient bien meilleurs (Glennerster, et al., 1994).

Beaucoup d'études ont montré que la distance entre le sujet et l'objet virtuel observé a une grande corrélation avec la façon dont est perçue sa forme, sa taille et sa profondeur (Paljic, et al., 2002) (Paljic, 2004) (Eggleston, et al., 1996) (Holway, et al., 1941). En effet, lorsque l'objet est plus éloigné du sujet, sa perception devient plus évasive (Figure 37).

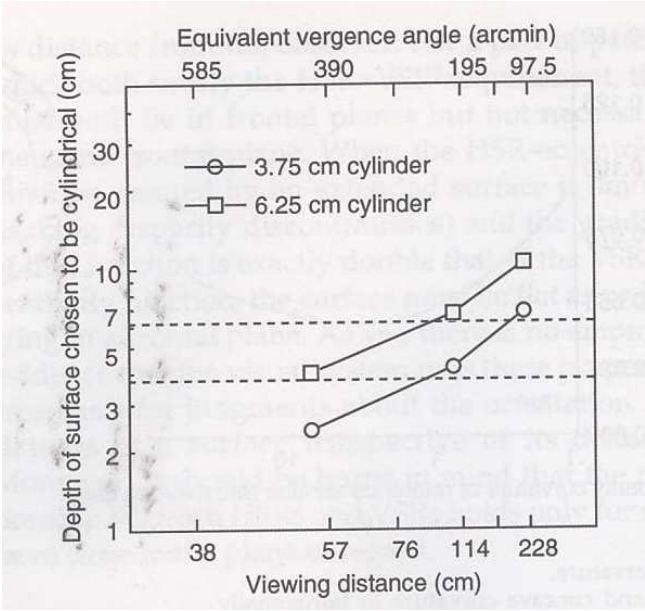


Figure 37 : variation de la perception de forme en fonction de la distance (Johnston, 1991)

Une autre explication a été donnée à cette distorsion en fonction de la distance d'observation. L'écran était plus ou moins éloigné de l'observateur pour limiter le conflit accommodation/convergence. Cette différence de position de l'écran a pu influer sur les disparités verticales ; celles-ci auraient modifié la perception globale des formes. Durgin a donc refait également la même expérience, mais cette fois, non plus en projetant un cylindre sur un écran, mais en utilisant des cônes réels. L'épaisseur du cône variait entre 50% et 200% de son diamètre et il y avait une distance de 1 à 3 mètres entre le cône et les sujets. Cette étude ne montrait aucune sous-estimation de la profondeur par rapport à la largeur de cône, même lorsque l'observateur se trouvait à plus de trois mètres de l'objet. Par contre, lorsqu'ils étaient proches, la surestimation restait présente, mais moins forte que celle relevée par l'expérience de Johnston. Durgin expliquait ces différences d'une part par la présence d'autres indices de profondeur présents dans la scène regardée, et par le fait que la perception des formes a pu être troublée par la technique de visualisation de Johnston (Durgin, et al., 1995).

2.1.2.2. Analyse des parallaxes de mouvements

Les parallaxes de mouvement ne varient pas en fonction de la distance d'observation. Ceci signifie que la courbure locale d'une surface peut être déduite de la courbure de disparité sans échelle de distance.

Durgin a refait ses expériences sur la perception des formes grâce aux disparités binoculaires (voir paragraphe 2.1.2.1) avec les parallaxes de mouvements. Dans son expérience, les sujets étaient fixes et les cônes en bois se déplacent devant ces derniers. Le résultat montrait que la perception de la forme était beaucoup moins précise avec les parallaxes de mouvements qu'à l'aide des disparités binoculaires (Durgin, et al., 1995). Cependant, la différence d'angle de vue était de 4,92°, une autre expérience a été menée avec une différence d'angle de vue de 7,2° et a donné une bien meilleure perception des courbures. Il faut toutefois noter que ces expériences ont été réalisées alors que le sujet était fixe et l'objet en mouvement, il est fort possible que si nous rajoutions l'information du système vestibulaire, nous obtiendrions de meilleurs résultats.

2.1.2.3. Analyse des textures

Une courbure peut être également perçue grâce aux gradients de texture, ou plus précisément, grâce aux variations de gradient de texture. En effet, les variations de tailles et de formes des éléments sont uniformes, notre cerveau perçoit un plan oblique ou incliné; par contre, si la variation n'est pas uniforme, il préférera percevoir une surface courbée plutôt que des éléments de texture non homogènes (voir Figure 38).

Figure 38 : Certaines surfaces complexes peuvent être reproduite de façon réaliste à l'aide de gradients de texture

Bajery et Liebermann ont établi que les variations dans le spectre des fréquences pouvaient être interprétées comme une distorsion d'une texture homogène (Bajery, et al., 1976). En effet, le rapprochement des éléments de la texture dû à l'éloignement ou au changement d'orientation de la surface, provoque une hausse de fréquences spatiales (voir Figure 39).

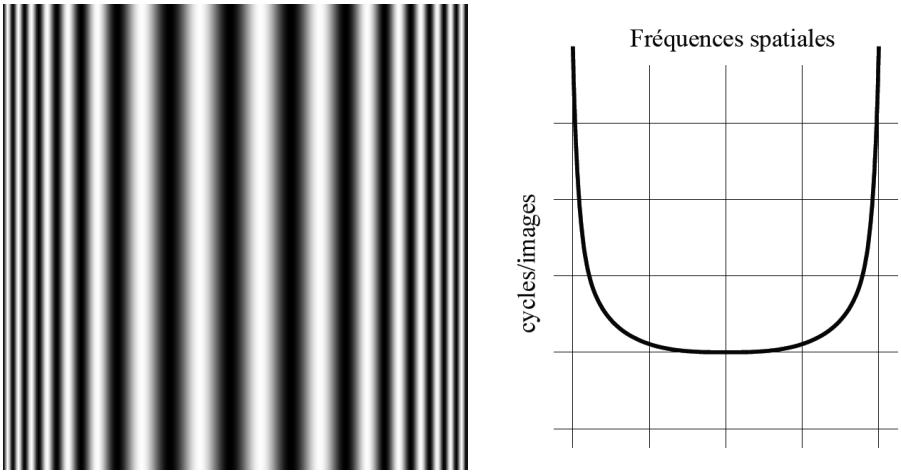


Figure 39 : Texture d'une surface courbe et traduction de la variation de celle-ci dans le domaine fréquentiel (Massot, 2006)

Sakai et Kinkel ont montré que le cerveau humain semblait suivre soit les variations des pics de fréquences, soit les variations de la moyenne des fréquences spatiales. La Figure 40 montre des

textures où ne varient que les pics de fréquences (à gauche), ou que la moyenne des fréquences (à droite). Les auteurs ont remarqué que lorsque les textures possèdent des pics de fréquences (comme sur la gauche de la figure) ce sont les variations de pic qui semblent donner plus d'information sur la courbure de la surface. Dans le cas contraire, ce sont les variations de la moyenne des fréquences qui permettent une meilleure perception d'une courbe. (Sakai, et al., 1995)

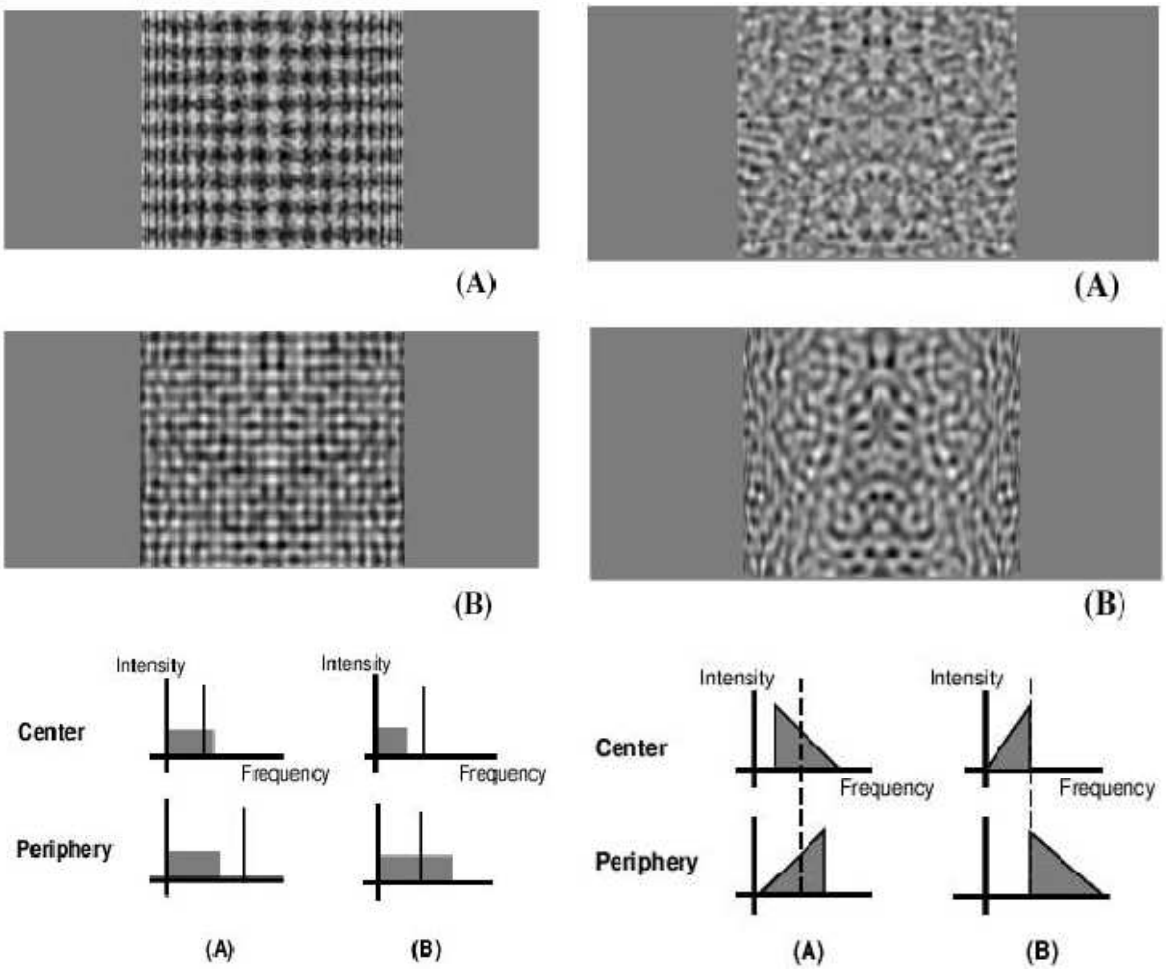


Figure 40 : Textures crées avec soit une variation des pics de fréquences (à gauche) soit une variation de la fréquence moyenne (à droite) (Sakai, et al., 1995)

2.1.2.4. Analyse de l'ombre

Une définition de l'ombre pourrait être « *variation de la quantité de lumière reflétée par une surface en fonction de l'orientation de cette surface par rapport à la source de lumière* » (Palmer, 1999). L'ombre est une source d'informations très puissante. Pour illustrer ceci, considérons une sphère faite d'un matériel qui reflète la lumière de façon homogène, comme représentée sur la Figure 41. Cette sphère est éclairée par un seul point de lumière situé à une certaine distance. Nous pouvons remarquer que l'intensité lumineuse reflétée par la sphère n'est pas la même en tout point. En effet, le point pour lequel la normale pointe vers la source lumineuse est beaucoup plus éclairé que le point pour lequel la normal pointe dans la direction opposée. C'est la variation de la lumière sur la surface qui nous permet d'interpréter la Figure 41 comme étant une sphère et non un disque.

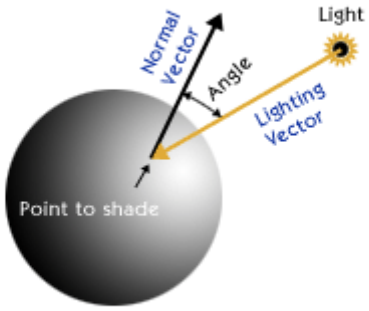


Figure 41 : l'ombre sur une sphère (upvector)

Koenderink a étudié la perception de l'orientation des surfaces (et donc des courbures) sur un torse humain en fonction des ombres apparaissant sur celui-ci (Koenderink, et al., 1992)(Koenderink, et al., 1996). Ils demandaient aux sujets de placer une jauge dans le sens de l'orientation perçue à certains points. La jauge est un ovale surmonté d'un petit segment, celui-ci représente la normale de la surface, l'ovale, le plan tangent à celle-ci. La Figure 42 représente le torse à gauche et l'orientation moyenne donnée par un sujet, à droite. Nous en déduisons que nous pouvons percevoir les courbures grâce aux ombres et à la façon dont la lumière se réfléchi sur la matière sans aide de la vision binoculaire ou des parallaxes de mouvement.

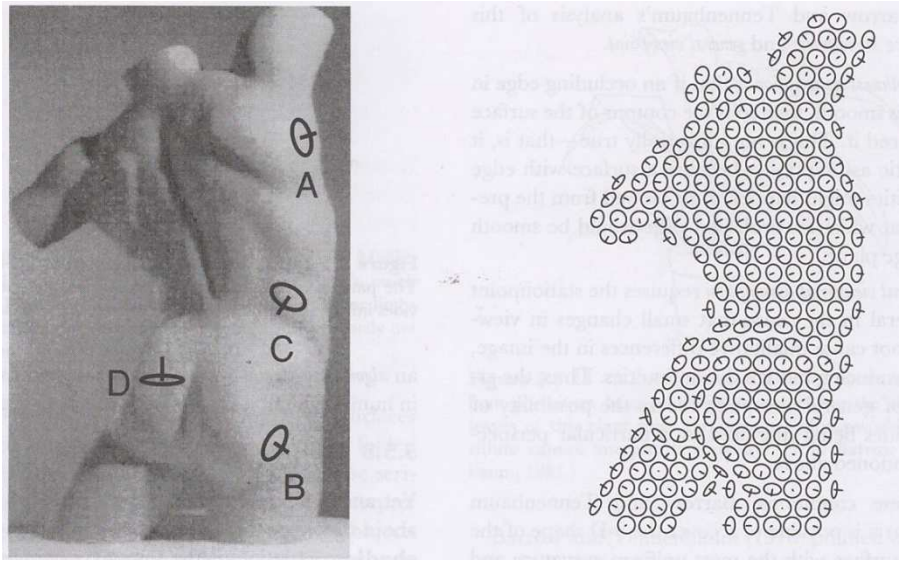


Figure 42 : orientation des surfaces de la peau en fonction des ombres (Koenderink, et al., 1996)

2.1.3. Perception des objets

Nous avons étudié la perception des surfaces, de leur orientation et de leur courbure. Cependant, puisque nos tests vont être réalisés en faisant une comparaison entre la forme d'un objet réel et celle d'un objet virtuel, il serait judicieux d'étudier également la perception des objets. En effet, notre cerveau va devoir, pour identifier la forme d'un objet, d'abord identifier l'objet lui-même.

Nous ne détaillerons pas tous les indices entrant en compte dans la perception des objets, mais il nous semble important d'évoquer les indices qui peuvent avoir une incidence sur notre expérience.

2.1.3.1. Les formes simples

Dans la résolution du conflit des formes possibles, le groupement ou la disjonction se fait dans le sens de la réalisation d'une forme privilégiée. Les formes privilégiées sont régulières, simples et symétriques.

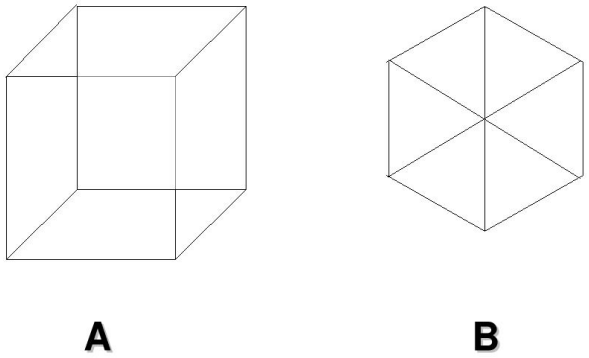


Figure 43 : différentes vues d'un cube

Lorsque nous voyons des figures géométriques, nous interprétons la forme de base pour qu'elle corresponde à une forme simple. Par exemple, sur la Figure 43, sont présentées différentes projections d'un cube. L'illustration A sera parfaitement interprétée comme un cube et nous y percevrons du relief. Par contre, l'illustration B sera interprétée comme un hexagone avec ses diagonales. Le cerveau ne trouve aucune forme simple dans le plan de la feuille qui corresponde à l'illustration A, donc il cherche cette forme simple en relief. Par contre pour l'illustration B, il trouve une (des) forme(s) simple(s) dans le plan de la feuille : un hexagone ou 6 triangles équilatéraux.

Nous pourrions penser que cette perception disparaît si la vision binoculaire est ajoutée, mais en réalité, il n'en est rien. Kopfermann a mis en conflit la vision binoculaire avec les facteurs d'organisation que nous venons de voir à l'œuvre. Des figures sont photographiées sur des clichés transparents qu'on peut disposer l'un derrière l'autre de manière à faire coïncider certaines lignes. Si ces dessins représentent les parties d'une figure qui, d'après les lois précédentes, tendrait à paraître plane, cette tendance persiste malgré la différence de profondeur objective (Guillaume, 1979).

2.1.3.2. L'invariance du point de vue

L'image des objets change en fonction de notre angle de vue. Cependant, le cerveau humain continue à percevoir le même objet ou environnement. Cette capacité est appelée invariance du point de vue (voir la Figure 44).

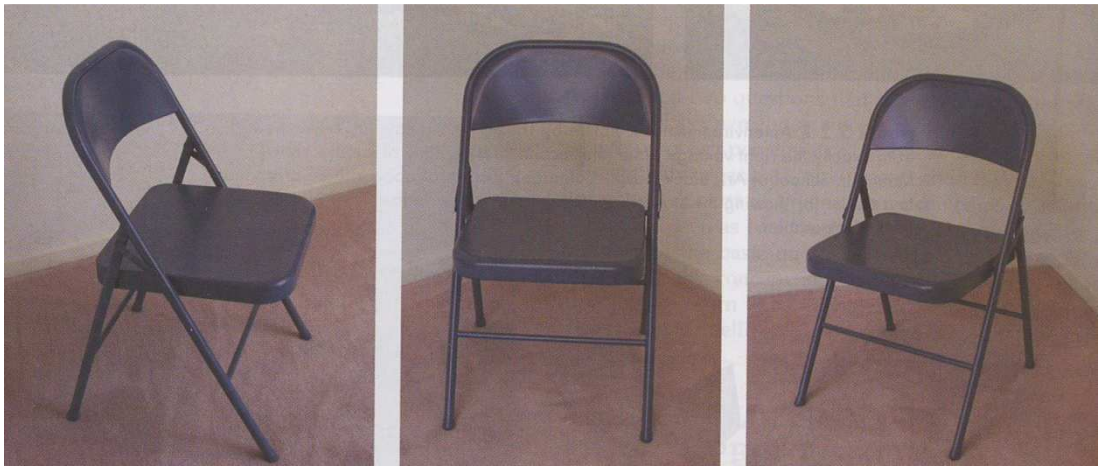


Figure 44 : nous voyons toujours la même chaise, bien que l'angle du point de vue change (Palmer, 1999)

2.1.3.3. Les couleurs et la luminosité

Dans une scène telle que celle représentée sur la Figure 45, il est nécessaire de déterminer quel changement de couleur est dû aux propriétés des objets et quel changement de couleur est dû aux modifications d'illumination. Par exemple, La différence de luminance entre l'aire (a) et l'aire (c) est produite par des peintures au mur différentes tandis que celle entre la zone (a) et (b) est provoquée par l'ombre du toit sur le mur. Ces affirmations nous semblent relativement faciles, mais pour des traitements d'image effectués sur ordinateur, il est encore extrêmement difficile de faire la différence.

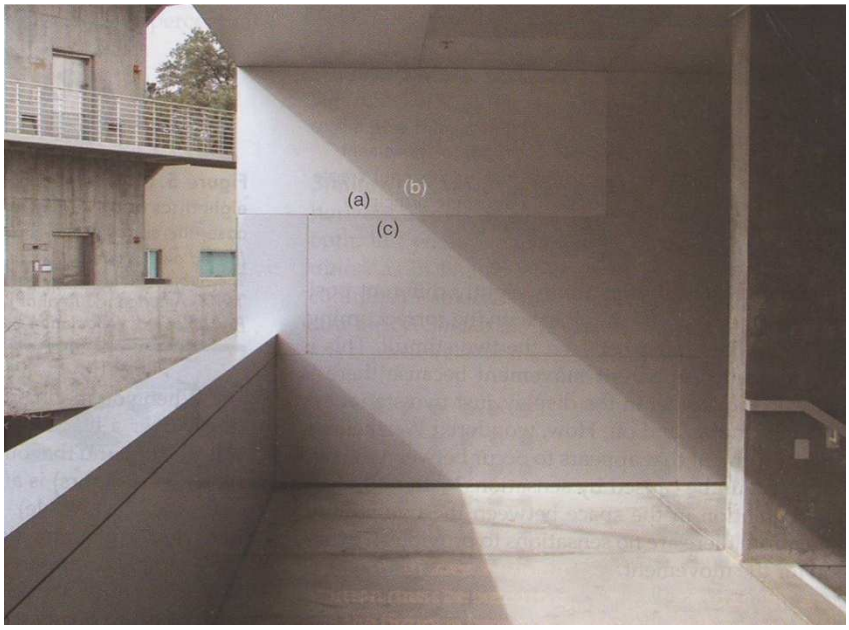


Figure 45 : nous distinguons les modifications de matière et de luminosité (Palmer, 1999)

2.1.3.4. L'objet et le fond

Pour que la discrimination entre l'objet et le fond soit bonne, il faut que leur intensité lumineuse soit différente. Un objet placé devant un décor uni, de couleur très différente mais de même

clarté sera très peu visible, leurs limites devenant fluctuantes, par contre, une différence de clarté permet de bien différencier les objets, même avec une teinte uniforme. (Guillaume, 1979)

Lorsque nous étudions les seuils, il importe de savoir si c'est un seuil sur l'objet ou sur le fond que nous évaluons. En effet, ceux-ci, bien que dans la même partie du champ, n'ont pas la même valeur suivant qu'ils soient vus comme figure ou comme fond. Une ombre légère est plus visible sur le fond que sur la figure; autrement dit, la figure offre plus de stabilité, plus de résistance à la variation que le fond (Guillaume, 1979).

2.2. Perception des formes en environnement immersif

Nous avons étudié la perception de formes en relief et survolé la perception des objets en vision naturelle. Nous allons maintenant nous intéresser aux études déjà réalisées en environnement virtuel. En effet, nous avons vu grâce aux critiques de l'étude de Johnston (voir paragraphe 2.1.2.1) que déduire des vérités générales sur la perception des formes sur la base d'expériences réalisées en environnement immersif peut être dangereux. L'affirmation inverse nous paraît vraisemblable : se baser uniquement sur des expériences étudiant la perception des formes en relief dans le monde réel sans prendre en compte les spécificités des environnements virtuels nous semble tout aussi dangereux. Nous passerons en revue les principaux facteurs influençant la perception de formes en immersion.

2.2.1. Influence du fond d'écran

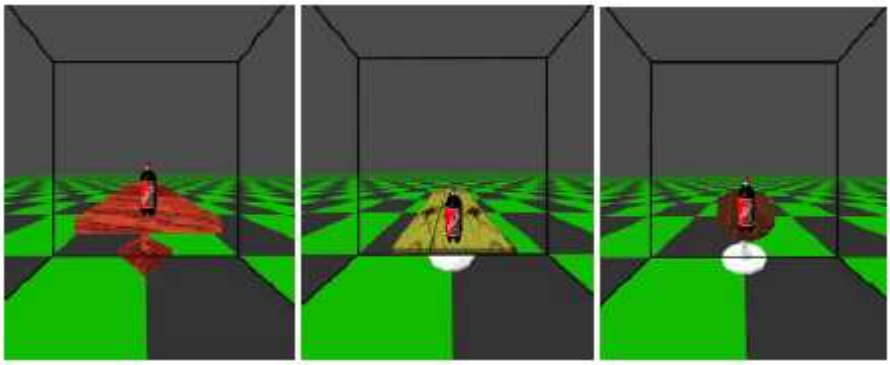


Figure 46 : les sujets de Luo devaient redimensionner une bouteille de Coca-Cola®, en fonction du fond d'écran et de la table sur laquelle elle est posée (Luo, et al., 2007)

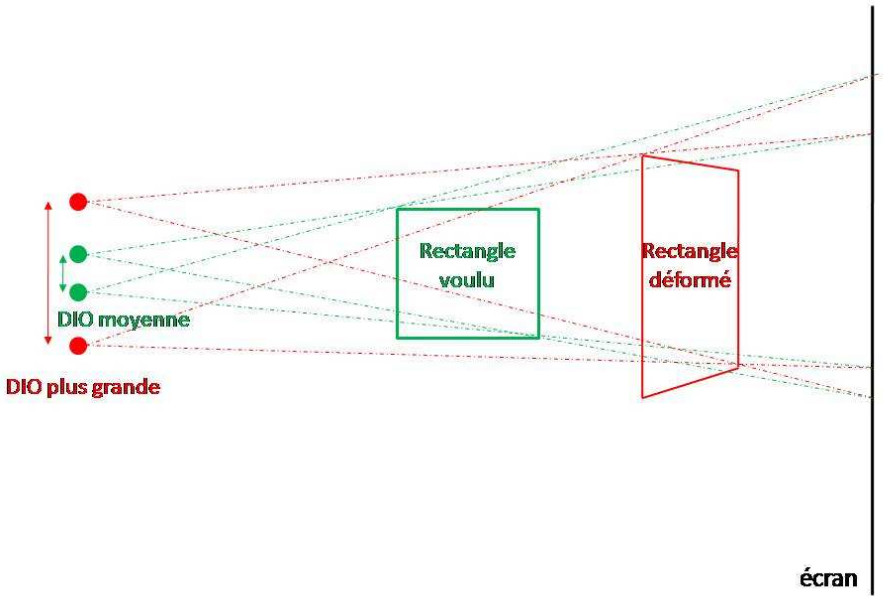
Luo a montré que le fond d'écran a une très grande importance dans la perception des formes et de la taille des objets (Luo, et al., 2007). Les indices de texture fournis par un fond d'écran nous permettent de mieux percevoir les formes de l'objet étudié (Figure 46).

Ces expériences confirment celles menées par Glennerster, la perception de formes et des objets est beaucoup plus simple lorsque plusieurs indices de profondeur concordants sont présents (Glennerster, et al., 1994). Dans ce cas présent, la texture du fond nous donne un indice de perspective très important. La table présente dans le décor aide aussi notre cerveau à percevoir correctement la forme de la bouteille par le jeu de textures qu'elle offre et permet une comparaison de taille avec un objet familier.

2.2.2. Influence de la distance inter-oculaire

La Distance Inter-Oculaire (DIO, en anglais IPD, Inter-Pupillary Distance) est variable pour l'être humain. En moyenne, elle est de 65 mm pour l'homme adulte, légèrement inférieure pour la femme, les écarts pouvant être significatifs, 50 à 70 mm environ, pour la population européenne (Fuchs, 2006).

Wann a publié la relation entre la perception des distances et la distance inter-pupillaire, mais il n'a pas évoqué la perception des formes (Wann, et al., 1995). Cependant nous avons déjà évoqué le lien intime entre perception des distances et la perception des courbures. Il est donc logique de supposer que lorsque la perception des distances est perturbée, la perception des formes l'est également.



47 : déformation d'un rectangle due à une DIO du sujet plus grande (vue du dessus)

Lorsque nous affichons un rectangle en stéréoscopie en faisant les calculs de projection en ne tenant compte que de la distance inter-oculaire moyenne, un observateur ayant une distance inter-oculaire plus grande, ne verra pas exactement le même rectangle : les points sont déplacés, les distances sont modifiées et donc les formes distordues. Ceci est illustré (de manière amplifiée) sur la 47. Nous voyons que les profondeurs sont écrasées et que le rectangle s'est transformé en trapèze. Le raisonnement inverse pourrait être fait pour une distance inter-oculaire de projection plus grande que la distance inter-oculaire du l'observateur ; le rectangle serait modifié dans ce cas aussi, mais dans l'autre sens. Il serait plus proche de l'observateur et la petite base du trapèze résultant serait face à l'observateur.

Lorsque la distance entre les caméras virtuelles est égale à la distance inter-pupillaire du sujet, nous parlons d'ortho-stéréoscopie.

2.2.3. Influence de l'accommodation et le flou dans la vision périphérique

Lorsque des images stéréoscopiques sont projetées, certains paramètres peuvent perturber la perception des formes et de l'environnement en général. Par exemple, en général, tout l'écran est vu net s'il est suffisamment petit (puisque nous accommodons sur l'écran), alors que dans la réalité, l'objet que nous regardons est net, mais le reste du monde est vu flou. Et cela peut changer un peu notre perception (Butts, et al., 1988)(Hillaire, et al., 2008)(Hillaire, et al., 2008)(Hillaire, et al., 2007).

2.3. Conclusion

Nous avons vu dans ce chapitre les différents facteurs influant sur la perception des formes, des inclinaisons et des courbures en immersion. Nous avons vu que les dérivées première et seconde des disparités binoculaires et des parallaxes de mouvement permettent de percevoir ces plans inclinés ou obliques ainsi que les courbures. Les gradients de textures ainsi que les variations de lumière et d'ombre sur les formes nous permettent également de percevoir les courbures. Dans le chapitre suivant, nous étudierons certains facteurs pouvant influer dans notre système d'immersion, nous expliquerons notre protocole ainsi que les résultats de nos expériences.

Chapitre 3 : Expérimentations et résultats sur la perception de formes en immersion

Ce chapitre traite des expérimentations faites pour déterminer les paramètres influant sur la perception des formes en immersion. Il commence par des explications de psychophysique, continue sur une brève introduction mathématique aux comparaisons statistiques de populations pour arriver à l'exposé de notre protocole de test. Nous terminons enfin par nos résultats commentés.

3.1. Psychophysique et statistique

3.1.1. Fondement

La psychophysique est une science qui étudie le rapport entre les stimuli et le ressenti de la personne, la façon dont elle perçoit ces stimuli. Gustave Théodore Fechner dans « Eléments de Psychophysique » a posé la base de cette matière en 1860 (Fechner, 1860). Nous n'aborderons pas ici les considérations mystiques qui accompagnaient les idées scientifiques, mais nous nous concentrerons sur les techniques et les méthodes qu'il a énoncées pour pouvoir mieux comprendre ce qu'est la perception.

Expliquons d'abord la différence entre une sensation et une perception. La sensation est le stimulus original, tel qu'il est appliqué sur les récepteurs nerveux, alors que la perception est la façon dont le cerveau l'interprète. Par exemple, la lumière est une sensation, une forme est une perception (Tiberghien, 1984). Nous appellerons donc « stimulus » l'ensemble des événements physiques qui déclenchent l'activité des récepteurs sensoriels. Ils sont ainsi à l'origine des réponses observées. Or, la description de ces événements n'est pas toujours exhaustive, puisque nous ne sommes jamais certains de tout connaître sur le fonctionnement de l'organisme que nous étudions. Pour pallier cette lacune, nous prendrons soin de mettre en œuvre des situations expérimentales très édulcorées où l'on tentera de ne faire varier qu'une seule dimension physique. Nous devrons donc exprimer toutes les caractéristiques de la situation susceptibles d'avoir une influence sur la réponse du sujet au stimulus étudié (Bonnet, 1986).

Il faut notamment faire attention à ne pas utiliser un stimulus qui puisse représenter autre chose que ce que nous souhaitons étudier. Par exemple, pour évaluer la capacité du sujet à donner le diamètre d'un cercle, il faut éviter d'utiliser des pièces de monnaie : des études ont montré que plus la valeur monétaire de la pièce est grande, plus sa dimension géométrique est surévaluée. Le niveau social du sujet influence également cette surévaluation (Tiberghien, 1984).

L'approche scientifique d'une question se caractérise par quatre étapes :

Etape 1 : Choisir, formuler et poser une question de recherche ;

Etape 2 : Formuler une (ou plusieurs) hypothèse(s) relative(s) à cette question ;

Etape 3: Conduire une étude empirique pour tester cette (ces) hypothèse(s) ;

Etape 4: Répondre à la question de recherche à la lumière des données recueillies (Lemaire, 2006) .

3.1.2. Détection de seuil

La notion de seuil est fondamentale en psychophysique. Elle renvoie à l'idée d'un passage entre deux états (l'état [0] et l'état [1]). Le seuil est la limite de perception de stimulation par un sujet, dans une situation expérimentale pour une tâche bien définie. Il y a quatre types de tâches : la détection, la discrimination, la reconnaissance et l'identification (Bonnet, 1986).

Pour un sujet donné et une dimension de stimulus donnée, la valeur du seuil n'est pas constante. De même, elle varie selon les individus. Les fluctuations du seuil peuvent être liées à plusieurs phénomènes : la fatigue, la modification de vigilance, les artéfacts de mesures. Pour l'acuité visuelle, un entraînement à la tâche influence très fortement le niveau du seuil critique (Tiberghien, 1984). Si nous considérons qu'il existe un niveau de stimulus critique (S_c) en dessous duquel l'individu considère que la réponse est l'état [0] (non détecté, non discriminé, non reconnu ou non identifié selon la tâche) et au-dessus duquel il estime que le résultat est plutôt l'état [1] (détecté, discriminé, reconnu ou identifié). Soit S_i le stimulus présenté :

Si $S_i<S_c \rightarrow R=\{0\}$

Si $S_i>S_c \rightarrow R=\{1\}$

La variabilité des réponses tient du fait que le niveau de S_c n'est pas constant. Nous postulerons donc le niveau S_c comme étant la limite définie statistiquement (Bonnet, 1986). Soit nous le définissons comme étant la médiane des seuils donnés : 50% des valeurs des seuils exprimés par le sujet sont en dessous du seuil critique. Soit nous le définissons comme étant la moyenne et nous donnons l'écart type. Nous verrons plus tard que nous utilisons la moyenne et l'écart type.

Il peut exister un certain nombre de fausses détections, c'est à dire que le sujet perçoit le stimulus alors qu'il n'existe pas. Plusieurs hypothèses peuvent exister :

- il existe des sensations « négatives » (l'idée d'inconscient apparaît) ;
- le sujet a répondu au hasard. Il faut donc prendre en compte la théorie de la décision ;
- il y a une hésitation entre deux réponses (en tout ou rien). Il faudra donc mesurer en plus la latence ou la certitude (Tiberghien, 1984).

Le Point d'égalisation subjectif (Point of subjective equalization – PSE) se définit comme étant la valeur du stimulus pour laquelle les sujets d'un test de comparaison considèrent que le stimulus est équivalent au stimulus présenté.

Enfin, la plus petite différence distinguable (Just noticeable difference – JND) est la plus petite différence de valeur de stimulus par rapport au stimulus présenté qui est perçue comme étant effectivement différente par 75% des sujets d'un test de comparaison.

3.1.2.1. Seuil absolu

La méthode du point central

Appelée aussi méthode d'autoréglage ou méthode de l'ajustement, cette technique consiste à demander au sujet de régler lui-même la valeur du stimulus pour que sa valeur soit celle du seuil. Néanmoins, il faut effectuer la mesure plusieurs fois pour compenser la variabilité du seuil critique expliquée plus haut. La rapidité de l'obtention des mesures permet un grand nombre de répétitions.

Ce procédé est très efficace lorsque le stimulus peut varier de façon continue, mais il peut être utilisé avec un stimulus discret. Dans ce cas, le choix du pas d'augmentation ou de diminution doit prendre en compte la précision que nous cherchons à atteindre.

Il est possible d'autoriser ou non les retours en arrière dans la zone de seuil pour que le sujet puisse affiner son résultat. Cependant, pour un seuil de détection, cela conduit souvent à une surestimation de la valeur du seuil. D'autre part, un phénomène d'hystérèse est constaté lorsque la valeur du stimulus ne peut que croître ou décroître (Bonnet, 1986).

La méthode des limites

Elle est appliquée lorsque le stimulus varie de façon discrète. La valeur du stimulus affichée au départ est largement en dessous du seuil. Elle augmente ensuite progressivement jusqu'à ce que le sujet exprime la perception du stimulus. La même progression doit être faite dans l'autre sens en partant d'une valeur nettement au dessus du seuil et en la diminuant jusqu'à ce que le sujet ne la perçoive plus.

La méthode des stimuli constants

Cette méthode est très semblable à la précédente, mais propose des stimuli dans un ordre aléatoire pour éviter l'effet d'hystérèse (Bonnet, 1986)(Tiberghien, 1984)(Goldstein, 2003) .

A priori, cette méthode présente les meilleures garanties de validité et de précision pour la mesure des seuils absolus ou différentiels, surtout parce qu'elle permet des estimations non biaisées de la fonction psychométrique (voir 3.1.3). Cependant, la stabilité des estimations requiert un grand nombre d'essais ainsi qu'une grande stabilité de l'état du sujet (Goldstein, 2003).

Il y a un certain nombre de règles à respecter pour que cette technique soit appliquée correctement :

1. Un niveau seuil doit être choisi. Il pourra être déterminé en procédant par exemple à un test selon la méthode des limites.
2. Un ensemble de 5 à 7 niveaux du stimulus centré sur le niveau déterminé est choisi. Le choix de ces niveaux doit être tel qu'en fin d'expérience, la probabilité d'une des deux réponses soit proche de zéro pour le niveau le plus bas et proche de l'unité pour le niveau le plus élevé.
3. Une règle de progression de ces niveaux doit être choisie. Nous pouvons choisir une règle arithmétique, ou une progression géométrique. Ce choix est basé sur les connaissances préalables que nous possédons sur la perception de la dimension étudiée.
4. Définir a priori un nombre d'essais (N) constant et identique pour chacun des n niveaux retenus. Plus N est élevé, meilleure est l'estimation de la fonction psychométrique.
5. La règle de présentation des différents niveaux au cours des essais successifs est basée sur un tirage au hasard sans remise : aucun niveau n'est représenté tant que les n niveaux de la série n'ont pas été présentés (Bonnet, 1986).

Les procédures adaptatives

Les procédures adaptatives ont une règle de présentation des niveaux successifs du stimulus qui définissent le niveau suivant en fonction des réponses données aux niveaux précédents. Elles présentent de nombreux avantages :

- elles ne nécessitent pas de connaître a priori le seuil du sujet ;
- elles cherchent à obtenir un maximum d'informations du sujet en un minimum d'essais. Elles présentent donc plus de stimulus dans la zone de seuil que dans les zones éloignées ;

- elles permettent une estimation de la fonction psychométrique ;
- elles ne supposent pas nécessairement que le sujet reste stable tout au long de l'expérience.

Cependant, elles sont relativement lentes et sont plus difficiles à mettre en œuvre.

Les procédures à contraintes faibles

Ce sont des dérivées de la méthode des limites. Nous présentons des essais de niveaux croissants du stimulus tant que la réponse est toujours la même, dès qu'elle change, le sens de présentation des valeurs est inversé.

L'escalier psychophysique : L'expérience commence avec la méthode des limites, généralement de niveau ascendant. Lorsque la réponse « perçu » est donnée, le sens s'inverse. Il faut choisir une règle d'arrêt, traditionnellement, on opte pour 10 inversions.

La méthode de poursuite des seuils dite méthode de Békésy : Elle est utilisée notamment lors de dépistage de surdité sélective (le patient n'entend pas une bande de fréquences particulière). Le son est diffusé en continu dans les haut-parleurs à la limite auditive. Tant que le patient appuie sur un bouton, l'intensité du son diminue, elle augmente lorsque le bouton est relâché. La fréquence du stimulus augmente tout le long du test, indépendamment des réponses du sujet.

Les procédures à contraintes fortes

Ces méthodes consistent principalement à rajouter à l'escalier psychophysique une règle de dépendance portant sur la précision de l'estimation du seuil au cours des essais successifs.

3.1.2.2. Mesure différentielle

Le stimulus étalon est soit présenté avant ou après, soit présent en même temps que le stimulus à étudier. Il peut également être présenté aléatoirement avant ou après pour limiter l'influence de la surestimation ou sous-estimation. Pour chaque échantillon, le sujet indique « plus grand » ou « plus petit ». Chaque stimulus apparaissant plusieurs fois, il est possible que le sujet donne des réponses différentes pour un stimulus. Si la personne répond deux fois « plus grand » et trois fois « plus petit », nous considérons que le stimulus est perçu « plus petit », si le nombre de réponses est identique, il est considéré comme « identique ». Il y a deux seuils à définir : la valeur liminaire supérieure et inférieure. Ce sont respectivement les limites « plus grand » et « plus petit ». Le seuil global est la demi-somme de ces deux seuils.

3.1.3. Fonction psychophysique

La fonction qui lie la probabilité de réponse « plus grand » (« détecté », « discriminé », identifié »…) au continuum des stimuli est appelée la « fonction psychophysique » ou « fonction psychométrique ».

Classiquement, on suppose que la fonction psychométrique

$$p_i(1) = F(S_i)$$

est une fonction normale cumulée dont l'équation est

$$F(S_i) = \int f(S_i)dS_i = \frac{1}{\sigma\sqrt{2\pi}}\int \exp\left[\frac{-(S_i - m)^2}{2\sigma^2}\right]dS_i$$

3-1

Où

- *m* est la moyenne arithmétique de la distribution ou espérance mathématique ;
- σ l'écart type ou écart quadratique moyen ;
- S_i la moyenne aléatoire, c'est-à-dire les niveaux du stimulus.

Pour obtenir ces estimations, il faut :

1. Pour chaque niveau S_i du stimulus, on calcule la probabilité des réponses « oui » : *p*

2. En utilisant une table de la loi normale réduite, on transforme chaque valeur *pi(1)* en valeur de z_i

3. On estime par la méthode des moindres carrés les paramètres de la droite $z_i = aS_i + b$

4. Des paramètres *a* et *b* de la droite de régression de z_i en S_i, on peut obtenir le seuil de détection qui n'est autre que la moyenne de la fonction (m_{si})

$$m_{S_i} = \frac{-b}{a}$$

3-2

En effet

$$z_i = \frac{S_i - m_{S_i}}{\sigma_{S_i}}$$

Lorsque $S_i = m_{si}$, alors $z_i = 0$ et en remplaçant S_i par m_{si}, on a

$$a \,.\, m_{S_i} + b = 0$$

L'écart type de la distribution normale réduite vaut z=1. Pour trouver l'écart type de la distribution du Si, nous cherchons la valeur de σ_{S_i}, qui est la distance de (1-b)/a à la moyenne m_{si}. En effet :

$$1 = a \,.\, S_i + b$$

D'où $\sigma_{S_i} = m_{S_i} - \frac{(1-b)}{a}$

Ou encore $\sigma_{S_i} = \frac{1}{a}$ (Bonnet, 1986)

3-3

3.1.4. Comparaison statistique de populations

Considérons un échantillon aléatoire simple contenant n composantes $\{x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}\}$ de moyenne $\overline{x}_k$ et d'écart-type s_k^2. Il a été prélevé dans une population globale U_k de moyenne μ_k et

de variance σ_k où *k=1,2* est le numéro de l'échantillon (Droesbeke, 1992) (Amzallag, et al., 1978) (Ramousse, et al., 1996).

Test d'égalité de moyenne

Nous voulons tester l'hypothèse *Ho* : $\mu_1 = \mu_2$ ou l'hypothèse inverse $H_1 : \mu_1 \neq \mu_2$

Posons

$$D = \bar{x}_1 - \bar{x}_2 \qquad \textbf{3-4}$$

l'estimation de la différence des deux moyennes des populations.

	Populations quelconques ($n \geq 30$)	Populations normales
σ_1, σ_2 connus	$T = \dfrac{D}{\sqrt{\dfrac{\sigma_1}{n_1} + \dfrac{\sigma_2}{n_2}}} \approx N(0,1)$ **3-5**	$T = \dfrac{D}{\sqrt{\dfrac{\sigma_1}{n_1} + \dfrac{\sigma_2}{n_2}}} \approx N(0,1)$ **3-6**
σ_1, σ_2 inconnus	$T = \dfrac{D}{\sqrt{\dfrac{s_1^2}{n_1 - 1} + \dfrac{s_2^2}{n_2 - 1}}} \approx N(0,1)$ **3-7**	Si $\sigma_1^2 = \sigma_2^2$ $T^* = \dfrac{D}{\sqrt{\dfrac{s_1^2}{n_1 - 1} + \dfrac{s_2^2}{n_2 - 1}}} \approx N(0,1)$ **3-8**

Nous pouvons rejeter l'hypothèse Ho et donc dire que les deux populations sont significativement différentes si :

$$\begin{cases} |T| > z_{1-\alpha/2} \\ |T^*| > t_{n1+n2;1-\frac{\alpha}{2}} \end{cases} \qquad \textbf{3-9}$$

où :

- α est la probabilité acceptable d'erreur ;
- z est la variable de la distribution normale réduite ;
- *t* est la variable de la distribution de Student.

Test d'égalité de deux variances

Nous voulons tester l'hypothèse *Ho* : $\sigma_1 = \sigma_2$ ou l'hypothèse inverse $H_1 : \sigma_1 \neq \sigma_2$

Posons

$$S_k^2 = \frac{n_k^2 s_k^2}{n_k - 1} \quad \text{Avec k=1,2}$$

3-10

Populations quelconques ($n_k \geq 50$)	$$T = \frac{(S_1^2 - S_2^2)}{\sqrt{\frac{\mu_4^{(1)} - S_1^4}{n_1} - \frac{\mu_4^{(2)} - S_2^4}{n_2}}} \approx N(0,1)$$ Si μ_4^k est inconnu, on l'estime par m_4^k (*k*=1,2) **3-11**
Populations normales	$$T^* = \frac{S_1^2}{S_2^2} = F_{n_1-1;n_2-1}$$ Si n_1 et $n_2 > 30$: $$T = \frac{\frac{1}{2}\log_e\left(\frac{S_1^2}{S_2^2}\right)}{\sqrt{\frac{1}{2}\left(\frac{1}{n_1-1} + \frac{1}{n_2-1}\right)}} \approx N(0,1)$$ **3-12**

Nous pouvons rejeter l'hypothèse Ho et donc dire que les deux populations sont statistiquement différentes si :

$$\begin{cases} |T| > z_{1-\alpha/2} \\ |T^*| > F_{n_1-1;n_2-1;1-\frac{\alpha}{2}} \end{cases}$$

3-13

où :

- α est la probabilité acceptable d'erreur ;
- z est la variable de la distribution normale réduite ;
- *F* est la variable de la distribution de Fisher-Snédecor.

Nous verrons par la suite que la plupart de nos distributions sont normales nous utiliserons donc généralement les équations 3-12 et 3-8.

En Annexe D, nous présentons les formules de comparaison de plusieurs distributions.

3.2. Protocole de test

Pour établir ce protocole, nous suivrons exactement le plan défini au paragraphe 3.1.1. Ce canevas prévoit 4 étapes, l'établissement des questions de recherche, l'émission des hypothèses, la description de l'étude expérimentale en elle-même et enfin, l'analyse des résultats.

3.2.1. Question de recherche

Il nous est possible, grâce aux techniques de stéréovision, de visualiser les objets réalisés par la conception par ordinateur. Les concepteurs peuvent ainsi modifier l'objet virtuel autant qu'ils le souhaitent. Lorsque cet objet leur convient, ils décident de faire construire cet objet en réel. Mais est-ce que cet objet réel coïncidera vraiment avec la forme qu'ils avaient perçue sur leur écran ? La question n'est évidemment pas de savoir si le service de construction a bien fait son travail, mais de savoir si la vision stéréoscopique a rendu la bonne perception de l'objet virtuel pour la personne qui l'a observé.

La perception d'une forme dépend entre autres des dimensions relatives dans les trois dimensions (longueur, hauteur, profondeur) de l'objet. Leur rapport conditionne les courbures et l'orientation des surfaces de l'objet. Nous pouvons donc nous interroger sur la conservation de ces proportions

Nous nous demanderons quels sont les paramètres qui influencent la perception d'un objet virtuel. Est-ce que le fait de disposer l'objet virtuel sur l'écran pour minimiser les disparités peut influencer le jugement du sujet ? Est-ce que la vision stéréoscopique est réellement plus efficace que la vision monoscopique ? Dans quelle mesure le fait de traquer le point de vue du sujet peut-il améliorer ses performances ? Nous essayerons de répondre à toutes ces questions

3.2.2. Hypothèses relatives aux questions

Nous souhaitons valider plusieurs hypothèses :

1. Il est préférable d'avoir le point de vue traqué en vision monoscopique ;
2. Il est préférable d'avoir le point de vue traqué en vision ortho-stéréoscopique ;
3. Il est préférable d'avoir le point de vue traqué en vision non-ortho-stéréoscopique ;
4. Il est préférable d'avoir la vision ortho-stéréoscopique que la vision non-ortho-stéréoscopique avec le point de vue traqué ;
5. Il est préférable d'avoir la vision ortho-stéréoscopique que la vision non-ortho-stéréoscopique sans le point de vue traqué ;
6. Il est préférable d'avoir la vision monoscopique avec le point de vue traqué que la vision ortho-stéréoscopique sans le point de vue traqué.

Nous vérifierons chacune de ces hypothèses pour une forme primitive (sphère) et pour une forme quelconque et a priori non connue des sujets. Nous les vérifierons également lorsque ces deux formes sont situées devant, derrière ou sur l'écran.

3.2.3. Etude empirique

3.2.3.1. Les sujets

15 sujets ont passé les tests de perception d'une sphère et 18 personnes, ceux de la forme quelconque. Les sujets doivent être en bonne santé. Ils peuvent avoir des problèmes de vision (myopie, astigmatisme, hypermétropie, presbytie), à condition qu'ils portent leurs verres correcteurs. Ils doivent être capables de se déplacer dans la pièce sans aucune gêne. La distance interoculaire est mesurée pour chaque sujet et intégrée dans les calculs des images virtuelles lorsqu'elles sont stéréoscopiques.

Il faut vérifier que les sujets ont une bonne vision binoculaire. Un simple test de contrôle est effectué avant le test de perception en lui-même : on montre au sujet deux cubes adjacents de dimensions différentes, l'un est nettement plus proche de l'autre. La personne doit déterminer lequel est plus proche et ce trois fois de suite pour éviter les réponses aléatoires.

3.2.3.2. La méthode employée

La méthode que nous allons employer est la méthode des stimuli constants. Elle est en effet la plus fiable et la plus précise (Bonnet, 1986). Nous avions vu au paragraphe 3.1.2.1 qu'il existe des règles à respecter qui conditionnent la fiabilité des résultats. Nous les respecterons donc toutes.

Nous cherchons un seuil de différentiation (« Percevez-vous le même objet virtuel que l'objet physique ? »), nous utilisons donc comme forme de départ l'objet virtuel correspondant à l'objet réel. Nous déterminerons un ensemble de 7 variations dans chaque dimension (largeur, hauteur, épaisseur) centrée autour des dimensions de l'objet de départ. Notre règle de progression est arithmétique pour plus de facilité. Chaque stimulus est présenté 3 fois par un tirage au sort sans remise.

3.2.3.3. Les objets

La sphère

La sphère réelle

La sphère réelle (Figure 49) est en fait un ballon en plastique lisse d'un diamètre de 20 centimètres peint en vert pour correspondre à la couleur du modèle virtuel. Il est présenté devant un drap noir et éclairé de la même façon que le modèle virtuel (lampe disposée au même endroit et d'intensité ajustée à la main par l'expérimentateur pour que la correspondance soit la meilleure possible - Figure 48). Nous avons tenté de prendre des films ou photo des deux sphères pour avoir la meilleure comparaison de luminosité. Cependant, comme l'écran est rétro-projeté, le capteur est saturé et ne peux prendre de photo ressemblant à ce que l'œil perçoit, le biais induit par ces capteurs est trop important.

La distance entre les deux lampes du bas est de 0,9m, elles sont situées 0,98m en dessous de la forme. La lampe du dessus est située 0,62m au dessus de la forme.

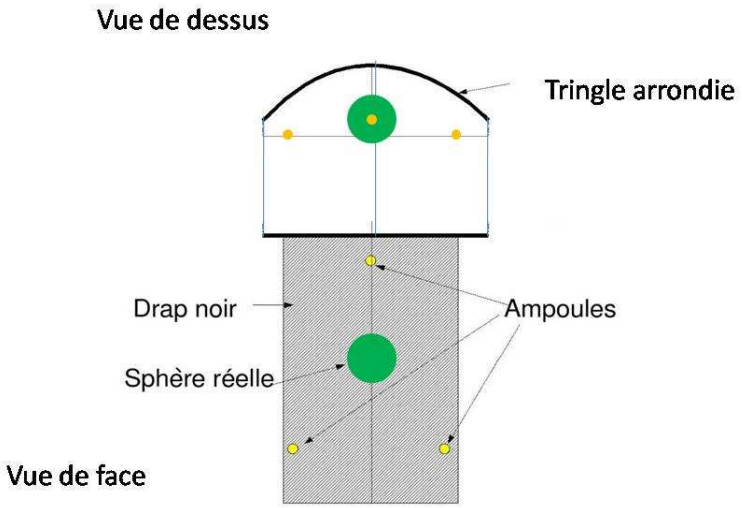


Figure 48 : schéma du dispositif réel

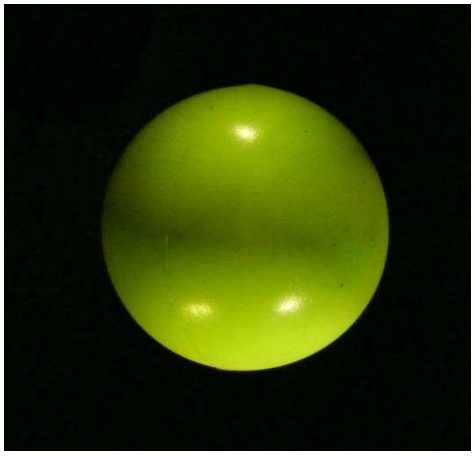


Figure 49 : sphère réelle

Les sphères virtuelles

La sphère étalon est donc la reproduction relativement exacte du ballon (supposé sphérique). Elle est verte et également disposée sur un fond noir (Figure 50). Nous faisons varier dimension par dimension par pas de 3% du diamètre initial (6mm) de façon centrée sur la valeur d'origine pour les dimensions horizontale et verticale (appelées pour plus de facilité dans la suite l'exposé la largeur et la hauteur) et par pas de 12% du diamètre initial (24mm) pour les variations de profondeur. Ces pas ont été définis lors de tests préliminaires, effectués sur 3 personnes. Lorsqu'une dimension est différente de la dimension de la sphère d'origine, les autres gardent la valeur d'origine, pour ne changer qu'une seule dimension à la fois.

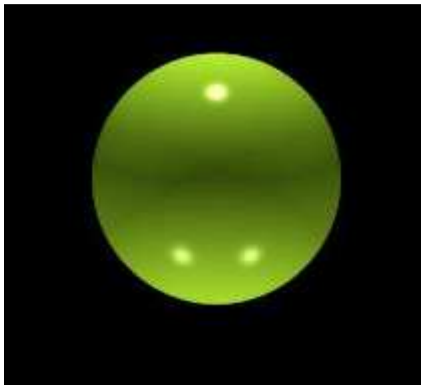


Figure 50 : sphère virtuelle

La forme inconnue

La forme inconnue a pour but de savoir si les sujets ne sont pas influencés par la forme simple que constitue une sphère. Il serait intéressant de connaître la réaction des sujets face à une forme totalement inconnue, dont le rayon de courbure n'est pas fixe. Ce cas est très intéressant car il se rapproche beaucoup plus des formes rencontrées dans la vie de tous les jours, en particulier dans l'industrie automobile. Néanmoins, les résultats sont également plus compliqués à analyser.

La forme réelle

Cette forme inconnue est une sphère « déformée ». Elle a la même dimension en moyenne que la sphère du paragraphe précédent, mais ne présente aucune symétrie (Figure 51). Elle a été imaginée et dessinée dans le logiciel CATIA (logiciel de CAO) et fabriquée en prototypage rapide dans de la mousse de polystyrène rigide à l'aide d'une machine-outil numérique CharlyRobot. Cette forme est peinte en vert avec la même peinture que celle utilisée pour la sphère réelle et est maintenue dans la même position que la forme virtuelle par trois câbles pour éviter les rotations..

Elle est également disposée devant un drap noir (le même que pour la sphère) et utilise le même éclairage (réglé de la même façon). Sur la Figure 52, nous voyons l'ensemble du dispositif. Notons que, sur la photo, les lampes sont apparentes. Lors des tests, elles sont cachées (celle du dessus par un drap, celles du dessous par une planche peinte en noir).

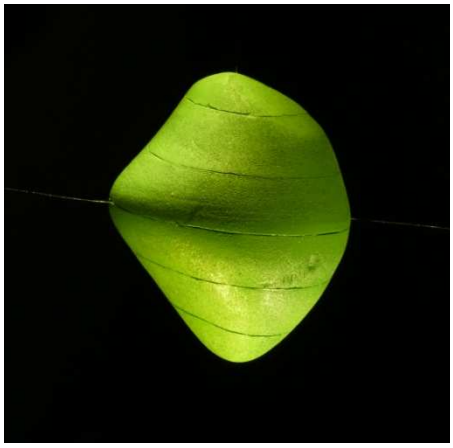


Figure 51 : forme inconnue réelle vue de près

Figure 52 : dispositif réel complet vue de loin

La forme virtuelle

La pièce virtuelle de base est la même que celle qui a servi à l'élaboration de la pièce réelle (sous CATIA) : elle est verte et placée devant un fond noir (Figure 53).

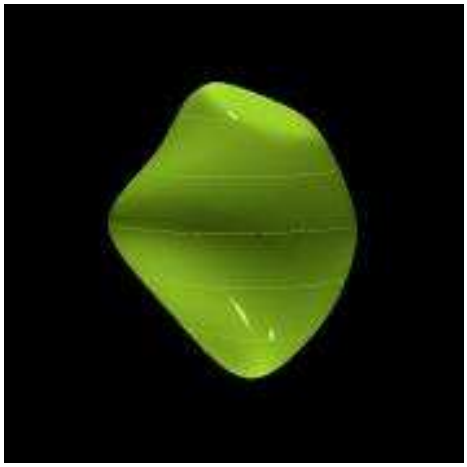


Figure 53 : forme inconnue virtuelle

3.2.3.4. Les expériences menées

Les séquences

Nous faisons varier plusieurs paramètres :

- Ecart entre les deux caméras virtuelles :
  - L'écart entre les deux caméras virtuelles est nul (vision monoscopique),
  - L'écart entre les deux caméras virtuelles est ajusté sur la distance inter-pupillaire du sujet,

  - L'écart entre les deux caméras virtuelles est doublé par rapport à la distance inter-pupillaire du sujet. Ce écartement a été choisi pour avoir la même déformation pour chaque sujet, quelque soit la distance entre ses deux yeux.
- Point de vue du sujet traqué
  - Le sujet est fixe dans la pièce et n'a pas le droit de bouger, la caméra virtuelle est fixe à l'endroit où est sensé se trouver la tête du sujet,
  - Le sujet peut se déplacer et la caméra virtuelle suit la position et l'orientation de sa tête.
- Position de la forme par rapport à l'écran
  - La forme est sur l'écran ce qui minimise les parallaxes horizontales,
  - La forme est devant l'écran ce qui permet au sujet de tourner davantage autour de celle-ci et diminue la pixellisation
  - La forme est derrière l'écran.
- La forme en elle-même
  - La forme est une sphère, une forme primitive bien connue de tous les sujets
  - La forme est a priori non connue des sujets.

Nous testons toutes les combinaisons de ces paramètres, ce qui nous donne 36 combinaisons. Lorsque le point de vue du sujet n'est pas traqué, celui-ci a alors une position fixe aussi bien en réel qu'en virtuel. Celle-ci dépend de la position de l'objet virtuel. Nous les avons définies pour que le sujet soit toujours à 1m20 aussi bien de la forme virtuelle que de la forme réelle. Cette distance a été choisie en fonction des résultats de Johnston (voir paragraphe 2.1.2.1) et en fonction de la disposition de notre matériel. Les croix au sol définissant ces positions sont mises en évidence pour le sujet avant de commencer la toute première séquence. Comme représenté sur la Figure 54, les écartements de l'objet par rapport à l'écran sont de 50 centimètres.

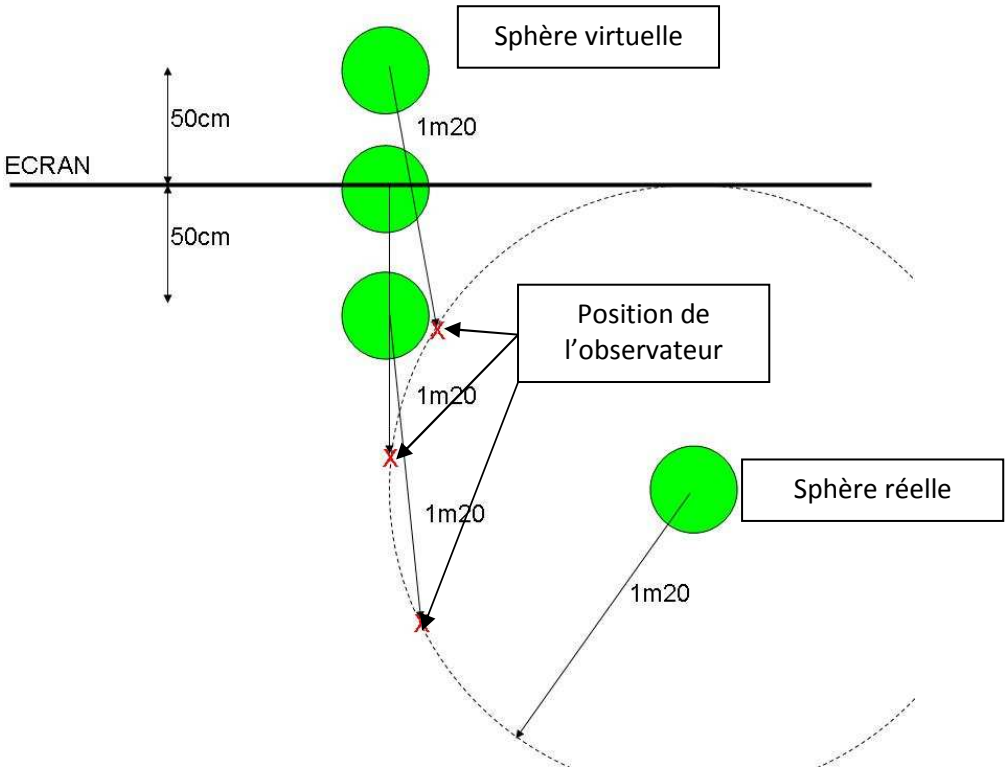


Figure 54 : disposition des objets virtuels et réel ainsi que de l'observateur lors du déplacement de l'objet virtuel

Lorsque la forme est derrière l'écran et que le sujet est fixe, la projection de la forme étalon fait 14 cm de diamètre, ce qui correspond à 87 pixels. Donc une variation de 3% de son diamètre modifie 2,61 pixels. Lorsque l'objet étalon est sur l'écran, sa projection fait bien 20 cm de diamètre, donc 124 pixels. Une modification de son diamètre de 3% change alors 3,72 pixels. Enfin, lorsque la forme étalon se trouve devant l'écran, sa projection mesure 34cm, autrement dit 211 pixels. Une modification de 3% supprime ou ajoute 6,33 pixels.

Au sein d'une séquence

Au début de chaque séquence, des consignes sont données au sujet en haut à gauche de l'écran. Les consignes peuvent être :

« Vous pouvez vous déplacer dans la pièce »

ou

« Placez-vous sur la croix la plus proche de l'écran »

Au milieu de l'écran, une phrase invite le testeur à appuyer sur la touche « A » de la Wiimote (le gros bouton rond sur le devant, Figure 55) dès qu'il aura lu les consignes.

« Appuyer sur A pour commencer»

Figure 55 : boîtier de réponse ©Nintendo

Lorsqu'une séquence débute, nous commençons à faire varier la largeur de la sphère virtuelle. Le sujet est prévenu tout le long du test par la phrase :

« La largeur de la sphère varie. La forme virtuelle est-elle moins large (-) ou plus large (+) ?»

Le sujet est alors invité à appuyer sur la touche « - » ou « + » de la manette (Figure 56) pour donner son avis. Lorsqu'une touche est enfoncée, une demande de confirmation apparaît à l'écran :

« Vous avez appuyé sur « -/+ » appuyez sur « A » pour confirmer »

Figure 56 : capture d'écran lors du test : demande de confirmation d'une réponse

Lorsqu'il a validé sa réponse, l'objet disparaît une demi-seconde pour éviter les comparaisons entre les formes virtuelles avant d'afficher la variation suivante. Si le sujet appuie sur « A » sans avoir donné une réponse, le programme ne valide pas et affiche de nouveau le même objet.

Il y a 3 séries de 7 variations aléatoires sans remise de la largeur. Lorsque ces 21 questions sont posées, le sujet est prévenu du changement de la grandeur étudiée par la phrase

« Changement de grandeur : la hauteur varie »

Une nouvelle série de 21 questions sur la hauteur s'ensuit, puis, une autre sur la profondeur. Chaque fois la question change :

« La hauteur de la sphère varie : la sphère virtuelle est-elle moins grande (-) ou plus grande (+) ? »

« La profondeur de la sphère varie : la sphère virtuelle est-elle moins profonde (-) ou plus profonde (+) ? »

Lors des tests avec la forme inconnue, il va de soi que le mot « sphère » des précédentes questions est remplacé par le mot « forme ».

3.2.4. Dispositif physique

Le dispositif de projection est un écran LUMIN d'une largeur de 3,10 m pour une hauteur de 1,74 m. Le projecteur est de type Christie Mirage 3, la distance entre l'écran et l'optique du projecteur 3.5m. Nos pixels font 1,61 mm de coté. La fréquence de projection est de 60 images par seconde. Le système de tracking est un ART2 fonctionnant également à 60 Hz d'une précision millimétrique (Figure 57).

Figure 57 : sujet en train de réaliser le test

3.3. Analyses des données

3.3.1. Collecte et prétraitement des données

Nous calculons les probabilités de réponses « plus grand » pour chaque variation et pour tous sujets confondus. Cette probabilité, comprise entre 0 et 1, est représentée sur un graphe et suit bien une normale cumulée. Sur la Figure 58 nous présentons le graphe pour l'expérience en vision ortho-stéréoscopique, point de vue traqué, avec une sphère virtuelle sur l'écran.

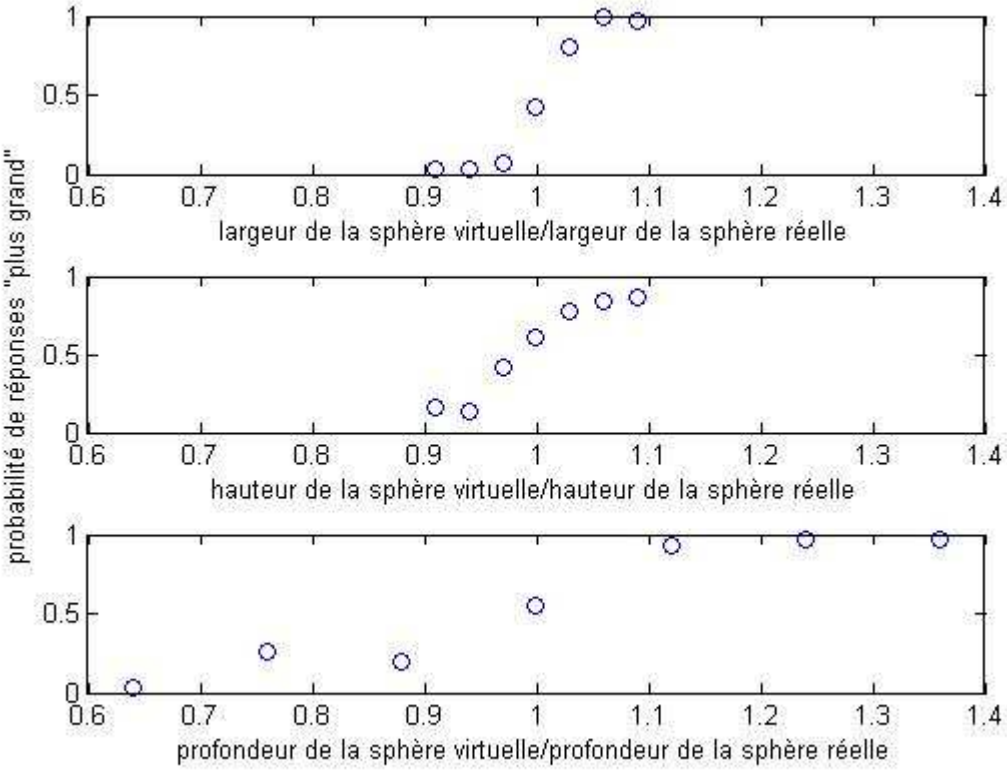


Figure 58 : exemple de collecte de données en vision ortho-stéréoscopique, point de vue traqué avec la sphère sur l'écran

A partir de ces données, nous calculons la droite à l'aide des tables des fonctions normales cumulées. Grâce à l'algorithme des moindres carrés, nous trouvons la droite de régression qui approxime notre nuage de points.

Nous savons que l'ordre de grandeur des modifications horizontales et verticales sont semblables, nous présenterons donc les variations des perceptions des formes en largeur et en hauteur ensemble, sauf lorsqu'une différence significative est présente.

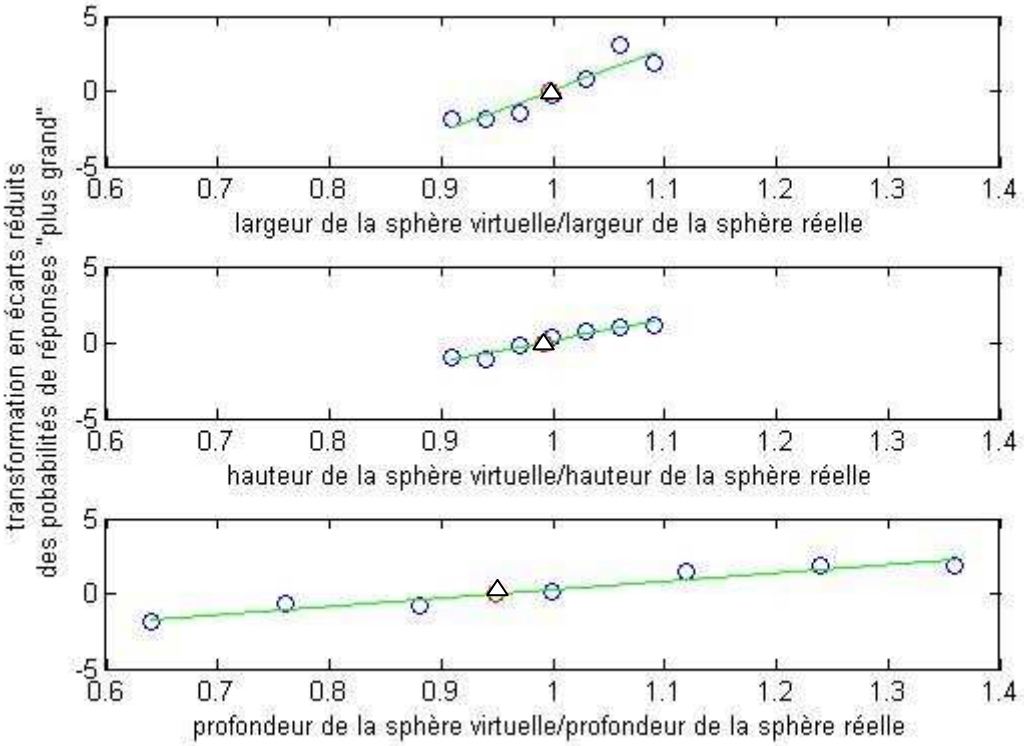


Figure 59 : transformation en écarts réduits des probabilités de réponse "plus grand" dans le cas de la vision ortho-stéréoscopique, point de vue traqué avec la sphère sur l'écran

Nous en déduisons le point d'égalisation subjective (la pente) et le JND (la racine de la droite). Sur la Figure 59, le point d'égalisation subjective est représenté (après calcul) par le triangle. Et nous pouvons retracer la courbe qui approxime le nuage de points dans l'espace des probabilités pour s'assurer visuellement que cela approxime bien nos données de départ (voir Figure 60).

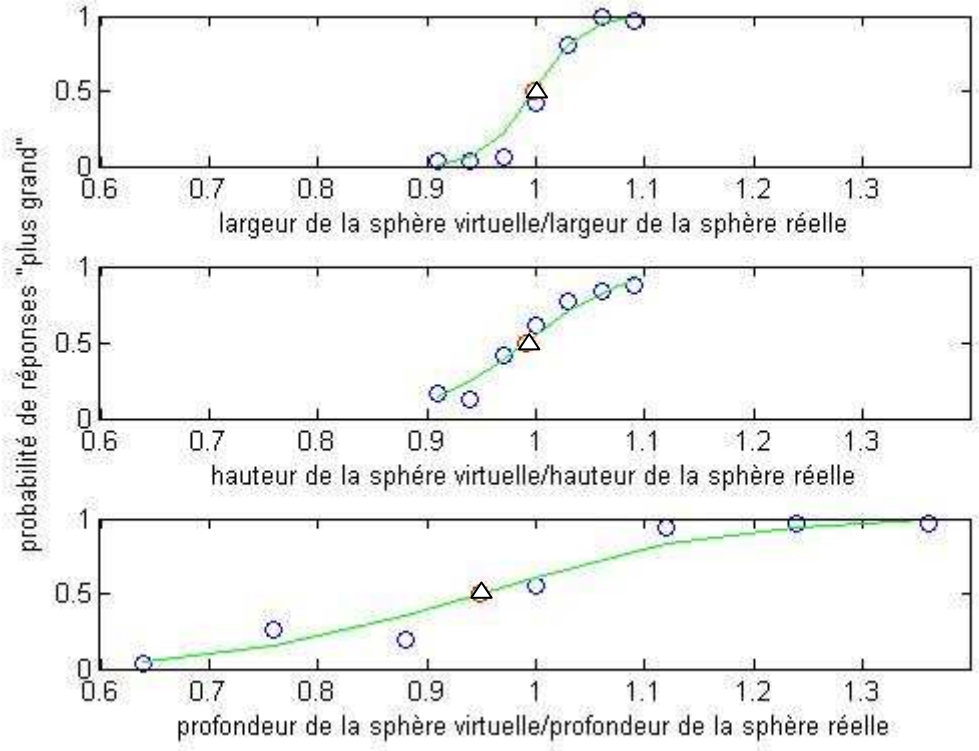


Figure 60 : probabilité de réponse "plus grand" pour la vision ortho-stéréoscopique, point de vue traqué avec la sphère sur l'écran et son lissage

A coté de chaque graphique, il y aura un petit tableau qui résume les conditions de l'expérience.

tracking	oui/non
projection	mono
forme	sphère
position	sur

Pour gagner de la place, nous utilisons des abréviations :

- Tracking : point de vue traqué ?
  - Oui : le point de vue est traqué
  - Non : le point de vue n'est pas traqué
- Projection : Est-ce que la distance entre les caméras virtuelles correspond à la distance inter-pupillaire ?
  - Mono : vision monoscopique
  - Ortho : vision ortho-stéréoscopique
  - Non-ortho : vision non-ortho-stéréoscopique
- Forme : type de forme utilisé.
  - Sphère : le test a été réalisé avec la sphère
  - Forme : le test a été réalisé avec la forme non connue a priori
- Position : la position de la forme par rapport à l'écran.

3.3.2. Hypothèse n° 1 : Il est préférable d'avoir le point de vue traqué en vision monoscopique

Pour tous les graphes qui suivent, pour chaque cas considéré, nous représentons le point d'égalisation subjective et le JND de part et d'autre. Plus le JND est grand, plus l'incertitude des sujets est grande.

3.3.2.1. Lorsque la sphère est sur l'écran

Nous commençons par étudier la perception des formes en profondeur (donc la variation de la courbure en profondeur) lorsque la sphère virtuelle est positionnée sur l'écran. Nous avons exposé les résultats de cette expérience sur la Figure 61. Le JND et le PSE, lorsque la tête du sujet est traquée, sont respectivement 18,63% et 86,29% du diamètre de la sphère réelle ; ceux lorsque le sujet et les caméras virtuelles sont fixes, sont respectivement 27,65% et 101,19%. Nous pouvons donc calculer la significativité des différences de JND qui sont, rappelons-le, assimilés aux écart-type des distributions. La population étant normale, nous pouvons utiliser la formule 3-11.

$$T = \frac{\sigma_{\max}^2}{\sigma_{\min}^2} = \frac{0{,}2765^2}{0{,}1853^2} = 3{,}3709$$

Ceci nous donne une probabilité de significativité de 99,76%. Elle est supérieure à 95%, nous considérons donc cet écart comme significatif. Nous pouvons également comparer la significativité de l'écart des moyennes avec la formule 3-5 puisque les populations sont normales :

$$T = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1}{n_1} + \frac{\sigma_2}{n_2}}} = \frac{1{,}0119 - 0{,}8629}{\sqrt{\frac{0{,}2765}{15} + \frac{0{,}1863}{15}}} = 0{,}84 \approx N(0{,}1)$$

Ce qui fait une probabilité de significativité de 80% et n'est donc pas suffisant pour être considéré comme significatif.

tracking	oui/non
projection	mono
forme	sphère
position	sur

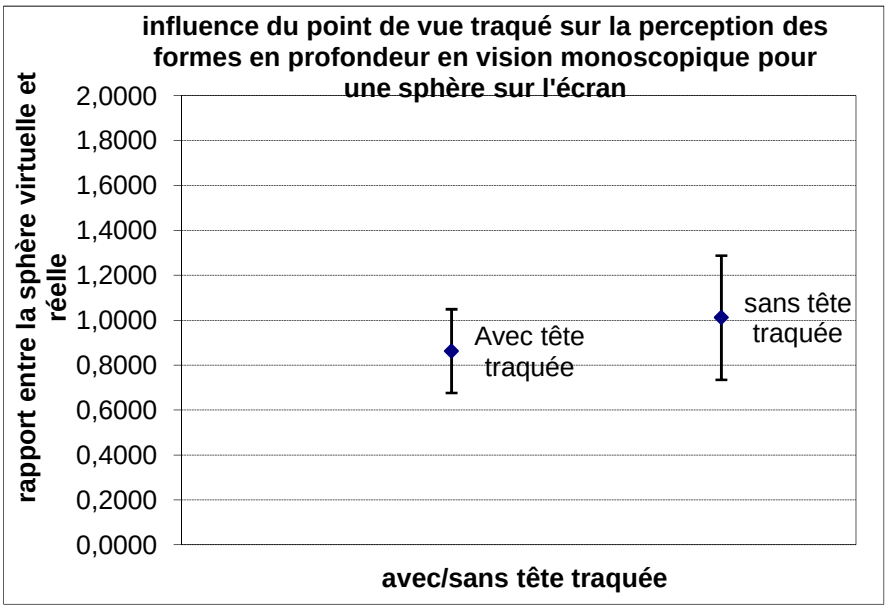


Figure 61 : Influence du point de vue traqué sur la perception des formes en profondeur en vision monoscopique pour une sphère sur l'écran

Nous voyons que le JND est plus important lorsque le sujet et les caméras virtuelles sont fixes. Cela signifie qu'un sujet détecte plus facilement une petite variation de courbure en profondeur lorsque sa tête est traquée et que les caméras virtuelles sont situées à l'endroit de sa tête.

Le diamètre de la sphère est 20cm, nous en déduisons donc que si le JND avec point de vue traqué est égal à 18%, il faudra 18%*20cm = 3.6cm de variation de profondeur pour que celle-ci soit perçue. Autrement dit, une « sphère » dont le diamètre en profondeur est compris entre 16.4cm et 23.6cm sera toujours vue comme une sphère (nous avons centré sur le diamètre étalon puisqu'il n'y a pas de différence significative des PSE). De même, lorsque le point de vue n'est pas traqué, il faudra 27%*20cm=5.4cm de variation en profondeur pour que celle-ci soit perçue. Si une « sphère » a un diamètre en profondeur compris entre 14.6cm et 25.4cm, elle sera toujours perçue comme une sphère.

Nous pouvons également nous intéresser aux variations de formes fronto-parallèles. Les résultats sont exposés sur la Figure 62. Les JND et PSE lorsque le point de vue est traqué sont 3.5% et 99% du diamètre de la sphère étalon. Lorsque le point de vue n'est pas traqué, ils sont égaux à 4% et 99%. La probabilité de significativité de l'écart des JND est 84%, la significativité de l'écart des PSE est 50%. La différence entre la variation entre la largeur et la hauteur a une probabilité de significativité de 98%. La légère amélioration due à la tête traquée n'est donc pas significative.

tracking	oui/non
projection	mono
forme	sphère
position	sur

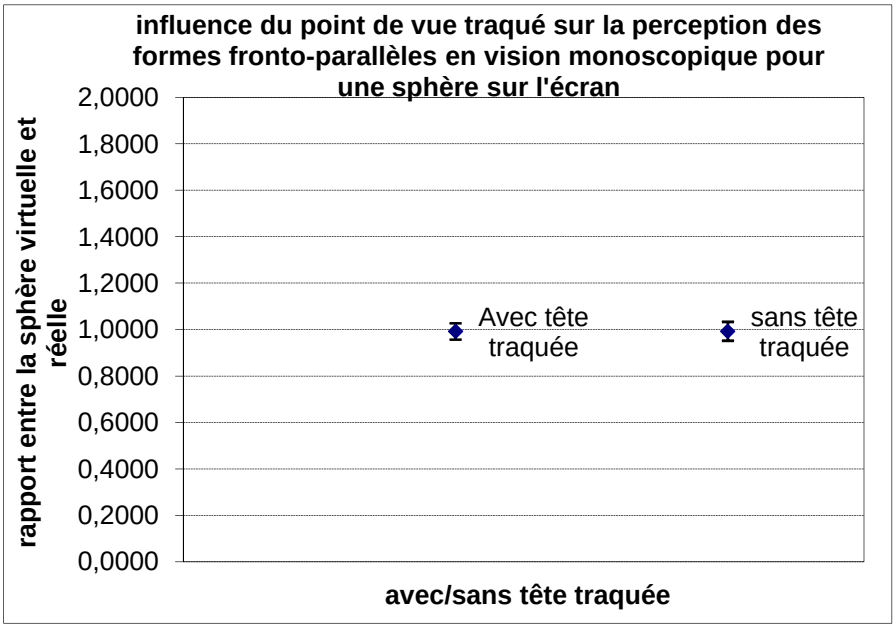


Figure 62 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique avec une sphère sur l'écran

Puisque la probabilité de significativité des différences entre la perception des courbures en largeur et en hauteur indique qu'elles sont différentes, nous présentons le graphe détaillé de ces différences sur la Figure 63. Cependant, aucun des écarts n'est significatif.

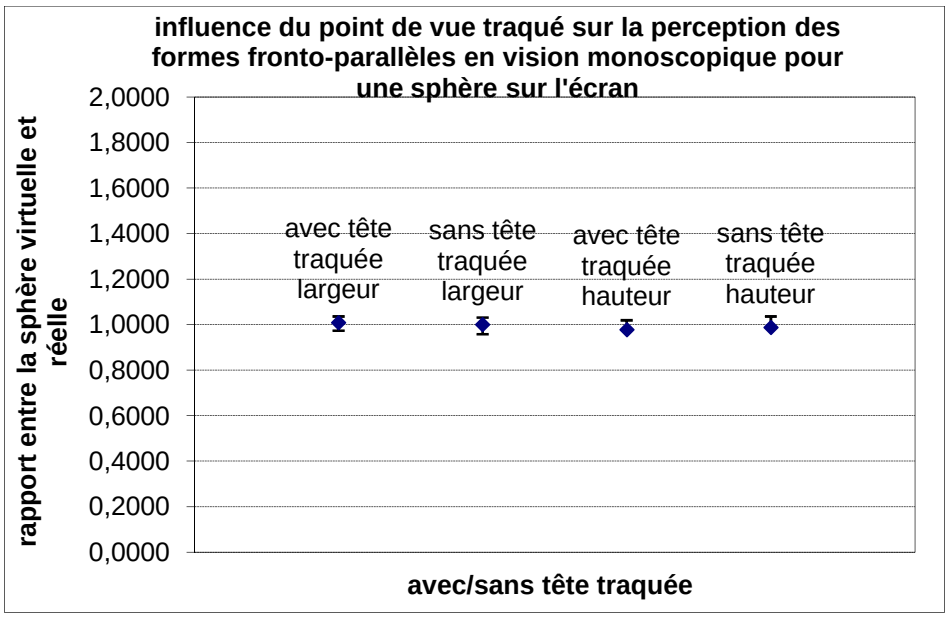


Figure 63 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique pour une sphère sur l'écran, détail

Il faudra donc une déformation de 4%*20cm=0.8cm pour percevoir la déformation de la sphère, avec ou sans point de vue traqué (puisque la différence de JND n'est pas significative).

3.3.2.2. Lorsque la sphère est derrière l'écran

Les résultats pour la perception des formes en profondeur sont exposés sur la Figure 64. La probabilité de significativité pour l'écart des JND est 99,9%. Cette différence est donc significative. La probabilité de significativité de l'écart des PSE est 63%, nous ne pouvons donc pas considérer cette différence comme significative. Les JND et PSE, lorsque la tête est traquée, sont respectivement 35% et 90% du diamètre de la sphère réelle. De même, les JND et PSE lorsque le sujet et les caméras sont fixes sont respectivement 90% et 100% du diamètre de la sphère réelle. Nous voyons donc qu'une modification de courbure en profondeur doit être beaucoup plus importante pour être perçue lorsque les caméras virtuelles ne suivent pas le point de vue du sujet.

tracking	oui/non
projection	mono
forme	sphère
position	derrière

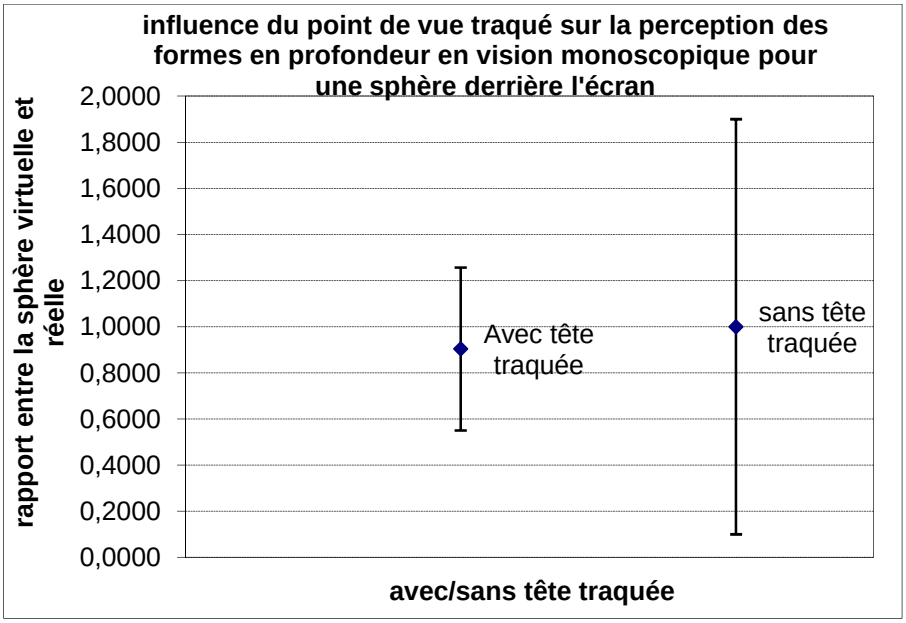


Figure 64 : influence du point de vue traqué sur la perception des formes en profondeur en vision monoscopique pour une sphère derrière l'écran

Nous pouvons faire la même analyse pour la perception de formes fronto-parallèles. Les résultats sont exposés sur la Figure 65. La probabilité de significativité des écarts entre les PSE, JND et entre la largeur et la hauteur sont respectivement 52%, 80% et 79%. Aucun de ces écarts n'est donc significatif. Le fait de traquer le point de vue du sujet n'influe donc pas sur la perception des formes fronto-parallèles pour des formes situées derrière l'écran.

tracking	oui/non
projection	mono
forme	sphère
position	derrière

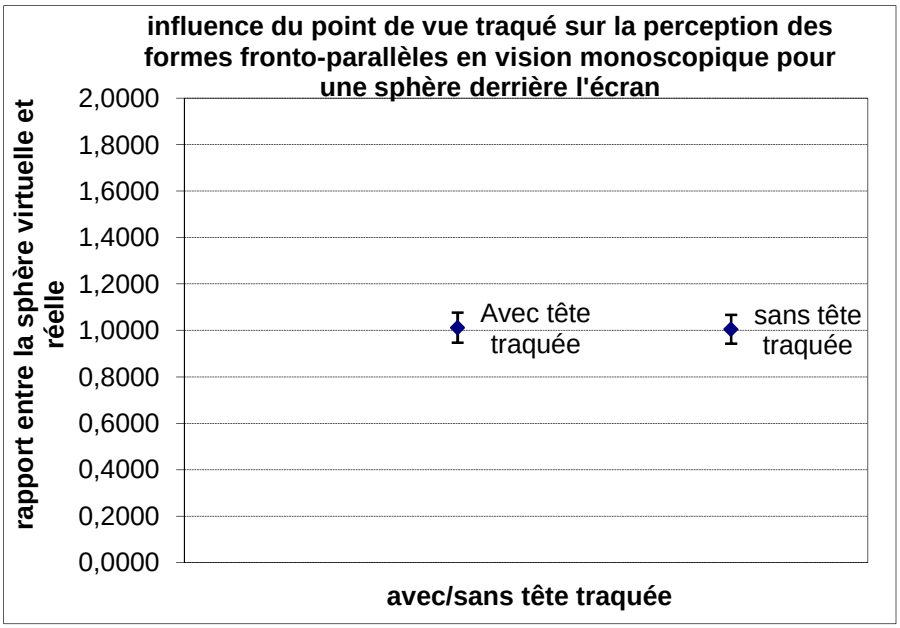


Figure 65 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique pour une sphère derrière l'écran

3.3.2.3. Lorsque la sphère est devant l'écran

Les résultats pour la perception des formes en profondeur sont présentés sur la Figure 66. Les PSE et JND lorsque le point de vue est traqué sont respectivement 86% et 35% du diamètre de la sphère réelle. Lorsque le point de vue n'est pas traqué, ils sont respectivement égaux à 96% et 100%.

La probabilité de significativité de l'écart des JND est 99,9% cette différence est donc significative. Par contre, la probabilité de significativité de la différence de PSE est 63% ; ce qui est insuffisant pour que cette différence soit significative. Nous voyons sur cette figure qu'une petite modification de la courbure en profondeur sera relativement mieux perçue lorsque le point de vue est traqué. Par contre, lorsque le point de vue du sujet n'est pas traqué, il faut une énorme modification de cette courbure pour qu'elle soit perçue.

tracking	oui/non
projection	mono
forme	sphère
position	devant

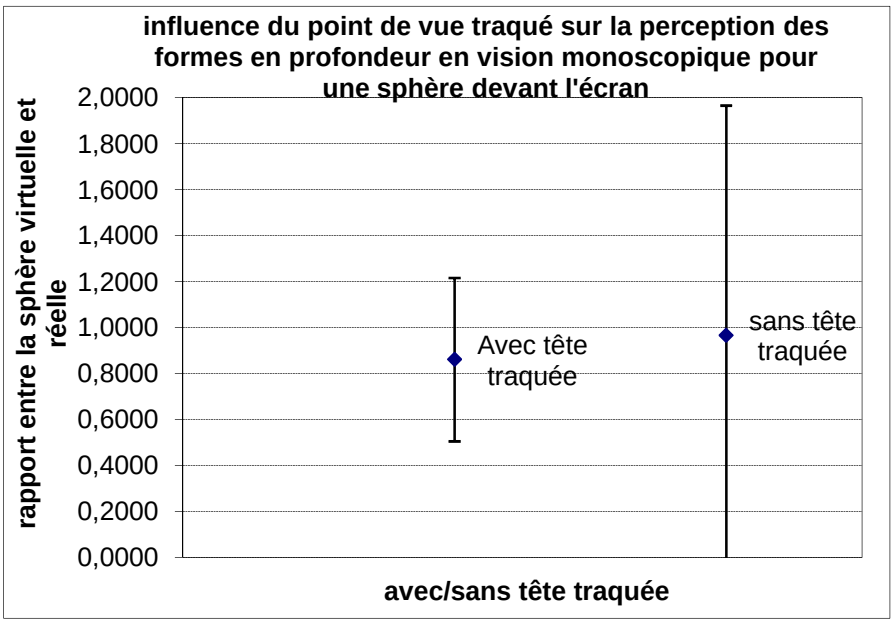


Figure 66 : influence du point de vue traqué sur la perception des formes en profondeur en vision monoscopique pour une sphère devant l'écran

Nous pouvons donner les résultats pour la perception des formes fronto-parallèles. Ceux-ci sont visibles sur la Figure 67. La probabilité de significativité des écarts entre les PSE, les JND et la différence entre la perception de courbures horizontales et verticales sont respectivement 51%, 98% et 80%. Nous voyons donc que lorsque la sphère est devant l'écran, le point de vue traqué a une incidence sur les JND de la perception des formes fronto-parallèles. Cependant, la détection de changement de forme est légèrement plus fine lorsque la tête est immobile.

tracking	oui/non
projection	mono
forme	sphère
position	devant

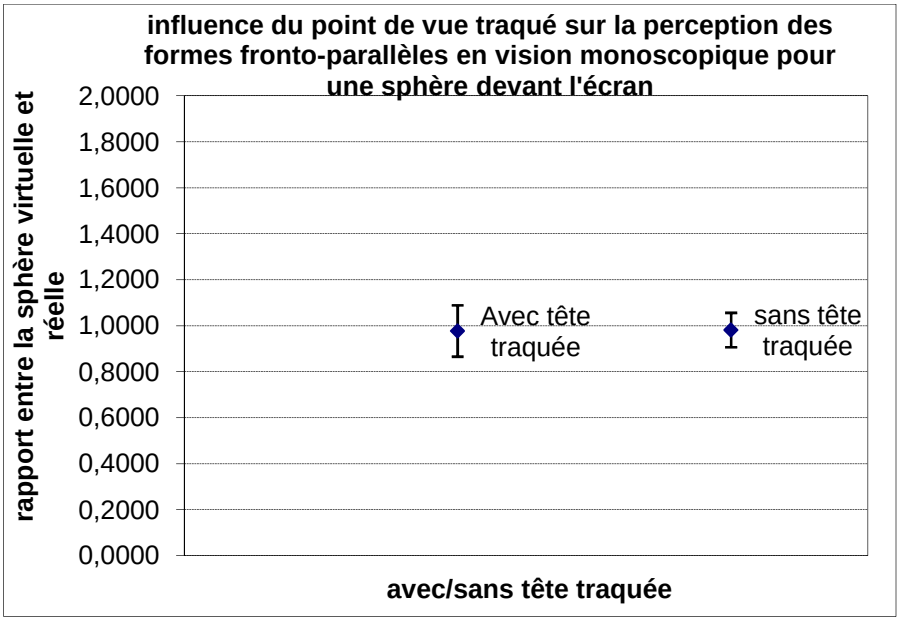


Figure 67 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique pour une sphère devant l'écran

3.3.2.4. Lorsque la forme non connue est sur l'écran

Les résultats de l'expérience sur la perception des courbures en profondeur sont présentés sur la Figure 68. Les JND et PSE avec point de vue traqué sont respectivement 25% et 97% du diamètre initial de la forme. Sans point de vue traqué les PSE et JND sont égaux à 100% et 113% du diamètre initial. La probabilité de significativité de l'écart entre les JND est de 100% et donc hautement significative, tandis que celle de l'écart entre les PSE est de 71% et n'est donc pas significative. Nous voyons qu'une modification de la courbure en profondeur sera plus finement perçue lorsque le point de vue de l'utilisateur est traqué.

tracking	oui/non
projection	mono
forme	Forme
position	sur

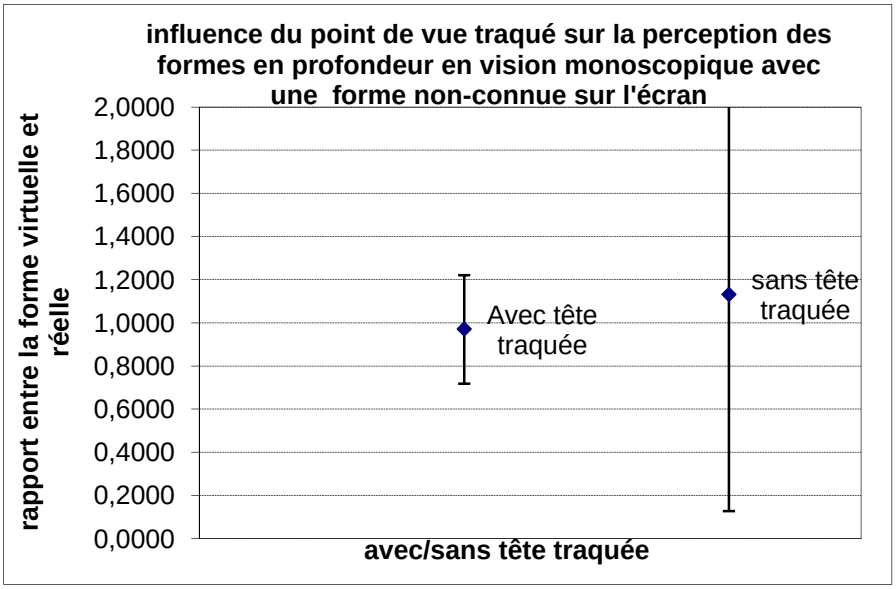


Figure 68 : influence du point de vue traqué sur la perception des formes en profondeur en vision monoscopique avec une forme non-connue sur l'écran

Nous pouvons refaire cette expérience pour les formes fronto-parallèles. Les résultats sont exposés sur la Figure 69. Les PSE et JND avec le point de vue traqué sont égaux à 100% et 6,3% de la

dimension étalon, sans point de vue traqué, ils sont égaux à 101% et 4,8%. La probabilité de significativité de l'écart entre les PSE et JND sont respectivement de 52% et 93%. Ni l'un ni l'autre de ces écarts ne sont significatifs.

tracking	oui/non
projection	mono
forme	Forme
position	sur

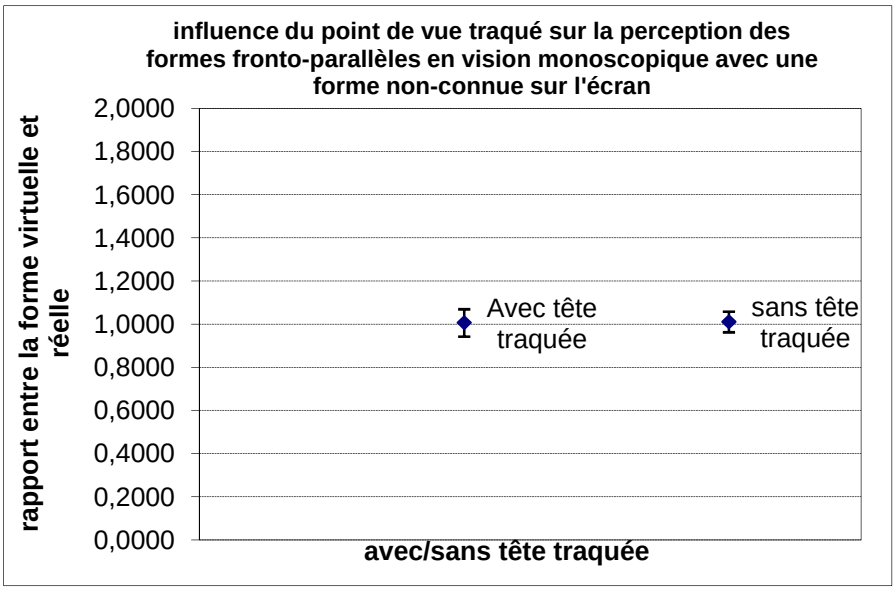


Figure 69 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique avec une forme non-connue sur l'écran

3.3.2.5. Lorsque la forme non connue est derrière l'écran

Les résultats de l'expérience sur la perception des courbures en profondeur sont exposés sur la Figure 70. Le PSE et JND avec point de vue traqué sont respectivement égaux à 84% et 64% du diamètre initial de la forme. Sans point de vue traqué, les PSE et JND sont 91% et 108%. La probabilité de significativité de l'écart entre les PSE est 59% ; ce n'est donc pas significatif. Par contre, la probabilité de significativité de l'écart des JND est 99% ; cet écart-ci est significatif. Une modification de la courbure sera donc mieux perçue lorsque le point de vue est traqué.

tracking	oui/non
projection	mono
forme	Forme
position	derrière

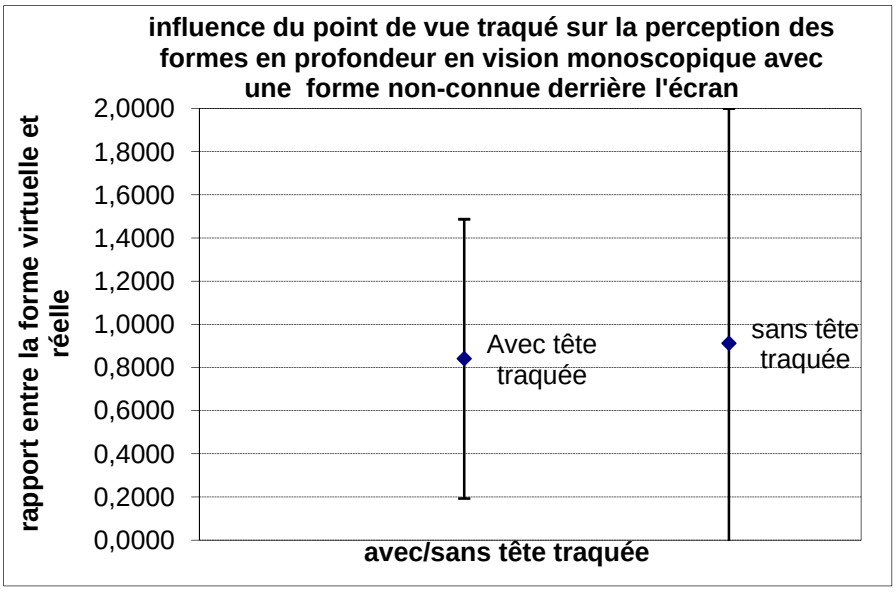


Figure 70 : influence du point de vue traqué sur la perception des formes en profondeur en vision monoscopique avec une forme non-connue derrière l'écran

Nous pouvons faire de même avec les formes fronto-parallèles. Le résultat est sur la Figure 71. Les PSE et JND avec le point de vue traqué sont 100% et 7% tandis que ceux sans le point de vue traqué sont respectivement 99% et 9,4% de la dimension initiale de la forme. Les probabilités de significativité des écarts entre les PPSE et les JND sont 54% et 94%, aucun n'est donc significatif, même si la différence entre les JND s'approche très près du seuil de significativité. La probabilité de différence de perception des courbures horizontales et verticales s'élève à 99%, cette différence est donc significative.

tracking	oui/non
projection	mono
forme	Forme
position	derrière

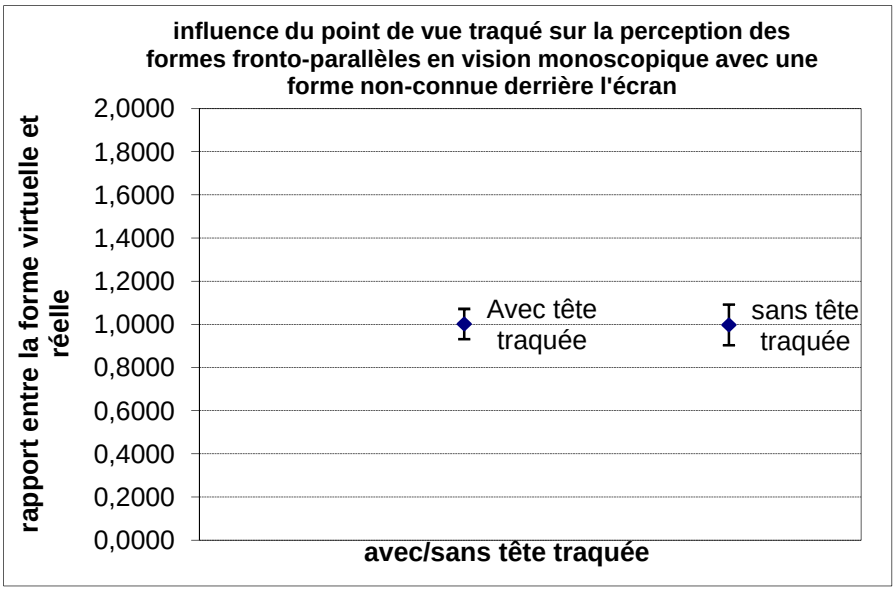


Figure 71 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique avec une forme non-connue derrière l'écran

La différence entre la perception de la largeur et de hauteur étant significative, nous détaillons chacune d'entre elle (Figure 72). Les PSE et JND en largeur avec point de vue traqué sont 100% et 6% ; sans point de vue traqué, ils sont également égaux à 100% et 6%. Les PSE et JND en hauteur avec point de vue traqué sont égaux à 99% et 7% tandis que sans point de vue traqué, ils sont respectivement égaux à 99% et 13%. Les probabilités de significativité des écarts de PSE sont toutes négligeables, par contre, la probabilité de significativité de l'écart de JND en hauteur est 99%, cet écart est donc loin d'être insignifiant. Les modifications de la courbure verticale sont plus facilement perçues lorsqu'il n'y a pas de point de vue traqué.

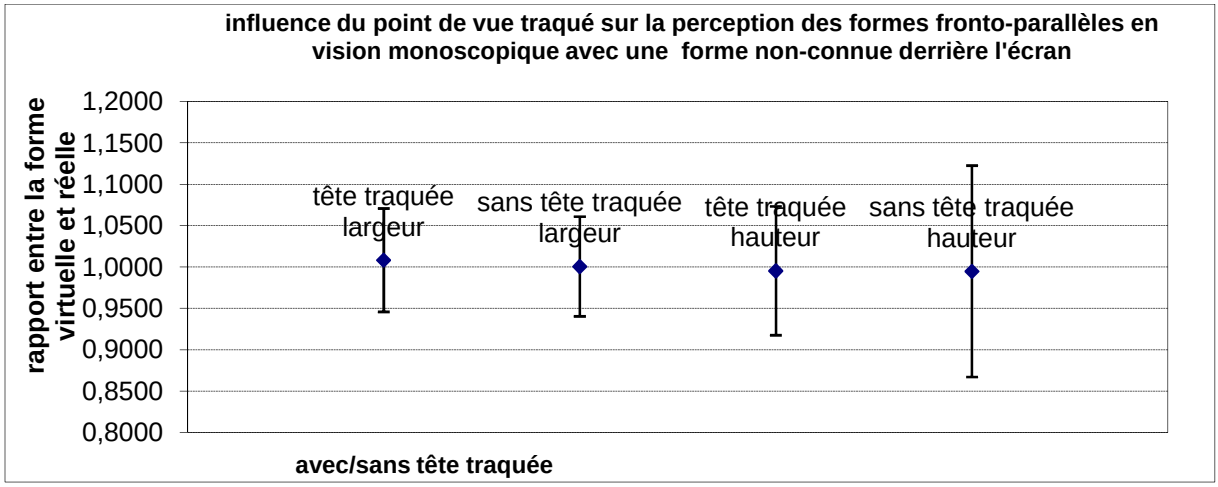


Figure 72 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique avec une forme non-connue derrière l'écran, détails

3.3.2.6. Lorsque la forme non connue est devant l'écran

Les résultats de l'expérience sur la perception des formes en profondeur sont présentés sur la Figure 73. Les PSE et JND avec point de vue traqué sont 86% et 37%, tandis que sans point de vue traqué ils sont égaux à 52% et 153% de la dimension étalon. La probabilité de significativité de l'écart entre les PSE est 82%, donc insignifiante (probablement à cause des très grands JND), par contre, la probabilité de significativité de l'écart des JND est 100%, comme nous pouvions nous y attendre en regardant les valeurs. Traquer le point de vue des sujets est très important pour la perception des formes en profondeur lorsque la forme est devant l'écran en vision monoscopique.

tracking	oui/non
projection	mono
forme	Forme
position	devant

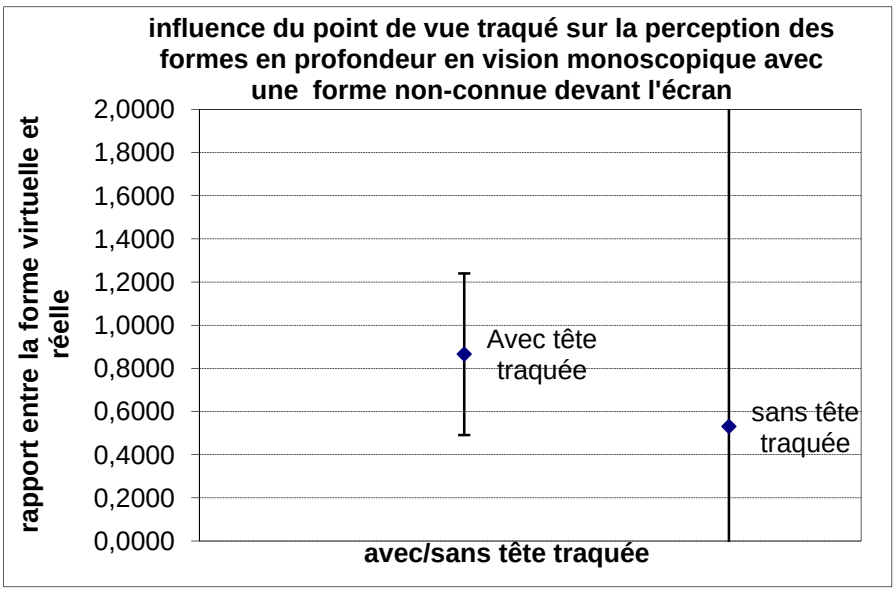


Figure 73 : influence du point de vue traqué sur la perception des formes en profondeur en vision monoscopique avec une forme non-connue devant l'écran

Faisons de même avec les formes fronto-parallèles, les résultats en sont présentés sur la Figure 74. Les PSE et JND lorsque le point de vue n'est pas traqué sont égaux à 99% et 6%, de même, lorsque le point de vue est traqué ils sont également de 99% et 6% des dimensions de la forme étalon. Il n'y a donc aucune différence avec ou sans point de vue traqué pour les formes fronto-

parallèles. Les probabilités de significativité des écarts entre les PSE et JND sont 58% et 78%, donc toutes les deux insignifiantes. De même la probabilité de significativité de différence entre la perception en largeur et en hauteur est 91% donc insignifiante également.

tracking	oui/non
projection	mono
forme	Forme
position	devant

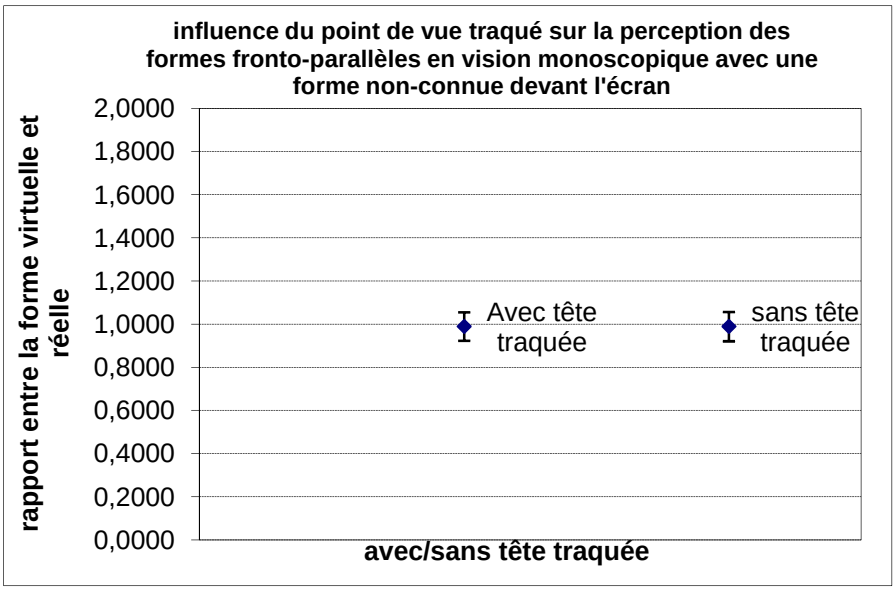


Figure 74 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique avec une forme non-connue devant l'écran

3.3.2.7. Conclusion sur l'hypothèse n°1

Nous avons donc vu tout au long des paragraphes précédents que le point de vue traqué est extrêmement important pour la perception des formes en profondeur en vision monoscopique. Nous constatons une amélioration de **40%** (derrière l'écran 55%, sur l'écran 8% et devant l'écran 65%) du JND pour la sphère et **80%** (derrière 44%, sur 75% et devant 116%) pour une forme non connue. Ces améliorations sont des moyennes calculées sur toutes les positions des formes. Nous constatons que lorsque celles-ci sont placées devant l'écran, les améliorations sont plus fortes, probablement parce que le sujet peut tourner plus aisément et techniquement autour de l'objet. Cela signifie que la différence de modification nécessaire pour être perçue est 40%*20cm=8cm plus grande lorsque le point de vue n'est pas traqué pour la sphère et 16cm pour une forme non connue. Il faut une modification de 30% (dans notre cas 6cm) du diamètre de la sphère en profondeur pour qu'une modification avec tête traquée soit perçue et 70% (14cm) lorsque le point de vue n'est pas traqué. Pour la forme non connue a priori, il faut 53% (10cm) de modification du diamètre pour qu'elle soit perçue sans point de vue traqué, et 119% (autrement dit, il n'y a presque aucune modification perçue).

Les JND pour la perception des formes fronto-parallèles ne sont que très peu modifiés avec point de vue traqué. Le point d'égalisation subjective ne varie pas en fonction du point de vue traqué. Aucun n'écart n'était significatif, que ce soit pour la perception des courbures en profondeur ou fronto-parallèles.

3.3.3. Hypothèse n° 2 : Il est préférable d'avoir le point de vue traqué en vision ortho-stéréoscopique

3.3.3.1. Lorsque la sphère est sur l'écran

Les résultats de l'expérience sur la perception des courbures en profondeur sont exposés sur la Figure 75. Lorsque le point de vue est traqué, les PSE et JND sont respectivement 95% et 18% du diamètre de la sphère et 106% et 52% lorsque les caméras virtuelles sont fixes. La probabilité de significativité de l'écart des JND est 99%, cet écart est donc significatif. Par contre, la probabilité de significativité des écarts de PSE étant de 94%, ce n'est pas significatif. Nous voyons donc que les modifications de la courbure en profondeur se voient plus facilement lorsque le point de vue est traqué que lorsque celui ne l'est pas.

tracking	oui/non
projection	ortho
forme	Sphère
position	sur

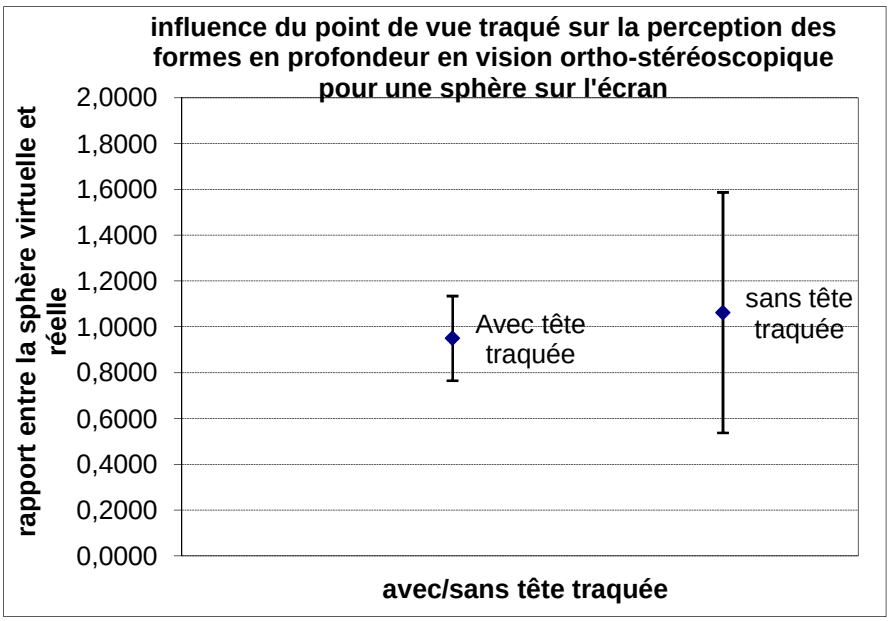


Figure 75 : influence du point de vue traqué sur la perception des formes en profondeur en vision ortho-stéréoscopique pour une sphère sur l'écran

Il n'y a pas de différences significatives pour la perception des formes fronto-parallèles, donc nous plaçons ces résultats en annexe A.

3.3.3.2. Lorsque la sphère est derrière l'écran

Les résultats de l'expérience sur la perception des courbures en profondeur sont présentés sur la Figure 76. Les PSE et JND lorsque le point de vue est traqué sont respectivement de 93% et 17% du diamètre de la sphère. Tandis que lorsque le point de vue n'est pas traqué, ils sont égaux à 105% et 86%. La probabilité de significativité de l'écart des JND est égale à 100%, c'est donc significatif. La probabilité de significativité des écarts de PSE est égale à 67%. Nous voyons donc que lorsque la sphère est derrière l'écran, une modification de courbure en profondeur sera mieux perçue lorsque le point de vue est traqué que lorsque celui-ci ne l'est pas.

tracking	oui/non
projection	ortho
forme	Sphère
position	derrière

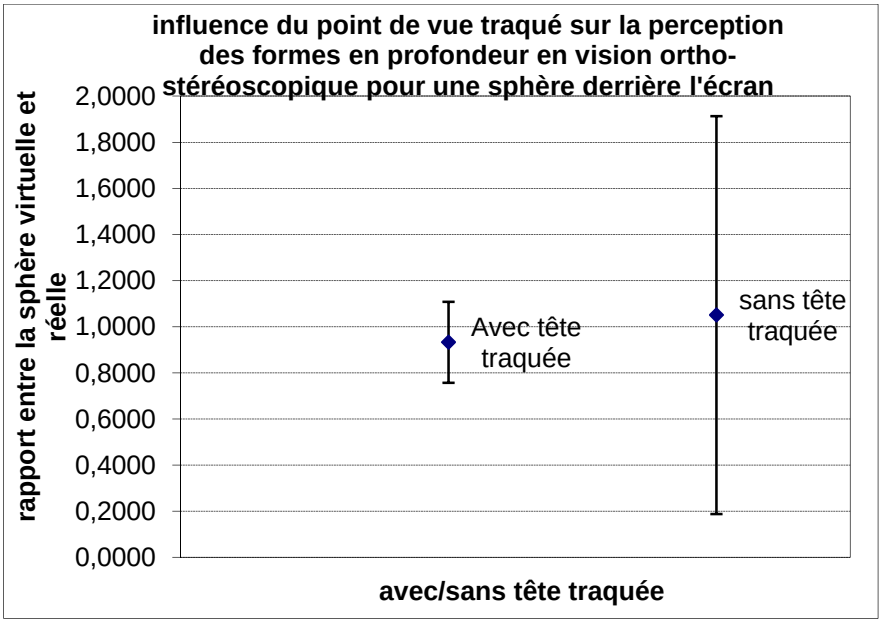


Figure 76 : influence du point de vue traqué sur la perception des formes en profondeur en vision stéréoscopique pour une sphère derrière l'écran

Les formes fronto-parallèles sont exposées en annexe A.

3.3.3.3. Lorsque la sphère est devant l'écran

Les résultats de l'expérience sur la perception des courbures en profondeur sont exposés sur la Figure 77. Les PSE et JND lorsque le point de vue est traqué sont respectivement 100% et 23% du diamètre de la sphère réelle. Tandis que lorsque le point de vue n'est pas traqué, ils sont égaux à 99% et 25%. Nous voyons donc que, cette fois-ci, la différence entre les JND n'est pas très grande. D'ailleurs, la probabilité de significativité de cet écart est 81%, il n'est donc pas significatif. La probabilité associée à la significativité de l'écart des PSE est 52% et n'est pas non plus significative.

Le fait que la pixellisation soit plus fine pour une sphère positionnée devant l'écran explique probablement cette non-significativité. Les disparités représentées peuvent être suffisamment fines pour pouvoir se suffire à elles-mêmes.

tracking	oui/non
projection	ortho
forme	Sphère
position	devant

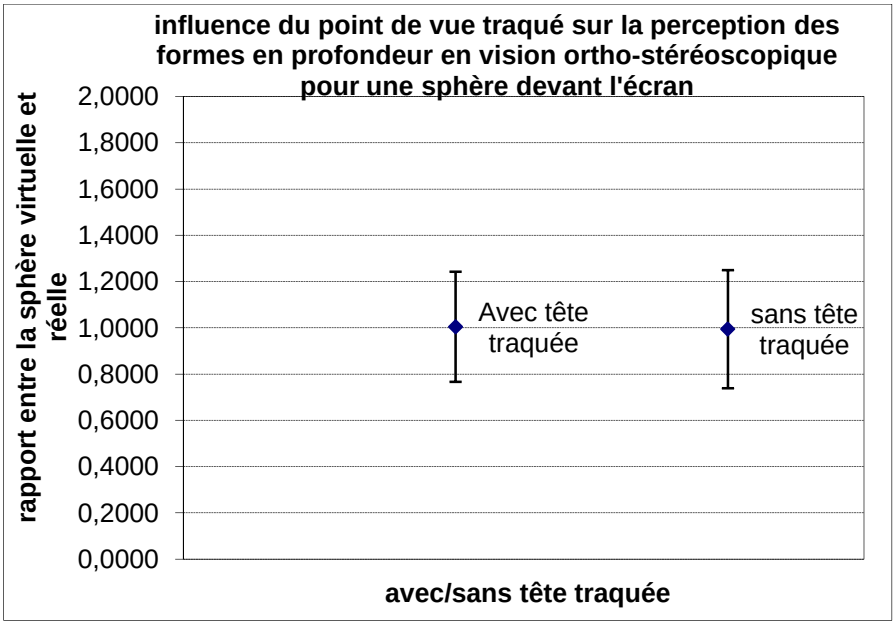


Figure 77 : influence du point de vue traqué sur la perception des formes en profondeur en vision stéréoscopique pour une sphère devant l'écran

Les différences de perception des formes fronto-parallèles n'étant pas significatives, les résultats sont placés en annexe A.

3.3.3.4. Lorsque la forme non connue est sur l'écran

Les résultats de l'expérience sur la perception des formes en profondeur se trouvent sur la Figure 78. Les PSE et JND lorsque le point de vue est traqué sont respectivement 94% et 18% de la dimension initiale de la forme, tandis que sans point de vue traqué ils sont égaux à 90% et 52%. Les probabilités de significativité de l'écart de PSE est 59% donc non-significatif. Par contre, celle de l'écart des JND est 100%, donc significative. Il est important de traquer le point de vue du sujet pour la perception des formes en profondeur lorsque le sujet a une vision ortho-stéréoscopique et que l'objet est placé sur l'écran.

tracking	oui/non
projection	ortho
forme	Forme
position	sur

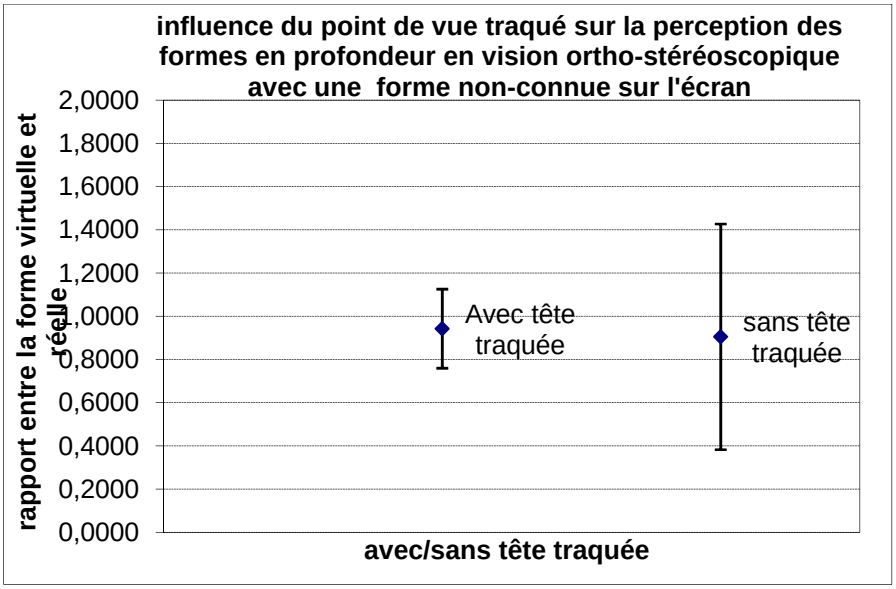


Figure 78 : influence du point de vue traqué sur la perception des formes en profondeur en vision ortho-stéréoscopique avec une forme non-connue sur l'écran

Les résultats pour la perception des formes fronto-parallèles sont en annexe A.

3.3.3.5. Lorsque la forme non connue est derrière l'écran

Nous présentons les résultats de l'expérience sur la perception des courbures en profondeur sur la Figure 79. Les PSE et JND avec point de vue traqué sont 113% et 30% des dimensions de la forme réelle tandis que sans point de vue traqué, ils s'élèvent à 129% et 123%. Nous pouvons dire que sans le point de vue traqué les sujets n'arrivent pas à déterminer la moindre modification de forme en profondeur. Les probabilités de significativité des écarts entre les PSE et les JND sont respectivement 71% et 100%. Donc l'écart de JND est significatif. Le point de vue traqué est important pour une meilleure perception des formes en profondeur en vision ortho-stéréoscopique pour un objet non-connu derrière l'écran.

tracking	oui/non
projection	ortho
forme	Forme
position	derrière

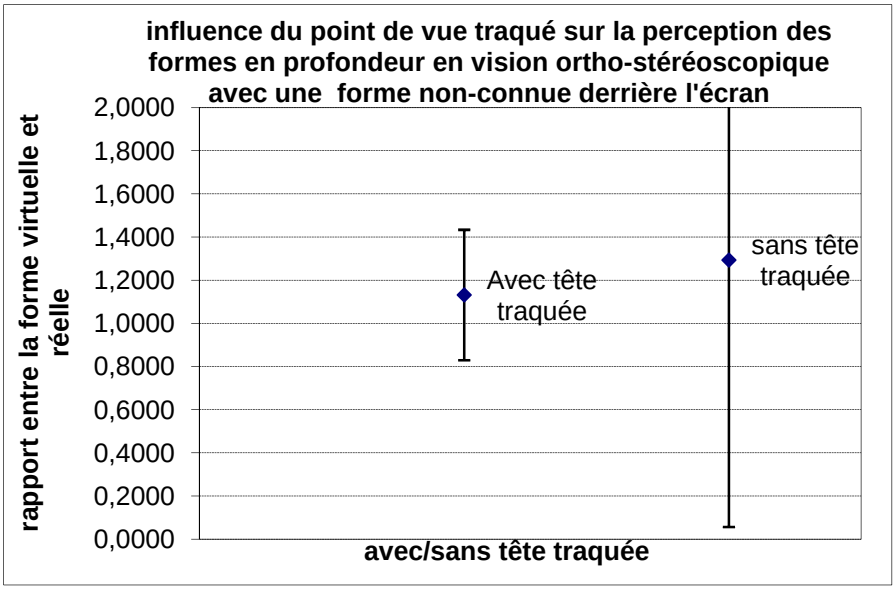


Figure 79 : influence du point de vue traqué sur la perception des formes en profondeur en vision ortho-stéréoscopique avec une forme non-connue derrière l'écran

Les différences de perception des formes fronto-parallèles n'étant pas significatives, nous les plaçons en annexe A.

3.3.3.6. Lorsque la forme non connue est devant l'écran

Les résultats de l'expérience sur la perception des formes en profondeur sont présentés sur la Figure 80. Les PSE et JND avec le point de vue traqué sont égaux à 100% et 20% des dimensions de la forme étalon. Sans point de vue traqué, ils sont égaux à 96% et 36%. Les probabilités de significativité des écarts de PSE et de JND sont 57% et 99%. Donc, l'écart entre les JND est significatif. Il vaut mieux traquer le point de vue du sujet si la perception des formes en profondeur est importante.

tracking	oui/non
projection	ortho
forme	Forme
position	devant

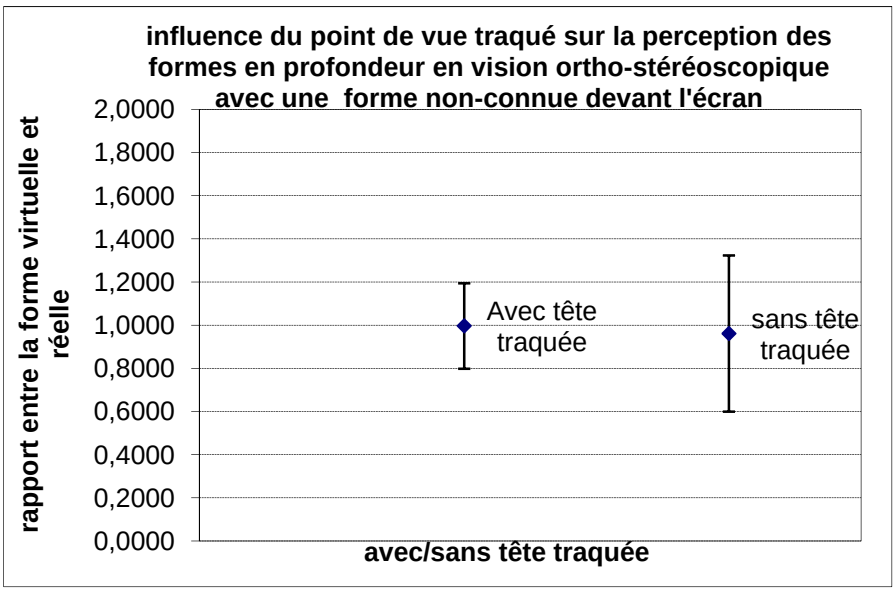


Figure 80 : influence du point de vue traqué sur la perception des formes en profondeur en vision ortho-stéréoscopqiue avec une forme non-connue devant l'écran

Les différences de perception des formes fronto-parallèles n'étant pas significatives, nous avons placé les résultats en annexe A.

3.3.3.7. Conclusion sur l'hypothèse n°2

Nous avons vu au cours des paragraphes précédents que le point de vue traqué est très important pour la perception des formes en profondeur en vision ortho-stéréoscopique. L'amélioration des JND pour la sphère est 35% (7cm) en moyenne (69% derrière, 34% sur et 2% devant l'écran) et 52% (10,4cm) pour la forme non connue a priori (93% derrière, 48% sur et 16% devant). Contrairement à ce que nous avions obtenu en vision monoscopique, l'amélioration lorsque la forme est devant l'écran est bien moindre que lorsque celle-ci se trouve derrière l'écran. Il est fort probable qu'avec l'amélioration de la pixellisation, les disparités horizontales ressortent bien mieux, et donc les sujets ont moins besoin des parallaxes de mouvement.

Lorsque le point de vue est traqué, une modification de 16% du diamètre de la sphère (3cm) peut être perçue et 17% pour une forme non connue a priori (3 cm). Lorsque le point de vue n'est pas traqué, une modification de 53% du diamètre en profondeur (10cm) est perceptible pour une sphère et 73% (15cm) pour une forme non-connue à priori.

Il n'y a pas de véritable amélioration du point de vue traqué sur la perception des formes fronto-parallèles, ni du point d'égalisation subjective.

3.3.4. Hypothèse n° 3 : Il est préférable d'avoir le point de vue traqué en vision non-ortho-stéréoscopique

3.3.4.1. Lorsque la sphère est sur l'écran

Les résultats de l'expérience sur la perception des courbures en profondeur sont exposés sur la Figure 81. Les JND et PSE lorsque le point de vue est traqué sont respectivement 15% et 90% du diamètre de la sphère réelle. Lorsque celui-ci n'est pas traqué, ils sont égaux à 90% et 84% du diamètre de cette dernière. Les probabilités de significativité des écarts entre les PSE et JND sont 58% et 100%. Nous voyons donc qu'il faudra une plus grande modification de la courbure en profondeur lorsque le point de vue n'est pas traqué.

tracking	oui/non
projection	Non-ortho
forme	Sphère
position	sur

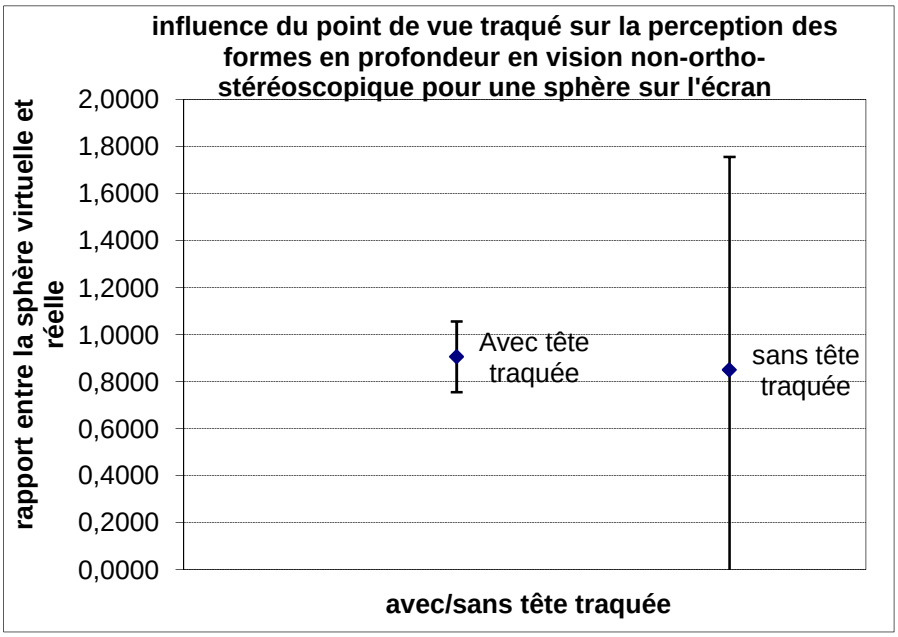


Figure 81 : influence du point de vue traqué sur la perception des formes en profondeur en vision non-ortho-stéréoscopique pour une sphère sur l'écran

Si nous nous intéressons à la perception des formes fronto-parallèles, nous obtenons le graphe exposé sur la Figure 82. La probabilité de significativité des écarts entre les JND, PSE et de la différence entre la perception des courbures en largeur et en hauteur sont respectivement 96%, 52% et 83%. Nous le voyons donc, lorsque la sphère est sur l'écran, le point de vue traqué peut aider à percevoir de façon plus subtile les variations de courbures fronto-parallèles.

tracking	oui/non
projection	Non-ortho
forme	Sphère
position	sur

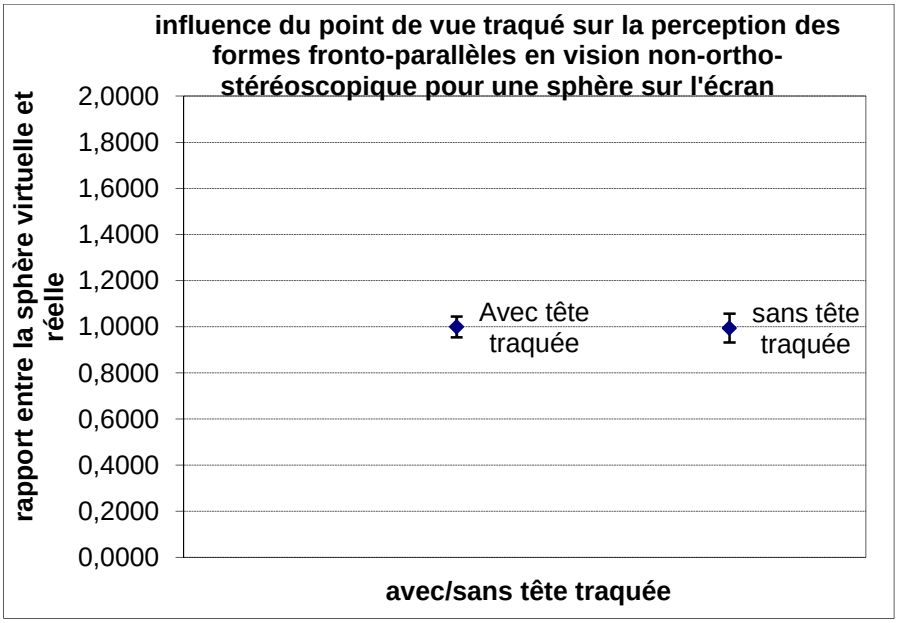


Figure 82 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique pour une sphère sur l'écran

3.3.4.2. Lorsque la sphère est derrière l'écran

Les résultats de l'expérience sur la perception des courbures en profondeur sont exposés sur la Figure 83. Les JND et PSE lorsque le point de vue est traqué sont respectivement 11% et 98% du diamètre de la sphère réelle, tandis que lorsque le point de vue n'est pas traqué, ils sont égaux à 100% et 97%. Nous voyons donc une différence de JND, d'ailleurs la probabilité de significativité de l'écart des JND est 100%. La probabilité de significativité des PSE est seulement 65% et n'est donc pas significative.

tracking	oui/non
projection	Non-ortho
forme	Sphère
position	derrière

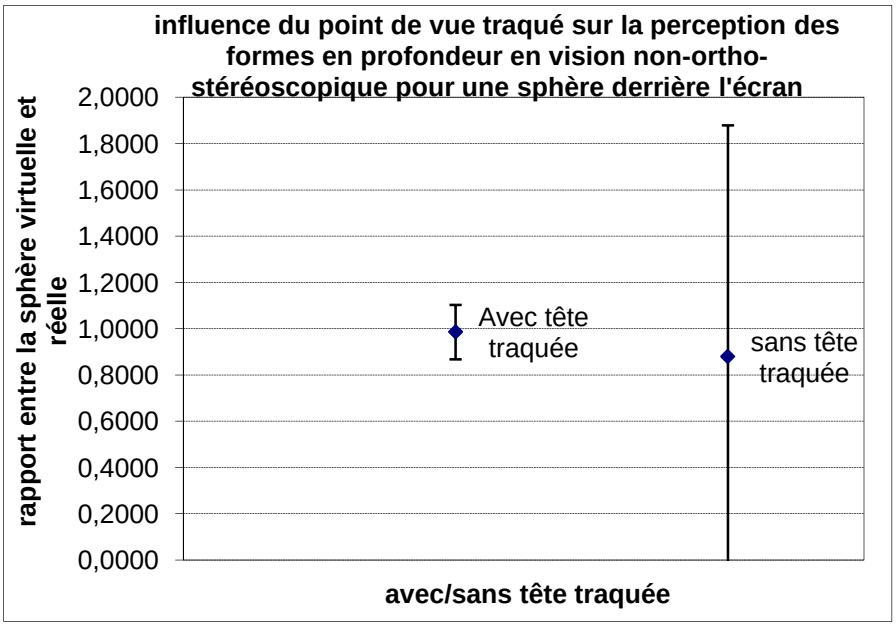


Figure 83 : influence du point de vue traqué sur la perception des formes en profondeur en vision non-ortho-stéréoscopique pour une sphère derrière l'écran

Nous pouvons refaire cette expérience pour la perception des formes fronto-parallèles. Cette fois encore, la probabilité de significativité de l'écart entre les JND est supérieure à 95% (99%). Le point de vue traqué a donc une influence non négligeable sur la perception des formes fronto-parallèles lorsque la vision n'est pas ortho-stéréoscopique et que la sphère est derrière l'écran. Les probabilités de significativité d'écart entre les PSE et de différence entre la largeur et la hauteur sont négligeables (54% et 71%). Le graphique étant sensiblement le même que lorsque la sphère est sur l'écran, il est exposé en Annexe A.

3.3.4.3. Lorsque la sphère est devant l'écran

Les résultats de l'expérience sur la perception des formes en profondeur sont exposés sur la Figure 84. Les PSE et JND lorsque le point de vue est traqué sont respectivement 96% et 11% contre 91% et 42% lorsque le point de vue n'est pas traqué. La probabilité de significativité de l'écart des PSE est 62% et n'est pas significatif. Par contre, la probabilité de significativité de l'écart des JND est 100%. Donc cette différence est significative.

tracking	oui/non
projection	Non-ortho
forme	Sphère
position	Devant

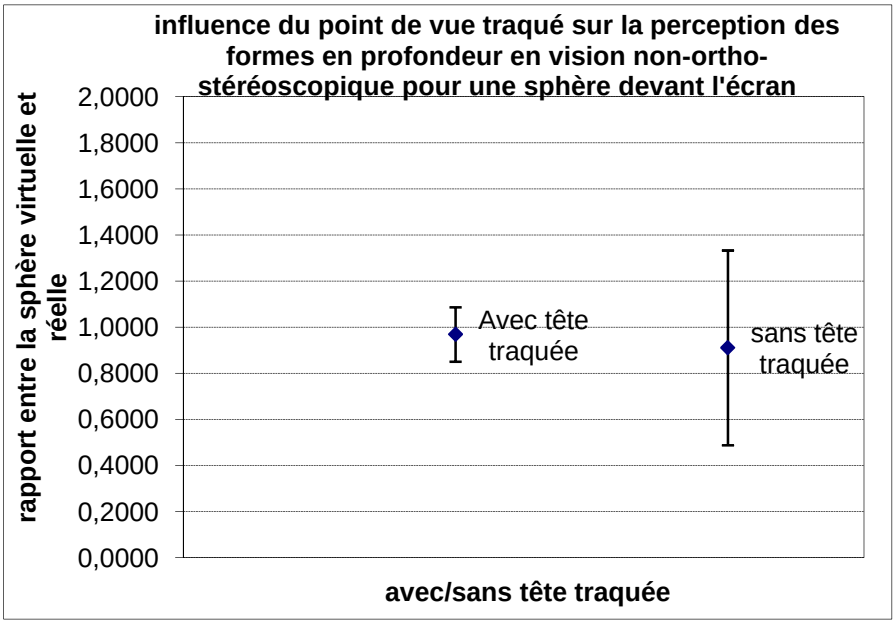


Figure 84 : influence du point de vue traqué sur la perception des formes en profondeur en vison non-ortho-stéréoscopique pour une sphère devant l'écran

Nous pouvons faire de même avec les formes fronto-parallèles. Les PSE et JND lorsque le point de vue est traqué sont 101% et 3% du diamètre initial de la sphère. Lorsque le point de vue n'est pas traqué, ils sont égaux à 98% et 8%. La différence entre les JND est significative (99,9%). Par contre, l'écart entre les PSE et la différence entre la perception des courbures en horizontal et en vertical sont négligeables (61% et 90%). Les résultats graphiques de cette expérience sont proposés en Annexe A, puisqu'il ressemble énormément au graphique pour la sphère sur l'écran.

3.3.4.4. Lorsque la forme non connue est sur l'écran

Nous présentons les résultats de l'expérience sur la perception des courbures en profondeur sur la Figure 85. Les PSE et JND avec le point de vue traqué sont 94% et 24% des dimensions de la forme étalon ; tandis que sans le point de vue traqué, ils sont égaux à 90% et 52 %. Les probabilités de significativité des écarts de PSE et de JND sont respectivement 57% et 99,9%, la différence de JND est donc significative. Cette fois encore, le fait de traquer le point de vue est important pour la perception de forme en profondeur.

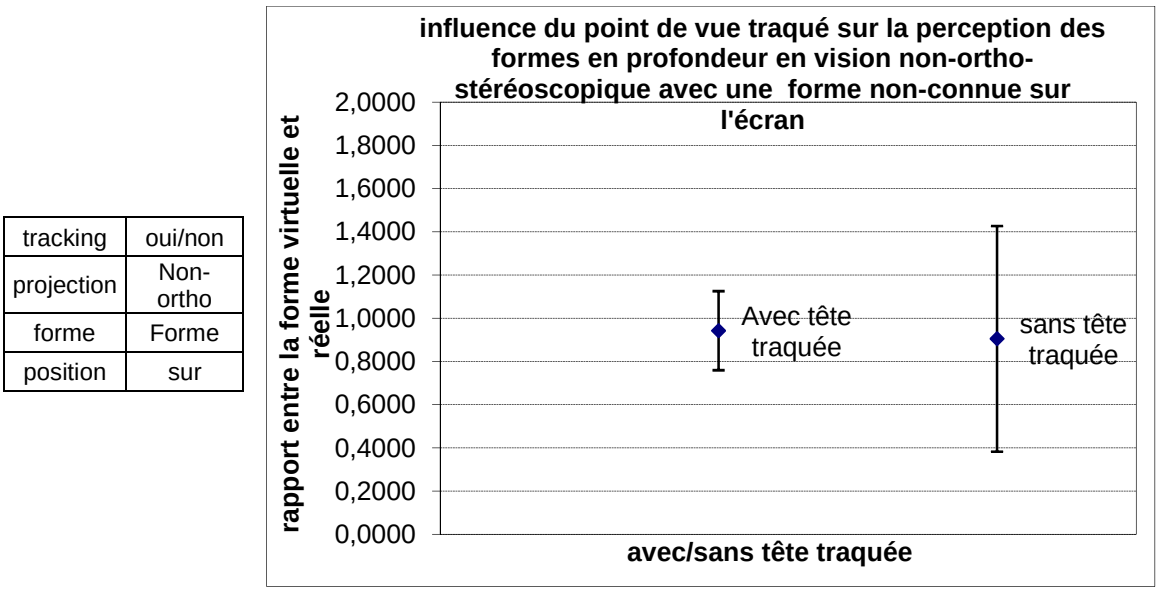


Figure 85 : influence du point de vue traqué sur la perception de formes en profondeur en vision non-ortho-stéréoscopique avec une forme non-connue sur l'écran

Nous pouvons nous intéresser aux formes fronto-parallèles, les résultats graphiques sont en Annexe A. Les PSE et JND avec point de vue traqué sont respectivement 101% et 4% du diamètre initial de la forme. Sans point de vue traqué, ils sont égaux à 100% et 8%. Le point de vue traqué est donc important dans cette configuration-ci également puisque la probabilité de significativité de l'écart des JND est 99%. Par contre, l'écart de PSE et la différence entre les perceptions des courbures en largeur ou en hauteur ne sont pas significatives.

3.3.4.5. Lorsque la forme non connue est derrière l'écran

Les résultats de l'expérience sur les formes en profondeur sont exposés sur la Figure 86. Les PSE et JND avec point de vue traqué sont 97% et 22% de la dimension initiale tandis que sans point de vue traqué, ils sont égaux à 92% et 67%. Les probabilités de significativité des écarts de PSE et de JND sont 60% et 99,9%. Donc l'écart de JND est significatif. Une modification de courbure en profondeur sera perçue plus finement avec le point de vue traqué.

tracking	oui/non
projection	Non-ortho
forme	Forme
position	derrière

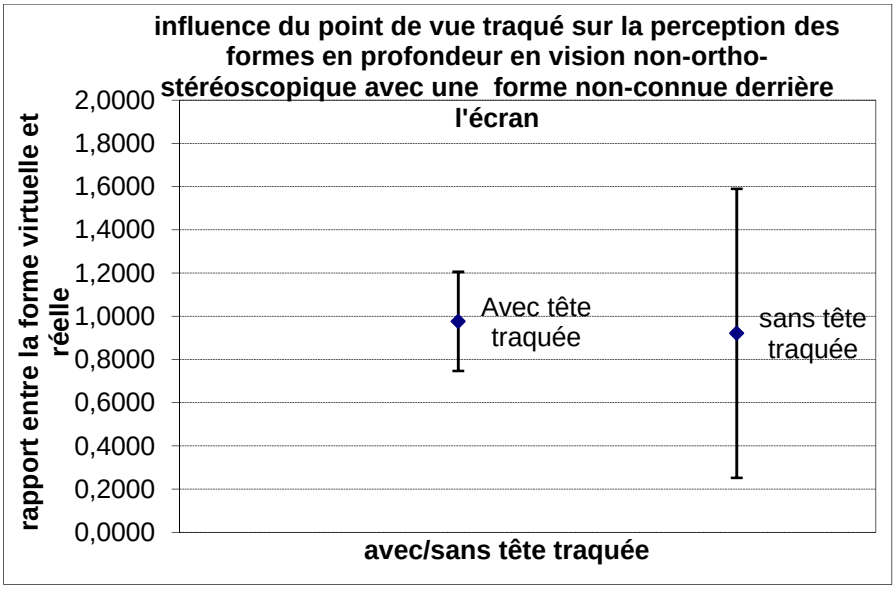


Figure 86 : influence du point de vue traqué sur la perception des formes en profondeur en vision non-ortho-stéréoscopique avec une forme non-connue derrière l'écran

Nous refaisons les mêmes expériences pour les formes fronto-parallèles, les résultats en sont exposés en Annexe A. Les PSE et JND avec point de vue traqué sont 100% et 3,5% des dimensions initiales de la forme, tandis que sans point de vue traqué, ils sont égaux à 99% et 7%. Comme nous pouvions nous y attendre au vu des résultats, l'écart des JND est significatif : 99% Donc pour les formes fronto-parallèles également, il vaut mieux avoir le point de vue traqué pour une meilleure perception des formes. Les autres écarts sont négligeables.

3.3.4.6. Lorsque la forme non connue est devant l'écran

Les résultats de l'expérience sur la perception des formes en profondeur sont présentés sur la Figure 87. Le PSE et le JND lorsque le point de vue est traqué sont respectivement 105% et 14% des dimensions de la forme étalon contre 87% et 38% lorsque le point de vue n'est pas traqué. La probabilité de significativité de l'écart de JND est 99,9%, tandis que celle des PSE est 86%. Nous voyons donc que la différence de JND est significative.

tracking	oui/non
projection	Non-ortho
forme	Forme
position	devant

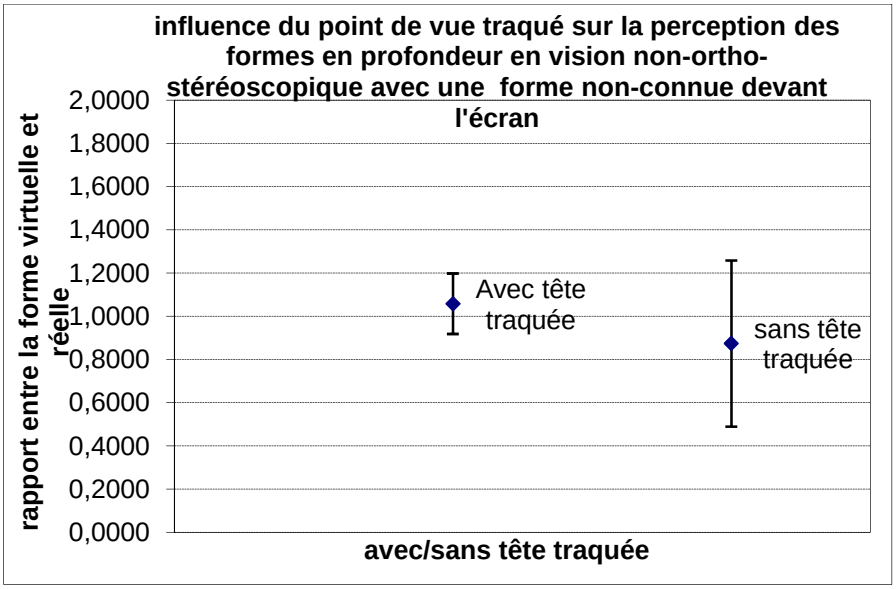


Figure 87 : influence du point de vue traqué sur la perception des formes en profondeur en vision non-ortho-stéréoscopique avec une forme non-connue devant l'écran

Nous donnons également les résultats sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique en annexe A. Les JND et PSE avec le point de vue traqué sont 102% et 4,5% des dimensions de la forme réelle, tandis que sans point de vue traqué, ils sont égaux à 100% et 7%. Cette différence de JND est significative (99% de probabilité de significativité), cependant, aucun autre écart n'est significatif.

3.3.4.7. Conclusion sur l'hypothèse n°3

Les améliorations des JND sont flagrantes ici également lorsque le point de vue est traqué : 62% (12cm) des dimensions des objets pour la sphère (86% derrière, 6% sur et 32% devant l'écran) et 32% (7cm) pour la forme non connue (45% derrière, 28% sur et 22% devant l'écran). Nous en déduisons que quelque soit le mode de projection que nous employons, il est très important que le point de vue soit traqué pour la perception des formes en profondeur. La progression des améliorations en fonction de la position de l'objet par rapport à l'écran est à peu près la même que celle lorsque les sujets sont en vision ortho-stéréoscopique, ce qui est cohérent avec le fait que les sujets remplacent le manque de parallaxes de mouvement par les disparités horizontales lorsque les premières sont absentes.

Les modifications perçues en profondeur font 18% du diamètre de la sphère (3,7cm) lorsque le point de vue est traqué, et elles font 76% sans point de vue traqué (15cm). Pour la forme non connue a priori, les différences perçues avec et sans point de vue traqué sont 20% (4cm) et 52% (17cm).

Contrairement aux hypothèses n°1 et n°2, cette fois ci, les formes fronto-parallèles sont également améliorées lorsque le point de vue est traqué, de quelques pourcents quelque soit la position de l'objet. Par contre, il n'y aucune modification notable des points d'égalisation subjective.

3.3.5. Hypothèse n° 4 : Il est préférable d'avoir la vision ortho-stéréoscopique que la vision non-ortho-stéréoscopique avec le point de vue traqué

3.3.5.1. Lorsque la sphère est sur l'écran

Les résultats de ces expériences sont exposés sur la Figure 88. Les JND et les PSE en vision ortho-stéréoscopique sont respectivement 95% et 18% du diamètre de la sphère étalon, tandis que ceux obtenus en vision non-ortho-stéréoscopique sont 90% et 15%. Aucun de ces écarts n'est significatif : 90% de probabilité de significativité pour l'écart de JND et 61% pour celui de PSE.

tracking	Oui
projection	Ortho/Non-ortho
forme	Sphère
position	sur

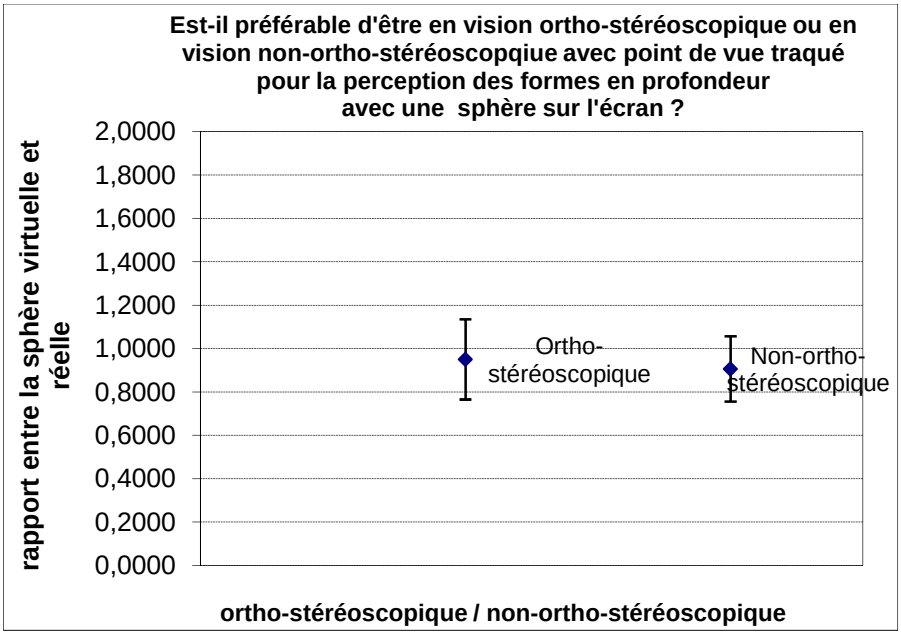


Figure 88 : Comparaison de la perception des formes en profondeur avec une sphère sur l'écran avec point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique

Aucune différence pour la perception des formes fronto-parallèles n'est significative, donc nous exposons les résultats de cette expérience en annexe A.

3.3.5.2. Lorsque la sphère est derrière l'écran

Les résultats de l'expérience sur la perception des courbures en profondeur sont exposés sur la Figure 89 Les PSE et JND lorsque le sujet est en vision ortho-stéréoscopique sont 93% et 17%, ceux lorsque le sujet est en vision non-ortho-stéréoscopique sont 98% et 11% du diamètre initial de la forme. La probabilité de significativité de l'écart des JND est 97,7% ; cet écart est donc significatif. Par contre, celui entre les PSE ne l'est pas (64%).

tracking	Oui
projection	Ortho/Non-ortho
forme	Sphère
position	derrière

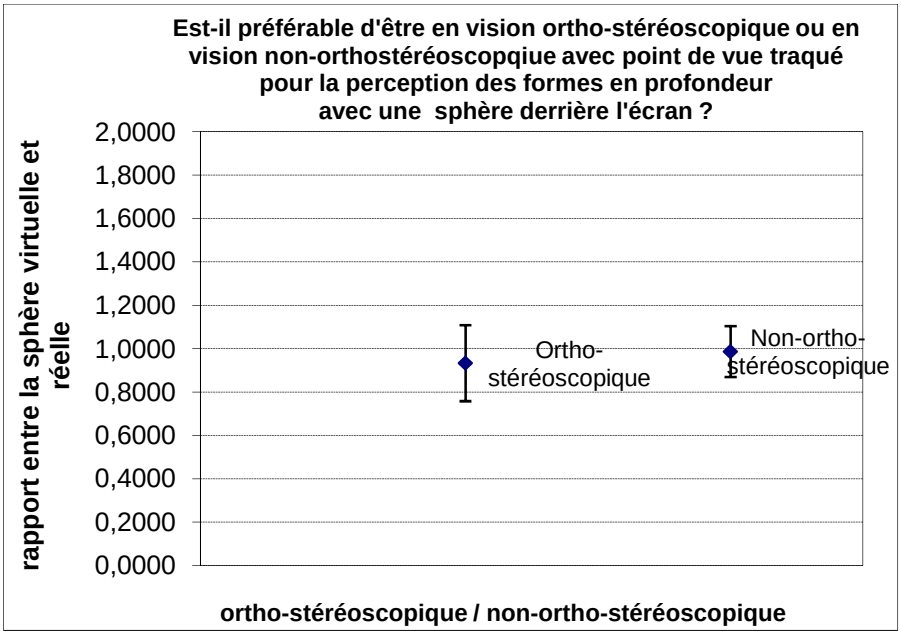


Figure 89 : Comparaison de la perception des formes en profondeur avec une sphère derrière l'écran avec point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique

Les résultats pour la perception des forme fronto-parallèles sont en annexe A.

3.3.5.3. Lorsque la sphère est devant l'écran

Les résultats de l'expérience sur la perception des courbures en profondeur sont visibles sur la Figure 90. Les PSE et JND en vision ortho-stéréoscopique sont 100% et 24% du diamètre de la sphère réelle, tandis qu'en vision non-ortho-stéréoscopique ils sont égaux à 96% et 11%. La probabilité de significativité de l'écart des PSE est 59%, donc, cet écart n'est pas significatif. Par contre, l'écart de JND est, lui, significatif : sa probabilité de significativité s'élève à 99,9%.

tracking	Oui
projection	Ortho/Non-ortho
forme	Sphère
position	devant

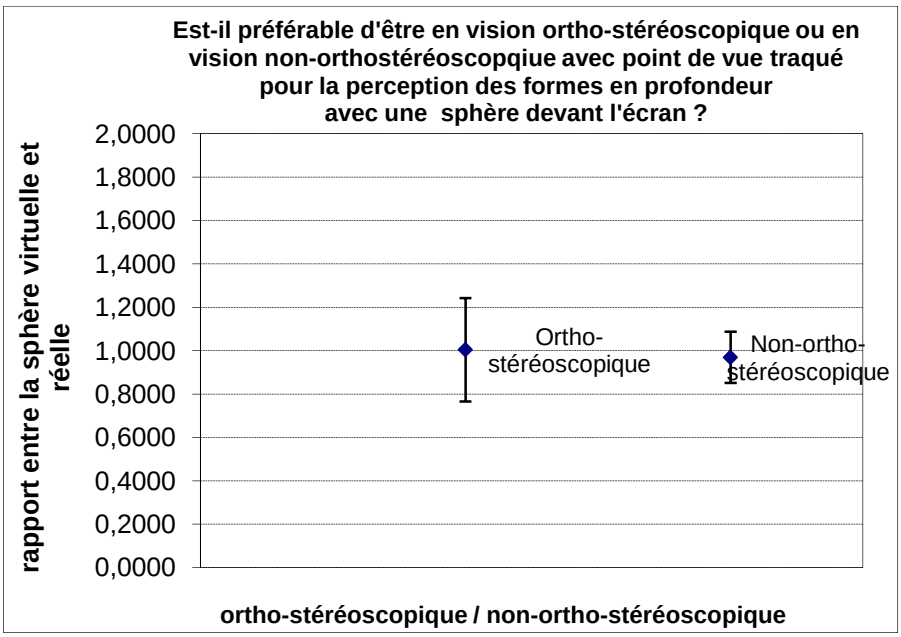


Figure 90 : Comparaison de la perception des formes en profondeur avec une sphère devant l'écran avec point de vue traqué en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique

Les résultats de l'expérience sur les formes fronto-parallèles sont en annexe A.

3.3.5.4. Lorsque la forme non connue est sur l'écran

Nous présentons les résultats de l'expérience sur la perception des courbures en profondeur sur la Figure 91. Les PSE et JND en vision ortho-stéréoscopique sont respectivement 100% et 15% des dimensions de la forme réelle. Tandis que ceux en vision non-ortho-stéréoscopique sont 94% et 18%. Aucun des écarts de PSE ou de JND n'est significatif (67% et 91%). La vision ortho-stéréoscopique n'aurait donc que peu d'influence sur la perception des courbures en profondeur dans ce cas-ci. Ceci peut s'expliquer par le fait que lorsque l'objet est sur l'écran les disparités sont très faibles, donc une différence légère entre les caméras virtuelles ne se traduit pas par une grande différence de disparité, cela n'influe donc pas énormément sur la perception des formes. Surtout lorsque le point de vue est traqué, nous avons vu que le point de vue traqué a une très grande influence, donc les sujets se reposent plus sur lui que sur les disparités horizontales.

tracking	Oui
projection	Ortho/Non-ortho
forme	Forme
position	sur

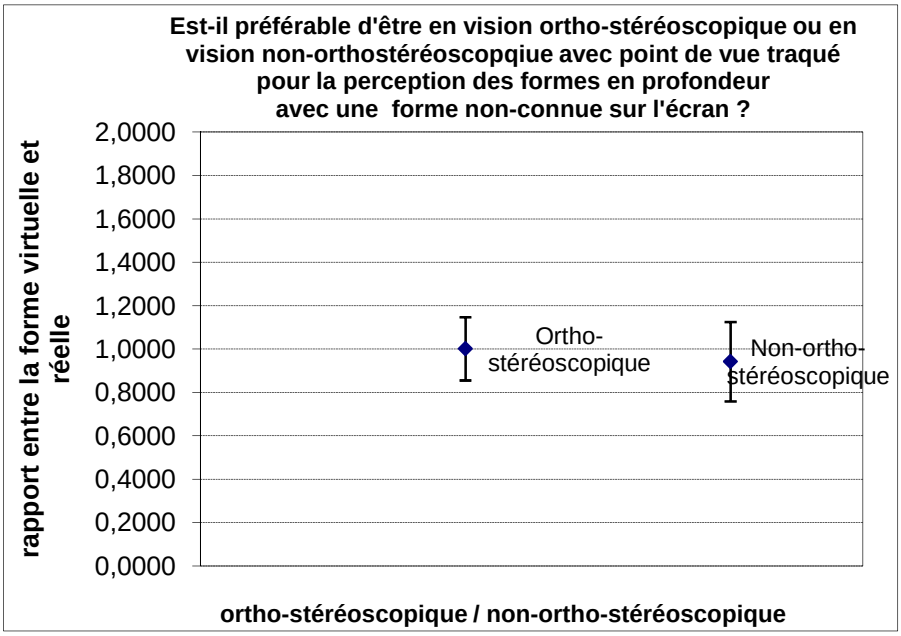


Figure 91 : Comparaison de la perception des formes en profondeur avec une forme inconnue sur l'écran avec point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique

Nous présentons les résultats pour la perception des formes fronto-parallèles en annexe A.

3.3.5.5. Lorsque la forme non connue est derrière l'écran

Les résultats de l'expérience sur la perception des courbures en profondeur sont présentés sur la Figure 92. Les PSE et JND en vision ortho-stéréoscopique sont 113% et 30% des dimensions initiales tandis qu'en vision non-ortho-stéréoscopique, ils sont égaux à 97% et 23%. Cette fois encore, les écarts ne sont pas significatifs : 94% de probabilité de significativité pour les écarts de JND et 81% pour ceux de PSE. Donc lorsque la forme est derrière l'écran, la distance entre les caméras virtuelles influence peu. Ceci est probablement dû à la pixellisation, les écarts étant petits, il est plus facile de se fier au point de vue traqué qu'aux disparités.

tracking	Oui
projection	Ortho/Non-ortho
forme	Forme
position	Derrière

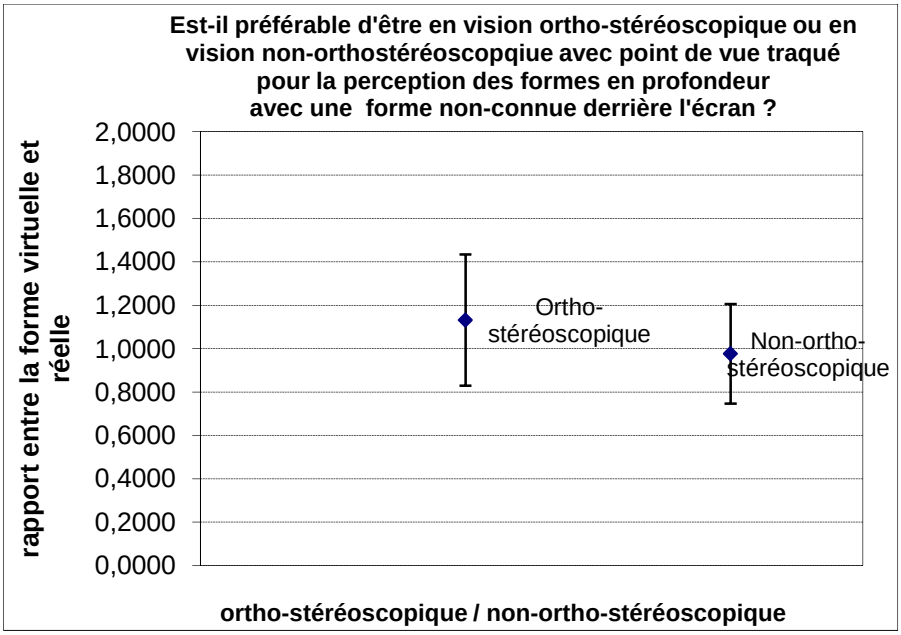


Figure 92 : Comparaison entre la perception des formes en profondeur avec une forme inconnue derrière l'écran avec point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique

Les résultats n'étant pas significatifs pour la perception des formes fronto-parallèles, ils sont présentés en annexe A.

3.3.5.6. Lorsque la forme non connue est devant l'écran

Les résultats de l'expérience sur la perception des formes en profondeur sont exposés sur la Figure 93. Les PSE et JND avec la vision ortho-stéréoscopique sont respectivement 99% et 18% de la dimension de base et en vision non-ortho-stéréoscopique ils sont égaux à 105% et 14% du diamètre de la forme initiale. Cette fois, la différence de JND est significative : 96%. Les sujets voient une modification de la courbure en profondeur plus rapidement lorsque la distance entre les caméras virtuelles est plus grande que la distance entre leurs yeux. Cela s'explique probablement parce que le relief est augmenté, les modifications sont plus importantes, donc plus facilement visibles.

tracking	Oui
projection	Ortho/Non-ortho
forme	Forme
position	Devant

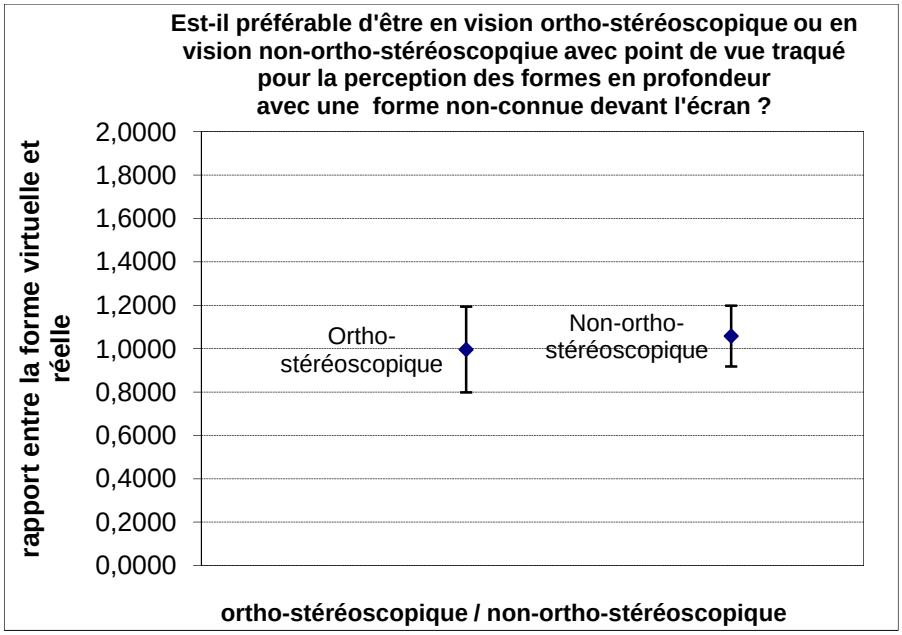


Figure 93 : comparaison de la perception des formes en profondeur avec une forme inconnue devant l'écran avec point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique

Les résultats de l'expérience sur les formes fronto-parallèles sont présentés en annexe A.

3.3.5.7. Conclusion sur l'hypothèse n°4

Il vaut mieux être en vision non-ortho-stéréoscopique pour la perception des formes en profondeur. L'amélioration des JND pour la sphère est 7,3% de ses dimensions (6% derrière, 3% sur et 13% devant l'écran), et pour la forme non connue a priori, 4% (7% derrière, 3% sur et 4% devant l'écran). Nous supposons qu'en vision non-ortho-stéréoscopique, les différences de profondeur sont plus marqués, donc les sujets perçoivent plus rapidement lorsque la profondeur est modifiée. De façon assez logique, il n'y a aucune différence significative lorsque les formes sont positionnées sur l'écran. Les parallaxes horizontales ne sont pas suffisamment différentes pour être visibles (masquées par la pixellisation). Notons tout de même que nos expériences portent sur la perception des formes, et non sur la perception de la taille, en effet, beaucoup de sujets ont déclaré voir les objets devant l'écran globalement plus grands que l'objet réel en vision non ortho-stéréoscopique.

Pour une sphère, une modification de 20% du diamètre (4cm) sera détectable en vision ortho-stéréoscopique, en vision non-ortho-stéréoscopique, il faudra 12% (2,5cm) pour que cette modification soit percevable. Pour une forme, il faudra une modification de 21% (4cm) en vision ortho-stéréoscopique et 18% (3,5cm) en vision non-ortho-stéréoscopique.

Il n'y a aucun changement notable de point d'égalisation subjective, ni d'amélioration de perception des formes fronto-parallèles dans un sens ou dans un autre.

3.3.6. Hypothèse n° 5 : Il est préférable d'avoir la vision ortho-stéréoscopique que la vision non-ortho-stéréoscopique sans le point de vue traqué

3.3.6.1. Lorsque la sphère est sur l'écran

Les résultats de l'expérience sur la perception des courbures en profondeur sont présentés sur la Figure 94. Les PSE et JND en vision ortho-stéréoscopique sont 106% et 52% du diamètre de la sphère étalon, tandis qu'en vision non-ortho-stéréoscopique, ils sont égaux à 84% et 90%. La probabilité de significativité de l'écart des JND est 99%, donc cet écart est significatif. Par contre, l'écart des PSE n'est pas significatif (75%).

tracking	non
projection	Ortho/Non-ortho
forme	Sphère
position	sur

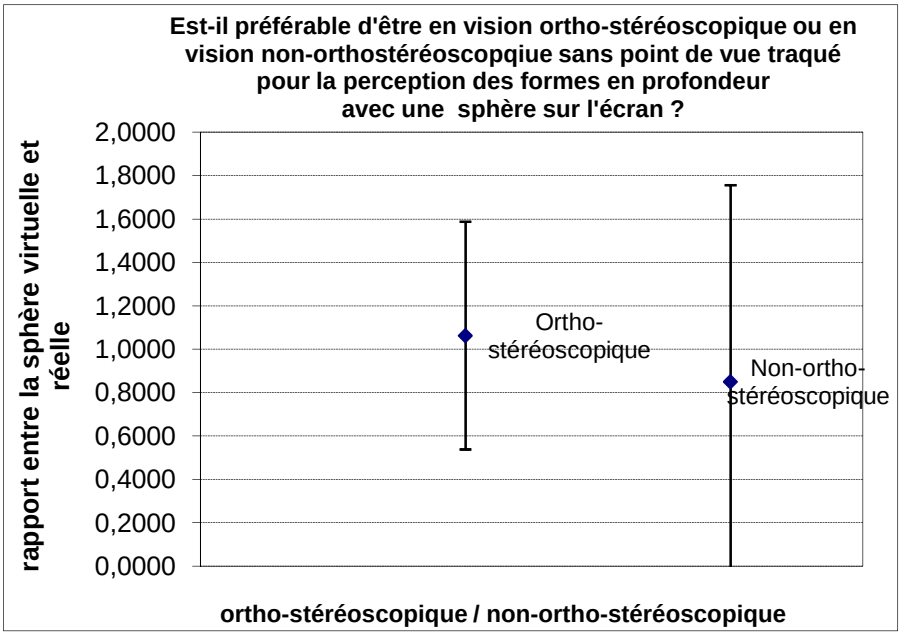


Figure 94 : Comparaison entre la perception des formes en profondeur avec une sphère sur l'écran sans point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique

Aucune différence de perception des formes fronto-parallèles n'est significative, nous exposons donc les résultats en annexe A.

3.3.6.2. Lorsque la sphère est derrière l'écran

Les résultats de cette expérience sur les courbures en profondeur sont présentés sur la Figure 95. Les PSE et JND en vision ortho-stéréoscopique sont 105% et 86% du diamètre initial, tandis qu'en vision non-ortho-stéréoscopique, ils sont égaux à 97% et 100%. Les écarts ne sont pas significatifs, celui entre les JND a une probabilité de significativité associée de 87% et celui entre les PSE, 68%.

tracking	non
projection	Ortho/Non-ortho
forme	Sphère
position	derrière

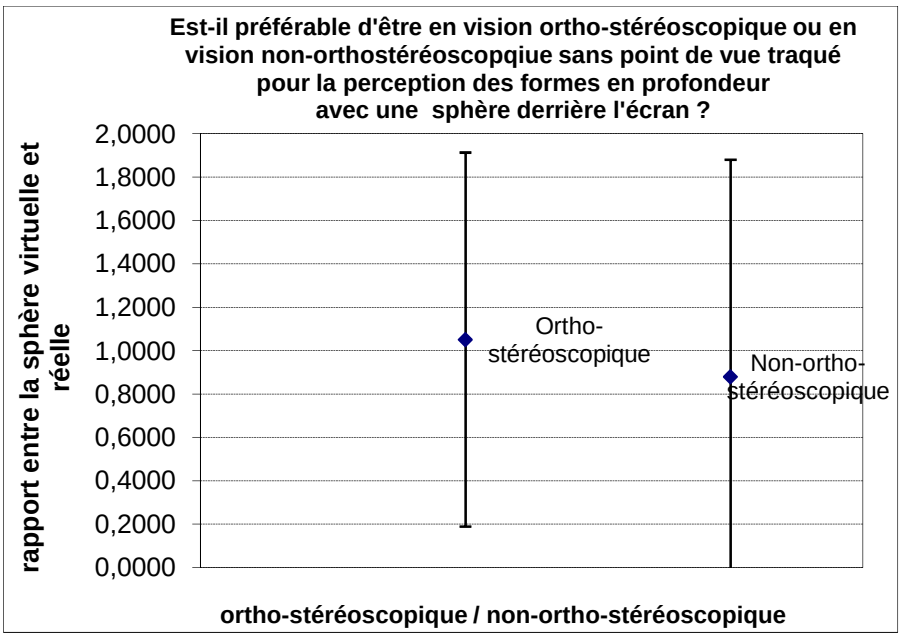


Figure 95 : Comparaison entre la perception des formes en profondeur avec une sphère derrière l'écran sans point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique

Nous pouvons réaliser les mêmes expériences pour les formes fronto-parallèles, ces résultats sont proposés en Figure 96. La probabilité de significativité et l'écart entre les PSE et celle associée à la différence de perception des courbures en largeur et en hauteur sont trop petites pour que ces écarts soient significatifs (51% et 91%). Par contre, celle associée à l'écart des JND est suffisamment grande (99%) pour que cette différence soit considérée comme significative.

tracking	non
projection	Ortho/Non-ortho
forme	Sphère
position	derrière

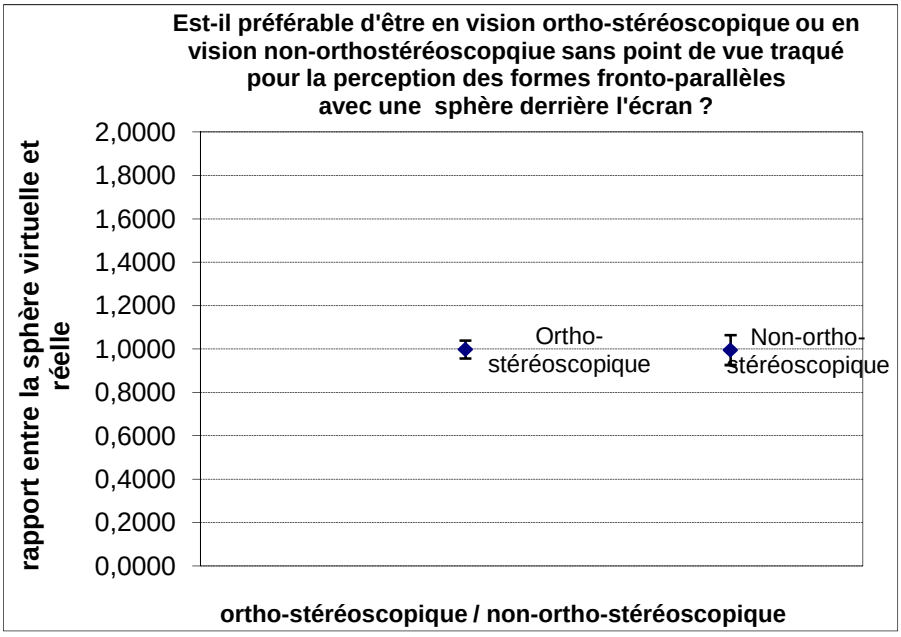


Figure 96 : Comparaison entre la perception des formes fronto-parallèles avec une sphère derrière l'écran sans point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique

3.3.6.3. Lorsque la sphère est devant l'écran

Les résultats des expériences sur la perception des courbures en profondeur sont exposés sur la Figure 97. Les PSE et JND en vision ortho-stéréoscopique sont 99% et 26% du diamètre de la

sphère réelle. En vision non-ortho-stéréoscopique, ils sont égaux à 91% et 42%. Les probabilités de significativité des écarts des PSE et de JND sont respectivement 65% et99%. Nous voyons donc que la différence entre les JND est significative.

tracking	non
projection	Ortho/Non-ortho
forme	Sphère
position	devant

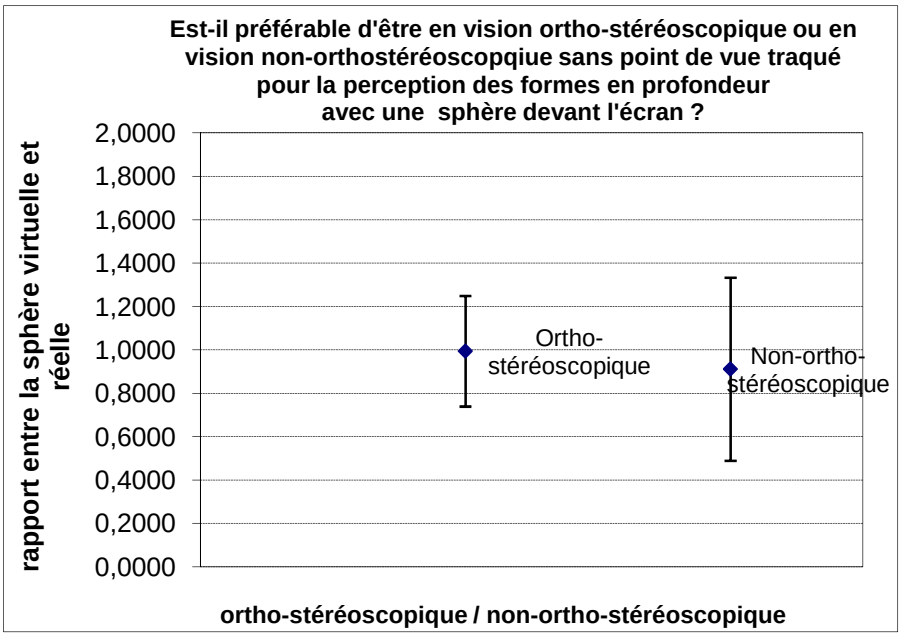


Figure 97 : Comparaison entre la perception des formes en profondeur avec une sphère devant l'écran sans point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique

Les résultats de l'expérience sur les formes fronto-parallèles sont visibles en annexe A (les résultats graphiques sont très semblables au cas précédant). Les PSE et JND en vision ortho-stéréoscopique sont égaux à 98% et 4% du diamètre étalon, tandis qu'en vision non-ortho-stéréoscopique, ils sont égaux à 99% et 7%. La probabilité de significativité de la différence des JND étant 99,9% nous pouvons considérer cet écart comme significatif. Par contre, il n'y a pas de différence suffisamment significative pour les PSE ni de différence entre la perception des courbures horizontales et verticales.

3.3.6.4. Lorsque la forme non connue est sur l'écran

Les résultats de l'expérience sur la perception des formes en profondeur sont présentés en Figure 98. Les PSE et JND en vision ortho-stéréoscopique sont respectivement 105% et 62% du diamètre initial de la forme. En vision non-ortho-stéréoscopique, ils sont égaux à 90% et 52%. Ces écarts ne sont pas significatifs puisque la probabilité de significativité pour l'écart de JND est 91% et pour l'écart de PSE, 71%.

tracking	non
projection	Ortho/Non-ortho
forme	Forme
position	sur

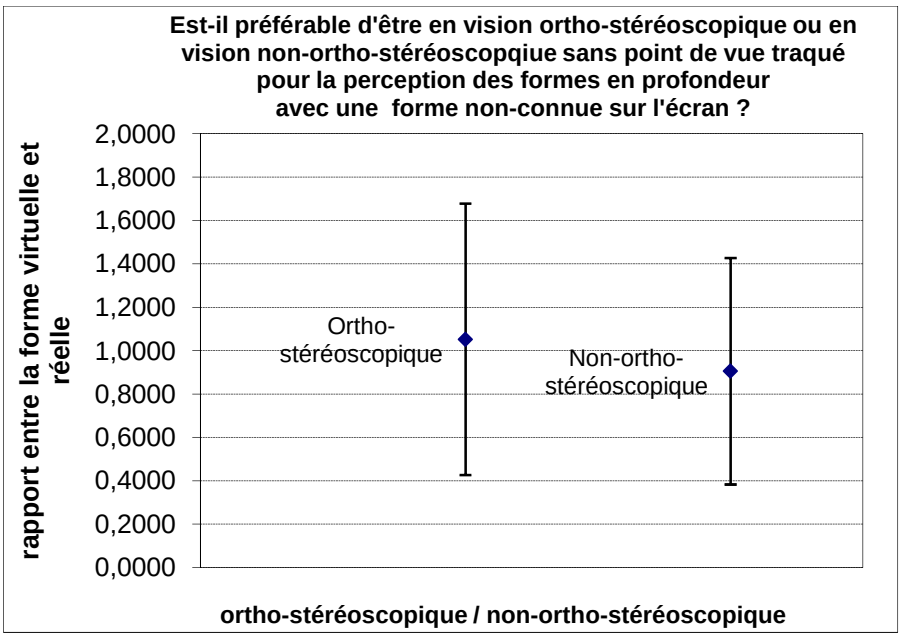


Figure 98 : Comparaison entre la perception des formes en profondeur avec une forme non connue l'écran sans point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique

La même expérience a été menée pour les formes fronto-parallèles, les résultats en sont présentés en annexe A (les résultats graphiques étant très semblables aux précédents). Les PSE et JND lorsque les sujets sont en vision ortho-stéréoscopique sont 100% et 4% des dimensions de la forme réelle et lorsqu'ils sont en vision non-ortho-stéréoscopique, ils sont égaux à 99% et 7%. L'écart de JND est significatif (99%). Donc, la perception des formes fronto-parallèles est perturbée lorsque l'objet est sur l'écran et que la distance entre les caméras virtuelles est double par rapport à celle entre ses yeux.

3.3.6.5. Lorsque la forme non connue est derrière l'écran

Les résultats de l'expérience sur la perception des courbures en profondeur sont présentés sur la Figure 99. Les PSE et JND en vision ortho-stéréoscopique sont 129% et 123% du diamètre initial de la forme. En vision non-ortho-stéréoscopique, nous avons un PSE et un JND égaux à 92% et 66%. Contrairement à l'écart de JND lorsque le point de vue de l'utilisateur était traqué, cet écart-ci est significatif : 99%.

tracking	non
projection	Ortho/Non-ortho
forme	Forme
position	derrière

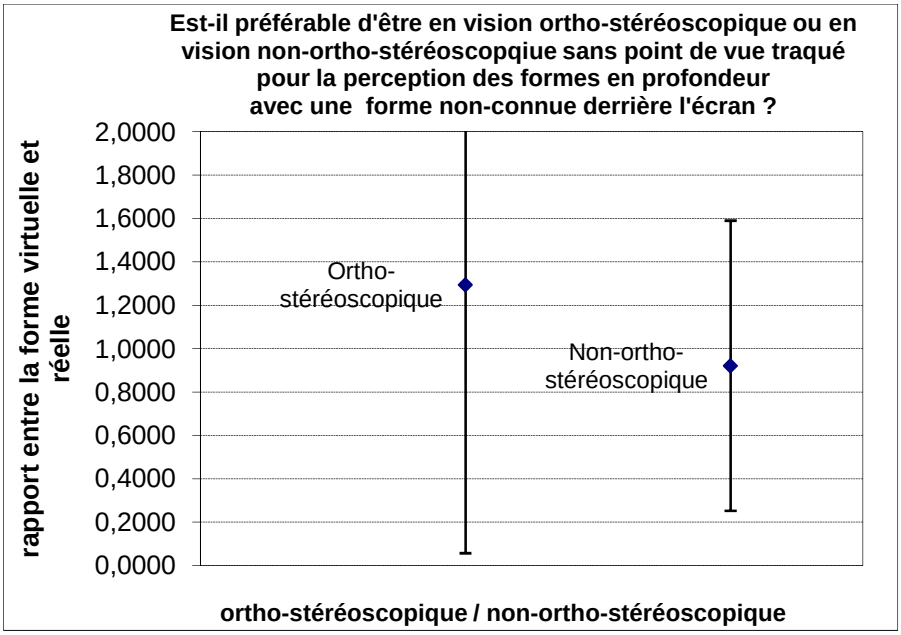


Figure 99 : Comparaison entre la perception des formes en profondeur avec une forme non connue derrière l'écran sans point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique

Nous faisons de même pour les formes fronto-parallèles, les résultats graphiques en sont exposés en annexe A. Les PSE et JND en vision ortho-stéréoscopique sont 101% et 4% des dimensions de la forme étalon tandis qu'en vision non-ortho-stéréoscopique ils sont égaux à 99% et 7%. Comme lorsque la forme était sur l'écran, la perception des formes fronto-parallèles est également perturbée de façon significative (99%) lorsque l'objet est derrière l'écran. L'écart entre la perception des modifications de largeur et de hauteur, ainsi que celui des PSE son négligeables (83% et 61%).

3.3.6.6. Lorsque la forme non connue est devant l'écran

Nous présentons les résultats de l'expérience sur l'influence de la distance entre les deux caméras virtuelles sans point de vue traqué sur la perception des formes en profondeur sur la Figure 100. Les PSE et JND en vision ortho-stéréoscopique sont 96% et 36% du diamètre de la forme initiale ; en vision non-ortho-stéréoscopique, ils sont égaux à 87% et 38%. Les probabilités de significativité de l'écart des JND et des PSE sont respectivement 81% et 60%. Ces deux écarts sont donc non-significatifs.

tracking	non
projection	Ortho/Non-ortho
forme	Forme
position	Devant

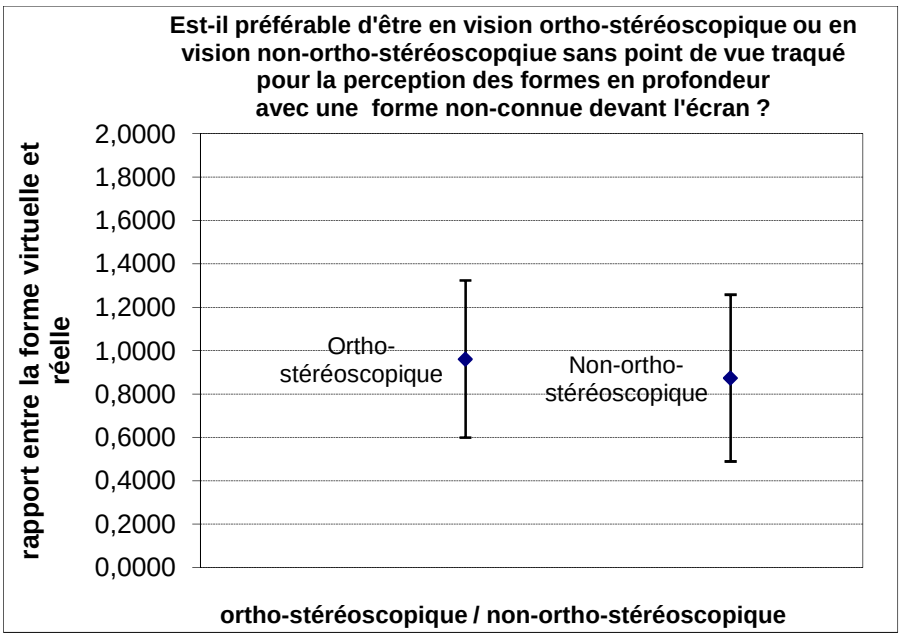


Figure 100 : Comparaison entre la perception des formes en profondeur avec une forme non connue devant l'écran avec point de vue traqué en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique

Nous pouvons également présenter les résultats de l'expérience similaire pour la perception des formes fronto-parallèles en annexe A (les résultats graphiques étant très semblables à ceux pour la forme sur l'écran). Les PSE et JND en vision ortho-stéréoscopique sont 100% et 5% des dimensions de la forme étalon et en vision non-ortho-stéréoscopique, 100% et 7%. La probabilité de significativité de l'écart des JND est 98%, mais les autres écarts sont non-significatifs.

3.3.6.7. Conclusion sur l'hypothèse n°5

Les conclusions pour l'hypothèse n°5 sont plus compliquées à interpréter. Pour une sphère, la vision ortho-stéréoscopique améliore la perception des formes en profondeur de 22% (14% derrière, 38% sur et 16% devant l'écran), mais pour une forme non connue a priori, c'est la vision non-ortho-stéréoscopique qui améliore cette perception de 21% (57% derrière, 10% sur et -2% devant l'écran). Nous pensons que c'est à cause du fait que sans point de vue traqué, les résultats sont très mauvais quelque soit la vision proposée, et donc ce ne sont pas des améliorations très grandes (d'ailleurs peu d'entre elles sont significatives). De ce fait, nous ne proposons pas plus de résultats chiffrés.

Par contre, il semblerait que la vision non-ortho-stéréoscopique améliore de façon significative la perception des formes fronto-parallèles. Les points d'égalisation subjective n'ont subi aucune modification significative.

3.3.7. Hypothèse n° 6 : Il est préférable d'avoir la vision monoscopique avec le point de vue traqué que la vision ortho-stéréoscopique sans le point de vue traqué

3.3.7.1. Lorsque la sphère est sur l'écran

Les résultats de l'expérience sur les formes en profondeur sont proposés sur la Figure 101. Les PSE et JND en vision monoscopique avec le point de vue traqué sont 86% et 19% du diamètre de la sphère étalon. Ceux en vision ortho-stéréoscopique sans point de vue traqué sont égaux à 106% et 52%. Les probabilités de significativité associées aux écarts de PSE et de JND sont respectivement 82% et 99,9%. Nous voyons donc que la différence des JND est significative et qu'il est préférable, pour une meilleure perception des formes en profondeur, d'être en vision monoscopique avec le point de vue traqué qu'en vision ortho-stéréoscopique sans point de vue traqué.

tracking	Oui/non
projection	Mono/ortho
forme	Sphère
position	Sur

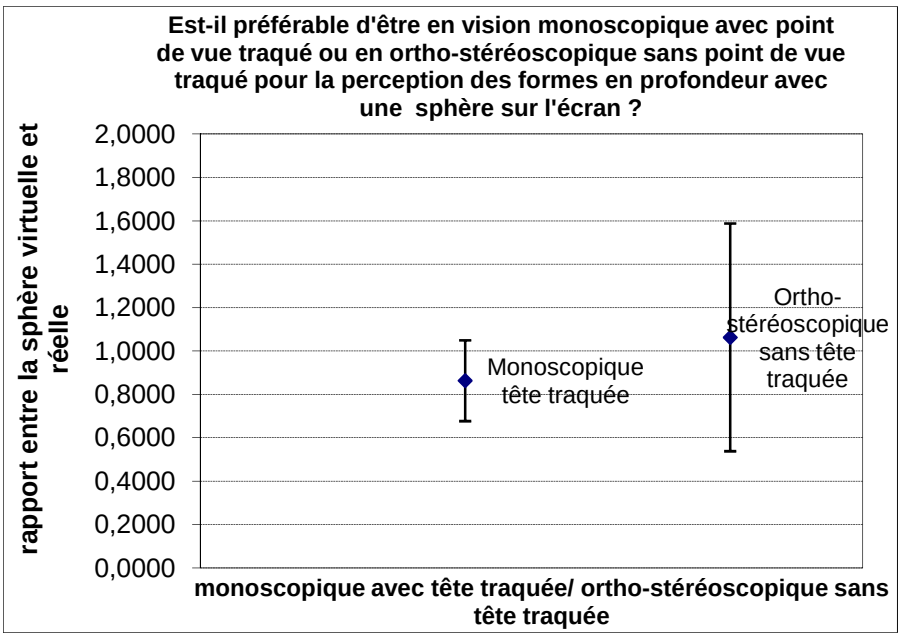


Figure 101 : comparaison entre la perception des formes en profondeur avec une sphère sur l'écran sans point de vue traqué en vision monoscopique et avec point de vue traqué en vision ortho-stéréoscopique

Les différences de perception des formes fronto-parallèles sont non significatives, nous les plaçons donc en annexe A.

3.3.7.2. Lorsque la sphère est derrière l'écran

Les résultats de l'expérience sur la perception des formes en profondeur sont présentés sur la Figure 102. Les PSE et JND en vision monoscopique avec point de vue traqué sont 90% et 35% du diamètre initial de la sphère, en vision ortho-stéréoscopique sans point de vue traqué, ils sont égaux à 105% et 86%. La probabilité de significativité de l'écart de JND est 99,9%, donc cet écart est significatif. Par contre, celui entre les PSE n'est pas significatif (69%).

tracking	Oui/non
projection	Mono/ortho
forme	Sphère
position	Derrière

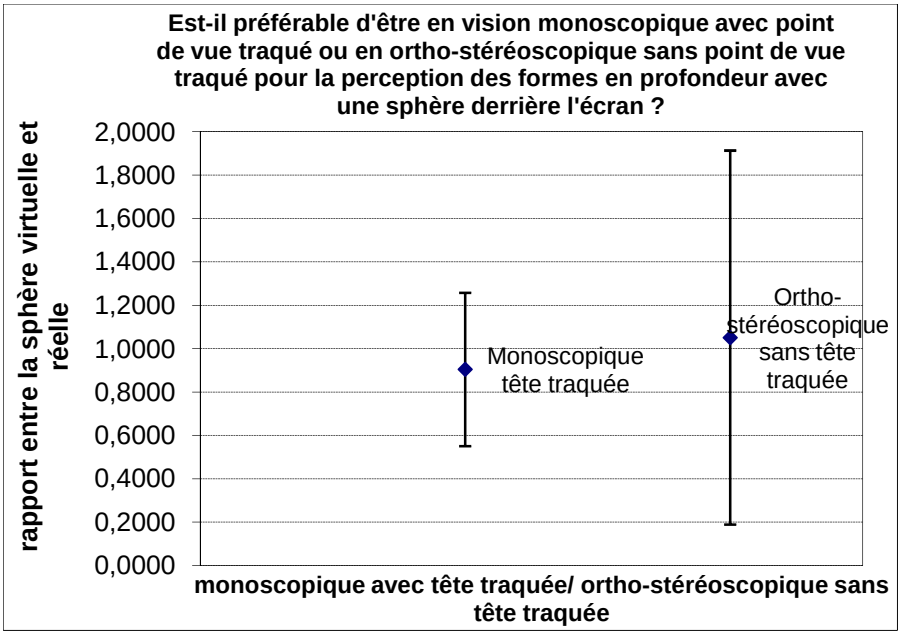


Figure 102 : comparaison entre la perception des formes en profondeur avec une sphère derrière l'écran en vision monoscopique avec point de vue traqué et en vison ortho-stéréoscopique sans point de vue traqué

Les résultats des expériences sur les formes fronto-parallèles sont visibles sur la Figure 103. Les PSE et JND en vision monoscopique avec point de vue traqué sont respectivement 101% et 6% du diamètre de la sphère réelle. Ceux en vision ortho-stéréoscopique sans point de vue traqué sont égaux à 99% et 4%. L'écart de JND est significatif (98%) par contre, aucun autre écart n'est significatif.

tracking	Oui/non
projection	Mono/ortho
forme	Sphère
position	Derrière

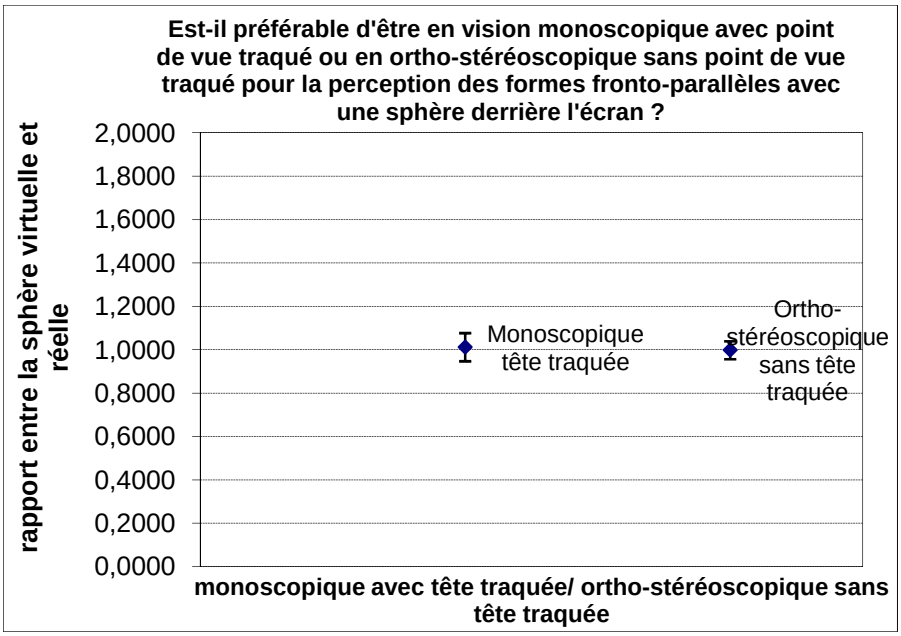


Figure 103 : Comparaison entre la perception des formes fronto-parallèles avec une sphère derrière l'écran en vision monoscopique avec point de vue traqué et en ortho-stéréoscopique sans point de vue traqué

3.3.7.3. Lorsque la sphère est devant l'écran

Les résultats des expériences sur les courbures en profondeur sont exposés sur la Figure 104. Les PSE et JND en vision monoscopique avec point de vue traqué sont 86% et 35% du diamètre initial, tandis qu'en vision ortho-stéréoscopique sans point de vue traqué, ils sont égaux à 99% et 25%. Les probabilités de significativité associées aux écarts entre les JND et les PSE sont respectivement 95% et 74%. Nous voyons donc que l'écart entre les JND est juste significatif.

tracking	Oui/non
projection	Mono/ortho
forme	Sphère
position	Devant

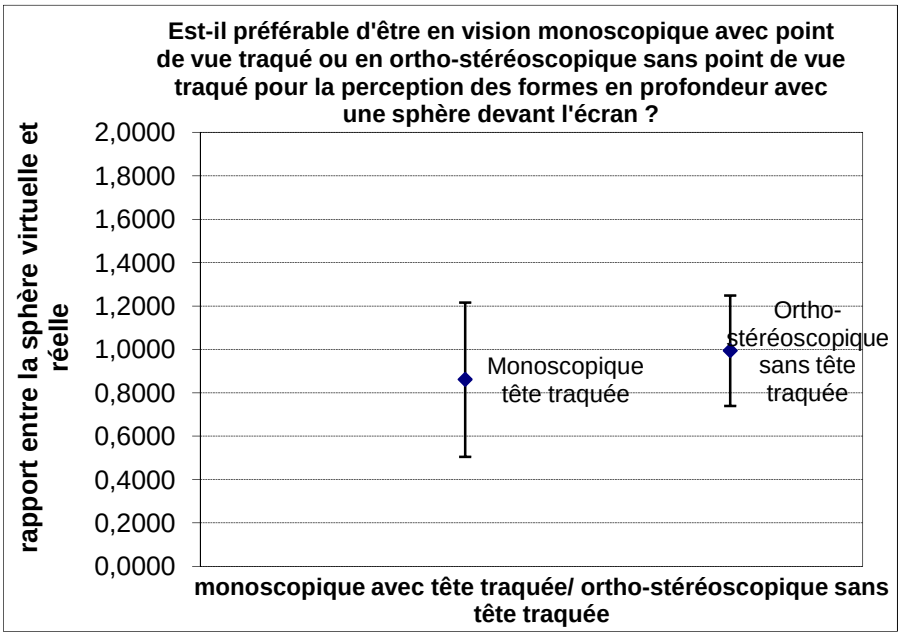


Figure 104 : comparaison entre la perception des formes en profondeur avec une sphère devant l'écran en vision monoscopique avec point de vue traqué et en ortho-stéréoscopique sans point de vue traqué pour

Les résultats des expériences sur les formes fronto-parallèles sont visibles en annexe A. Les PSE et JND en vision monoscopique avec point de vue traqué sont 97% et 11% du diamètre de la sphère réelle, tandis qu'en vision ortho-stéréoscopique sans point de vue traqué, ils sont égaux à 98% et 4%. La différence entre les JND est significative (99,9%), par contre, les autres écarts sont négligeables.

3.3.7.4. Lorsque la forme non connue est sur l'écran

Les résultats de l'expérience sur la perception des formes en profondeur sont présentés sur la Figure 105. Les PSE et JND en vision monoscopique avec le point de vue du sujet traqué sont 97% et 25% ; tandis que lorsque les sujets sont en vision ortho-stéréoscopique sans point de vue traqué, le PSE et le JND sont égaux à 105% et 62% du diamètre initial de la forme. La probabilité de significativité pour l'écart des PSE est 61%, cet écart est donc négligeable. Par contre, celle associée à l'écart de JND est 99,9%, cet écart est donc significatif. Il vaut mieux être en vision monoscopique avec le point de vue traqué qu'en vision ortho-stéréoscopique sans point de vue traqué.

tracking	Oui/non
projection	Mono/ortho
forme	Forme
position	Sur

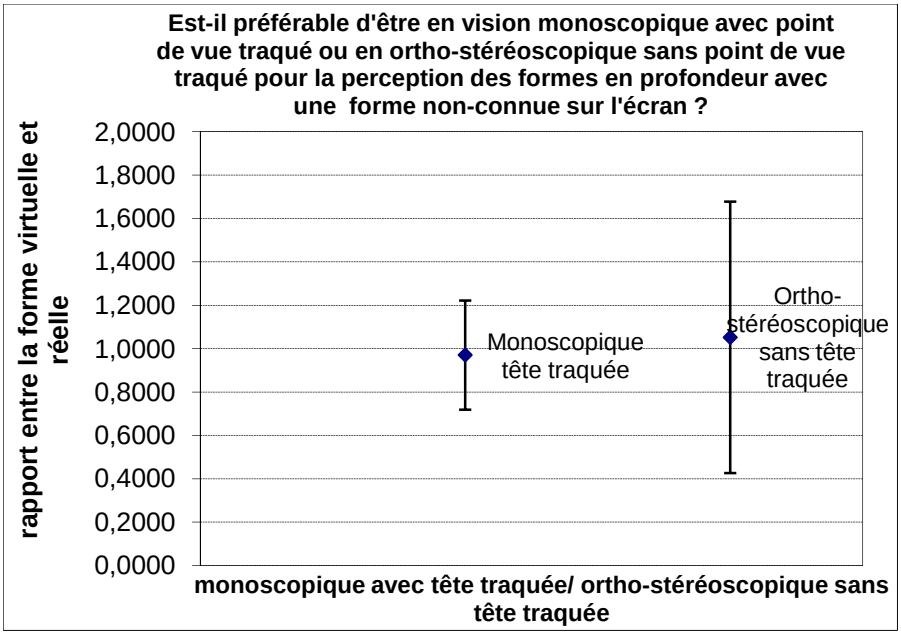


Figure 105 : comparaison entre la perception des formes en profondeur avec une forme non connue sur l'écran en vision monoscopique avec point de vue traqué et en ortho-stéréoscopique sans point de vue traqué

Nous faisons de même pour les formes fronto-parallèles. Les résultats de cette expérience sont exposés en annexe A. Les PSE et JND en vision monoscopique avec point de vue traqué sont 100% et 6% des dimensions étalon, tandis qu'en vision ortho-stéréoscopique sans point de vue traqué, ils sont égaux à 100% et 4%. L'écart de JND est significatif (98%). Pour la vision des formes fronto-parallèles, il vaut mieux être en vision ortho-stéréoscopique sans point de vue traqué qu'en vision monoscopique avec point de vue traqué. Les autres écarts sont négligeables.

3.3.7.5. Lorsque la forme non connue est derrière l'écran

Nous présentons les résultats de l'expérience sur la perception des formes en profondeur sur la Figure 106. Les PSE et JND en vision monoscopique avec point de vue traqué sont respectivement 84% et 64% des dimensions de la forme étalon. En vision ortho-stéréoscopique sans point de vue traqué, les PSE et JND sont égaux à 123% et 123%. La probabilité de significativité de l'écart des JND est 99,9%, c'est dont très significatif. La différence de PSE quant à elle n'est pas significative (91%).

tracking	Oui/non
projection	Mono/ortho
forme	Forme
position	Derrière

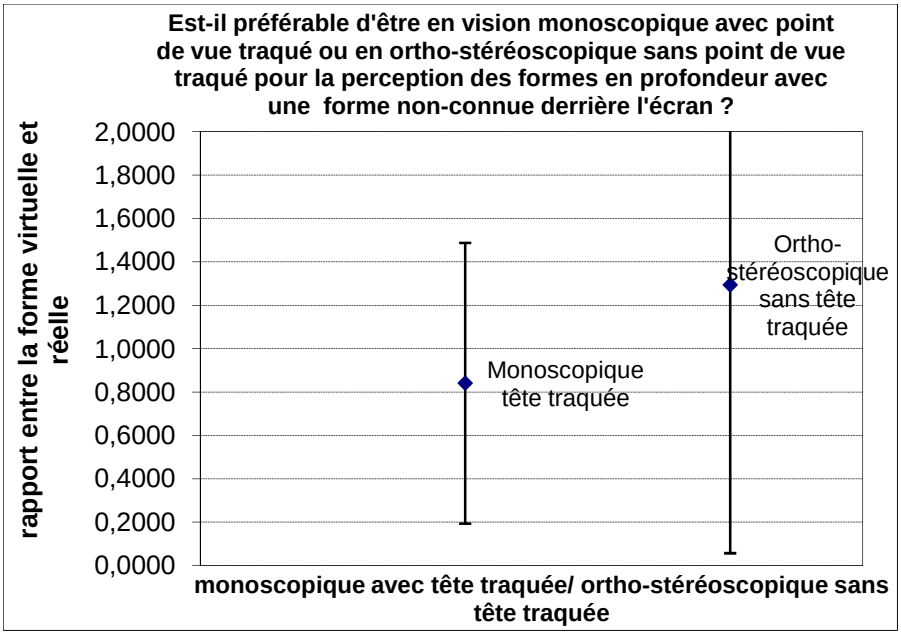


Figure 106 : comparaison entre la perception des formes en profondeur avec une forme non connue derrière l'écran en vision monoscopique avec point de vue traqué et en ortho-stéréoscopique sans point de vue traqué

Nous pouvons faire la même expérience pour les formes fronto-parallèles. Les résultats sont exposés en annexe A. Les PSE et JND en vision monoscopique avec point de vue traqué sont respectivement 100% et 7% des dimensions initiales tandis qu'avec une vision ortho-stéréoscopique, les PSE et JND sont 101% et 4,5%. La perception des formes fronto-parallèles est meilleure lorsque les sujets sont en vision ortho-stéréoscopique sans point de vue traqué et cet écart est significatif (98%). Les écarts de PSE et entre la perception des courbures horizontales et verticales ne sont pas significatifs.

3.3.7.6. Lorsque la forme non connue est devant l'écran

Les résultats de l'expérience sur la perception des formes en profondeur sont présentés sur la Figure 107. Les PSE et JND en vision monoscopique avec point de vue traqué sont égaux à 86% et 37% du diamètre initial de la forme. En vision ortho-stéréoscopique sans point de vue traqué, les PSE et JND sont respectivement égaux à 96% et 36%. Ces écarts ne sont pas significatifs. Il est probable que les disparités horizontales soient suffisamment fines pour permettre de se passer de parallaxe de mouvement.

tracking	Oui/non
projection	Mono/ortho
forme	Forme
position	Devant

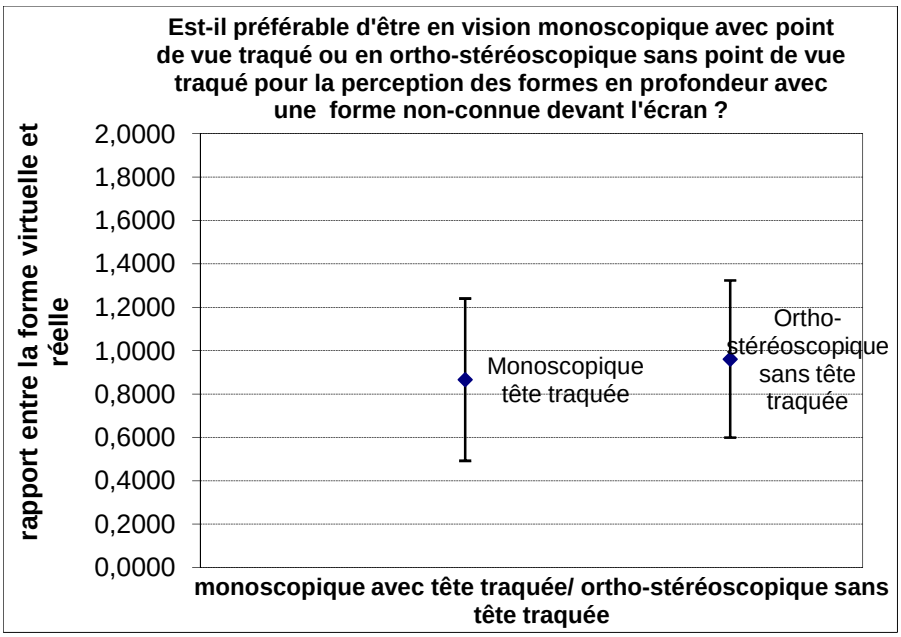


Figure 107 : est-il préférable d'être en vision monoscopique avec point de vue traqué ou en ortho-stéréoscopique sans point de vue traqué pour la perception des formes en profondeur avec une forme non connue devant l'écran ?

La différence de perception des formes fronto-parallèles n'étant pas significative, nous plaçons nos résultats en annexe A.

3.3.7.7. Conclusion sur l'hypothèse n°6

Globalement, il vaut mieux être en vision monoscopique avec le point de vue traqué qu'en vision ortho-stéréoscopique sans point de vue traqué pour la perception des formes en profondeur. Nous relevons une amélioration du JND de 24% (51% derrière, 33% sur et -10% devant l'écran) de la dimension de la sphère et 31% (59 derrière, 37% sur et -1% devant l'écran) pour la forme non connue a priori. Cependant, lorsque les formes sont devant l'écran, il semblerait que cette amélioration soit moins évidente, les disparités horizontales sont probablement mieux définies (puisque la pixellisation est beaucoup plus fine), et les sujets arrivent à se passer plus facilement des parallaxes de mouvement.

Les déformations décelables en vision monoscopique avec point de vue traqué sont, pour une sphère, 30% (donc 6cm) et pour la forme non connue 42% (8,4cm). Lorsque les sujets sont en vision ortho-stéréoscopique sans point de vue traqué, ces déformations percevables sont 54% (10cm) pour une sphère et 73% (15cm).

La perception des formes fronto-parallèles semble meilleure en vision ortho-stéréoscopique sans point de vue traqué. Les points d'égalisation subjective quant à eux ne sont pas modifiés dans un cas comme dans l'autre.

Nous pourrions penser que les sujet mets plus de temps à percevoir les modifications de forme lorsqu'il a le point de vue traqué puisqu'il tourne autour du l'objet. Cependant, les sujet mettaient généralement plus de temps à se décider lorsque le point de vue n'était pas traqué qu'à se déplacer pour avoir un autre point de vue. Comme nous le voyons sur la Figure 108, la moyenne du temps passé à la tâche lorsque le point de vue est traqué en vision monoscopique est plus faible qu'en vision stéréoscopique avec point de vue traqué. Cette différence est significative (probabilité

de significativité : 99%). Il est donc préférable d'être en vision monoscopique avec point de vue traqué qu'en vision stéréoscopique sans point de vue traqué, si l'on veut minimiser le temps d'une tâche.

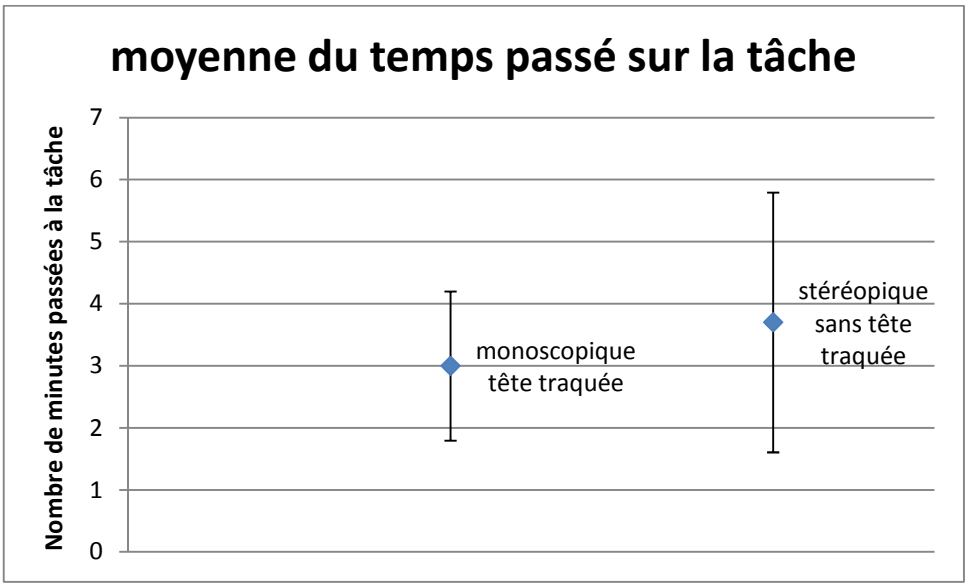


Figure 108 : moyenne de temps passé sur la tâche

3.4. Conclusion

Nous avons exposé dans ce chapitre notre protocole d'expérimentation, certaines méthodes de calcul des données psychophysiques et statistiques ainsi que tous nos résultats sur la perception des formes fronto-parallèles et en profondeur selon les différents modes de projection, les différentes formes présentées et les différentes positions de celles-ci. Dans le chapitre suivant, nous discutons de ces résultats.

Chapitre 4 : Discussion des résultats sur la perception des formes en relief en immersion

Dans ce chapitre, nous étudions et nous discutons les résultats obtenus lors de nos expériences. Nous les comparons aux autres expériences fondamentales déjà menées. Nous finissons en discutant des limitations de nos résultats en fonction de l'ordre des séquences ou des limitations techniques.

4.1. Interprétation des résultats

4.1.1. Interprétation générale.

4.1.1.1. Influence du point de vue traqué

Pour bénéficier de parallaxes de mouvement, la caméra virtuelle doit se situer au même endroit que le point de vue de l'observateur réel, nous parlons alors de point de vue traqué. Nous avons vu au chapitre précédant que ce paramètre est très important pour la perception de formes et spécifiquement pour celles des courbures quelque soit la distance entre les caméras, la forme observée et la position de celle-ci. En effet, nous avons remarqué une amélioration du JND moyenne de 46% du diamètre en profondeur d'une sphère et 54% pour une forme non connue a priori lorsque le point de vue était traqué. Ceci signifie que lorsque le point de vue n'est pas traqué, une modification beaucoup plus grande peut être appliquée sur la courbure en profondeur sans que le sujet ne le remarque. A l'opposé, lorsque le point de vue de l'observateur est traqué, il détectera beaucoup plus rapidement la moindre modification sur cette courbure.

Cette amélioration de la perception des formes grâce au point de vue traqué est due au fait que le sujet bénéficie des parallaxes de mouvement. Celles-ci permettent une perception des profondeurs très rapide et améliorent la perception des courbures. De plus, les informations visuelles concordent avec les informations du système vestibulaire. Ceci permet une meilleure perception, grâce aux informations supplémentaires. Un système de point de vue traqué permet aux sujets de « tourner » autour de la forme virtuelle. Celui-ci peut donc avoir plusieurs points de vue tout en « sachant » que c'est le même objet, comme nous l'avions évoqué au paragraphe 2.1.3.2. Ainsi, lorsqu'ils sont traqués, les sujets contournent presque systématiquement la forme virtuelle pour voir les modifications de profondeur. Ce changement de point de vue leur permet de percevoir plus facilement les différences de courbures.

Nous avons également remarqué que lorsque nous étudions la perception de la forme non symétrique, les sujets sont moins capables de percevoir une modification de la forme en profondeur sans point de vue traqué qu'avec la sphère. En effet, sur la sphère, les reflets des lumières sont symétriques et forment des points de repère aisés à utiliser pour la perception parce qu'ils sont réguliers (voir paragraphe 2.1.2.3). Ceux de la forme non primitive ne sont pas symétriques, les sujets ont beaucoup plus de mal à les utiliser.

Le point de vue traqué est donc très important pour la perception des formes. Nous avons par ailleurs déterminé, grâce aux coefficients de significativité, que c'est le paramètre le plus important, que ce soit pour la sphère, ou pour la forme non symétrique. Les personnes souhaitant que leur système d'immersion reflète correctement la perception des formes doivent donc se concentrer en priorité sur le suivi du point de vue de leurs utilisateurs par la caméra virtuelle.

Nous avons également remarqué que lorsque la sphère est devant l'écran et que le sujet est en vision stéréoscopique, il a tendance à se servir plus des disparités horizontales (plus marquées à cause de la technique) et à moins utiliser les parallaxes de mouvements, nous développerons cet aspect dans les paragraphes suivants.

Pour ce qui concerne la dernière hypothèse (Est-il préférable d'être en vision monoscopique avec point de vue traqué ou en vision ortho-stéréoscopique sans point de vue traqué ?), nous avons vu que les parallaxes de mouvements peuvent très bien « remplacer » les disparités horizontales tout en laissant intacte (voire en l'améliorant) la perception des courbures en profondeur. Lorsque deux personnes sont dans une salle immersive équipée d'un système de traqueur capable de suivre deux têtes en même temps et d'un système de projection stéréoscopique mono-utilisateur, il serait préférable de ne pas projeter les images stéréoscopiques liées au point de vue d'un des deux utilisateurs, mais de traquer les deux personnes et de projeter les deux points de vue de chaque utilisateur en monoscopique. Cela améliorerait la perception des courbures de la personne dont le point de vue n'était pas traqué, et cela lui éviterait un cyber-malaise parfois très prononcé si l'observateur traqué bouge beaucoup (voir paragraphe 5.1.2.3).

4.1.1.2. Influence de la distance inter-oculaire de projection

Nous remarquons qu'il n'y a pas une grande différence de perception des formes en relief entre la vision ortho-stéréoscopique et la vision non-ortho-stéréoscopique lorsque le point de vue est traqué. Il semble que la perception des courbures en profondeur soit légèrement améliorée (en moyenne 5,61% sur les deux formes, cette amélioration étant légèrement plus forte pour la sphère) lorsque le sujet est en vision non-ortho-stéréoscopique. Il est probable que l'augmentation artificielle des disparités horizontales permette dans une certaine mesure de percevoir un changement de profondeur en chaque point plus rapidement. Notons que lorsque l'objet est sur l'écran, il n'y a logiquement aucune différence de perception entre les deux visions, les différences de disparités étant masquées par la pixellisation de notre projecteur. Lorsque le point de vue n'est pas traqué, l'interprétation est plus difficile. En effet, nous avons vu au paragraphe précédant que la perception des formes en relief est extrêmement compliquée lorsque le point de vue n'est pas traqué, il semblerait donc que les légères modifications dues au changement de distance entre les caméras virtuelles soient masquées par cet effet. Cependant, nous avons remarqué que la perception des courbures en profondeur semblait améliorée pour la sphère en vision ortho-stéréoscopique et au contraire, détériorée pour la forme non connue a priori.

Cependant, il faut tout de même noter que nous avons étudié la perception des formes et non celle des distances ou des tailles. En effet, lorsque les objets étaient présentés en vision non-ortho-stéréoscopique, beaucoup de testeurs (plus de la moitié) nous ont demandé de confirmer que nous étudiions bien la courbure, donc la relation entre les dimensions et non la taille (ou les dimensions absolues). Cette question était presque systématique lorsque l'objet était présenté devant l'écran : les sujets le percevaient beaucoup plus grand que lorsque la vision stéréoscopique était bien réglée. Ceci semble logique au vu de la relation entre la distance inter-oculaire et les distances ou les tailles évoquées au paragraphe 2.2.3. Mais nous constatons que notre expérience n'a pas relevé de différence pour la perception des formes.

Peut-être que la distance inter-oculaire n'était pas suffisamment différente de celle des sujets pour voir apparaître une différence très marquée comme celle représentée sur la 47. Mais nous

avons estimé que, dans le contexte qui nous concerne, l'espacement des caméras virtuelles n'était pas si différent de la distance entre les yeux des observateurs. En effet, une erreur courante en immersion, est de régler la distance entre les caméras virtuelles sur la DIO moyenne (6,3cm, parfois même arrondi à 6,5cm). Or, il est extrêmement rare qu'une personne ait une DIO très différente de 6,5cm (en général entre 5 et 7cm). En conséquence, lorsque beaucoup de personnes sont en immersion en même temps, nous pouvons régler la distance entre les caméras sur la DIO moyenne sans compromettre la perception des formes. Par contre, cela modifiera la perception des tailles et des distances pour les personnes dont la DIO est éloignée de la moyenne. Pour une personne seule en immersion, nous conseillons évidemment d'adapter la distance entre les caméras pour conserver la perception des tailles (en prenant la DIO pour la convergence la plus probable pour la scène observée).

4.1.1.3. La forme en elle-même

Nous avons vu que les modifications sur la sphère sont globalement beaucoup plus rapidement perçues, les sujets se référant probablement à leur standard interne. Ils « savent » ce qu'est une sphère et donc lorsque celle-ci est modifiée, ils le perçoivent plus rapidement. Cependant, lorsque nous trouvons qu'un paramètre améliore la perception des courbures pour la sphère, nous trouvons presque à chaque fois la même amélioration pour la forme non connue.

4.1.1.4. Influence de la position de la forme virtuelle par rapport à l'écran de projection

Lorsque nous projetons une forme située devant l'écran, l'image projetée est plus grande que la taille réelle de la forme. Inversement, lorsque nous projetons un objet virtuel derrière l'écran, l'image projetée est au contraire beaucoup plus petite que sa taille réelle. Considérons le cas où le diamètre de la projection de la sphère est de 34 cm lorsqu'elle est située devant l'écran, 20 cm lorsque la sphère est sur l'écran et 14 cm derrière l'écran. Nos pixels font 1,61mm de coté, donc la projection de la sphère placée devant l'écran a un diamètre constitué de 340/1.61 pixels, soit 211 pixels, tandis que la sphère sur l'écran a un diamètre de projection de 124 pixels et celle qui est derrière l'écran, 87 pixels, pour un même diamètre perçu par l'observateur. Ceci signifie que la pixellisation de la forme qui est devant l'écran est moindre que celle de la forme sur l'écran et a fortiori que celle qui se situe derrière l'écran. Il est possible que cette pixellisation, qui n'existe pas dans la vision naturelle, modifie la perception des formes et détériore celle qui a la plus forte pixellisation. En effet, nous avons vu que lorsque nous étudions l'influence des disparités horizontales sur la perception de la courbure, le fait de placer l'objet devant l'écran permet au sujet de mieux percevoir les disparités horizontales, voire même ne plus avoir trop besoin des parallaxes de mouvement (Figure 109).

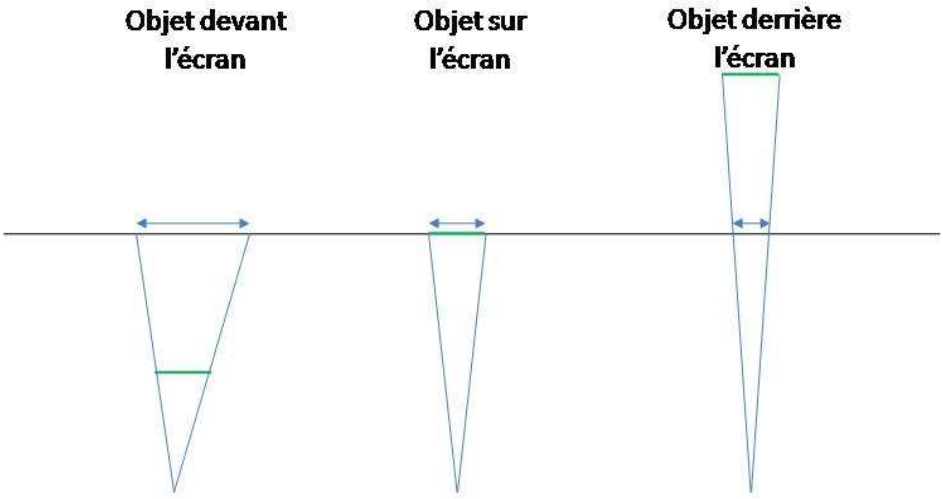


Figure 109 : différence de taille de projection en fonction de la position de l'objet virtuel

D'autre part, lorsque le point de vue est traqué, un sujet peut plus facilement tourner autour d'un objet qui se situe devant l'écran qu'autour d'un objet qui se situe derrière celui-ci. En effet, si un sujet se déplace sur la gauche de l'objet, lorsque celui-ci est situé devant l'écran, il n'est limité que par la projection de l'objet qui sort de l'écran s'il va trop loin. Par contre, lorsque l'objet est derrière l'écran, le sujet ne peut évidement pas se placer sur sa gauche, puisque celle-ci se trouve derrière l'écran où le sujet ne peut se positionner physiquement. Cela signifie que, physiquement, un objet placé derrière l'écran (ou même sur l'écran) diminue les angles de point de vue différents par rapport à un objet placé devant l'écran (Figure 110). Or, plus l'éventail d'angles de points de vue est grand, plus la perception des formes est performante (voir paragraphe 2.1.2.2).

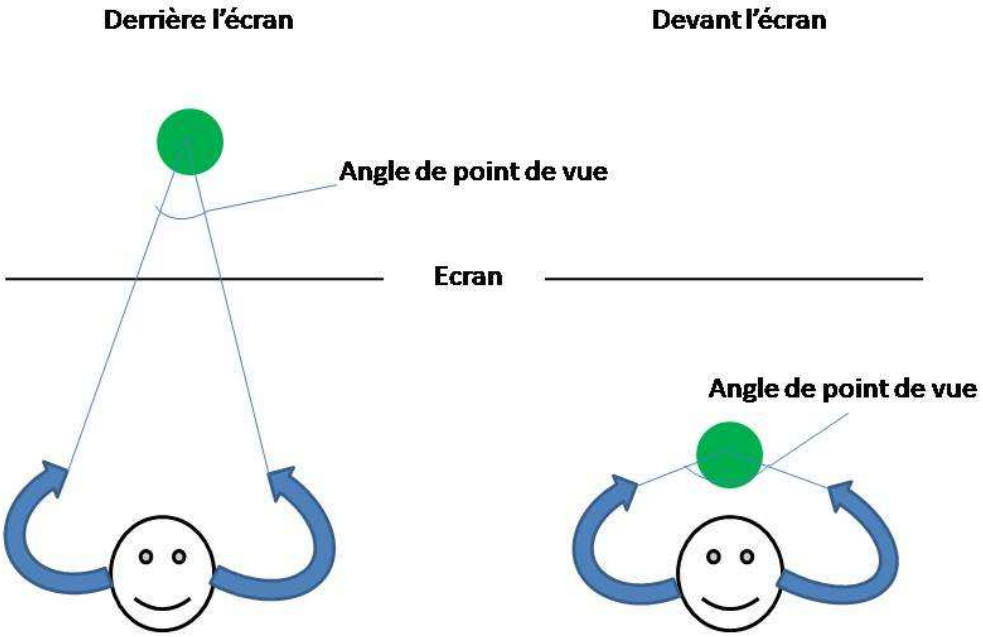


Figure 110 : différence d'angle de point de vue pour une forme devant ou derrière l'écran

Cependant, il faut toutefois relever que dans ces conditions, si la perception des formes est améliorée pour une forme devant l'écran, la fatigue visuelle associée à la vision de celle-ci est également augmentée. En effet, un objet qui n'est pas présenté sur l'écran a forcément des disparités horizontales plus grandes et génèrent un conflit accommodation-vergence plus grand. Ce sont deux causes bien connues de fatigue visuelle (voir le chapitre 5). En conséquence, nous conseillons de placer les objets devant l'écran pour une meilleure perception des formes lorsque l'immersion n'est pas trop longue. Au-delà d'une certaine durée d'immersion, il faudra choisir entre placer l'objet devant l'écran et donc garder une bonne perception tout en fatiguant les yeux de l'observateur, ou placer l'objet sur l'écran, perdre un petit peu de perception des formes mais en stressant moins le système visuel du sujet.

4.1.2. Comparaison avec d'autres expériences

4.1.2.1. Expérience de Johnston

Nous voyons quelques similitudes entre nos expériences et l'expérience de Johnston (Johnston, 1991): nous comparons la courbure présumée circulaire dans un environnement immersif sur fond noir. Cependant, il y a néanmoins quelques différences : Johnston n'étudiait que les modifications de la courbure en profondeur, nous comparons également les modifications de courbures horizontales et verticales. En effet, ce ne sont pas les mêmes mécanismes psychophysiques qui entrent en jeu pour la courbure en profondeur et celle en largeur ou en hauteur, mais les deux sortes de courbures sont importantes pour percevoir une forme globale dans un environnement immersif ou non.

D'autre part, une grande différence par rapport à l'expérience de Johnston porte sur la position relative de la forme virtuelle par rapport à l'écran En effet, Johnston reculait l'écran en même temps que son cylindre virtuel pour garder des disparités horizontales faibles. Par contre, ce faisant, il modifiait ses disparités verticales. Nous ne modifions pas la position de l'écran par rapport à celle de la forme virtuelle, mais nous avons exclu toute disparité verticale artificielle en mettant les caméras virtuelles parallèles (voir 1.2.3.5). D'autre part, dans nos expériences, ce n'est pas la distance entre la forme virtuelle et l'observateur qui est modifiée lorsque l'observateur est fixe, mais bien la distance entre l'écran et la forme virtuelle. Malgré cette différence, nous avons tenu compte de ses résultats pour placer l'observateur fixe. Selon son étude, un objet situé à un mètre est perçu de façon optimale : à moins d'un mètre, il est surévalué, à plus d'un mètre il est sous-évalué. Cependant, pour notre expérience, 1 mètre entre le sujet et la forme aurait provoqué de très grandes disparités, donc nous avons choisi de mettre la forme à 1,20m. De plus, cette valeur permet de tourner plus aisément autour de la forme. Cette augmentation n'a pas eu de grande influence puisque nous voyons que les points d'égalisation subjective de nos expériences sont tous proches de la dimension étalon.

4.1.2.2. Expérience de Glennerster et Luo

Nous avons choisi un fond noir, sans aucun indice de profondeur. Or, Glennerster avait critiqué Johnston pour la même raison, évoquant le fait que les indices supplémentaires contenus dans un fond texturé améliorent la perception des courbures (Glennerster, et al., 1994). Cependant, nous avons préféré garder un fond noir parce qu'un fond texturé est extrêmement difficile à reproduire en réel. Si Glennerster ne faisait de comparaison que sur des cylindres circulaires pour se servir du standard interne des sujets, la moitié de notre expérience est justement basée sur le fait que nos sujets n'avaient pas de standard interne pour notre forme inconnue. Par conséquent, il faut que le monde virtuel et le monde réel de comparaison soient les plus proches possible, ce qui nous restreint sur la texture du fond et donc sur les indices de profondeur supplémentaires.

Luo, dans son expérience avec une bouteille de coca-cola, utilisait également le standard interne de ses sujets (Luo, et al., 2007). En effet, la plupart des sujets (sinon la totalité) connaissaient la forme d'une bouteille de coca-cola avant de faire le test. Il n'avait donc pas besoin d'une comparaison forte entre le monde réel et le monde virtuel. Cependant, si nous pouvions recréer exactement le même fond texturé en réel et en virtuel, il serait très intéressant de refaire ces expériences avec un fond texturé pour pouvoir comparer les résultats avec les nôtres.

4.1.2.3. Expérience de Durgin

L'expérience de Durgin montrait que la perception des courbures en profondeur était plus précise et plus aisée avec les disparités binoculaires qu'avec les parallaxes de mouvement (Durgin, et al., 1995). Or nous arrivons au résultat inverse. Ceci peut s'expliquer de deux façons. Les expériences de Durgin étaient réalisées entièrement en réel avec des cônes réels. Les nôtres sont une comparaison entre des formes réelles et virtuelles, mais les variations de formes ne se font qu'en virtuel. Cela implique d'une part que Durgin n'a, à aucun moment, de conflit entre l'accommodation et la vergence et qu'il n'a pas non plus de problème de disparités verticales modifiées. Puisque nos variations étudiées sont virtuelles, nous avons des disparités verticales inexistantes (voir 1.2.3.5) et un conflit accommodation-vergence qui dépend de l'emplacement de la forme virtuelle (plus la forme virtuelle est loin de l'écran, plus le conflit sera grand). Cette différence explique probablement que la perception des formes basée sur les disparités binoculaires de notre expérience est moins bonne que celle relevée lors de son étude.

D'autre part, dans ses expériences, Durgin ne se servait pas des informations données par le système vestibulaire. En effet, ses sujets étaient fixes et les cônes se déplaçaient devant eux selon un angle de vue assez limité. Dans notre expérience, nos sujets se déplacent et donc se servent de leur système vestibulaire pour leur apporter des informations supplémentaires permettant d'avoir une

vision plus cohérente avec ce qu'ils ressentent. D'autre part, ils peuvent se déplacer autour des objets comme ils le souhaitent. Donc, généralement, les différences de point de vue sont beaucoup plus grandes que dans l'expérience de Durgin et donc beaucoup plus informatives. Ces deux différences fondamentales expliquent probablement pourquoi nous obtenons de telles différences de résultats.

4.2. Les limites de l'expérience

4.2.1. Ordre d'exécution des séquences

Malheureusement, les séquences n'étaient pas présentées aléatoirement et, pour certains résultats, cela peut être gênant. En effet, les séquences présentant les formes avec une distance inter-pupillaire double par rapport à celle du sujet, sont toujours à la fin du test. Le sujet a déjà été confronté à la forme et la connaît donc mieux que lors des séquences avec une distance inter-pupillaire normale, ou nulle (vision monoscopique). Cela peut modifier légèrement les résultats.

Cependant, il faut remarquer que les séquences sans point de vue traqué se déroulent toujours après les séquences avec tête traquée et que nous voyons bien que, malgré cela, les JND de ces séquences sont beaucoup plus grands. Il serait bon de refaire ces expériences en présentant les séquences de façon totalement aléatoire sans repêchage. Ainsi, nous éviterions tout phénomène d'apprentissage et les résultats ne favoriseraient plus une configuration plutôt qu'une autre.

4.2.2. Interface visuelle

4.2.2.1. Validité pour une autre configuration d'immersion

Ces résultats sont obtenus pour des stimuli présentés dans notre salle immersive. Il est envisageable que ce ne soit pas exactement les mêmes avec une autre technique de présentation. Par exemple, nous ne voyons pas comment reproduire ce test avec un casque immersif. En effet, la forme réelle doit être toujours visible, or, avec un casque, l'utilisateur ne voit pas le monde réel qui l'entoure. Il pourrait l'enlever chaque fois qu'il souhaite se référer à la forme réelle, mais dans ce cas, les réglages ne seraient plus corrects (un casque immersif est très délicat à régler pour un sujet). Nous pourrions prendre un casque semi-transparent qui permette de voir le monde réel en même temps que le monde virtuel.

L'expérience dans un CAVE (une immersion dans un cube à 5 faces) serait probablement comparable à celle que nous avons menée. Cependant, il faut noter qu'il serait alors plus compliqué de séparer les lumières du monde virtuel de celles du monde réel. En effet, comme les 5 faces du cube renvoient un petit peu de lumière, certains rayons lumineux arriveraient sur la forme réelle. Il faudrait donc faire attention aux ombres et jeux de lumières sur la forme réelle qui pourraient être modifiés. Par contre, dans une telle installation, nous pourrions étudier la composante du point de vue traqué plus en détail : le sujet pourrait vraiment tourner autour de la forme virtuelle. En effet, avec un seul écran (comme dans notre installation), pour compenser les effets pseudoscopiques, la forme virtuelle devrait être projetée hors de l'écran lorsque le sujet tourne trop autour de la forme. Mais lorsque nous avons des écrans latéraux, la forme virtuelle peut être projetée hors de l'écran frontal. Il est alors possible de tourner autour de la forme virtuelle, pour autant que la projection de cette dernière ne soit pas sur la forme réelle.

4.2.2.2. Influence de la taille des pixels

Nous avons vu qu'une variation de forme sur objet virtuel situé derrière l'écran est moins bien perçue que sur un objet situé sur ou devant l'écran. Il est fort possible que cette différence soit due à la pixellisation, inévitable, de l'image. Nous devrions refaire des tests identiques avec une réso-

lution d'écran beaucoup plus fine pour vérifier si ce paramètre influence réellement la perception des objets éloignés de l'écran.

Des expériences ont été menées par Alexis Paljic pour savoir si la finesse de la pixellisation de la sphère virtuelle, placée au niveau de l'écran, pouvait influencer la perception de sa forme. Il a comparé cette perception pour des pixels de 1,61mm et pour des pixels de 0,67mm. Il semblerait qu'il n'y ait aucune différence significative entre ces deux projections. Nous pouvons donc dire que la différence de perception entre la forme devant l'écran ne dépend pas de la pixellisation. Cependant, nous avions évoqué que la pixellisation pouvait surtout influer sur la perception de la forme virtuelle lorsqu'elle était derrière (ou sur) l'écran, il serait donc intéressant de refaire la même expérience avec une sphère située en retrait par rapport à l'écran de projection.

Partie 2 : Diminution de la Fatigue visuelle en vision stéréoscopique

Chapitre 5 : Etat de l'art de la fatigue visuelle en vision stéréoscopique

La fatigue visuelle est un terme sous lequel on regroupe un ensemble de phénomènes réversibles qui se manifestent généralement par des symptômes oculaires et visuels ainsi que des symptômes généraux peu spécifiques :

- Symptômes oculaires et visuels :
  - sensation de lourdeur,
  - picotements, brûlures,
  - rougeur des globes oculaires ou des paupières,
  - larmes,
  - paupières qui tremblent,
  - vision trouble ou voilée,
  - myopie temporaire,
  - quelques fois vision double,
  - éblouissement.
- Symptômes généraux :
  - céphalées frontales,
  - sensations vertigineuses,
  - mal à la nuque et à la tête pendant et après le travail

Généralement, ces symptômes disparaissent après une nuit de sommeil. Cependant, il peut arriver parfois que la fatigue visuelle s'accumule sur plusieurs jours (Albert, et al., 2002).

Il ne faut pas confondre la fatigue visuelle et la fatigue oculaire. Cette dernière donne lieu à des yeux qui piquent, des irritations et des sensations d'œil sec. Ces symptômes sont causés par une insuffisance de sécrétion lacrymale. Nous clignons de l'œil entre 12 et 20 fois par minute : ceci permet la formation d'un film lacrymal de façon continue protégeant la surface de l'œil. Le travail sur écran peut donner lieu à une diminution de cette fréquence de clignement et donc peut assécher la surface des yeux.

Lors d'une projection stéréoscopique, nous essayons de reproduire la vision binoculaire, nous en exploitons le principe, mais ce n'est pas exactement pareil. Il y a des différences et celles-ci peuvent provoquer des fatigues visuelles. Dans ce chapitre, nous abordons les différentes causes de la fatigue visuelle en stéréoscopie. Nous traitons d'abord les causes directement liées au principe de la vision en relief, ensuite les causes purement techniques et enfin nous étudierons en détail l'influence des hautes fréquences sur la fatigue visuelle.

5.1. Causes de la fatigue visuelle en stéréoscopie

5.1.1. Problèmes inhérents au principe de la vision stéréoscopique

5.1.1.1. Les problèmes d'accommodation – vergence

Nous avons vu au paragraphe 1.2.3.4 que la vergence et l'accommodation sont liées en vision binoculaire. Le modèle théorique d'interaction entre l'accommodation et la vergence le plus connu est celui énoncé par Schor en 1986. Il explique que la vergence et l'accommodation ont trois composantes : la première est une composante caractérisée par une réponse rapide, elle est dite phasique. La seconde est une réponse plus lente et appelée tonique. Ces deux composantes sont soumises à des rétroactions qui permettent de les contrôler. Ce sont précisément ces boucles de feedback dont nous parlions au 1.2.1.3 qui permettent de percevoir le relief. En plus de ces deux composantes, nous voyons sur la Figure 111 qu'il existe des composantes d'interaction : la vergence accommodative et l'accommodation de vergence (Schor, 1986).

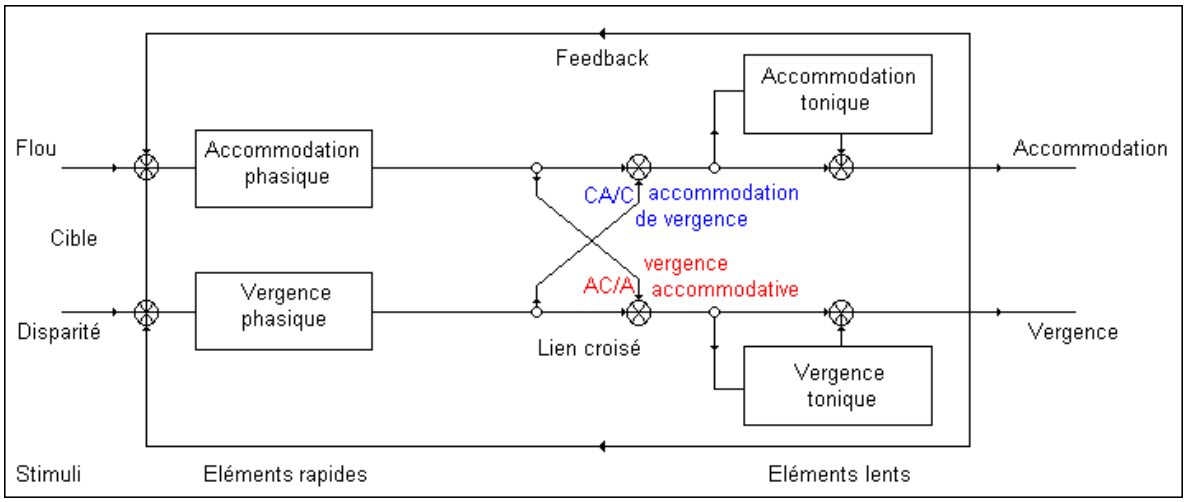


Figure 111 : interaction entre l'accommodation et la vergence, adaptation du modèle de Schor (Neveu, 2008)

L'accommodation et la vergence sont donc profondément liées l'une à l'autre. En effet, pour une accommodation imposée, notre système visuel ajuste sa vergence et vice-versa. L'accommodation entraîne la convergence et une baisse d'accommodation entraîne une divergence (Neveu, 2008).

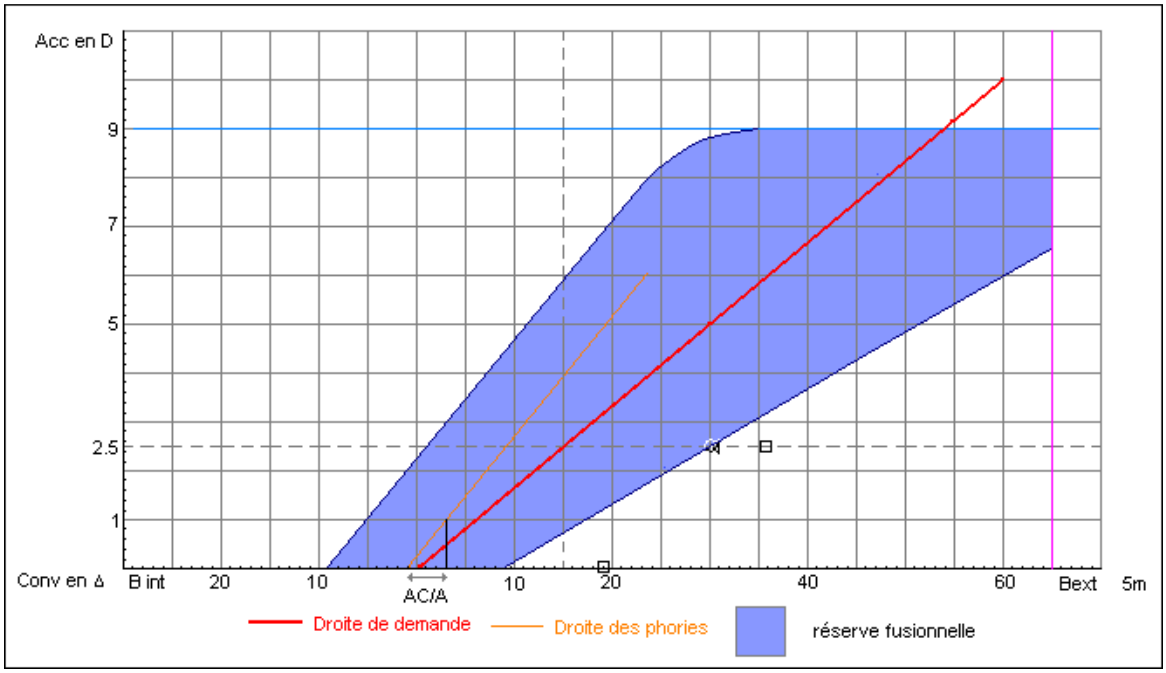


Figure 112 : relation accommodation - vergence. La droite de demande est la convergence qui devrait être mise en jeu pour une quantité d'accommodation donnée et réciproquement (Neveu, 2008)

Nous pouvons donc tracer une droite de demande en accommodation/vergence en situation normale (Figure 112). Autour de cette droite, nous trouvons les points dont l'accommodation ne correspond pas à la vergence demandée. Si de telles accommodation et vergence se trouvent dans la réserve fusionnelle (ou zone fusionnable), le point sera vu unique au prix d'un certain effort. Dans le cas contraire, l'accommodation et la vergence ne correspondent pas suffisamment pour que le cerveau fasse le lien : il ne fusionne pas l'image gauche et l'image droite.

La relation accommodation - vergence peut être mise en défaut artificiellement lorsque nous regardons des images stéréoscopiques sur un écran : un objet virtuel positionné en arrière de l'écran fait diverger les axes optiques sur lui tandis que les yeux accommodent sur l'écran (Figure 25). Cette modification de la relation accommodation - vergence est susceptible de créer des difficultés visuelles et de la fatigue pour l'observateur, sauf lorsque l'objet se situe au niveau de l'écran. Si nous nous referons à la Figure 112, l'accommodation fixe correspond à une droite horizontale. La vergence tolérable (donc la disparité tolérable) correspond aux deux points d'intersection entre la frontière de la zone fusionnable et cette ligne droite. Nous voyons donc que, en fonction de la distance de l'écran, l'écart de vergence toléré est plus ou moins grand. Notons que plus nous nous rapprochons de la limite de la zone fusionnable, plus nous fatiguons notre système visuel (Yano, et al., 2002) (Rushton, et al., 1999).

5.1.1.2. Les problèmes liés à la stéréopsie

Disparités horizontales

La parallaxe angulaire horizontale (et donc la disparité horizontale rétinienne) doit être petite pour ne pas créer de grandes difficultés pour la vision stéréoscopique, chaque point à fusionner devant être contenu dans l'aire de Panum. Dans le cas contraire, le cerveau est incapable de fusionner les deux images ou le fait avec une certaine fatigue. L'objectif est donc de créer des images stéréoscopiques avec de faibles parallaxes horizontales tout en gardant un effet de profondeur.

La limite de fusion n'est pas symétrique en avant ou en arrière de l'écran, mais nous ne pouvons ériger une règle générale. En effet, cette dissymétrie dépend des sujets, certaines personnes fusionnent plus facilement des points de fortes disparités négatives, d'autres fusionnent plus aisément des disparités positives. Elle dépend également de la disparité de fixation. C'est le fait que certains sujets n'arrivent pas à converger exactement sur un point précis. L'horoptère passe un peu devant ou derrière le point, mais pas suffisamment loin que pour le sujet s'en aperçoive. Cette disparité de fixation explique également la dissymétrie de l'aire de Panum (Richards, 1971) (Manan, et al., 2002).

Il y a encore beaucoup de discussion de l'effet du temps sur la limite de fusion. Mitchell et Palmer montrèrent que cette limite n'était pas affectée par un temps d'exposition de 10 à 120 ms (Mitchell, 1966) (Palmer, 1961). Woo a trouvé que la limite de fusion pour une barre verticale augmente de 2 à 4 min arc lorsque la durée d'exposition augmente de 5 à 100 ms (Woo, 1974). Les raisons de ces contradictions restent encore inconnues. Schor et Tyler ont exploré les liens entre la limite de fusion et les caractéristiques spatio-temporelles des stimuli présentés. Ils ont présentés deux lignes verticales ondulantes en opposition de phase à une fréquence de 0,1 à 5 Hz (Figure 113). L'aire de Panum décroît fortement lorsque la fréquence temporelle augmente (Schor, et al., 1981).

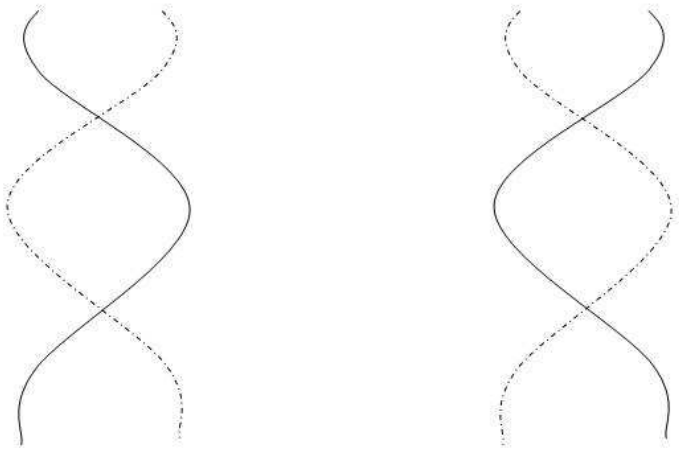


Figure 113 : deux lignes verticales ondulantes en opposition de phase

La limite de fusion est plus grande lorsque le stimulus est d'abord fusionnable et quitte l'aire de Panum doucement, plutôt que lorsque le stimulus n'est pas fusionnable et rentre progressivement dans l'aire de Panum. Nous voyons donc une hystérèse à la limite de fusion. Fender et Julesz en ont fait l'expérience : deux barres verticales augmentaient puis diminuaient en disparités. Lorsque la disparité augmente la limite de fusion est de 65 min arc, lorsqu'elle descend, la limite de fusion est de 42 min arc (Fender, et al., 1967). Cependant, cette hystérèse n'est pas particulière à la fusion binoculaire, elle existe dans beaucoup de phénomènes psychophysiques dans lesquels un seuil est mesuré.

Disparités verticales

Des études ont montré qu'il est difficile de fusionner pour des parallaxes verticales d'angle supérieur à 20 minutes d'arc (Julesz, 1971). Nielsen et Poggio ont mesuré les angles de disparités verticales qu'un sujet peut tolérer avant d'être incapable de percevoir la profondeur dans un stéréogramme de points aléatoires. Des disparités verticales de 3,5 min arc sont acceptées dans la région centrale du stéréogramme et jusque 6,5 min arc dans le reste du stéréogramme (Nielsen, et al.,

1984). Les tolérances de ces disparités verticales sont légèrement augmentées lorsque des indices monoculaires viennent compléter les indices binoculaires.

Il est donc souhaitable que les deux axes optiques des points de vue des images soient parallèles pour éviter les parallaxes verticales qui procurent des gênes à l'affichage. Les algorithmes permettant la création d'images stéréoscopiques de synthèse dans ce cas sont relativement simples et rapides (calculs de translations et de projections). Pour les images stéréoscopiques de synthèse, les points de vue sont presque toujours à axes parallèles pour éviter les parallaxes verticales.

Toutefois, il est possible d'avoir les axes des caméras virtuelles convergents. La position des deux pupilles et l'angle de convergence doivent être connus avec une grande précision. Cette technique permet de reproduire des disparités verticales plus fidèles à celles présentes en vision naturelle. Ces disparités verticales participent à une meilleure perception des distances et des formes et augmentent la cohérence entre les indices monoculaires et les indices stéréoscopiques. Cela réduit donc la fatigue visuelle. Ceci serait possible en traquant les yeux du sujet extrêmement rapidement et précisément. Mais ce procédé est cher et complexe à mettre en œuvre.

5.1.2. Les problèmes secondaires de technique

5.1.2.1. Qualité de l'image et de l'affichage

Tout ce qui fatigue les yeux en vision monoscopique fatigue également en vision stéréoscopique. Un écran ou projecteur de qualité est un facteur clef. Il est important que les caméras ou les appareils photographiques ne déforment pas les formes ou les couleurs des objets. Si le système est couplé avec des lunettes actives (les images sont présentées successivement à l'œil gauche et l'œil droit par occultation des verres synchronisée avec le changement d'images gauche et droite sur l'écran), le(s) projecteur(s) doiv(en)t être suffisamment rapide(s) pour afficher les deux images. Les images fantômes doivent être évitées (les images fantômes sont des images destinées à l'œil droit vue par l'œil gauche ou vice-versa suite à une limitation technique).

Lorsque les images sont projetées ou affichées sur un écran, il est préférable d'avoir une fréquence de rafraîchissement élevée. Certaines personnes sont plus sensibles que d'autres à la fréquence de balayage, mais même si le sujet ne le perçoit pas, plus cette fréquence est faible, plus son système visuel fatigue rapidement. Pour des écrans monoscopiques, on conseille souvent un minimum de 75 Hz, quitte à descendre la résolution. Ceci sera difficilement possible avec un écran stéréoscopique, néanmoins, il vaut mieux privilégier une haute fréquence de rafraîchissement (60 Hz).

Le contraste de l'affichage doit être correctement réglé. En effet, le noir doit être bien noir, et non gris foncé. De même la luminosité ne doit pas être trop forte : entre autres, une page blanche ne doit pas éblouir. Il existe des images qui facilitent le réglage de l'écran ou du projecteur comme celle représentée sur la Figure 114. Sur cette image, la nuance entre le A et le B ainsi qu'entre le Y et le Z doit être perceptible, la couleur A doit être bien noire.

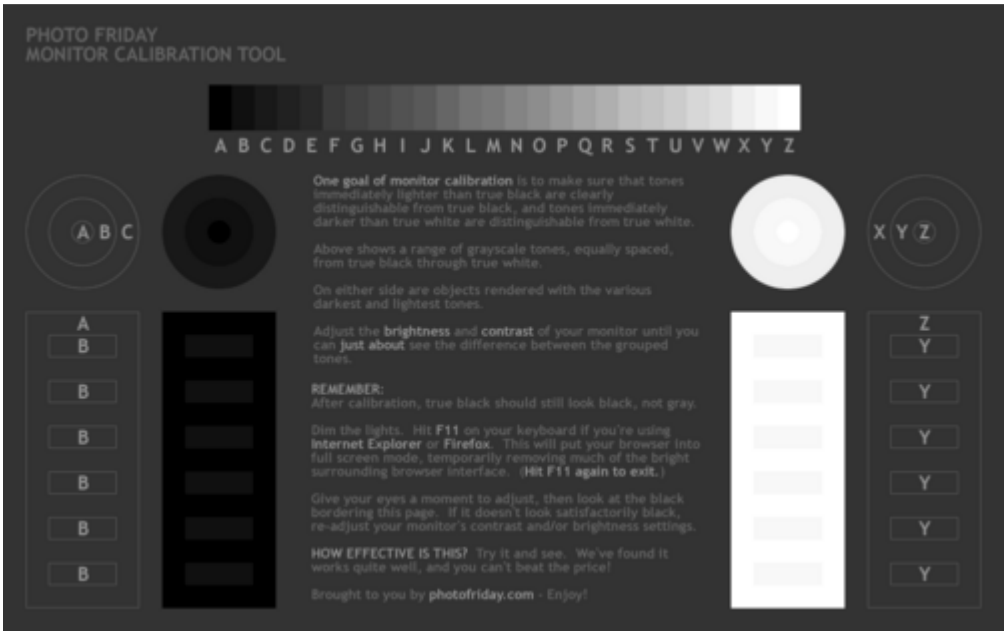


Figure 114 : image pour faciliter le réglage du contraste d'un écran ou d'un projeteur (htt092)

Si du texte est affiché et doit être lu, il faut veiller à ce que les caractères soient suffisamment grands pour que le système visuel n'ait pas à forcer. De même, les polices comprenant des empâtements sont à éviter. Il faut éviter aussi les reflets de la lumière naturelle ou artificielle sur l'écran.

5.1.2.2. Propriétés différentes entre l'image de gauche et l'image de droite

Techniquement, il est possible d'afficher des images différentes sur chaque œil. Elles peuvent différer en bien des aspects : formes ou objets présentés, orientations différentes, tailles d'objets différents, couleurs ou luminosités différentes...

Prenons un cas extrême : affichons une grille verticale sur l'œil gauche, et une grille horizontale sur l'œil droit. Nous percevons les deux grilles en surimposition, mais l'une d'entre elles s'imposera sur l'autre pendant quelques secondes, ensuite, l'autre prend le dessus pendant quelques autres secondes. La dominance de l'une ou de l'autre image fluctue constamment. C'est ce que l'on appelle la rivalité binoculaire.

Levlet (Levelt, 1965) (Levelt, 1966) a apporté des preuves que cette dominance est modifiée lorsque les contrastes et la luminance sont différents pour chaque image. Il proposait que l'augmentation de ceux-ci pour une des deux images diminue la durée de la période pour laquelle celle-ci est supprimée par le cerveau. En d'autres termes, un stimulus « fort » sera supprimé moins longtemps qu'un stimulus plus faible. Une image sans contours sera supprimée indéfiniment par le motif de l'autre image.

Évidemment, ce cas est extrême, cependant, il est courant de ne pas afficher la même image sur l'œil gauche et sur l'œil droit de façon involontaire. Une différence de luminosité générale peut être causée par un projecteur plus faible que l'autre dans le cas d'affichage stéréoscopique à plusieurs projecteurs. Le cerveau passe alors également d'une image à l'autre de façon plus discrète, mais néanmoins très fatigante (Kooi, et al., 2001). Il en va de même lorsque les couleurs des deux images ne sont pas identiques. Ceci est possible lors d'un dérèglement des projecteurs, lors de la projection d'images anaglyphes ou encore suite à des algorithmes d'augmentation de bande passante (Stelmach, et al., 1999) (Stelmach, et al., 2000). Le cerveau s'intéresse à une image puis à l'autre parce que les couleurs ne correspondent pas.

Une modification de la hauteur de projection d'une image par rapport à l'autre entraîne des disparités verticales sur tout l'écran. Ce qui, en plus d'être fatigant pour les yeux et plus difficile à fusionner (voir paragraphe 5.1.1.2), perturbe également la perception des formes. De même, un décalage horizontal d'une des deux images modifie les disparités horizontales en les augmentant ou les diminuant selon le point regardé. Si, pour un point, elles sont diminuées, la perception du relief et des distances est modifiée et ne correspond peut-être plus aux indices monoculaires et (ou) cinétiques. Si, au contraire, les disparités horizontales sont augmentées pour un autre point, cela entraîne également des modifications de perception, voire même une diplopie. En effet, si les disparités sont trop augmentées, le point en question sort de l'aire de Panum et est donc vu double. Dans tous les cas, cela provoque une fatigue visuelle. Un décalage en orientation d'une des projections modifie les disparités horizontales et verticales de la périphérie de l'écran (en supposant qu'il n'y ait pas de décalage ni horizontal ni vertical supplémentaire). Cela provoque également de la fatigue visuelle.

5.1.2.3. Les problèmes dans la rapidité de la correction des mouvements pseudoscopiques

Il est possible d'éviter les mouvements pseudoscopiques tout en permettant à l'observateur de se déplacer. Cependant, si cette technique n'est pas correctement réglée, elle peut également provoquer des pertes d'équilibre ou des nausées (Jones, et al., 2001) (Young, et al., 2006).

Il peut y avoir un mauvais alignement de la tête traquée et de la caméra virtuelle en l'absence de mouvement, ou une mauvaise réaction de celle-ci en présence de mouvement. Il peut également y avoir une certaine latence entre le moment ou l'observateur bouge la tête et le moment ou l'image affichée change pour prendre en compte ce déplacement. Ceci est dû au temps de réponse des capteurs, au temps de transfert des données de celui-ci à l'unité centrale et au temps de calcul pour prendre en compte les nouvelles coordonnées de l'observateur.

Cette latence, lorsqu'elle est supérieure à 75ms (MacKenzie, et al., 1993), provoque des tiraillements des yeux, maux de tête, pâleur, transpiration excessive, sécheresse buccale, désorientation, vertige, nausée, vomissement et ataxie (trouble de la coordination du mouvement volontaire non explicable par un déficit moteur) (LaViola, 2000) (Money, 1970). Ce phénomène a beaucoup de noms : mal du virtuel, cybermalaise, mal du cyberespace, mal du simulateur, mal des simulateurs, cyberopathie mais la plus part du temps, nous utilisons sa dénomination anglaise : cybersickness. La théorie du conflit entre le système vestibulaire et le système visuel est la plus ancienne et la plus largement acceptée (Reason, et al., 1975). Les deux sens renvoient des informations différentes au cerveau qui subit cette incohérence sensorielle.

Un des gros problèmes de ce cybermalaise est que les symptômes peuvent perdurer plusieurs heures après la stimulation (Kellog, et al., 1980), voire même plusieurs jours. Les pilotes d'avions ne peuvent pas voler pendant 24h après avoir travaillé sur un simulateur de vol (Gower, et al., 1989).

5.2. Les problèmes de hautes fréquences spatiales

5.2.1. Limite en fusion

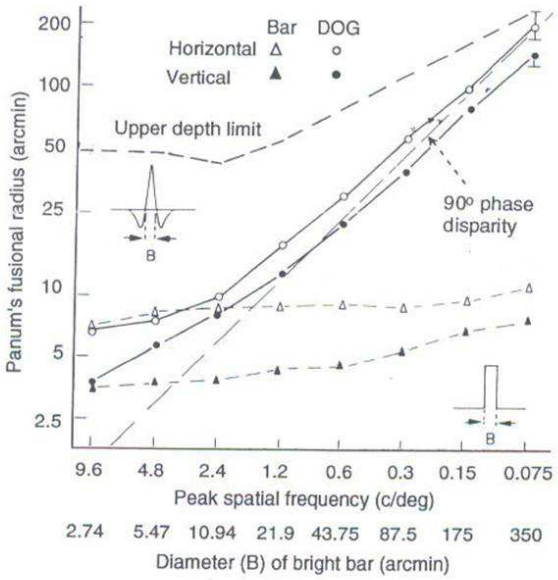


Figure 115 : relation entre les fréquences spatiales et l'aire de Panum (Schor, et al., 1986)

Plusieurs études ont montré que le seuil de fusion est plus grand lorsque nous regardons une image de faibles fréquences spatiales (Roumes, 1995) (Roumes, et al., 1992). Schor a utilisé une variation de luminance verticale suivant un profil en différences de Gaussiennes (DOG, difference of Gaussians). La bande de fréquences étudiée allait de 0.075 à 9.6 cpd. Les sujets devaient ajuster les deux images pour définir la disparité limite qu'ils n'arrivaient plus à fusionner. Il s'est avéré que le seuil de fusion diminue avec l'augmentation de la fréquence spatiale présentée, comme indiqué sur la Figure 115. La limite de fusion dans la direction verticale était inférieure à celle dans la direction horizontale (Schor, et al., 1986). Lorsque la fréquence spatiale était inférieure à 1,5 cpd, la limite de fusion correspond pratiquement à la phase de 90° (limite après laquelle le sujet ne fusionne pas les bandes correspondantes, également appelé la limite de Rayleigh). Ils en ont donc déduit qu'à cette fréquence-ci, l'acuité stéréoscopique n'était limitée que par l'acuité monoscopique. Pour des fréquences spatiales supérieures à 2,4 cpd, la limite de fusion s'étale entre 5 et 10 min arc. Donc pour les hautes fréquences, la limite de Rayleigh ne correspond plus à la limite de fusion stéréoscopique. La limite de fusion des plus hautes fréquences est entre 6 et 8 fois la largeur du centre de la DOG.

Ces expériences ont été répétées avec une barre verticale dont le profil ne suivait plus une DOG, mais des rectangles de bords nets dont la largeur et la luminance étaient égales aux DOG correspondantes. Nous pouvons voir sur la Figure 115 que la limite de fusion pour ces barres aux bords abruptes ressemble nettement à celle pour les DOG les plus étroites. Cela signifie que les sujets ont utilisé les bords des barres, donc les plus hautes fréquences, pour faire leur jugement.

En présence de hautes fréquences spatiales et pour un contraste donné, le changement de luminance est beaucoup plus abrupt que lorsqu'il n'y a que des basses fréquences. Il semblerait que c'est ce qui peut expliquer la dépendance entre la limite de fusion et la plus grande fréquence spatiale visible. Schor a étudié cette hypothèse en 1989, la logique était que si ce sont vraiment les fré-

quences spatiales le facteur clé, changer le contraste ou la luminance n'aurait aucun effet, mais si le contraste et la luminance sont les facteurs en question, alors les modifier aurait une influence sur les résultats (Schor, et al., 1989). Ces expériences ont montré que le contraste et la luminance n'ont aucun effet dans la plage de fréquences allant de 0,4 cpd à 3,2 cpd. Ce résultat a été confirmé par Heckmann (travaillant avec Schor) (Heckmann, et al., 1989). De plus, la limite de fusion n'était pas influencée par de la luminance ou du contraste ajouté grâce à de basses fréquences spatiales. Ils en ont conclu que les informations stéréoscopiques sont plus basées sur les fréquences spatiales que sur des données de luminance.

Si nous supprimons les plus hautes fréquences, nous pourrons augmenter l'aire de Panum. Pour chaque objet qui n'est pas vu dans l'aire de Panum, nous pouvons supprimer les hautes fréquences jusqu'à ce que l'aire de Panum englobe l'objet ou le point problématique.

5.2.2. Le confort et les hautes fréquences.

Wöpking a étudié le confort en fonction des fréquences spatiales présentes dans un couple d'images stéréoscopiques. Il a présenté à 12 sujets une paire d'images stéréoscopiques dont les contenus en fréquences spatiales et en disparités étaient calibrés. Les sujets devaient fixer un objet réel plat attaché au centre d'un écran. Sur l'écran, une photographie avec des disparités évoluant de 0 à 140 min arc était affichée. Le contenu en fréquences spatiales était contrôlé en réglant la focale du projecteur : elles s'étalaient de 1 à 23 cpd. Les sujets devaient alors évaluer leur inconfort de -2 (très incommodant) à 2 (imperceptible) (Wopking, 1995).

Une fonction de confort a été établie par Jérôme Perrin (Perrin, 1998) interpolée à partir des données de Wöpking. Elle établit une relation entre le confort ("*C(d,f)*"), les disparités horizontales ("*d*") et les fréquences spatiales ("*f* "). Comme pour Wöpking, l'inconfort est représenté par une valeur variant entre -2 (très incommodant) et 2 (imperceptible). La fonction de confort basée sur l'interpolation des données de Wöpking est représentée ci-dessous (Figure 116).

$$C(d, f) = a(d - d_0 - kf^{k'})$$

Les paramètres d'interpolation sont :

$a = -0.010$

$d_0 = 18.9$

$k = 221.1$

$k' = -0.74$

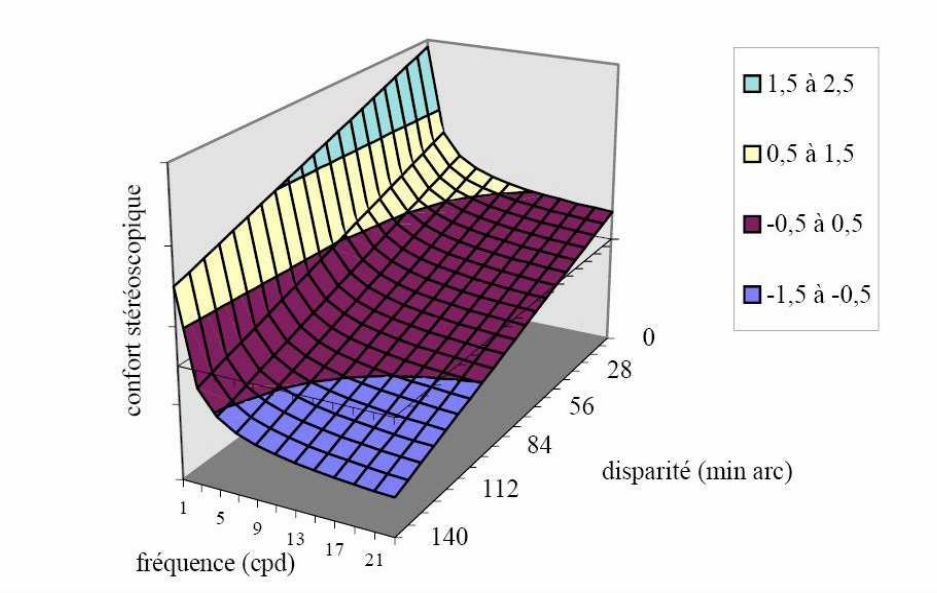


Figure 116 : fonction de confort (Perrin, 1998)

5.3. Conclusion

Nous avons évoqué dans ce chapitre les principales causes de fatigue visuelle en stéréoscopie. Nous allons dans les chapitres suivants nous concentrer sur la dernière cause étudiée : avoir des hautes fréquences spatiales et de grandes disparités au même endroit cause un inconfort visuel, voire même une diplopie. Nous allons donc supprimer les hautes fréquences gênantes. Le chapitre 6 explique les algorithmes étudiés pour faire ce traitement d'image. Le chapitre 7 retrace nos expériences sur la fatigue visuelle liées à l'association des hautes fréquences et de grandes parallaxes. Enfin, le chapitre 8 est une discussion des résultats obtenus dans le chapitre 7.

Chapitre 6 : Traitement adaptatif pour la diminution de la fatigue visuelle

Dans le chapitre sur la diminution de la fatigue visuelle, nous souhaitions supprimer les hautes fréquences lors de grandes disparités. Dans ce chapitre-ci, nous allons tenter d'apporter des éléments algorithmiques pour nous permettre de supprimer ces hautes fréquences.

Nous commencerons par une explication des algorithmes étudiés, à savoir, les transformées en ondelettes, BOX FILTER et les méthodes de flou obtenu par des moyennes de pixels. Nous poursuivrons l'exposé par des comparaisons entre tous ces algorithmes. Ensuite, nous proposerons d'en améliorer deux, à savoir BOX FILTER et l'algorithme des moyennes.

6.1. Explication de chaque algorithme

Tous les algorithmes qui suivent sont réalisés sur des images en niveau de gris. Les images que nous utiliserons en pratique seront des images en couleur, mais les traitements seront faits sur la luminance, qui correspond aux niveaux de gris pour ne pas faire trois fois le même traitement. Pour tous ces algorithmes, une contrainte très forte est que les calculs doivent être réalisés en « temps réel » pour nos applications de réalité virtuelle (en pratique le couple d'images stéréoscopiques doit être traité en 40 ms si possible).

6.1.1. Les ondelettes

6.1.1.1. Introduction aux ondelettes

Les transformées de Fourier sont bien connues dans la communauté scientifique. Elles permettent de connaître les fréquences présentes dans un signal temporel stationnaire ou une image. Mais aussi utile que soit cet outil, il ne permet pas de localiser ces fréquences. Il est vrai que calculer la fréquence en un point n'a pas de signification physique. Cependant, nous aimerions savoir où se situent les hautes fréquences de nos images, puisque ce sont ces fréquences qui posent problème lorsqu'elles sont au même endroit que les hautes disparités.

La transformée en ondelettes fait passer le signal dans une série de filtres pour en ressortir les fréquences présentes et surtout l'endroit où elles interviennent. Pour un signal en 2 dimensions (donnée, temps), nous obtenons soit un graphique en 3 dimensions (fréquence, valeur, temps) soit une succession de graphiques en 2 dimensions (valeur, temps) qui indiquent la présence d'une fréquence donnée à un temps donné.

6.1.1.2. La transformée de Fourier

Définition de la transformée de Fourier

La transformation de Fourier F associe à l'espace temps un espace dual dont la variable ω est la fréquence. Pour tout signal temporel f, on peut associer une transformée de Fourier $\hat{\omega}$

$$\forall f \in l^2(\Re),\ \hat{f}(\omega) = F[f](\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt$$

6-6-1

Et sa transformée inverse :

$$f(t) = F^{-1}[\hat{f}](t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \hat{f}(\omega) e^{-j\omega t} d\omega$$

6-2

Propriétés

La transformée conserve le produit scalaire :

$$\int f(t)\overline{g}(t)dt = \int \hat{f}(\omega)\overline{\hat{g}(\omega)}d\omega$$

6-3

Elle conserve l'énergie (égalité de Parseval) :

$$\int \left| f(t) \right|^2 dt = \int \left| \hat{f}(\omega) \right|^2 d\omega$$

6-4

Il existe aussi des propriétés de dilatation et de translation :

$$f(t-b) \longrightarrow e^{jb\omega} \hat{f}(\omega)$$

$$f(at) \longrightarrow \frac{1}{a} \hat{f}(\frac{\omega}{a})$$

$$f(-t) \longrightarrow \overline{\hat{f}(\omega)}$$

$$f'(t) \longrightarrow j\omega\hat{f}(\omega)$$

6-5

Avant d'énoncer la propriété suivante, rappelons la définition d'un produit de convolution

Le produit de convolution r de deux fonctions s et k est défini par :

$$r(t) = (s * k)(t) = \int s(\tau)k(t-\tau)d\tau$$

6-6

*Où l'opérateur de convolution est noté **

Le produit de l'espace temporel se transforme en produit de convolution dans l'espace fréquentiel et vice-versa.

$$F[s * k] = F[s]F[k]$$

et

$$F[sk] = F[s] * F[k]$$

6-7

La transformée de Fourier rapide

La transformée de Fourier rapide s'applique aux signaux numériques et accélère les calculs de la transformée de Fourier normale. En effet pour un signal à *n* échantillons, il faudra n^2 opérations pour avoir la transformée de Fourier par la méthode normale. Par contre avec la transformée de Fourier rapide, nous arrivons à *n.log(n)* opérations (Grenez, 2003).

Nous partons d'un échantillon *s* qui contient *n* valeurs, où *n* est une puissance de 2. Nous allons d'abord trier les valeurs selon leur position paire ou impaire dans deux vecteurs et nous refaisons la même opération avec ces deux vecteurs jusqu'à ce qu'il n'y ait plus qu'un élément dans chaque ensemble (Figure 117).

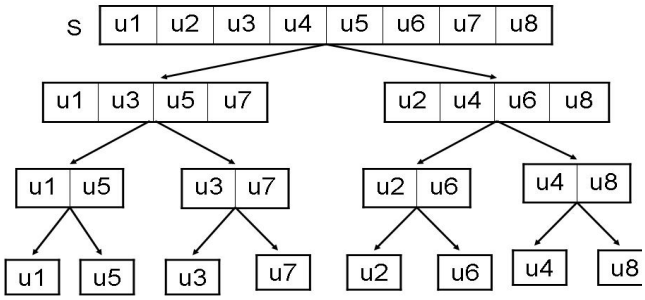


Figure 117 : décomposition paire-impaire

Ensuite, lorsque nous avons cette suite de nombre, nous additionnons les éléments deux par deux après avoir multiplié l'élément impair par $e^{2ki/\pi}$ où *k* est le rang qu'il doit obtenir dans le vecteur suivant (Figure 118).

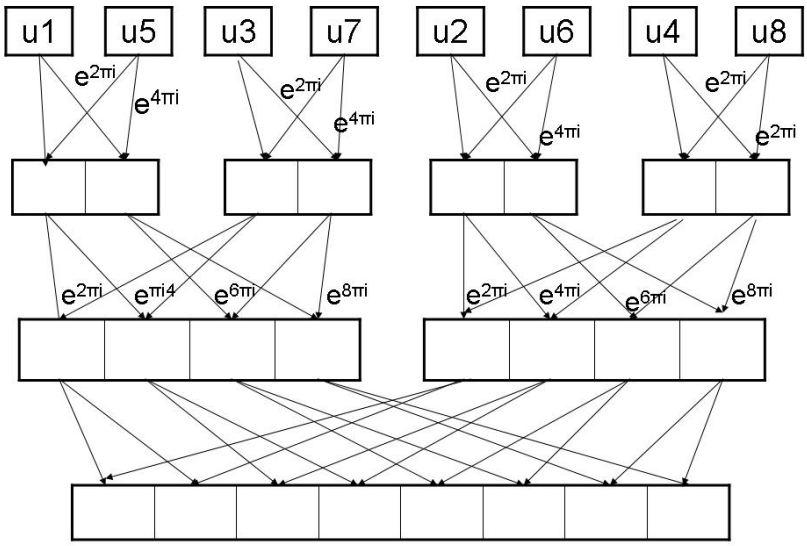


Figure 118 : fin de la FFT

Le vecteur ainsi obtenu est la transformée de Fourier du vecteur *s*. Pour la transformée inverse, c'est exactement le même algorithme sauf que lorsqu'on a le vecteur final de celui-ci, il faut en prendre le conjugué pour avoir le résultat de la transformée inverse.

Interprétation

Nous l'avons vu, la transformée est bijective, il y a exactement les mêmes informations dans la fonction et dans sa transformée puisque nous pouvons aussi bien construire l'une à partir de l'autre que l'inverse. Mais la manière de l'exprimer et leurs interprétations sont différentes. Dans l'espace temporel, il est possible de connaître la forme et l'amplitude du signal mais les fréquences présentes dans celui-ci ne sont pas lisibles explicitement. Par contre, dans l'espace fréquentiel, c'est l'inverse, nous connaissons les fréquences présentes dans le signal sans exactement en connaître la forme, ni la localisation temporelle de ses fréquences. Pour avoir les deux informations, il faut passer à une représentation temps-fréquence comme dans la transformée de Fourier glissante, ou temps-échelle comme dans la transformée en ondelettes.

La transformée de Fourier sur fenêtre glissante

Pour pallier ce manque, Gabor, en 1946 (Gabor, 1946), propose de limiter la transformée de Fourier à une fenêtre et de translater cette fenêtre le long du signal. Ceci permet de connaître les fréquences présentes dans un certain laps de temps.

6.1.1.3. Le principe d'incertitude d'Heisenberg

Nous ne pouvons pas décrire un signal aussi précisément que nous le souhaitons en fréquence et en temps simultanément. Il n'est pas possible d'avoir une fenêtre de Fourier toute petite et pouvoir trouver toutes les fréquences qui s'y trouvent.

L'énoncé

Pour toute fonction f(t) tel que

$$\int_{-\infty}^{+\infty} \left| f(t)^2 \right| dt = 1 \qquad \textbf{6-8}$$

le produit de la variance de t et de la variance de w est supérieur ou égale à $1/16\pi^2$

$$\left(\int_{-\infty}^{+\infty} (t - t_m)^2 \left| f(t) \right|^2 dt\right)\left(\int_{-\infty}^{+\infty} (\tau - \tau_m)^2 \left| \hat{f}(\tau) \right|^2 d\tau\right) \geq \frac{1}{16\pi^2} \qquad \textbf{6-9}$$

Les représentations temps-fréquence

Si nous voulons préciser le temps, nous devons nous contenter de fréquences moins précises et inversement. Nous pouvons représenter cette contrainte sur le plan temps-fréquence. Ce plan est séparé en rectangle d'aire $\Delta t \Delta \omega$ et le principe d'incertitude nous apprend que l'aire maximale de ce rectangle est de 1/4π mais suivant la base que nous utilisons, ces rectangles seront disposés différemment dans le plan (Hubbard, 1995)(Figure 119).

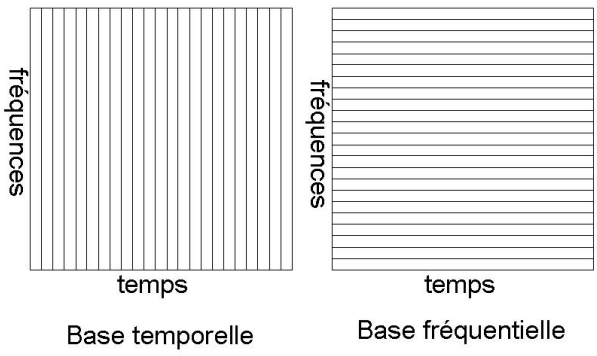


Figure 119 : plan temps-fréquence en temporel et en Fourier

Dans la transformée de Fourier sur fenêtre glissante, nous avons une sorte de compromis. Nous séparons le plan en rectangles d'aire égale mais dont la précision en temps égale celle de la fenêtre glissante, ce qui crée un quadrillage du plan (Figure 120).

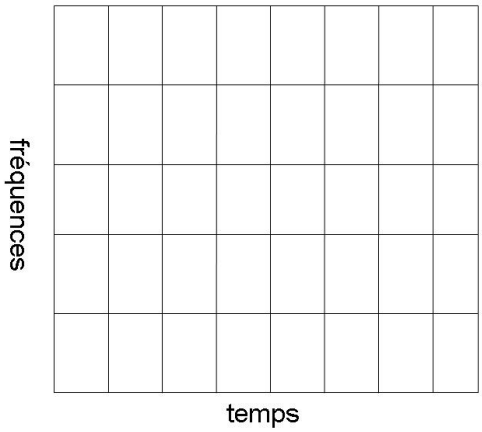


Figure 120 : plan temps-fréquence pour Fourier sur fenêtre glissante

Le plus efficace serait de faire un quadrillage progressif (Figure 121). Employer une grande fenêtre pour les basses fréquences et une fenêtre très petite pour les très hautes fréquences semble plus optimisé.

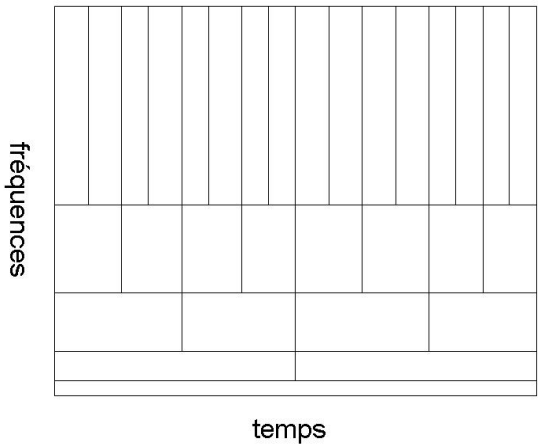


Figure 121 : plan temps-fréquence pour les ondelettes

6.1.1.4. Les transformées en ondelettes

Définition de la transformée en ondelettes

Les premières théories des transformées en ondelettes donnaient lieu à l'élaboration de représentations continues appelées **transformées en ondelettes continues** (Ranchin, 2001).

A partir d'une fonction de base ψ, appelée ***ondelette-mère*** *ou ondelette analysante, nous construisons une famille de fonctions élémentaires par dilatation et translation*

$$\psi_{a,b} = \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}), b \in \Re, a \neq 0$$

6-10

Les coefficients du signal sont alors les nombres

$$O[f](a,b) = \langle f, \psi_{a,b} \rangle = \int_{-\infty}^{+\infty} f(t)\overline{\psi_{a,b}(t)}dt$$

6-11

Prenons comme exemple la fonction f(x)=cos(x/100)², la décomposition est visible sur la Figure 122.

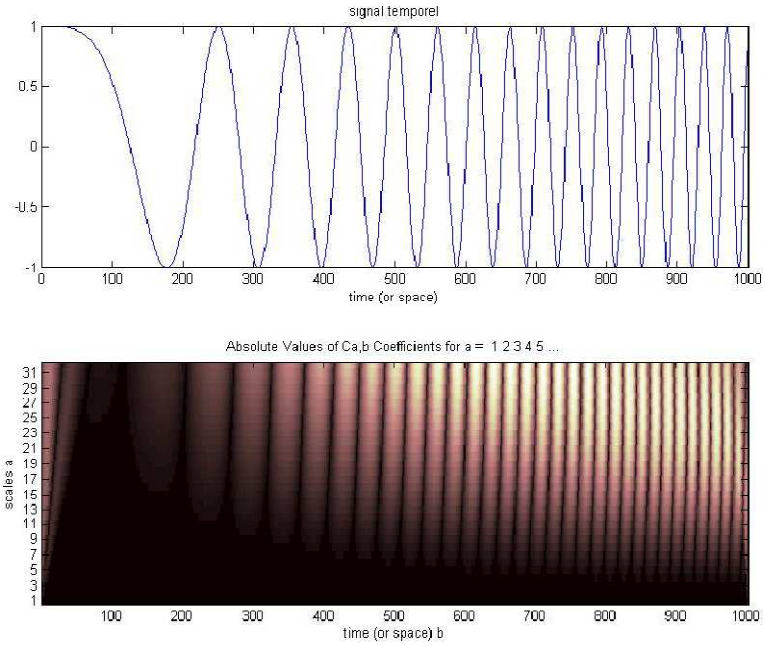


Figure 122 : décomposition en ondelettes continues d'un signal 1D

Nous voyons que plus les fréquences sont élevées dans le signal temporel, plus les hautes fréquences sont représentées en clair (donc élevées) dans le signal en ondelettes. Nous voyons aussi que plus le temps est précis, moins les fréquences le sont. Cela se voit particulièrement lorsque les fréquences sont très hautes, le temps est très précis, mais les fréquences élevées s'étalent sur presque toute la bande.

Nous pouvons reconstruire la fonction de départ.

Si la fonction $\psi \in L^1 \cap L^2$

1. $\int \frac{|\hat{\psi}(\omega)|^2}{|\omega|} = K \prec +\infty$, *avec* $\hat{\psi} = F[\psi]$

6-12

2. $\|\psi\|_2 = 1$

6-13

Alors le signal peut être reconstruit à partir de ses coefficients par la formule :

$$f(t) = \frac{1}{K} \iint_{\Re_2} O[f](a,b)\psi_{a,b}(t)\frac{dadb}{a^2}$$

6-14

Remarquons la conservation de l'énergie :

$$\frac{1}{K} = \iint_{\Re_2} |O[f](a,b)|^2 \frac{dadb}{a^2} = \int_{-\infty}^{+\infty} |f(t)|^2 dt$$

6-15

Transformée en ondelettes en 2 dimensions

En 2 dimensions l'équation de la définition du paragraphe précédant devient :

$$\psi_{a,b,R_\vartheta}(x) = \frac{1}{a}\psi(R_\vartheta \frac{x-b}{a}), b \in \Re, a \neq 0$$

6-16

Nous obtenons les résultats graphiques suivants. Le dessin de départ est représenté sur la Figure 123.

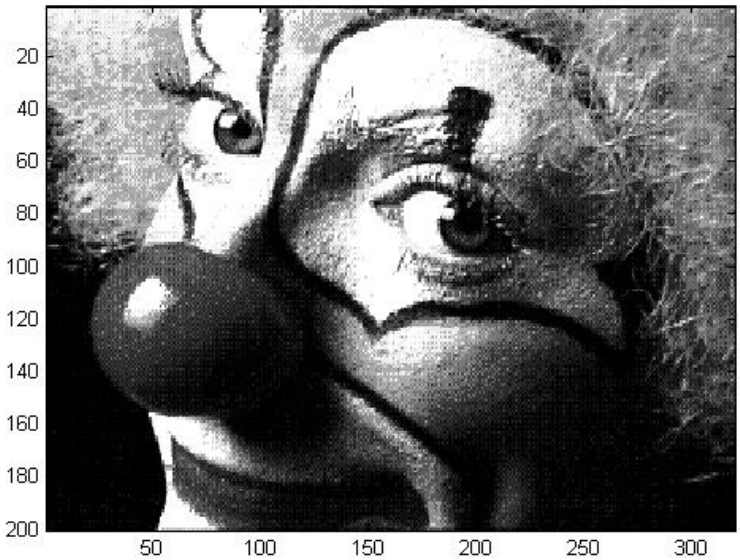


Figure 123 : figure à décomposer

Tandis que la première décomposition est exposée sur la Figure 124.

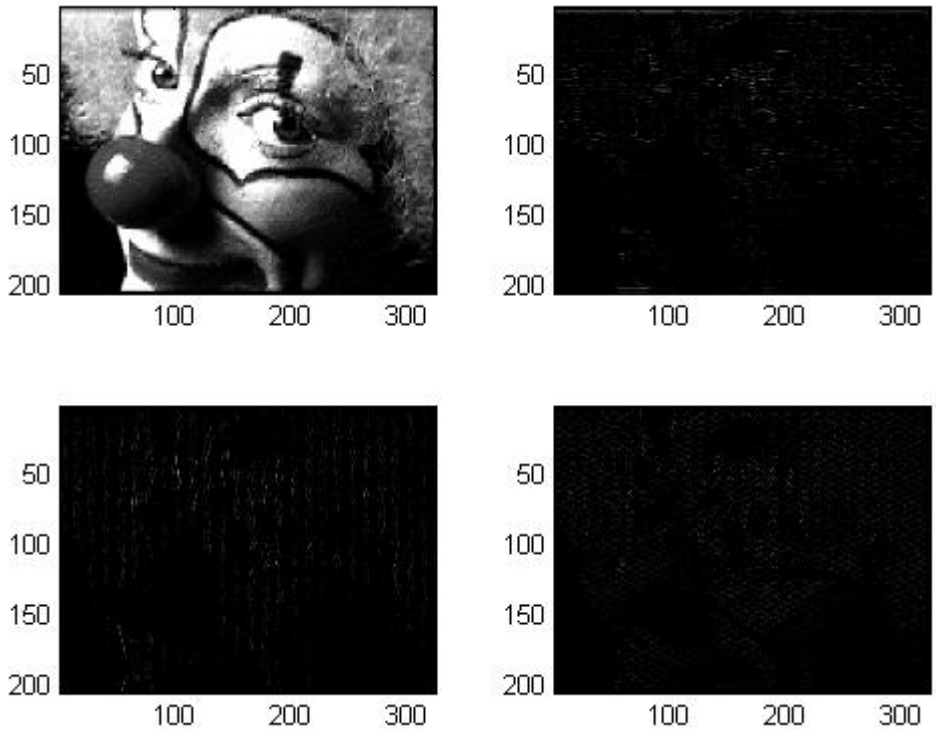


Figure 124 : décomposition en ondelettes d'une image

L'algorithme de Mallat

L'algorithme de Mallat, également appelé **transformée en ondelettes rapide** est celui utilisé dans la plupart des décompositions d'images (Benoit, et al.) (Mallat, 2000) (Meyer, 1992). Il s'agit de faire passer les signaux discrets dans des filtres passe-bas et passe-haut de façon récursive.

Transformation en ondelettes rapide en 1 dimension

Le signal de départ est simultanément étudié par un filtre passe-haut et son filtre passe-bas conjugué, suivi d'un échantillonnage, comme illustré sur Figure 125. Nous appliquons la récursivité illustrée sur la Figure 126. Le signal sortant du filtre passe-bas rentre dans le système de filtres suivant.

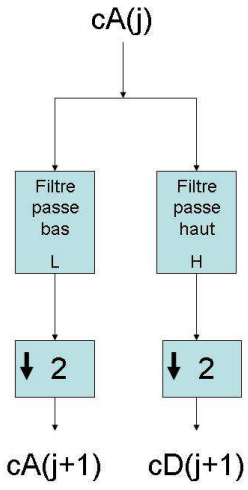


Figure 125 : décomposition d'un signal 1D par l'algorithme de Mallat

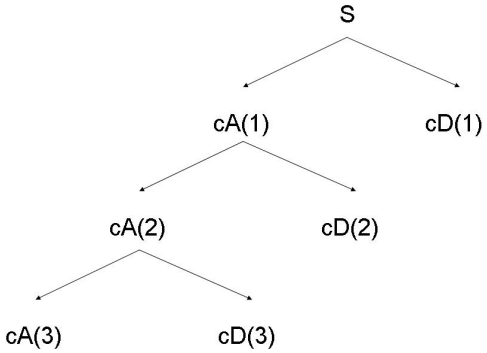


Figure 126 : la pyramide de l'algorithme de Mallat

Ce qui donne en reprenant l'exemple de fonction $f(x)=\cos(x/100)^2$, la transformée visible sur la Figure 127.

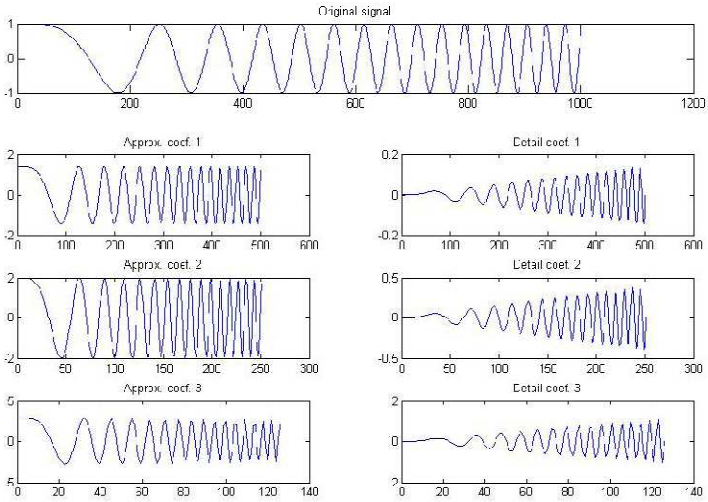


Figure 127 : transformée de Mallat pour un signal 1D

Remarquons au passage que chaque fois la taille du signal est divisée par deux. Il faut à présent recomposer le signal. Pour ce faire, nous partons de la dernière recomposition auquel nous interposons des 0 entre les données. Nous faisons de même avec les données de hautes fréquences, puis nous faisons passer les deux vecteurs résultants de cette opération dans des filtres de reconstruction passe-haut et passe-bas associés. Ensuite, nous additionnons les deux vecteurs sortant de ces filtres. Cette partie de l'algorithme est explicitée sur la Figure 128.

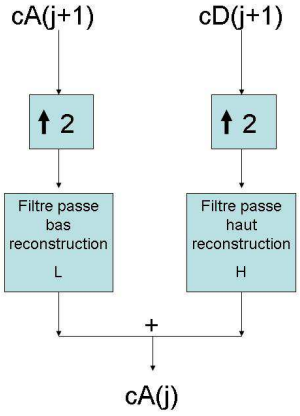


Figure 128 : algorithme de reconstruction de Mallat à une dimension

Il reste à calculer les filtres de reconstruction à partir des filtres de décomposition. Si nous appelons h_0 le filtre passe-bas et h_1 le filtre passe-haut de décomposition, nous appelons g_0 le filtre passe-bas et g_1 le filtre passe-haut de recomposition.

$$g_0[n] = \alpha^{-1}(-1)^n h_1[n]$$

6-17

et

$$g_1[n] = \alpha^{-1}(-1)^n h_0[n]$$

6-18

Avec

$$\alpha = \frac{1}{2}(h_0^{dc} h_1^{nyq} + h_1^{dc} h_0^{nyq})$$

6-19

Où h_b^{dc} correspond au gain du filtre pour la phase d'analyse à la fréquence DC et h_b^{nyq} correspond au gain du filtre pour la phase d'analyse à la fréquence de Nyquist.

$$h_b^{dc} = \sum_n h_b[n]$$

$$h_b^{nyq} = \sum_n (-1)^n h_b[n]$$

6-20

Pour *b*=0, 1

Transformation en ondelettes rapide en 2 dimensions

L'algorithme à deux dimensions n'est pas très différent de celui à une dimension, il faut juste l'appliquer deux fois de suite, une fois pour les lignes, une autre pour les colonnes. Nous prenons les données en ligne auxquelles nous appliquons l'algorithme de Mallat à une dimension. Ceci nous crée deux images dont la largeur est moitié moindre que celle de départ mais dont la hauteur n'est pas modifiée. Les colonnes de ces dernières passent à leur tour dans ce même système de filtre. Nous avons donc 4 images dont la hauteur et la largeur sont deux fois plus petites que celle de l'image de départ. Cette partie de l'algorithme est illustrée sur la Figure 129.

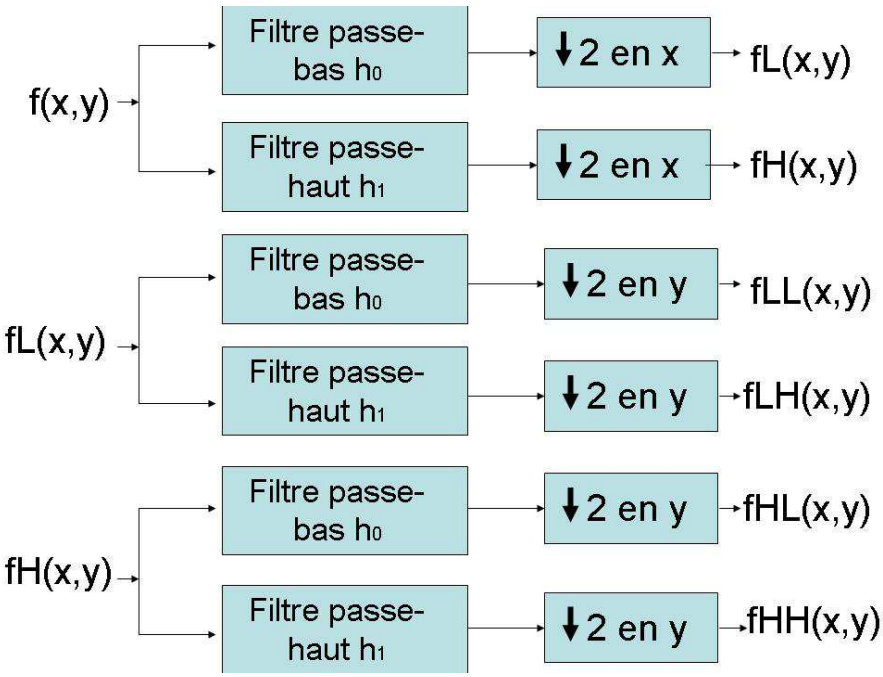


Figure 129 : décomposition en ondelettes rapide d'une image

La recomposition n'est pas très différente non plus. Il faut prendre soin de recomposer les deux images intermédiaires. Pour ce faire, nous utilisons les colonnes des 4 images finales et nous les faisons passer dans le système de filtres de reconstruction de l'algorithme de Mallat à une dimension. Nous retrouvons nos deux images dont la largeur est plus petite que la hauteur. Nous prenons donc les lignes de ces deux images pour reconstruire notre image de départ via le système de filtres de reconstruction. Cette partie de l'algorithme est illustrée sur la Figure 130.

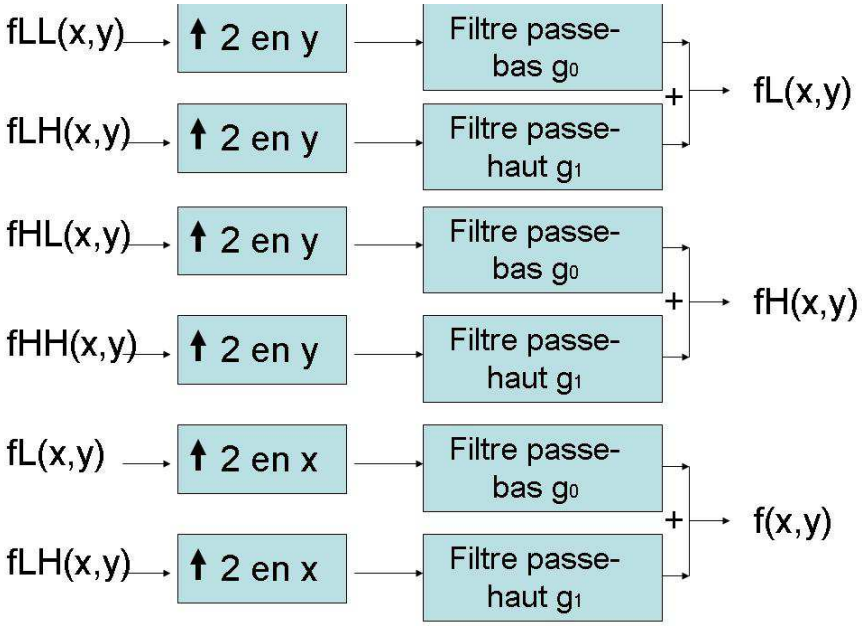


Figure 130 : recomposition en ondelettes rapide d'une image

6.1.1.5. Transformée de Fourier rapide ou produit de convolution

Lorsqu'un signal passe dans un filtre, un produit de convolution s'effectue entre le filtre et le signal. Mais nous avons vu au paragraphe 6.1.1.2 que le produit de convolution entre deux signaux équivaut au produit de leur transformée de Fourier. Dès lors, nous pourrions nous demander s'il n'est pas préférable d'effectuer ces opérations avec les transformées de Fourier rapides plutôt que de calculer le produit de convolution à chaque fois. D'autant plus que cette opération est la plus coûteuse en temps de calcul. Pour trouver quelle façon de procéder serait la plus rapide, faisons un calcul de complexité sur une image carrée de côté *n.*

La complexité d'une FFT pour une ligne est de

$$n\log(n)$$

6-21

Comme il faut une FFT et une FFT inverse, on arrive à

$$2.n\log(n)$$

6-22

Nous devons faire la multiplication avec la FFT du filtre, nous arrivons pour une ligne à

$$2.n\log(n)+n$$

6-23

Comme il faut faire toutes les lignes :

$$n(2.n\log(n)+n)=2n^2\log(n)+n^2$$

6-24

Et toutes les colonnes :

$$4n^2\log(n) + 2n^2$$

6-25

Comparons avec le produit de convolution : pour une ligne, chaque élément est multiplié par tous les éléments du filtre

$$n.n_{filtre}$$

6-26

Où n_{filtre} est le nombre d'échantillons nécessaires pour décrire le filtre.

Multiplié par le nombre de lignes

$$n^2.n_{filtre}$$

6-27

Puis toutes les colonnes

$$2n^2.n_{filtre}$$

6-28

Comparons les deux courbes sur la Figure 131. Nous voyons que pour le nombre d'opérations lors de l'utilisation du produit de convolution dépend très fortement de la longueur du filtre, et donc du type d'ondelettes utilisées. Lorsque nous utilisons la FFT, au contraire, le nombre d'opérations ne dépend pas de l'ondelette utilisée. Nous voyons donc que pour une certaine longueur de filtre, l'utilisation du produit de convolution ne sera plus rentable du point de vue du nombre d'opérations effectuées, mais que pour les ondelettes de Harr et de Daubechies, il vaut mieux utiliser le produit de convolution que la FFT.

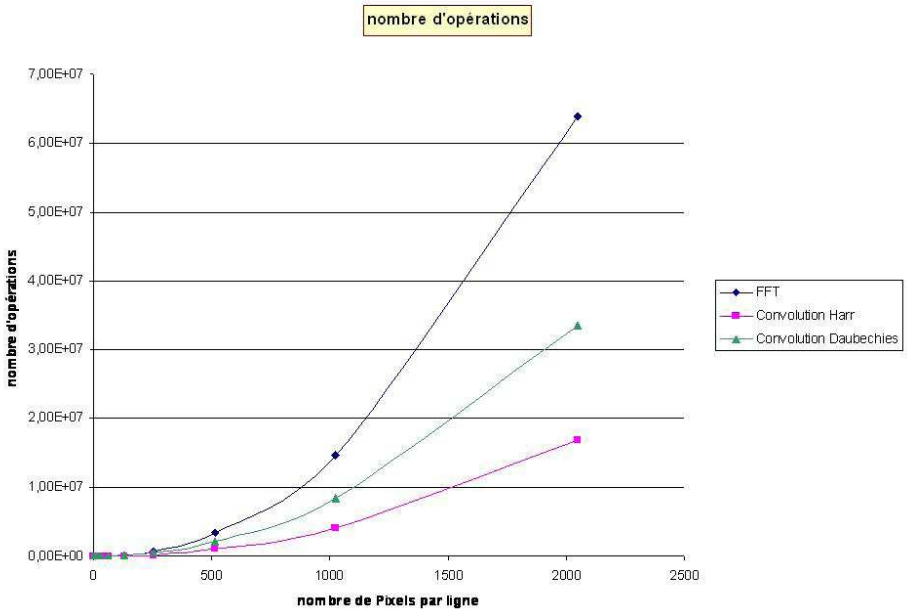


Figure 131 : comparaison théorique entre le nombre d'opérations pour un produit de FFT et un produit de convolution

L'algorithme de Mallat a été implémenté en C++ avec la FFT et avec le produit de convolution pour faire des tests de rapidité. Le produit de convolution est beaucoup plus rapide que sa rivale. Sur les figures suivantes le temps pour 1, 2, 3 et 4 décompositions recompositions avec l'ondelette de Haar est représenté (Figure 132, Figure 133).

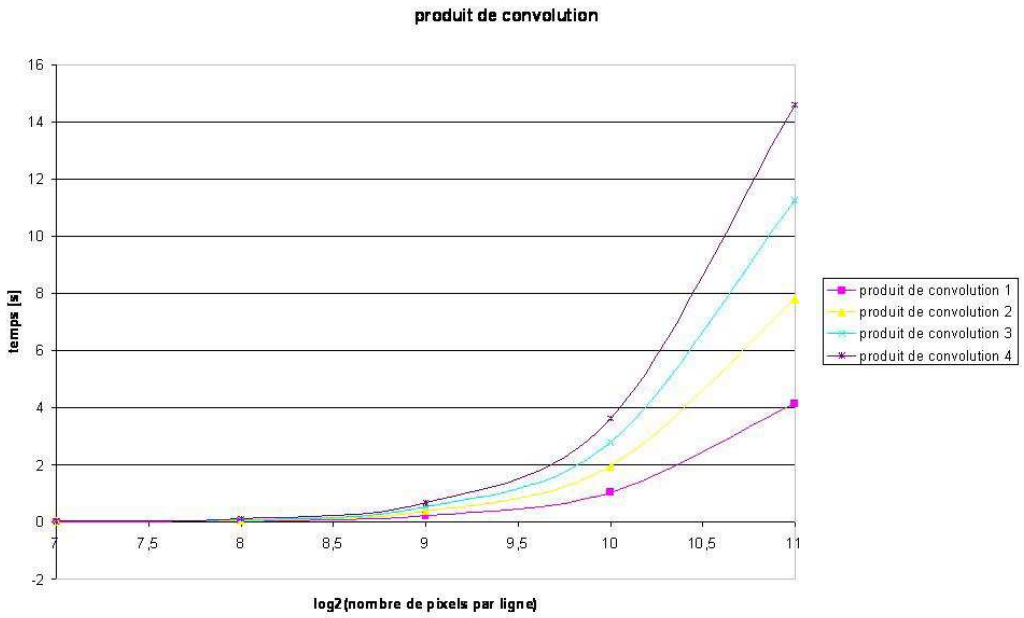


Figure 132 : temps de décomposition et recomposition du produit de convolution

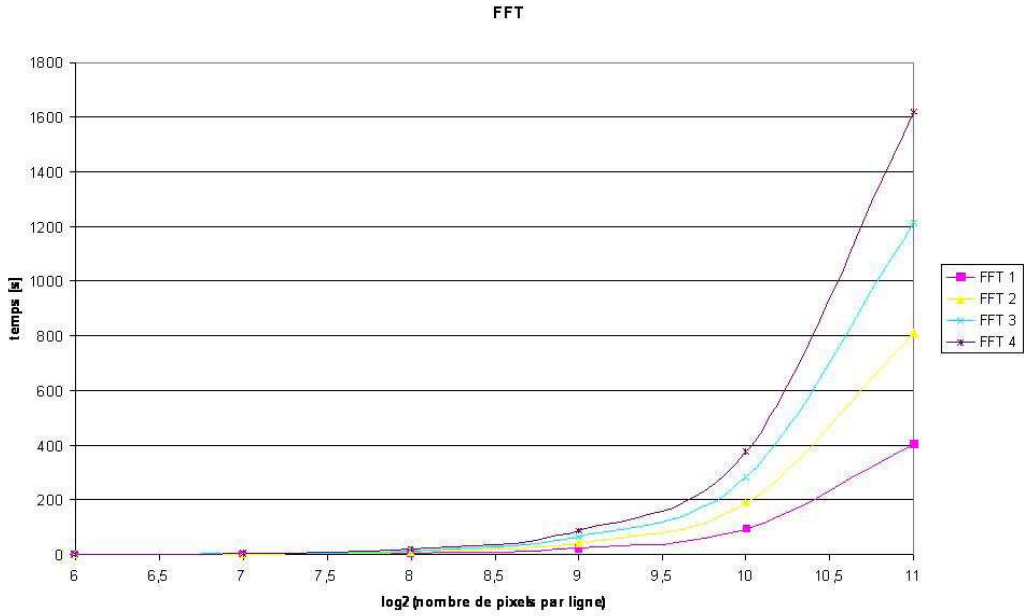


Figure 133 : temps de décomposition et recomposition avec la transformée de Fourier

Comme nous pouvons le voir sur les figures ci-dessus, le traitement ne peut se faire en temps réel pour des images de tailles courantes. En effet, pour une image de 512*512 pixels, 0.25 seconde est nécessaire pour effectuer une seule décomposition et recomposition en produit de convolution. Pour être en temps réel, il faudrait que ce traitement prenne moins de 0.016 seconde.

6.1.2. BOX FILTER

6.1.2.1. Introduction

La transformée en ondelettes permet de faire une analyse en multi-résolution très précise et très efficace. Malheureusement, le temps de calcul de cette méthode ne nous permet pas de traiter nos images stéréoscopiques en temps réel. Nous nous sommes donc orientés vers d'autres méthodes d'analyse multi-résolution plus rapides, sans toucher à l'efficacité de l'algorithme général. L'algorithme BOX FILTER permet cette analyse multi-résolution (Viola, et al., 2001). Il est utilisé en reconnaissance de points d'intérêt, mais il peut être légèrement détourné pour détecter les hautes fréquences.

6.1.2.2. Image intégrale

Au départ de l'algorithme BOX FILTER, il faut calculer « l'image intégrale ». La valeur d'un pixel de cette image intégrale est la somme des valeurs des pixels dans l'image de départ dont les indices de position (colonne et ligne) sont inférieurs ou égaux aux indices de position du pixel étudié.

Si

- i(x,y), la valeur des pixels de l'image de départ
- ii(x,y) la valeur des pixels de l'image intégrale

Alors

$$ii(x,y) = \sum_{x'<x,y'<y} i(x',y') \qquad \textbf{6-29}$$

Où i(x, y) représente le pixel situé sur la xième colonne et la y^{ième} ligne de l'image i (Bay, et al., 2006) (Viola, et al., 2001). Voyons avec un exemple très simple basé sur 16 pixels. Les valeurs de départ de ces pixels sont données sur la Figure 134.

1	1	1	1
1	4	1	1
1	1	1	4
1	1	1	1

Figure 134 : image de départ pour exemple de calcul BOX FILTER

Son image intégrale est représentée sur la Figure 135.

1	2	3	4
2	7	9	11
3	9	12	18
4	11	15	22

Figure 135 : image intégrale pour exemple de calcul BOX FILTER

Si nous prenons le deuxième pixel de la première ligne de l'image intégrale (2), il est la somme du premier pixel de la première ligne de l'image de départ (1) et du second pixel de la première ligne de l'image de départ (1).

De même le deuxième pixel de la deuxième ligne de l'image intégrale (7)

= premier pixel de la première ligne de l'image de départ (1)

+ Deuxième pixel de la première ligne de l'image de départ (1)

+ Premier pixel de la deuxième ligne de l'image de départ (1)

+ Deuxième pixel de la deuxième ligne de l'image de départ (4)

Et ainsi de suite pour tous les pixels de l'image intégrale.

Mais d'un point de vue informatique, il est préférable d'utiliser la version récursive qui introduit une nouvelle matrice s(x, y).

$$s(x,y) = s(x,y-1) + i(x,y)$$
$$ii(x,y) = ii(x-1,y) + s(x,y)$$

6-30

où s(x,-1)=0 et ii(-1,y)=0 (Viola, et al., 2001)

L'image intégrale peut donc se calculer en une lecture de l'image de départ.

6.1.2.3. La somme et la moyenne des pixels sur un rectangle

A partir de l'image intégrale, nous pouvons calculer la somme des pixels sur un rectangle. Si nous souhaitons calculer la somme des pixels du rectangle bleu sur la Figure 136, nous devons calculer :

D-C-B+A

(Viola, et al., 2001)

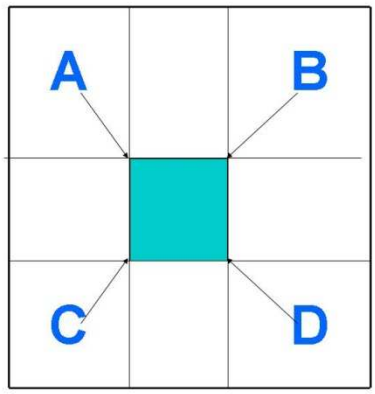


Figure 136 : représentation des points à sommer pour BOX FILTER

D étant la valeur en bas à droite du rectangle, sur l'image intégrale, C la valeur en bas à gauche, B la valeur en haut à droite et A la valeur en haut à gauche. Si nous reprenons notre exemple et que nous calculons la somme sur le carré 2x2 en haut à gauche, nous avons, sur l'image intégrale le carré correspondant affiché sur la Figure 137.

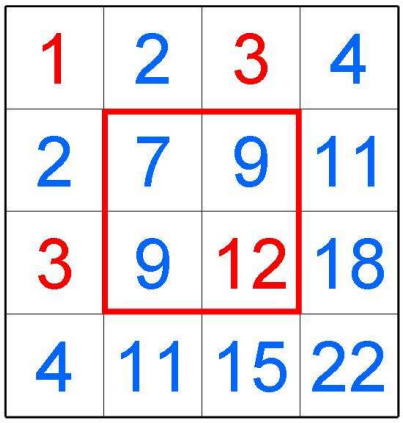


Figure 137 : représentation du processus de somme dans BOX FILTER 1

Où les valeurs de A, B, C et D sont affichées en rouge.

La somme des pixels sur ce carré est donc

D-C-B+A = 12-3-3+1 =7

Et nous pouvons vérifier avec l'image de départ : 4+1+1+1 =7 (Figure 138)

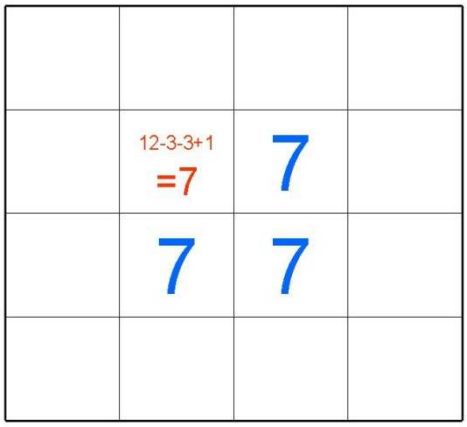


Figure 138 : représentation du processus de somme dans BOX FILTER 2

Bien sûr, pour un carré de 2x2 pixels, l'intérêt est assez mince, mais dès que ce rectangle est un petit peu plus grand, cette manœuvre prend tout son intérêt.

Maintenant que nous avons la somme de tous les pixels sur le rectangle, nous pouvons calculer simplement la moyenne sur ce rectangle en divisant par le nombre de pixels contenus dans ce dernier, ici 4 (Figure 139).

	7/4	1.75	
	1.75	1.75	

Figure 139 : explicitation des moyennes dans l'exemple BOX FILTER

6.1.2.4. Les hautes fréquences

A ce niveau ci, nous sommes en mesure de calculer les hautes fréquences de l'image. Nous retranchons la matrice de la moyenne à la matrice initiale (Figure 140).

1	1	1	1
1	4	1	1
1	1	1	4
1	1	1	1

—

	1.75	1.75	
	1.75	1.75	

=

	2.25	-0.75	
	-0.75	-0.75	

Figure 140 : hautes fréquences dans l'exemple de calcul de BOX FILTER

Nous voyons qu'il y avait une haute fréquence dans notre image de départ, puisque dans la matrice de hautes fréquences, il y a un nombre (2.25) relativement élevé.

6.1.2.5. Reconstruction

Normalement cet algorithme s'arrête ici. Nous voyons que cela pose problème, pour le but fixé : supprimer les hautes fréquences en certains endroits. En effet, nous savons où se situent les hautes fréquences, mais la reconstruction de l'image n'est pas encore possible. Pour cela, une légère modification a été apportée à cet algorithme, dont l'utilisation initiale n'avait pas besoin.

Puisque nous avions obtenu la matrice des hautes fréquences en retranchant la matrice des moyennes à l'image initiale, il était logique de supposer qu'en ajoutant la matrice des hautes fréquences à celle des basses fréquences (moyennes), nous retrouvions la matrice initiale (l'image initiale) aux effets de bords près.

Pour appliquer le traitement de l'image, nous supprimons donc les hautes fréquences gênantes avant d'ajouter l'image des hautes fréquences modifiée à la matrice des basses fréquences pour obtenir l'image originale modifiée.

6.1.2.6. Multi-résolution

Ci-dessus nous avons fait la décomposition et recomposition pour détecter les hautes fréquences qui montrent des variations sur des carrés de fréquences de 2 pixels de côté. Cependant, nous aimerions pouvoir le faire également pour des fréquences légèrement plus basses. En effet, nous souhaitons garder l'aspect « multi-résolution » que nous proposaient les transformées en ondelettes. Nous montrons ci-dessous qu'il est tout de même possible d'obtenir les mêmes informations.

Il est intéressant de considérer que nous avons obtenu cette matrice de hautes fréquences à partir d'une moyenne faite sur un carré de 2x2 pixels, donc, si nous prenons une matrice 4x4 pixels, voire 9x9 pixels nous obtenons des moyennes calculées sur un plus grand nombre de pixels et donc, une matrice des hautes fréquences révélant des fréquences de plus en plus basses. Plus nous prenons des rectangles de moyennes grands, plus la matrice de hautes fréquences montrera des détails grossiers.

6.1.2.7. Résultat des décompositions en images

En théorie donc, cette méthode devrait donner des images exploitables, mais quels sont exactement les résultats qu'elle peut donner, à quoi ressemblent ces matrices, de basses et de hautes fréquences?

Ci-dessous, une image que nous avions déjà utilisée comme exemple pour la transformée en ondelettes (Figure 141). Nous allons calculer son image intégrale (qui ne peut être représentée puisque les valeurs de ses « pixels » dépassent largement 255), Puis nous calculons la matrice de moyenne calculée sur 5 pixels et la matrice de détails associés.

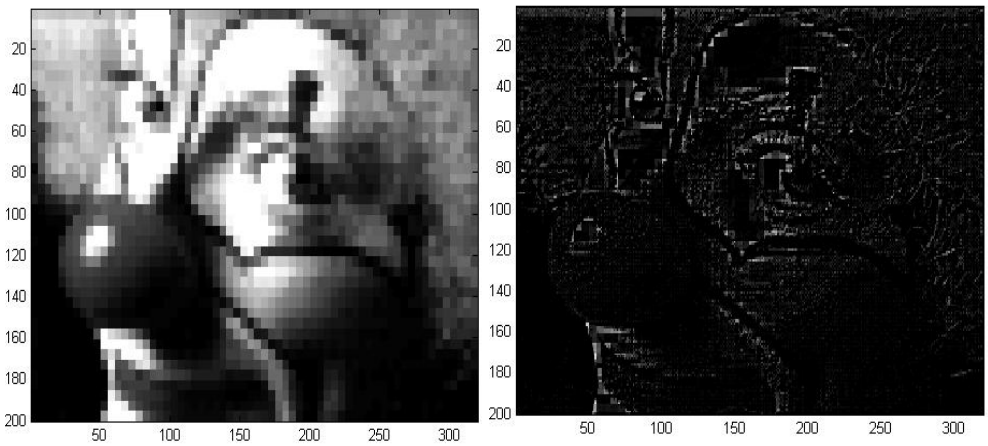


Figure 141 : décomposition en hautes et basses fréquences avec BOX FILTER sur une base de calcul carrée de 5 pixels de côté

Enfin, montrons la matrice de moyenne calculée sur 8 pixels et sa matrice de détails (Figure 142).

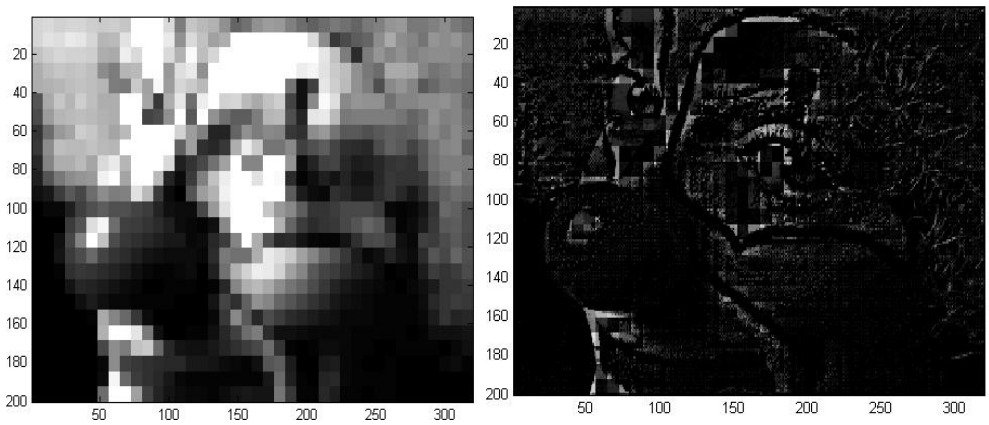


Figure 142 : décomposition en hautes et basses fréquences avec BOX FILTER sur une base de calcul carrée de 8 pixels de côté

Nous voyons que les images montrent bien une image floue et une autre montrant les détails qui ont été éliminés. Nous montrons en Annexe E le lien entre flou et hautes fréquences.

6.2. Comparaison des algorithmes

6.2.1. Comparaison ondelettes de Haar/BOX FILTER

6.2.1.1. Comparaison mathématique

Nous avons montré dans le paragraphe 6.1.1.5 que l'ondelette la plus rapide est l'ondelette de Haar. Elle est également la plus facile à utiliser et la plus facile à comprendre intuitivement.

L'ondelette de Haar a pour filtre passe-bas : *fb=[0.5, 0.5]* et pour filtre passe-haut *fh=[0.5,-0.5]*. Ce qui signifie que lorsque nous faisons la convolution numérique du signal s par le filtre passe-bas nous obtenons :

$$x[n]=s[n]*fb[2]+s[n+1]*fb[1] \quad \textit{6-31}$$

$$x[n]= s[n]*0.5+s[n+1]*0.5 \quad \textit{6-32}$$

Ce qui correspond à la moyenne de la donnée *s[n]* et de la donnée suivante *s[n+1]*.La convolution de ce filtre passe-bas et d'un signal arbitraire est représentée sur la Figure 143.

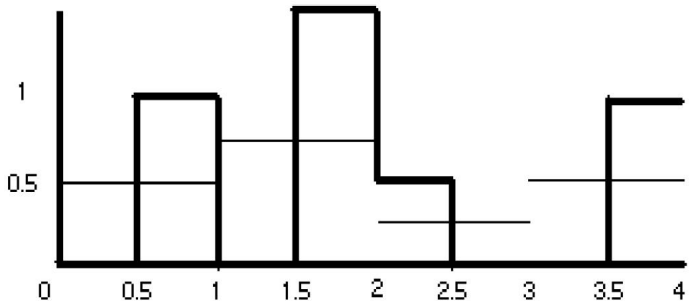


Figure 143 : exemple de calcul de transformée en ondelettes de Haar

Le signal est représenté en trait gras, tandis que la convolution du signal par le filtre passe-bas est représentée en trait fin. Nous voyons donc que la transformée en ondelettes de Haar correspond à faire une moyenne sur deux points consécutifs du signal.

6.2.1.2. Autre comparaison

Différences de principe

Rappelons que la transformée en ondelettes rapide est récursive. Pour avoir des fréquences montrant des détails sur 16 pixels, il faudra d'abord calculer les détails sur 2 pixels, puis sur 4, sur 8 avant de les calculer sur 16 pixels. Remarquons maintenant que ce n'est pas indispensable avec ce nouvel algorithme, puisque pour avoir des détails sur 16 pixels, il suffit de calculer la matrice des moyennes sur 16 pixels, ce qui ne prend pas plus de temps que la moyenne sur 2 pixels (grâce à l'image intégrale).

De même puisque, dans notre algorithme, l'image de départ est réobtenue en additionnant les matrices de moyennes et de détails, quel que soit le niveau de détails en question, la reconstruction n'est pas récursive non plus. Il n'est pas nécessaire de remonter niveau de détails par niveau de détails. Il suffit de calculer le niveau qui nous intéresse.

Notons également que pour notre reconstruction, il n'est pas nécessaire de reconstruire toute l'image. Imaginons, il n'y a qu'un petit morceau de l'image qui nous intéresse, disons un carré de 9 pixels sur 9. Nous ne modifions que celui-ci dans la matrice de détails, le reste de celle-ci restant

inchangée. Comme nous ne modifions jamais la matrice de basses fréquences, la somme de ces deux dernières est identique partout à l'image de départ, sauf sur le petit carré modifié. Il n'est donc pas utile de faire cette addition autre part que sur le carré en question, puisque, ailleurs, nous pouvons garder les pixels de l'image de départ.

Remarquons aussi que si nous souhaitons calculer des détails sur des rectangles de 9 pixels sur 11 ou sur 6, nous le pouvons parfaitement. Le rectangle de calcul s'ajustera tout de suite. Ce type de calcul n'est pas possible avec les ondelettes car les détails sont toujours calculés sur des carrés de $2^i, i \in \Re > 0$ pixels de côté pour des raisons d'orthogonalité.

Par contre, nous perdons certaines informations sur la direction de fréquences avec les bancs de filtres par rapport aux ondelettes. Par exemple, nous n'avons plus les trois matrices de haute fréquence en horizontal, vertical et diagonal. Nous pourrions retrouver les hautes fréquences verticales en calculant les matrices de basses fréquences sur des rectangles de largeur égale à celle de l'image et de hauteur correspondant au degré de détails. La matrice de détails horizontaux s'obtiendrait de la même façon, mais nous ne pouvons pas obtenir la matrice de hautes fréquences diagonales.

Différences de temps

Nous avions changé d'algorithme pour gagner du temps de calcul, observons maintenant si ce nouvel algorithme répond effectivement à cette attente. Pour pouvoir comparer les temps de calcul entre la transformation en ondelettes rapide et l'algorithme BOX FILTER, il faut comparer des traitements comparables. S'il faut 4 itérations à l'algorithme en ondelettes pour arriver à un niveau de détails de 16 pixels, nous obtenons tout de même les niveaux de détails de 2, 4 et 8 pixels en même temps. Donc, nous allons devoir calculer les niveaux de détails de 2, 4, 8 et enfin 16 pixels pour l'algorithme BOX FILTER pour obtenir les mêmes informations. Ci-dessous, le Tableau 1 des rapports de temps de calcul entre la transformée en ondelettes et BOX FILTER. Les deux algorithmes ont donc opéré 4 décompositions et recompositions chacun.

Taille de l'image	rapport de temps (ond/box)
216*216	1.937
512*512	17.733
1024*1024	9.945

Tableau 1 : rapport du temps de traitement d'image de la transformée en ondelettes et l'algorithme de BOX FILTER

Maintenant que nous savons que l'algorithme BOX FILTER est beaucoup plus rapide, il serait intéressant de savoir s'il calcule les décompositions et recompositions en temps réel. Cette fois, nous ne ferons qu'une décomposition et recomposition de toute l'image. Les résultats en images par seconde sont exposés dans la table ci-dessous.

Taille de l'image	nombre d'images par seconde
512*512	66
1024*1024	16

Tableau 2 : nombre d'images par seconde pour une décomposition et recompositon en BOX FILTER d'une image en fonction de la taille de celle-ci

Remarquons que pour une image 1024*024 nous ne sommes pas encore en temps réel. Par contre, pour une image 512*512, nous dépassons les 50 images par seconde requises pour la stéréoscopie. Cependant, il faut tenir compte du traitement et de la transformation du codage RGB en YUV De plus, dans la pratique, l'image doit être transférée de la carte graphique vers la carte mère (nous verrons pourquoi plus loin). Ces deux opérations quadruple le temps de calcul pour les images de 1024*1024.

6.2.2. Comparaison BOX FILTER/moyenne sur carte graphique

6.2.2.1. Comparaison mathématique et expérimentale

Nous avons vu que l'algorithme BOX FILTER calcule des moyennes sur des carrés de plusieurs pixels de côté. Il est donc normal que les résultats de ces deux algorithmes soient équivalents puisque par définition, l'algorithme tournant sur la carte graphique calcule des moyennes sur des carrés de pixels. Nous pouvons néanmoins montrer les résultats obtenus avec cet algorithme pour s'assurer qu'expérimentalement nous arrivons bien aux mêmes images. La Figure 144 est le résultat d'une moyenne effectuée sur des carrés de 2 pixels de côté.

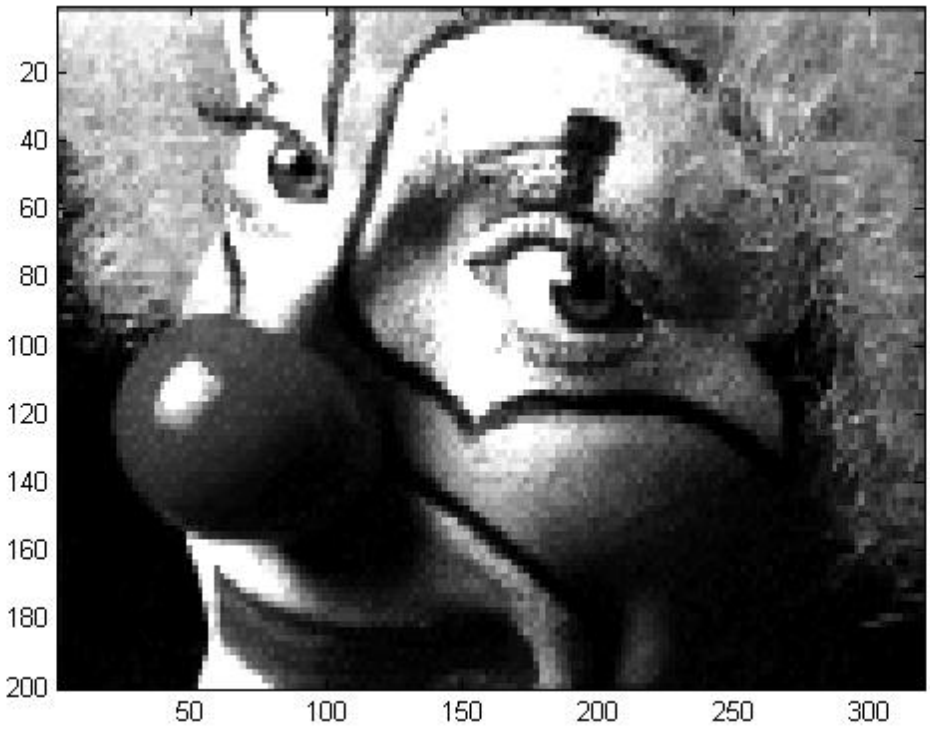


Figure 144 : décomposition avec une moyenne sur des carrés de 2 pixels de côté

La Figure 145 est le résultat d'une moyenne effectuée sur des carrées de 4 pixels de côté.

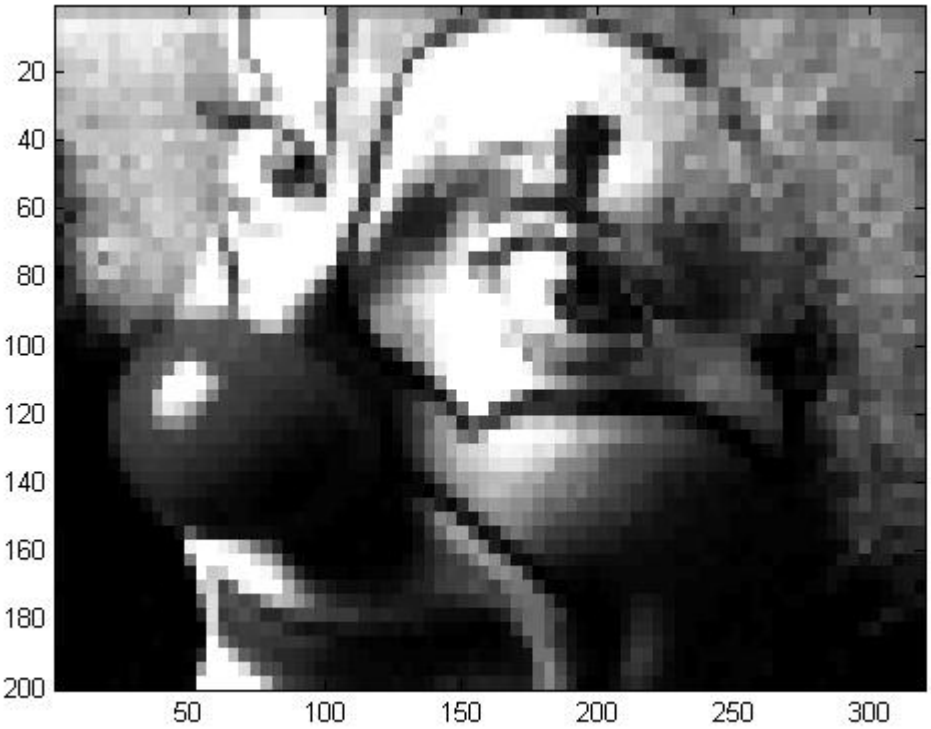


Figure 145 : décomposition avec une moyenne sur des carrés de 4 pixels de côté

Faisons le même calcul de similitudes entre ces images et celles obtenues par l'algorithme BOX FILTER pour vérifier que cette ressemblance n'est pas que visuelle : nous obtenons pour les moyennes réalisées sur des carrés de 2 pixels et 4 pixels comparées à l'algorithme BOX FILTER pour les mêmes précisions : 0 ! Ce qui signifie qu'il n'y a même pas de différence d'approximations numériques.

6.3. Utilisation de moyenne glissante et autres « flous »

6.3.1. Explication

Comme nous pouvons le voir sur les figures du paragraphe précédent, les images en basse résolution sont « pixellisées ». Pour supprimer ces crénelages, nous pourrions effectuer une moyenne glissante sur les pixels. Cela signifie que nous ne remplacerons plus les pixels d'un carré par la moyenne des valeurs de pixels sur ce même carré, mais que nous remplacerons chaque pixel par la moyenne effectué sur un carré entourant le dit pixel. La Figure 147 ci-dessous représente des moyennes sur 4 pixels calculées à parti de la Figure 146 comme précédemment.

1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24
25	26	27	28	29	30

31	32	33	34	35	36

Figure 146 : matrice de départ pour la différence de calcul des moyennes

4.5	4.5	6.5	6.5	8.5	8.5
4.5	4.5	6.5	6.5	8.5	8.5
16.5	16.5	18.5	18.5	20.5	20.5
16.5	16.5	18.5	18.5	20.5	20.5
28.5	28.5	30.5	30.5	32.5	32.5
28.5	28.5	30.5	30.5	32.5	32.5

Figure 147 : calcul des moyennes sur des carrés de pixels

En comparaison, la Figure 148 représente une moyenne glissante effectuée pour deux pixels :

4.5	5.5	6.5	7.5	8.5	6
10.5	11.5	12.5	13.5	14.5	12
16.5	17.5	18.5	19.5	12.5	18
22.5	23.5	24.5	25.5	26.5	24
28.5	29.5	30.5	31.5	32.5	30
31	32	33	34	35	36

Figure 148 : calcul des moyennes glissantes

Ce qui nous donne, appliqué à notre image de travail et pour une moyenne réalisée sur un carré de 2 pixels de côté, la Figure 149 qui, nous le voyons, est moins crénelée.

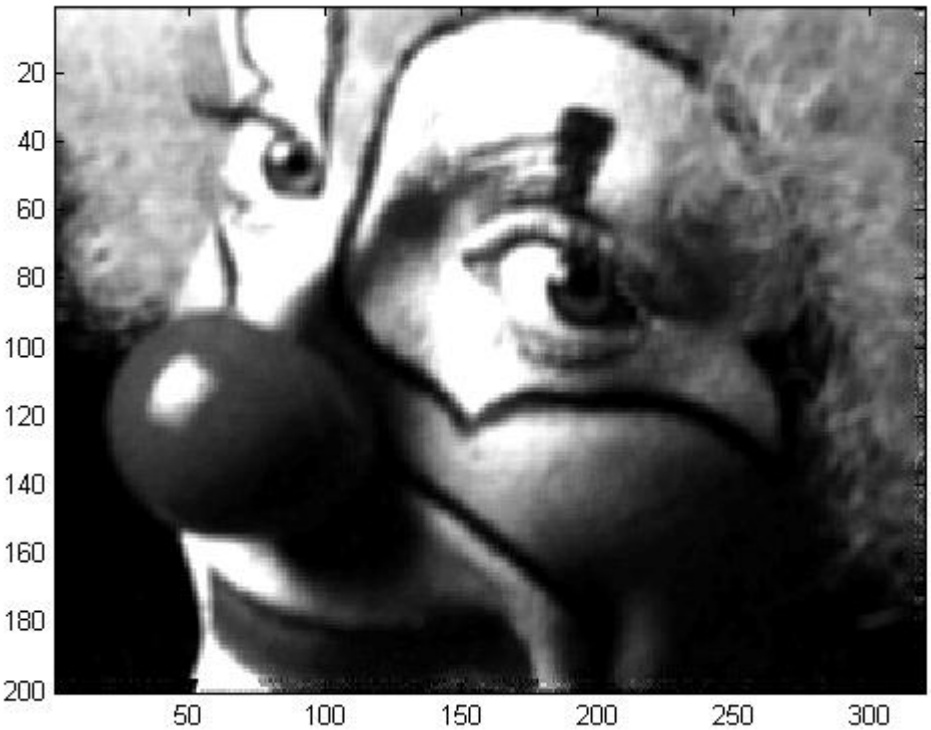


Figure 149 : moyennes glissantes sur des carrés de 2 pixels de côtés

6.3.2. Comparaison BOX FILTER glissant / moyenne glissante

Nous ne reviendrons pas sur la justification mathématique, nous avons vu que ce sont des moyennes totalement identiques (jusque dans les effets de bords). Par contre, nous développerons les différences de complexité et de rapidité.

6.3.2.1. Préambule : BOX FILTER / moyenne

Calculons la complexité pour des moyennes réalisées sur des carrés de *n* pixels de côté sur des images contenant *N*M* pixels.

Algorithme des moyennes :

Pour chaque petit carré, il faut faire *n*n* additions et une division :

$$n^2+1$$

Et cette opération est à réaliser autant de fois qu'il y a de carrés dans l'image : MN/n^2

Donc nous avons une complexité égale = $(n^2+1)*MN/n^2$

Algorithme BOX FILTER

Il faut calculer l'image intégrale : pour chaque pixel, il faut 2 additions (avec la formule de récursivité) : *2MN*

Et pour chaque carré de moyenne, il faut 4 additions et une division : *4+1*

Et ce multiplié par le nombre de carrés dans l'image : MN/n^2

Donc nous avons une complexité totale : $2MN+5MN/n^2$

Ces deux complexités pour une image de 1024*1024 pixels sont illustrées sur la Figure 150. La complexité de l'algorithme des moyennes en fonction de *n* est représentée en rond rouge, tandis que la complexité de BOX FILTER en fonction de *n* l'est en losange bleu.

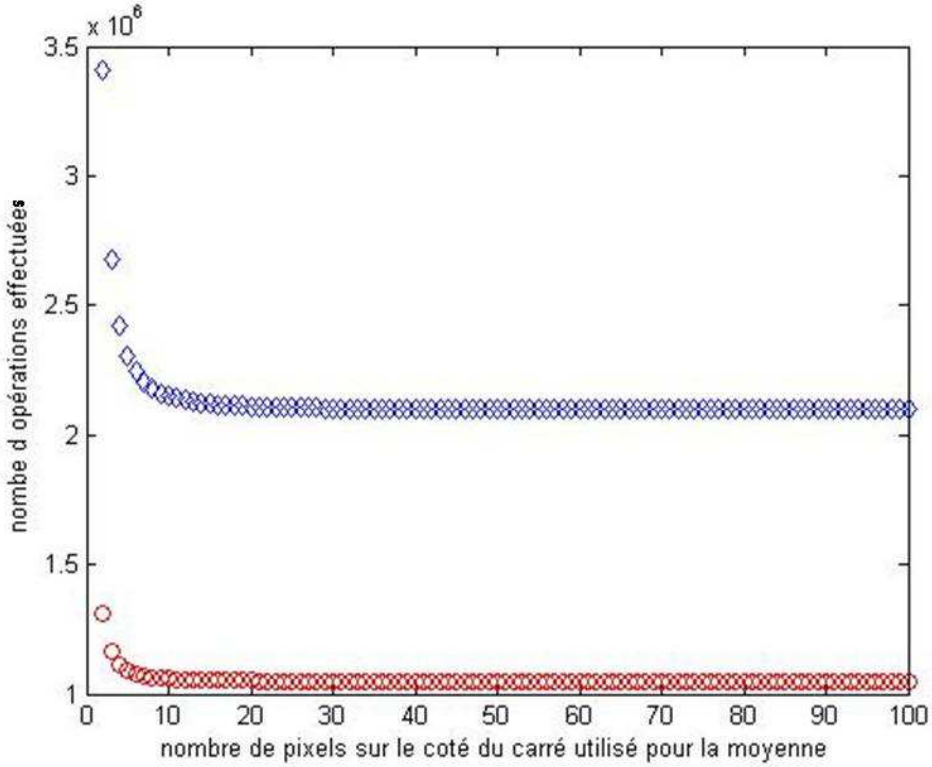


Figure 150 : complexité de BOX FILTER et des moyennes en fonction de la taille du carré de calcul pour une image de 1024*1024

Nous voyons que pour des images crénelées, donc correspondant à une décomposition de Haar, la complexité est moindre lors du calcul de la moyenne directement plutôt qu'avec BOX FILTER.

6.3.2.2. Complexité pour les algorithmes glissants

Algorithme des moyennes glissantes :

Pour chaque petit carré, il faut faire *n*n* additions et une division : n^2+1

Cette opération est à réaliser autant de fois qu'il y a de pixels dans l'image *: MN*

Donc nous avons une complexité globale : $(n^2+1)*MN$

Algorithme BOX FILTER :

Calculons l'image intégrale : pour chaque pixel, 2 additions sont nécessaires (avec la formule de récursivité) : *2MN*

Et pour chaque carré de moyenne: 4 additions et une division : *4+1*

Et ce, multiplié par le nombre de pixels dans l'image : *MN*

Donc nous avons une complexité globale : *2 MN +5MN =7MN*

Qui est une complexité constante.

Ces deux complexités pour une image de 1024*1024 pixels sont représentées sur la Figure 151. La complexité de l'algorithme des moyennes en fonction de n est représentée en rond rouge, tandis que la complexité de BOX FILTER en fonction de n est représentée en losange bleu.

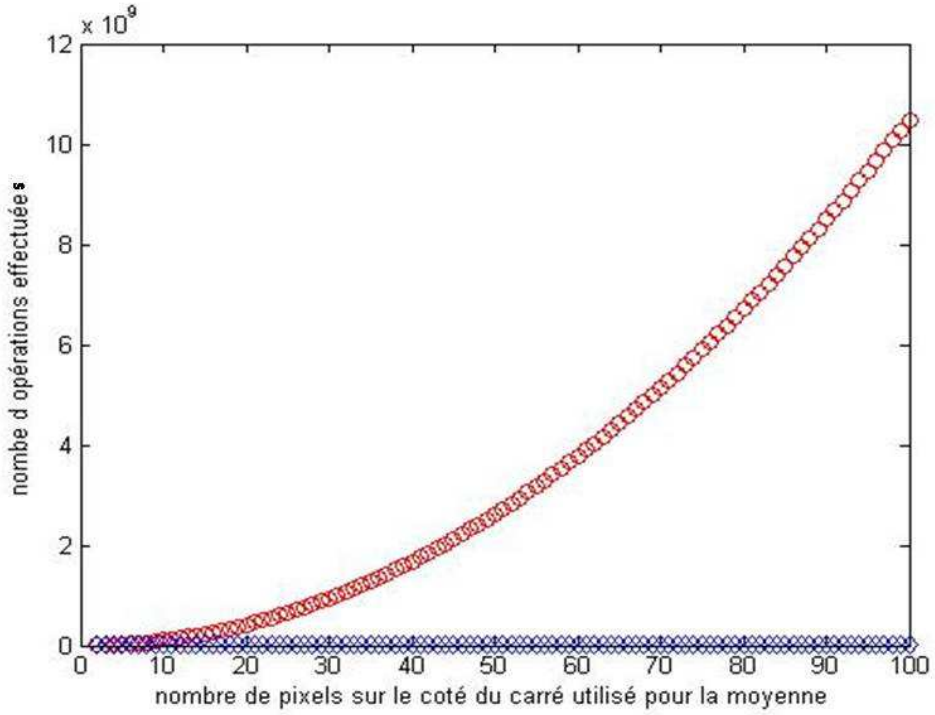


Figure 151 : complexité de BOX FILTER glissant et des moyennes glissantes en fonction de la taille du carré de calcul pour une image de 1024*1024

Nous voyons ainsi que très rapidement l'algorithme BOX FILTER est beaucoup plus performant en complexité et donc en rapidité.

6.3.2.3. Différence d'algorithme pour une différence de support

Mais alors, quel est l'intérêt de l'algorithme de la moyenne glissante si sa complexité explose en fonction du nombre de pixels sur lesquels est réalisée la moyenne ? Les complexités qui sont calculées ci-dessus donnent une idée fausse de la différence de temps pour calculer ces traitements car ils ne sont pas réalisés sur le même support ! En effet, l'algorithme BOX FILTER glissant est calculé sur la carte mère (le CPU) alors que l'algorithme des moyennes glissantes est calculé sur la carte graphique (le GPU). Or la rapidité d'accès mémoire à la carte graphique ou à la carte mère à partir d'un programme comme Virtools est sans comparaison. Pour faire le traitement sur le processeur de la carte mère, il faut d'abord transférer les images de la carte graphique vers la carte mère, faire le traitement et renvoyer l'image traitée vers la carte graphique pour qu'elle soit affichée. Or si le temps d'envoi de données du processeur vers la carte graphique est relativement rapide, il n'en est pas de même lors du trajet inverse !

Mais alors, pourquoi ne pas faire le traitement BOX FILTER sur la carte graphique ? Tout simplement parce qu'un programme de carte graphique s'exécute autant de fois qu'il y a de pixels en parallèle. Et qu'à partir d'un pixel, on accède facilement aux pixels voisins ou encore un peu plus loin, mais relativement difficilement aux pixels de toute l'image. Or, pour l'algorithme BOX FILTER, il faut calculer l'image intégrale avant de calculer les moyennes, donc, sur une carte graphique il faudrait recalculer une image intégrale autant de fois qu'il y a de pixels dans l'image ! Donc, la complexité devient : $2\ M^2N^2 + 5MN$. Ce qui est considérable. Pour obtenir la même complexité entre les moyennes glissantes et BOX FILTER glissant sur carte graphique, il faudrait alors calculer des moyennes glissantes sur des carrés de 1448 pixels de côté pour des images des 1024*1024 pixels ce qui est un non-sens.

6.3.2.4. Différence de temps

Comparons donc la rapidité des ondelettes de Haar sur le CPU, de BOX FILTER glissant sur le CPU et de la moyenne glissante sur le GPU vu de Virtools (donc en plein fonctionnement) :

- Ondelettes : 0.5 image par seconde
- BOX FILTER glissant : 4.6 images par seconde
- Moyenne glissante : 97 images par seconde

6.3.3. Autres « flous »

Nous voyons sur la Figure 149 qu'elle est encore légèrement moirée. Pour résoudre ce problème, nous pouvons lui appliquer des filtres moins « carrés ». Parmi eux, il existe les filtres quadratiques et gaussiens. Ils sont tous les deux calculés de la même façon que les moyennes glissantes et sont donc réalisés sur la carte graphique et non sur la carte mère (voir paragraphe 6.3.2.3). Ce qui les différentie des moyennes glissantes c'est le poids accordé aux pixels sur le bord de la base de calcul.

6.3.3.1. Filtres quadratiques

Equation

Ce filtre accorde un poids aux pixels de la base en fonction du carré de la distance au pixel traité. Nous calculons donc la distance entre le pixel traité et les pixels composant la base de calcul.

Si (X_{calcul}, Y_{calcul}) *est la position en ligne et en colonne du pixel* traité

Et (X, Y) *est la position en ligne et en colonne d'un pixel de la base de calcul,*

nous définissons la distance entre ces deux pixels :

$$d = \sqrt{(X_{calcul} - X)^2 + (Y_{calcul} - Y)^2}$$

6-33

Nous donnons à chaque pixel de la base de calcul un poids :

$$\begin{cases} 1 - d^2/r^2 & \text{si } 1 - d^2/r^2 > 0 \\ 0 & \text{Sinon} \end{cases}$$

Où *r* représente nombre de pixels sur le côté du carré de la base de calcul.

Ensuite nous additionnons les valeurs des pixels multipliées par leur poids et nous divisons cette valeur par l'addition de tous les poids des pixels en question. Ceci nous donne la valeur du pixel traité.

Résultat

La Figure 152 représente la moyenne quadratique glissante réalisée avec r=4 sur la Figure 123.

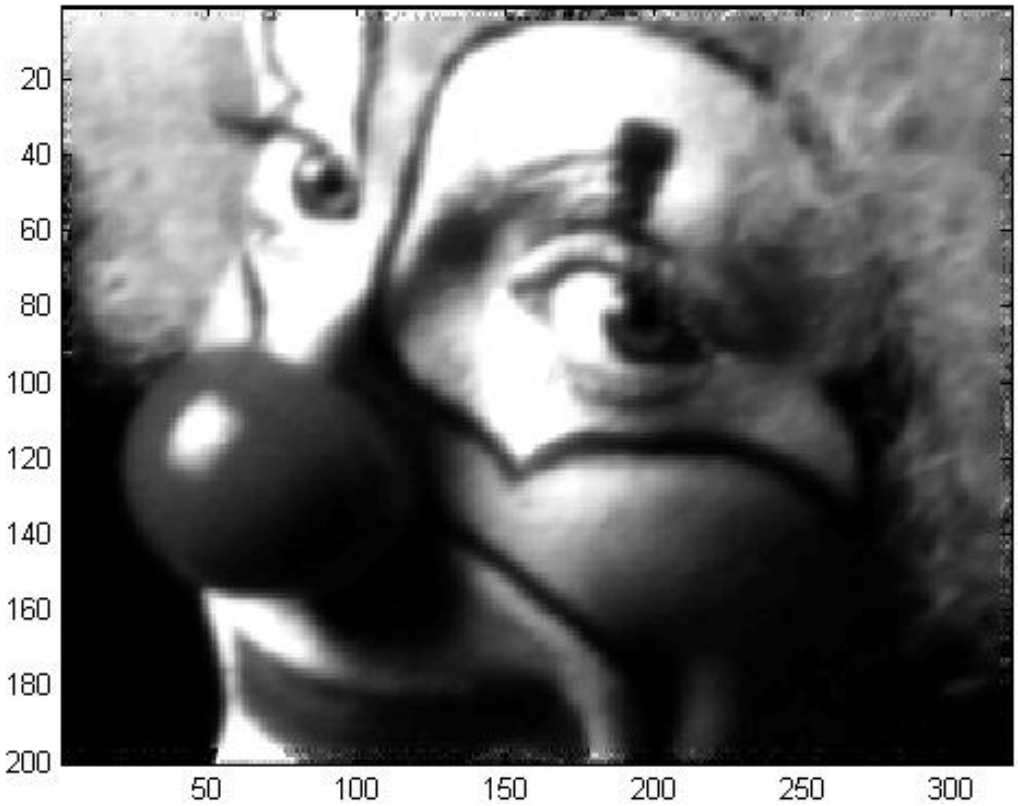


Figure 152 : moyenne quadratique glissante réalisée avec une base de calcul de 4 pixels de côté

Nous voyons que le moiré est très atténué, mais reste présent.

Remarque

Ce type de flou est assez pratique à utiliser, mais reste cependant assez lourd à calculer à cause notamment du calcul de la distance qui contient une racine carrée.

6.3.3.2. Filtres gaussiens

Equation

Ce filtre accorde un poids aux pixels de la base suivant une fonction gaussienne en 2 dimensions centrée sur le pixel traité. Si *r* représente le nombre de pixels sur le côté du carré de la base de calcul, le pixel traité a un poids de 2^{r-1}. Les autres poids sont également des puissances de deux. Les puissances les plus faibles sont nulles ($2^0 = 1$) et elles sont incrémentées en allant vers le centre de la base de calcul.

Nous pouvons écrire la répartition de poids pour une base de calcul de *r* pixels de côté :

$$\begin{pmatrix} 2^0 & 2^1 & & 2^1 & 2^0 \\ 2^1 & 2^2 & & 2^2 & 2^1 \\ & & 2^{r-1} & & \\ 2^1 & 2^2 & & 2^1 & 2^0 \\ 2^0 & 2^1 & & 2^2 & 2^0 \end{pmatrix}$$

Donc pour une base de calcul carré de 7 pixels de côté nous aurons une matrice de poids égale à :

1	2	4	8	4	2	1
2	4	8	16	8	4	2
4	8	16	32	16	8	4
8	16	32	64	32	16	8
4	8	16	32	16	8	4
2	4	8	16	8	4	2
1	2	4	8	4	2	1

Ensuite nous additionnons les valeurs des pixels multipliés par leur poids et nous divisons cette valeur par l'addition de tous les poids des pixels en question. Cela nous donne la valeur du pixel traité.

Résultat

La Figure 153 représente la moyenne quadratique glissante réalisée avec *r*=4 sur Figure 123.

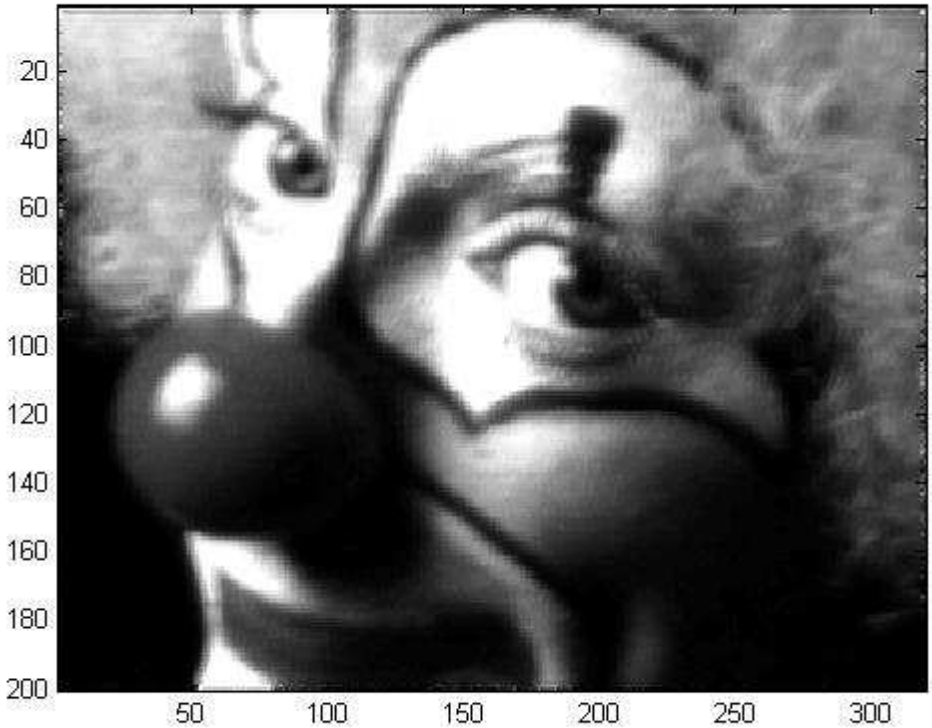


Figure 153 : moyenne gaussienne glissante réalisée avec une base de calcul de 5 pixels de côté

Nous voyons que le moiré est encore légèrement atténué.

Remarque

Ce type de flou est très utilisé dans le traitement d'image, notamment dans les logiciels industriels. Son calcul en base 2 le rend très aisé à calculer.

6.4. Comparaison globale

Dans le Tableau 3, nous récapitulons les différentes propriétés des trois algorithmes. Celles qui nous intéressent ou qui pourraient nous intéresser sont en vert. Nous avons donc décidé d'utiliser les calculs de moyennes glissantes sur GPU. En effet, c'est la seule méthode qui permet de faire nos calculs en temps réel. Cependant, la transformée en ondelettes aurait pu être très intéressante à cause de son lien direct avec les fréquences. Nous avons réussi à palier ce manque pour les calculs de « flou » sur GPU (voir ANNEXE E), mais un lien direct aurait été plus intuitif. De plus, nous avions montré dans le chapitre 1 le lien entre les ondelettes et notre système visuel, cela aurait été plus judicieux de garder ce lien. L'algorithme de Box filter aurait été une bonne alternative par rapport à la transformée en ondelettes pour la rapidité si nous étions restés sur CPU. Il nous aurait permis de ne traiter qu'une seule zone très facilement. En effet, si nous ne souhaitons traiter que la périphérie de l'écran, cela aurait très facilement possible avec cet algorithme en allégeant les traitements d'images.

Critère	Ondelettes	Box filter	GPU
Nombres d'images traitées dans Virtools par seconde	0.5	4.6	97
Temps réel	Non	Non	Oui
Décomposition	Image entière	Image entière	Par objet
Recomposition	Image entière	Zone	Par objet
Choix du « flou »	Oui	Non	Oui
Lien direct avec les fréquences	Oui	Non	Non

Tableau 3 : comparatif des trois algorithmes étudiés

6.5. Conclusion

Nous avons exposé dans ce chapitre trois algorithmes de suppression de hautes fréquences par endroit. Une comparaison mathématique et fonctionnelle a été réalisée entre eux. Nous avons décidé d'utiliser le traitement sur GPU pour tester l'influence de cette suppression sur la fatigue visuelle. Dans le prochain chapitre, nous étudions cette influence, nous exposons notre protocole d'expérimentation ainsi que nos résultats.

Chapitre 7 : Expérimentations et résultats bruts sur la diminution de la fatigue visuelle en vision stéréoscopique

7.1. Introduction

Dans ce chapitre, nous présentons nos expérimentations pour analyser la diminution de la fatigue visuelle. Nous commençons par expliciter les méthodes de mesure de la fatigue visuelle, ensuite nous expliquons en détail notre expérimentation et notre protocole. Nous terminons le chapitre en commentant nos résultats bruts.

7.2. Moyens de mesurer la fatigue visuelle

Il existe deux grandes sortes de tests de fatigue : les tests objectifs et les tests subjectifs. Les tests objectifs sont basés sur des mesures optométriques. Les tests subjectifs sont des questionnaires pour savoir si la personne est plus fatiguée, si elle a mal aux yeux, etc.

Quel que soit le test utilisé, il est très important de faire une mesure de fatigue visuelle avant et après l'immersion en vision stéréoscopique pour avoir un élément de comparaison. Ces deux mesures sont nécessaires. En effet, un sujet peut ressentir une certaine fatigue visuelle avant de commencer le test (manque de sommeil, stress, longue lecture…) et cette information est cruciale lorsque nous étudions la fatigue visuelle après le test.

7.2.1. Les tests plus objectifs

7.2.1.1. Acuité visuelle

L'acuité visuelle est définie comme la séparation angulaire minimum entre deux lignes nécessaire pour qu'elles soient vues séparées. Elle correspond à la performance monoculaire et est également utilisée par les médecins pour contrôler les effets d'une affectation oculaire et son évolution (Arden, 1988). Nous pouvons donc l'utiliser comme mesure de la fatigue visuelle : plus l'angle visuel requis est grand plus la fatigue visuelle est grande.

Différentes mesures possibles

- Mesure de détection : « voyez vous une ouverture dans le cercle ? »
- Mesure descriptive : « dans quel sens est l'ouverture du cercle ? »
- Mesure d'interprétation : « quelle est la lettre ou le chiffre ? »
- Mesure interactive : «Les lettres forment elles un mot ? Vous évoque-t-il quelque chose ? »

Les optotypes

Un optotype est un motif ou caractère de forme et de dimensions bien définies, destiné au contrôle de la qualité de la vue.

Les anneaux de Landlot

Appelés également le « C » de Landlot, ils constituent l'optotype de référence le plus largement accepté pour l'utilisation en laboratoire des examens visuels (Figure 154). Il s'agit d'un cercle non fermé dont la largeur du trait et la largeur d'ouverture sont égales à 1/5eme du diamètre extérieur. Les bords de l'ouverture doivent être parallèles. Il n'y a aucun empâtement. Son plus gros avantage est de n'avoir qu'un seul élément de discrimination : l'orientation de l'ouverture (Arden, 1988).

Figure 154 : cercles de Landlot

Cependant, ces optotypes posent plusieurs problèmes de communication qui détériorent la mesure de l'acuité visuelle. En effet, quand plusieurs optotypes sont présentés en ligne, il peut y avoir une erreur de sens de lecture (de droite à gauche au lieu de gauche à droite), de plus le sujet peut répéter deux fois le même optotype décalant alors les réponses suivantes. D'autre part, la lettre « C » étant la plus habituelle à lire, il est possible que la position de l'ouverture à droite soit plus facilement reconnaissable que les autres.

Les chiffres et les lettres

Ce sont ceux qui sont couramment utilisés par les opticiens et les ophtalmologues. C'est un test de reconnaissance de lettres noires sur fond blanc (Figure 155). Ces dernières sont classées par ligne d'acuité progressive (croissante ou décroissante). Cependant ce type de test rend relativement non maîtrisable les éléments extérieurs comme la mémorisation des lettres ou l'influence du décor autour du panneau (Neveu, 2008) ou la charge émotionnelle d'un mot (s'ils sont utilisés).

Figure 155 : optotypes lettrés et chiffrés

Autre type d'optotypes rencontrés

L'acuité peut également être mesurée avec des lignes noires et blanches (Costello, et al., 1996).

7.2.1.2. Fonction proximum d'accommodation

Ce test, également appelé « Donder's Push-up Test », consiste à placer une mire visuelle à 40cm des yeux du sujet et de la rapprocher de celui-ci avec une vitesse comprise entre 5 et 7,5 centimètres par seconde. Lorsque le sujet ne voit plus la mire nettement, il le signale et nous mesurons alors la distance entre la mire et les yeux du sujet. Ce test peut se faire plusieurs fois d'affilée pour obtenir une moyenne. Normalement il se fait avec les verres correcteurs du sujet s'il en porte et il se fait d'abord avec les deux yeux, l'œil gauche et enfin l'œil droit.

Plus la distance finale est grande, plus le sujet a de difficultés à accommoder et donc nous pouvons en déduire que s'il y a une différence entre deux mesures c'est que les yeux du sujet étaient plus fatigués pour la mesure la plus élevée (Benjamin, 2007).

7.2.1.3. Facilité d'accommodation

Connu aussi sous le nom de « Flipper Lens Test » ce test permet de savoir en combien de temps nos yeux peuvent accommoder à un nouveau stimulus de façon répétitive. Nous plaçons une mire visuelle à 40cm des yeux du sujet et nous plaçons très rapidement des lentilles de +2 dioptries entre les yeux du sujet et la mire (en plus des verres correcteurs du sujet si nécessaire - Figure 156). Dès que la personne voit la mire nettement, elle le signale et nous plaçons des lentilles de -2 dioptries et ainsi de suite pendant 30 secondes.

A la fin des 30 secondes, nous relevons le nombre de cycles. Un sujet normal obtient généralement 5 cycles. S'il en a moins, cela peut signifier que ses yeux sont fatigués (Benjamin, 2007).

Figure 156 : flipper Lens Test

7.2.1.4. Acuité stéréoscopique

L'acuité stéréoscopique est le seuil juste perceptible de la parallaxe stéréoscopique. Elle varie en fonction des individus et des conditions de l'expérience. Elle est donc un bon indicateur de la performance sensorielle et motrice du système visuel, le seuil augmentant avec la fatigue visuelle.

Tests en vision de loin

Le Polatest de Zeiss (zeiss) consiste en plusieurs lignes polarisées projetées sur un écran. Il est demandé au sujet de déterminer quel objet est représenté devant le plan (Figure 157 et Figure 158).

Figure 157 : le Polatest de Zeiss (zeiss)

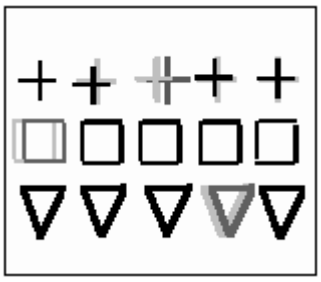


Figure 158 : projection de ligne avec le Polatest

Les acuités présentées sont 0.5’, 1’, 2’, 3’, 4’, 5’ d’arc. L’acuité stéréoscopique de loin du sujet est l’acuité correspondant à la première ligne pour laquelle il réussit à indiquer le relief.

Tests en vision de près

Lors de l’exécution des tests suivants, les mires de visions sont positionnées à 40 cm des yeux du sujet perpendiculairement à l’axe de la vision. Elles doivent être éclairées avec une lumière normale et suffisamment forte.

Test de la mouche

Le « test de la mouche » représente, comme son nom l’indique, une mouche en 3 dimensions vue d’en haut (page de droite de la Figure 159). Les ailes sont au-dessus du plan (2 ou 3 cm) et le sujet doit essayer de les attraper avec les doigts. Nous notons la hauteur désignée par le sujet ainsi que sa rapidité d’exécution.

Figure 159 : test de la mouche et test de Wirt (precision-vision)

Test de Wirt

C’est une photo polarisée d’un point parmi un ensemble de 4 points et dans une série de 9 figures (partie gauche de la Figure 160). Le sujet doit indiquer quel point se soulève par rapport aux autres. Les acuités mesurées pourront varier de 800 secondes d’arc pour le test n°1 à 40 secondes d’arc pour le test n°9.

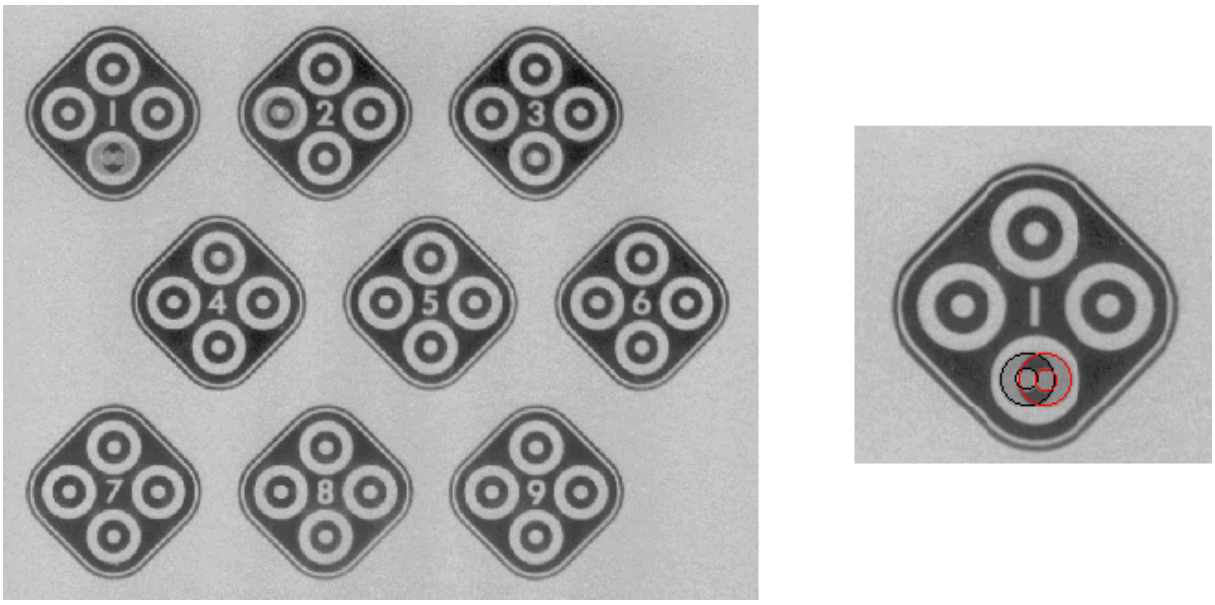


Figure 160 : points de Wirt

Test de Lang

Figure 161 : test de Lang

Ce test, développé par le professeur Lang, n'est pas réellement utilisé pour mesurer une acuité stéréoscopique précise, mais plus pour savoir si une personne possède la vision stéréoscopique. Il est basé sur des stéréogrammes de points aléatoires représentant des formes avec une acuité stéréoscopique. Les enfants doivent déterminer la forme qu'ils perçoivent (voiture, éléphant) (Figure 161).

Test de Randot

Figure 162 : test de Randot (precision-vision)

Ce test est basé sur les stéréogrammes de points aléatoires. Les sujets doivent identifier 6 formes géométriques pendant le test. Les acuités présentées vont de 400 à 20 secondes d'arc (Figure 162).

Test TNO

Ce test utilise des anaglyphes pour séparer les deux images gauche et droite (Figure 163). Il comporte 7 planches, les trois premières permettent d'établir si la vision stéréoscopique de près est présente, trois autres permettent une mesure quantitative, une planche sert à mettre en évidence une absence de stéréoscopie. Les acuités mesurées vont de 15 à 480 secondes d'arc. Il possède donc la plage la plus étendue des acuités présentées avec le test de Wirt.

Figure 163 : test TNO

7.2.1.5. Les hétérophories dissociées

L'hétérophorie est une déviation des axes visuels maintenue latente par le réflexe de fusion (Hugonnier, et al., 1965). L'hétérophorie ne peut donc être mise en évidence que lorsque la tendance à la fusion est momentanément supprimée. Il faudra donc placer le couple oculaire en position passive en brisant la fusion. Si la déviation est nulle avant l'expérience et marquée après, nous pourrons en déduire une certaine adaptation.

La seule méthode objective de mise en évidence de la phorie est l'occultation d'un œil. Lorsque l'on relâche cette occultation on relève la vergence de l'œil. Cependant cette méthode ne permet pas une mesure simple. Dans tous les autres tests on demandera au sujet de préciser ce qu'il voit. Il s'agit donc de méthodes subjectives. Cependant, ce test révèlera plus une adaptation à une phorie forcée qu'une fatigue visuelle.

7.2.1.6. Notre choix

Nous avons décidé de nous appuyer sur trois mesures objectives de fatigue visuelle : la facilité d'accommodation, la fonction proximum d'accommodation et les points de Wirt pour l'acuité stéréoscopique. Les deux premiers nous donnent un renseignement sur la fatigue générale de l'œil, sur sa vivacité de réaction et sur les limites qu'il est encore capable d'atteindre. Ils sont plus précis que l'acuité visuelle simple. Nous souhaitions évidemment avoir également une mesure de la fatigue stéréoscopique et donc une mesure de l'acuité stéréoscopique. Pour qu'elle soit la plus fine possible, nous avions le choix entre le test TNO et les points de Wirt. Cependant le fait que le Test TNO se fasse sur anaglyphe peut perturber légèrement la mesure. Il est donc préférable de mesurer l'acuité stéréoscopique grâce au point de Wirt. Nous n'avons pas retenu les mesures d'hétérophories dissociées car, comme nous l'avons dit dans le paragraphe précédant, elles révèlent plus une adaptation à une phorie forcée que la fatigue visuelle en elle-même.

7.2.2. Les tests subjectifs

Les tests subjectifs sont une série de questions à poser avant et après les tests : le ressenti de l'observateur, ses impressions, sa fatigue... Mais il faut faire très attention : on peut focaliser l'attention du candidat sur une impression qu'il n'aurait peut-être pas relevée sans le questionnaire avant le test (Costello, et al., 1996).

7.3. Le protocole

7.3.1. La tâche

7.3.1.1. Le monde Virtuel

Notre monde virtuel se compose de 5 cylindres de 10 cm de diamètre et de 20 cm de hauteur, ils sont à une hauteur de 80 cm et sont répartis de part et d'autre de l'écran réel. Il y a également une sphère de 5 cm de diamètre, au bout d'un tube de 60 cm « relié » à la WiiMote (Figure 55). Il y a également un carré rouge de 10 cm de côté qui vient se poser sur un cylindre choisi aléatoirement. Il y a également un fond d'écran et trois prismes. Tous ces objets (à part le carré rouge) sont couverts par des textures très riches en hautes fréquences (voir Figure 164 et Figure 165 pour une vue de coté sans texures).

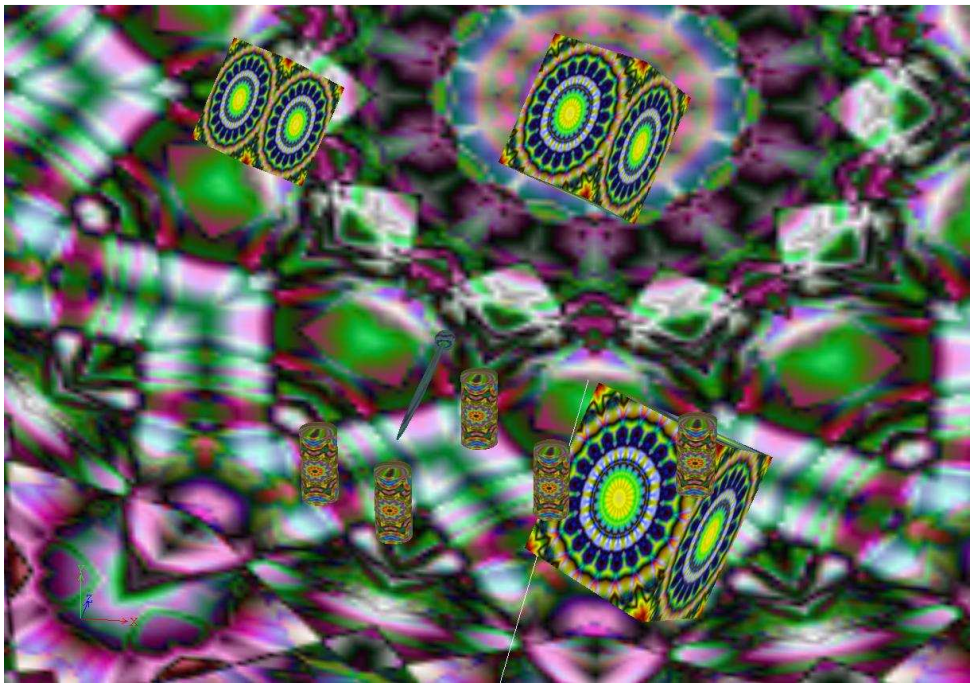


Figure 164 : le monde virtuel

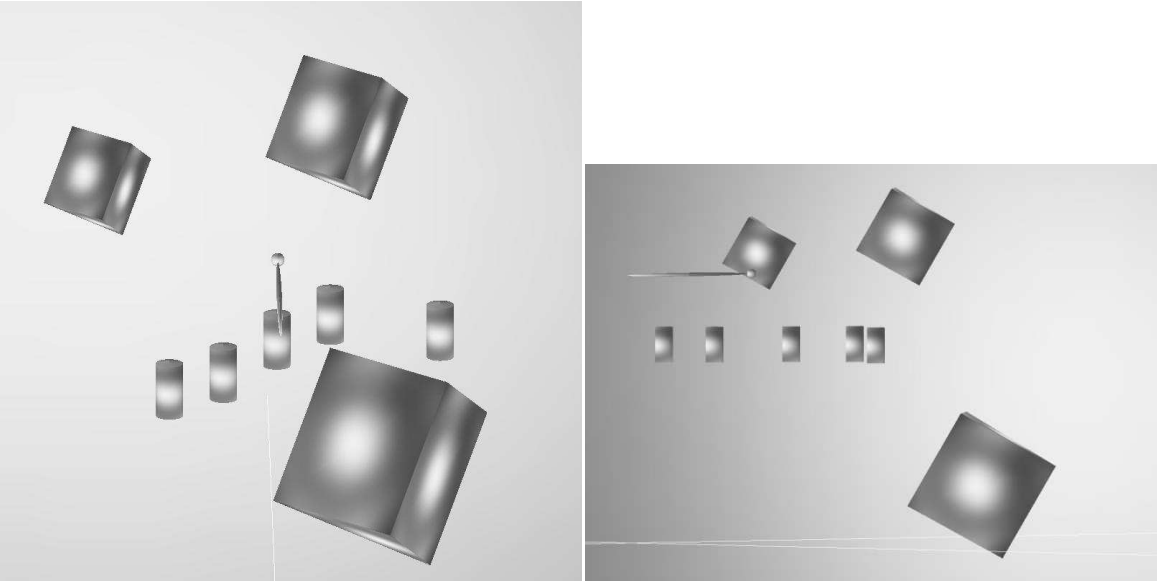


Figure 165 : monde virtuel sans texture vu de face et de coté

7.3.1.2. Le « but du jeu »

Le carré rouge va venir se poser sur un cylindre choisi de façon aléatoire (sans repêchage). Il est demandé au sujet de venir poser le bas de la sphère tangentiellement au carré rouge, de le maintenir dans cette position 2 secondes puis de placer le centre de la sphère au centre du carré rouge. Il attend de nouveau 2 secondes, puis appuie sur le bouton « A » de la WiiMote (Figure 55).

Lorsque le bouton « A » est pressé, le carré « choisit » un autre cylindre, tandis que le précédent translate horizontalement de gauche à droite avec une vitesse et un sens aléatoire, mais dans un temps donné (3 secondes). De cette façon, le sujet ne peut mémoriser l'emplacement des cylindres. Cependant, cette translation est annulée s'il y a un risque que le cylindre sorte du champ de

vision. Le carré rouge se pose sur chaque cylindre 50 fois. Ce jeu est une question de précision, pas de rapidité, le sujet peut donc prendre autant de temps qu'il le souhaite avant de compter ses 2 secondes pour être sûr d'être bien positionné. Ce jeu dure environ 30 minutes.

Cette tâche a été choisie parce qu'elle permet de concentrer l'attention du sujet un certain temps sur des objets situés de part et d'autre de l'écran, afin de fixer le regard sur différentes disparités. Elle comprend une mesure de précision qui nous permet de vérifier si notre traitement diminue l'efficacité d'un sujet lors d'une tâche. Enfin, ce « jeu » ressemble à une opération de manutention que pourrait effectuer un ouvrier dans un environnement virtuel (en apprentissage par exemple).

7.3.2. Les mesures de fatigues effectuées

7.3.2.1. Les mesures objectives

Efficacité de la tâche

Nous mesurons l'erreur de distance entre le centre de la sphère et le centre du carré lors de la pression du bouton « A » de la WiiMote, ainsi que le temps passé lors de la réalisation de la tâche. Ce n'est pas une mesure de fatigue à proprement parler, mais plus une mesure d'efficacité. Bien entendu, pour que cette mesure soit valable, le sujet ne doit pas avoir plus d'entraînement pour l'expérience avec le traitement que celle sans ce dernier. L'ordre des expériences sera donc laissé au hasard. Cette mesure est effectuée pendant chaque session.

Fonction proximum d'accommodation

La fonction proximum d'accommodation est mesurée selon les règles dictées au paragraphe **Erreur ! Source du renvoi introuvable.**. Elle est mesurée avant et après chaque session, à chaque ois, deux mesures sont prises (avec les deux yeux) pour faire la moyenne. Cette mesure a été choisie pour sa capacité à détecter de très faibles variations dans la fatigue visuelle (Lambooij, et al., 2009).

Facilité d'accommodation

La facilité d'accommodation est mesurée selon les règles dictées au paragraphe 7.2.1.3. Cependant, nous la pratiquons pendant 1minute pour plus de précision. Elle est mesurée avant et après chaque session.

Acuité stéréoscopique

L'acuité stéréoscopique est mesurée grâce aux points de Wirt (voir paragraphe 7.2.1.4). Elle est mesurée avant et après chaque session. Cette technique de mesure a été choisie pour la plage très étendue des acuités mesurées et leur finesse.

7.3.2.2. Les mesures subjectives

Les mesures subjectives sont présentées sous formes de questions posées au sujet après que les deux mondes aient été expérimentés. Elles sont au nombre de trois :

1. Quel monde était le plus esthétique ?
2. Est-ce que la tâche était plus facile que la dernière fois ?
3. Est-ce que la tâche était plus fatigante que la dernière fois ?

7.3.3. Déroulement

Le test se fait sur deux jours différents : un jour pour le test sans le traitement d'image, un autre avec ce dernier. Cet espacement est nécessaire pour que la fatigue du premier test n'influence pas le second, quel que soit l'ordre dans lequel ils sont présentés. Les deux jours ne sont pas forcément consécutifs.

Lors de l'accueil du sujet, il tire « à pile ou face » pour savoir s'il commence par la scène traitée contre la fatigue ou par la scène non traitée. Les tests accommodatifs et stéréoscopiques sont mesurés et la simulation est lancée. Les règles sont expliquées. Le sujet est invité à nous signaler lorsqu'il est prêt. Il effectue la tâche (Figure 166), puis les mesures accommodatives et stéréoscopiques sont reprises pour pouvoir comparer avec les mesures prises avant la tâche. Le second test se fait avec un repos oculaire d'au moins 24h, les acuités visuelles et stéréoscopiques sont remesurées avant de commencer le test.

Figure 166 : sujet en train de réaliser la tâche non floue

7.3.4. Les sujets

Les sujets choisis sont des personnes entre 20 et 40 ans. Cette limite supérieure a été fixée pour que la presbytie ne perturbe pas les mesures accommodatives, particulièrement la facilité accommodative (Benjamin, 2007). Ils pouvaient présenter des problèmes visuels tels que la myopie ou l'astigmatisme à condition qu'ils portent leur correction visuelle (lunettes ou lentilles). Aucun d'eux ne présentait de problème de stéréoscopie. Les résultats sont présentés sur un panel de 20 personnes, en majorité des hommes (85%).

7.3.5. Calcul et Algorithme choisi

Nous avons choisi l'algorithme de flou sur carte graphique réalisé avec une moyenne glissante pour pouvoir effectuer les traitements en temps réel. Le flou adaptatif a été calculé selon la fonction de Jérôme Perrin (voir paragraphe 5.2.2) : pour *C(d,f)* le confort visuel, *f* la fréquence spatiale présente et *d* la disparité,

$$C(d, f) = a(d - d_0 - kf^{k'})$$

7-1

Les paramètres d'interpolation sont :

$a = -0.010$

$d_0 = 18.9$

$k = 221.1$

$k' = -0.74$

Nous en déduisons donc la fréquence maximale pour un confort donné :

$$f = \sqrt[k']{\frac{1}{ka}(ad - ad_0 - C(f, d)}$$

7-2

Nous considérons que le confort doit au moins être égal à 0. Pour calculer la fréquence acceptable grâce à l'équation 7-1, il nous manque la disparité *d* ou la parallaxe horizontale. Celle-ci se calcule à partir de la position du sujet et de la position du point virtuel (Figure 167).

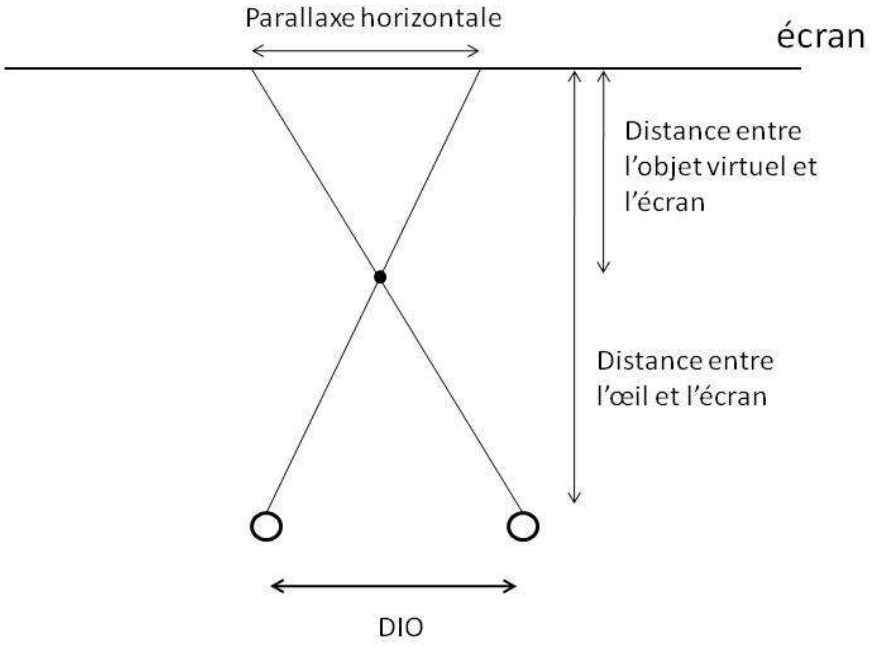


Figure 167 : calcul des parallaxes horizontales

Grâce aux relations des triangles semblables, nous pouvons écrire :

$$ParallaxeHorizontale = \frac{Distance\ ceObjet * DIO}{(distance\ ceObjet - Distance\ ceOeil)}$$

7-3

Nous connaissons donc maintenant la fréquence maximale acceptable en un point. Cependant, un programme de carte graphique ne considère que les pixels et pas la fréquence spatiale. Nous allons devoir exprimer notre fréquence spatiale (en cycle par degré d'angle visuel) en pixels (Figure 168).

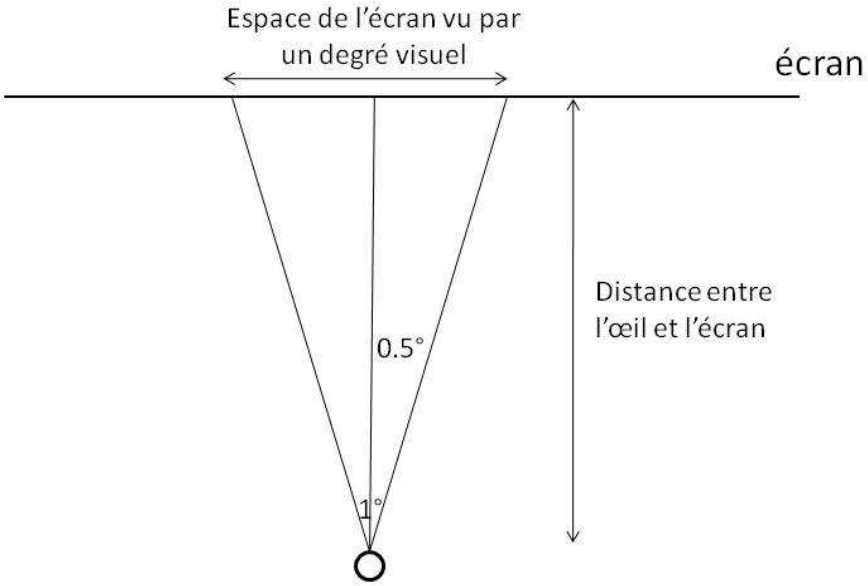


Figure 168 : calcul du nombre de pixel par degré d'angle visuel

Nous partons du demi-angle visuel pour utiliser les relations trigonométriques dans un triangle rectangle :

$$\tan(\frac{1°}{2}) = \frac{Distance\ ceSurEcran\ /2}{distance\ ceOeil}$$

7-4

Donc

$$Distance\ ceSurEcran = 0.0087 * distance\ ceOeil$$

7-5

Et comme nos pixels font 1,61mm, nous pouvons écrire :

$$nombrePixelParDegréVisuel = \frac{0.0087 * distance\ ceOeil}{0.00161}$$

7-6

Les moyennes seront donc calculées sur ce nombre de pixels divisé par la fréquence spatiale acceptée pour une disparité donnée.

7.3.6. Dispositif physique

7.3.6.1. Test de vision

Lors des tests de vision (accommodatifs ou d'acuité stéréoscopique), les mires devaient se situer à 40 cm des yeux des sujets. Un pupitre de musique solidaire d'un siège de voiture a donc été utilisé pour garantir cette distance (Figure 169).

Figure 169 : la mire est fixée à 40 cm des yeux du sujet grace à un pupitre fixé au siège

7.3.6.2. Projection

Le dispositif de projection est un écran LUMIN d'une largeur de 3,10 m pour une hauteur de 1,74 m. Le projecteur est de type Christie mirage 3. La distance entre l'écran et l'optique du projecteur est de 3,5m. Nos pixels font 1,61 mm de côté. La fréquence de projection est de 60 images par seconde. Le système de tracking est un ART2 fonctionnant également à 60 Hz et d'une précision millimétrique.

7.4. Analyse des résultats

7.4.1. Fonction proximum d'accommodation

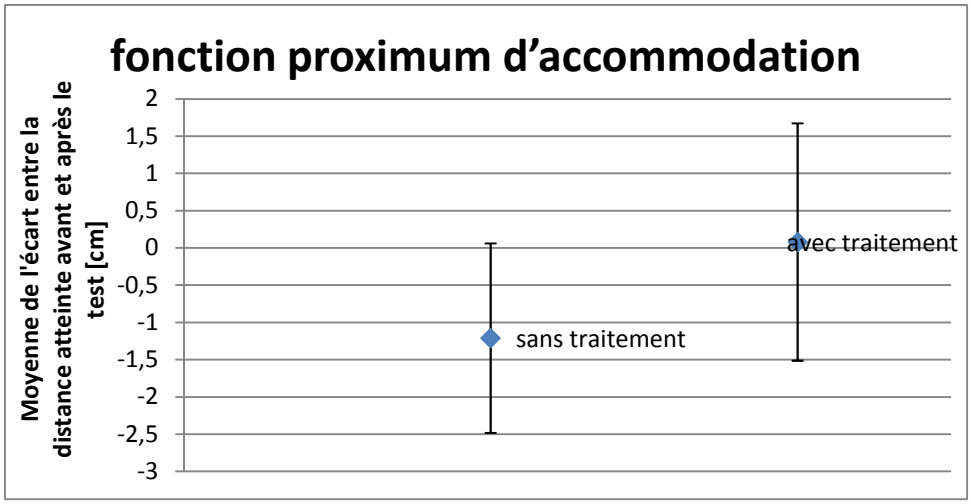


Figure 170 : modification de la fonction proximum d'accommodation

Sur la Figure 170 sont présentées les différences de mesures de la fonction proximum d'accommodation prises avant et après chaque tâche. Nous voyons que la fonction proximum d'accommodation n'est presque pas modifiée lorsque nous appliquons le traitement de suppression progressive des fréquences spatiales. Lorsque ce dernier n'est pas appliqué, les sujets perdent 1,21cm. Leurs yeux sont donc plus fatigués lorsque notre traitement d'image n'est pas appliqué.

Nous détaillons le calcul de significativité dans l'annexe C. La probabilité associée à la significativité des différences de moyenne est de 99.6%, donc cette différence n'est pas une simple erreur statistique.

7.4.2. Facilité d'accommodation

La Figure 171 représente les différences des écarts relatifs entre le nombre de demi-cycles du support de lentille présenté sur la Figure 156 avant et après la tâche. Comme nous pouvons le voir, la différence est en moyenne de 1,77 demi-cycle par minute. La répartition des écart relatif de demi-cycle suit une loi normale. Cette différence est significative puisque la probabilité de significativité associée à celle-ci est 99.45%.

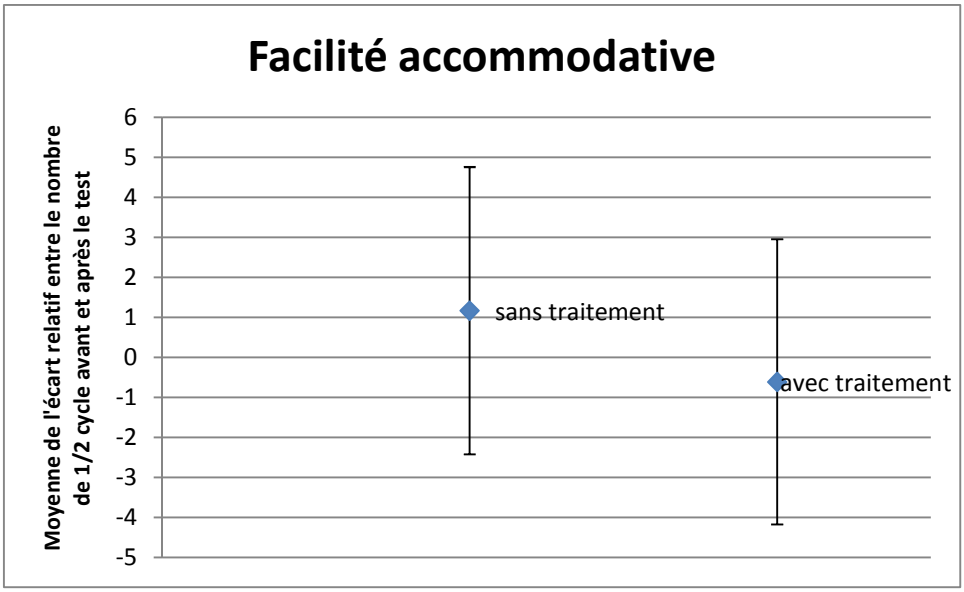


Figure 171 : modification de la facilité accommodative

7.4.3. Acuité stéréoscopique

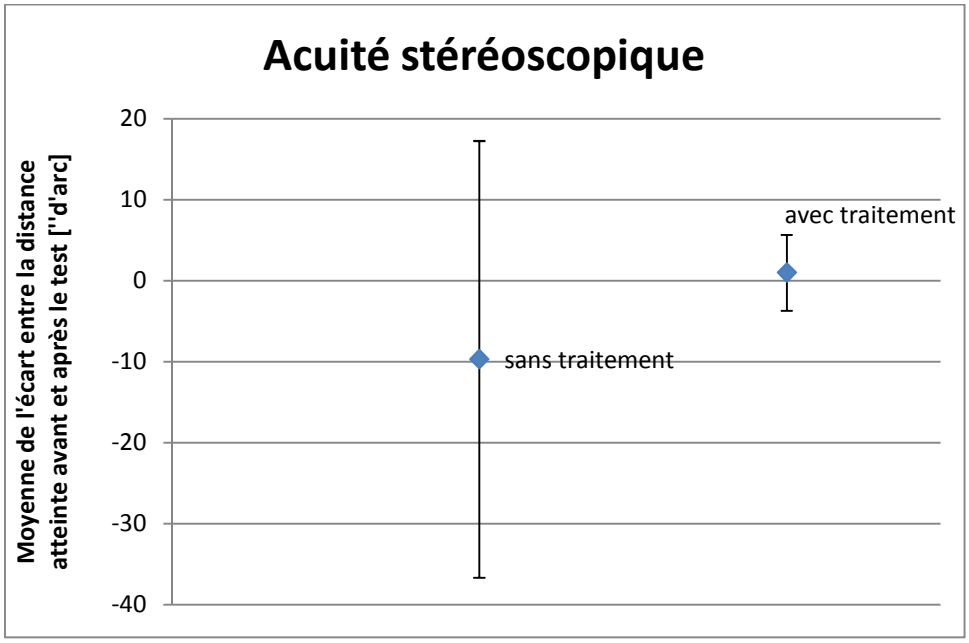


Figure 172 : modification de l'acuité stéréoscopique

Sur la Figure 172, nous voyons que la différence d'écart d'acuité stéréoscopique est assez importante. L'acuité stéréoscopique semble être nettement plus diminuée lorsque notre traitement n'est pas appliqué et à l'inverse, elle ne semble pas modifiée lorsque notre algorithme modifie l'image.

Cependant, il faut noter que cette distribution n'est malheureusement pas normale. La différence entre les deux écarts-types est significative et n1 et n2 sont inférieurs à 30. Cependant, lors-

que nous calculons la significativité avec ces tests, nous obtenons à chaque fois une probabilité de risque <0.00001.

7.4.4. Efficacité lors de l'exécution de la tâche

Sur la Figure 173, nous présentons un graphique sur lequel les barres bleues représentent la moyenne des erreurs commises sans traitement (en mètres) tandis que les barres rouges représentent les moyennes des erreurs commises avec traitement pour chaque cylindre. Les barres d'erreurs représentent les écart-types. Cependant, aucun écart de moyennes n'est significatif, mais les différences d'écart type sont toutes augmentées de façon significative lorsqu'il n'y a le traitement, sauf pour le cylindre n°3, qui est positionné sur l'écran.

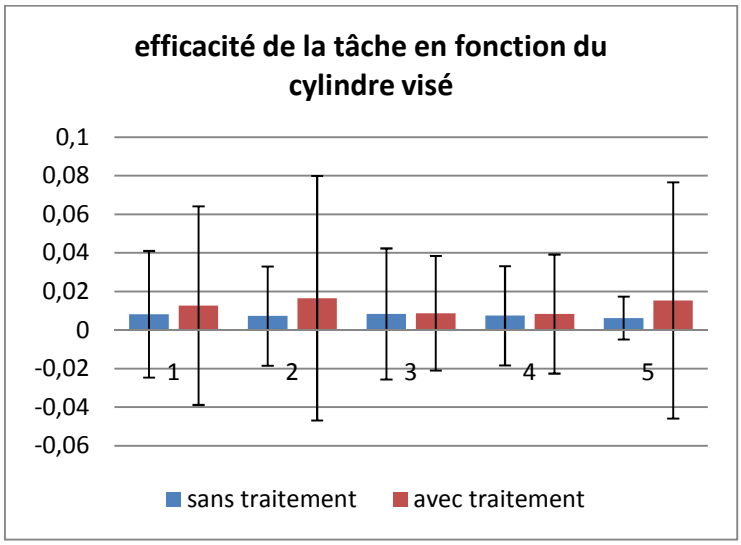


Figure 173 : efficacité de la tâche en fonction du cylindre pointé

7.4.5. Mesures subjectives

Sur la Figure 174, nous voyons que les sujets préfèrent le monde qui n'est pas traité d'un point de vue esthétique. Notons que presque le tiers des sujets interrogés n'ont pas vu la différence d'un jour sur l'autre.

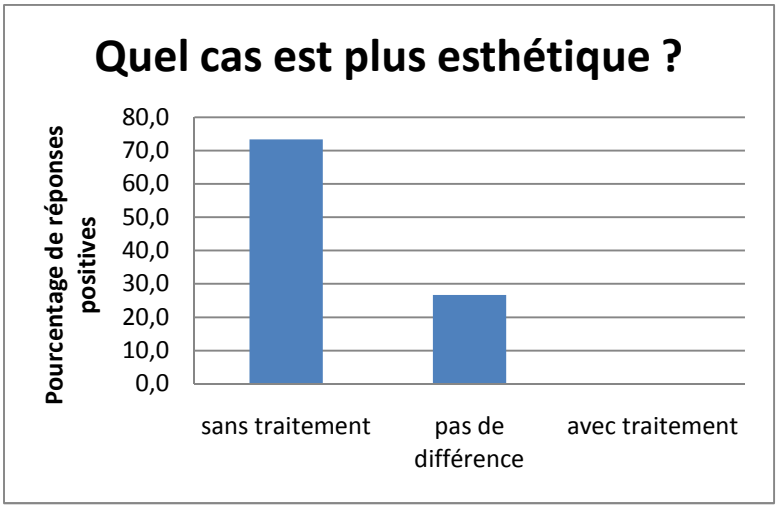


Figure 174 : quel cas est plus esthétique?

Sur la Figure 175 nous pouvons voir que deux fois plus de sujets estiment que le monde sans notre traitement adaptatif est plus fatiguant que le monde sans ce dernier. Ceci tend plutôt à confirmer les résultats des tests de fatigue visuelle que nous avons réalisés dans les paragraphes précédents.

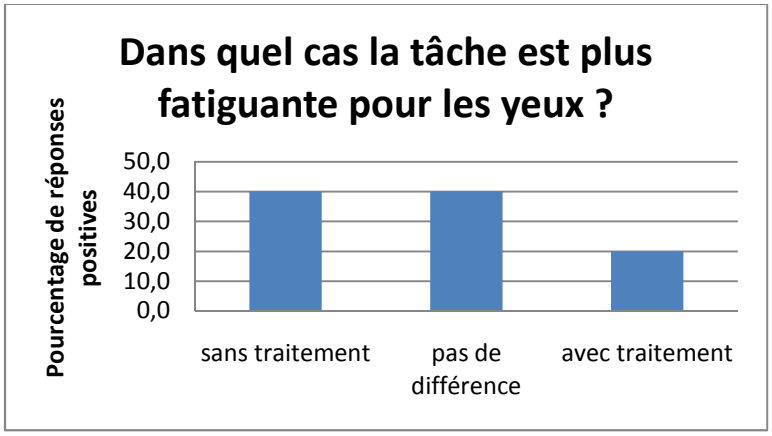


Figure 175 : dans quel cas la tâche est plus fatigante pour les yeux?

Sur la Figure 176 nous voyons que les sujets trouvent également que la tâche est plus facile lorsque que le monde virtuel est traité avec notre traitement de l'image.

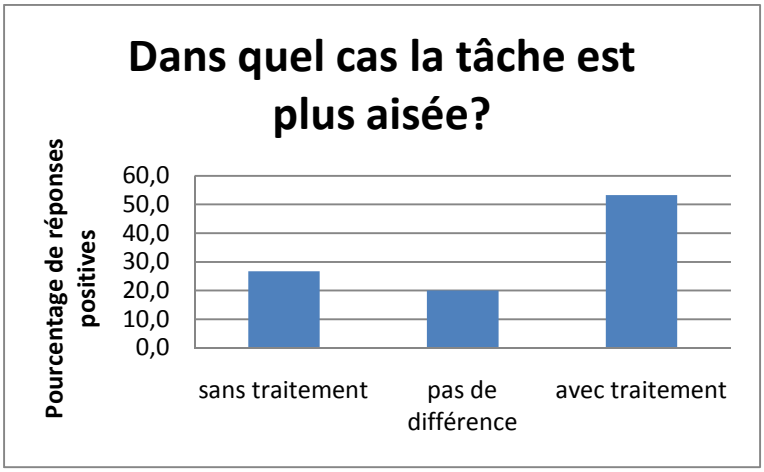


Figure 176 : dans quel cas la tâche est plus aisée?

7.5. Conclusion

Dans ce chapitre, nous avons exposé les méthodes de mesure de la fatigue visuelle ainsi que notre protocole de test. Nous avons montré que notre traitement d'image par flou adaptatif diminue la fatigue visuelle en vision stéréoscopique. Cependant, il diminue quelque peu l'efficacité de la tâche ainsi que l'aspect esthétique du monde virtuel. Dans le chapitre suivant, nous discutons de nos résultats et de notre protocole de test.

Chapitre 8 : Discussion des résultats et de la méthode

Dans ce chapitre, nous discutons des résultats obtenus aux chapitres précédents. Nous discutons également de la méthode de traitement d'image pour limiter la fatigue. Nous discutons notamment le fait de traiter objet par objet ou toute l'image en une fois, les avantages et inconvénients de chacune des méthodes.

8.1. Discussion des résultats

Nous avons vu que notre traitement diminue de façon significative la détérioration de la fonction proximum d'accommodation ainsi que la facilité d'accommodation après immersion. La diminution de l'acuité stéréoscopique après immersion est très nettement réduite, voire même annulée. Ceci tend à montrer que notre traitement diminue réellement la fatigue visuelle en immersion. De plus nos sujets ont remarqué que la tâche était moins fatigante.

Par contre, les images traitées sont moins jolies et moins bien acceptées par les sujets, même s'ils reconnaissent qu'elles sont moins fatigantes. Cela pourrait être gênant lorsque l'esthétique entre en ligne de compte. L'esthétique peut être importante lors de l'inspection d'un intérieur de voiture par un designer dans notre application. Il est également important pour une œuvre artistique, une démonstration au public, lorsque l'utilisateur a besoin d'apprécier la finition de son produit…

Nous voyons que l'efficacité de la tâche n'a pas été amoindrie. L'erreur de précision n'a pas été modifiée. Par contre, son écart-type a été augmenté. Cela pourrait signifier qu'en moyenne les sujets n'ont pas fait plus d'erreur, mais que par rapport au monde virtuel sans traitement, certains ont fait plus d'erreurs avec traitement et d'autres en ont fait moins. D'autre part, beaucoup de sujets affirment que la tâche est plus aisée lorsque l'image est traitée contre la fatigue visuelle.

8.2. Traiter les matériaux ou traiter l'image ?

Nous avons vu au paragraphe 6.4 que le traitement réalisé sur la carte graphique, est réalisé sur chacun des objets, alors que les traitements en ondelettes et en BOX Filter étaient réalisés sur toute l'image. Nous sommes donc en droit de nous interroger sur les implications d'une telle différence. Nous avons donc réalisé deux programmes sur carte graphique, l'un travaille matériaux par matériaux (ou objet par objet), l'autre traite toute l'image en post-traitement (comme une texture). Il est assez logique de penser qu'il n'y aura pas de différences notables s'il y a très peu d'objets (comme dans le monde utilisé pour notre test sur l'influence de notre traitement), les différences entres les deux seront très faibles. Par contre, dans un monde avec beaucoup d'objets différents (ou des objets composés de beaucoup de matériaux différents) les différences sont beaucoup plus visibles. Notons que c'est souvent le cas dans les mondes virtuels : par exemple un intérieur de voiture virtuel est constitué de beaucoup de petits objets (boutons, poignées, grilles…) Il est donc intéressant de faire cette comparaison pour un monde plus compliqué que le monde « de test ». Néanmoins, il serait tout de même plus aisé de comparer ces deux méthodes sur un monde relativement régulier.

Pour illustrer cette différence, nous appliquons les deux méthodes sur un monde ou le nombre de matériaux est très grand. Ce monde est représenté sur la Figure 177 : il s'agit d'une flotte

de vaisseaux spatiaux. L'avantage de ce monde virtuel et qu'il est composé de beaucoup de petits vaisseaux identiques, eux-mêmes composés de 3 matériaux différents : le cockpit, la carlingue et les ailes. Ces deux derniers matériaux sont couverts de textures contenant des très hautes fréquences. Nous pouvons donc apprécier les différences sur les petits objets rapprochés. Il contient également un plus gros vaisseau, qui est composé d'un matériau unique.

Figure 177 : image de départ

8.2.1. Traiter les matériaux

Nous commençons par appliquer l'algorithme matériau par matériau. Nous voyons sur la Figure 178 que les hautes textures des objets ont été floutées, mais les contours des matériaux restent très nets : par exemple, le cockpit du premier petit vaisseau (très flouté) a un contour très marqué. les disparités nulles se situent sur le quatrième petit vaisseau qui, lui, n'est pas flouté. Nous voyons également que nous n'avons pas appliqué notre traitement sur le fond étoilé.

8.2.1.1. Avantages

Un des avantages est le fait que si un des objets est important, nous pouvons décider de ne pas lui appliquer le traitement. Dans notre cas, nous avons décidé de ne pas l'appliquer sur le fond pour l'exemple, mais n'importe quel objet aurait pu ne pas être traité. C'est notamment important lorsque qu'il y a un objet qui doit être modifié en détail par un opérateur (dans un contexte industriel) ou un objet qui doit être mis en évidence pour guider les yeux de l'observateur (dans un contexte ludique ou artistique).

Un autre avantage de cette technique est de garder les contours nets. En effet, si la tâche est de dénombrer le nombre de vaisseaux, elle reste possible avec cette méthode : nous pouvons faire la

différence entre chaque objet. D'autre part, il est plus facile de percevoir la profondeur d'un objet si ses contours sont nets, les disparités des bords n'étant pas complètement effacées. De plus, le fait que ce soit moins « esthétique » est moins marqué.

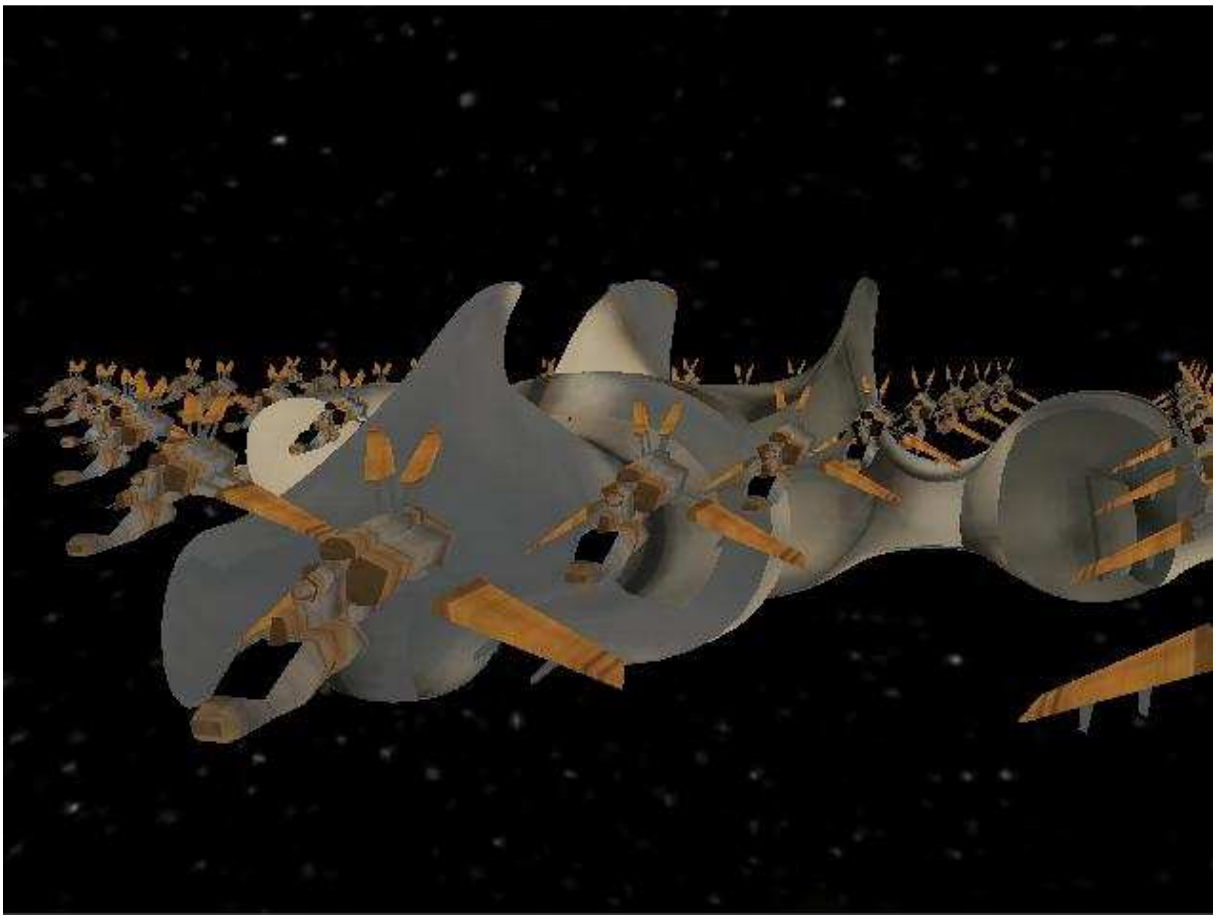


Figure 178 : traiter objet par objet

8.2.1.2. Inconvénients

Il y a malheureusement quelques inconvénients à cette méthode. Comme la différence entre les objets reste marquée, s'il y a beaucoup de petits objets, il y a beaucoup de hautes fréquences dues aux « ruptures » entre ces objets. Dans notre cas, nous avons beaucoup de petits vaisseaux dans le fond et leurs ailes provoquent des hautes fréquences.

Un autre désavantage est que ce traitement est plus difficile à gérer pour un utilisateur. En effet, ce traitement doit être appliqué sur tous les matériaux. Donc chaque fois que l'utilisateur introduit un nouveau matériau, il doit appliquer le traitement. Cela peut vite s'avérer contraignant lorsqu'il y a beaucoup de matériaux comme dans les mondes virtuels un petit peu complexes. De plus, plus il y a de matériaux, plus le traitement est lourd. En effet, il traite également les parties des objets qui sont occultés.

8.2.2. Traiter l'image

Sur la Figure 179, nous exposons le traitement effectué sur toute l'image. L'endroit qui n'est pas flouté se trouve également au niveau du quatrième petit vaisseau. Contrairement à l'autre méthode, les petits vaisseaux ne sont plus discernables après le cinquième voire le sixième vaisseau. Ceci correspond mieux à une situation réelle du flou de profondeur.

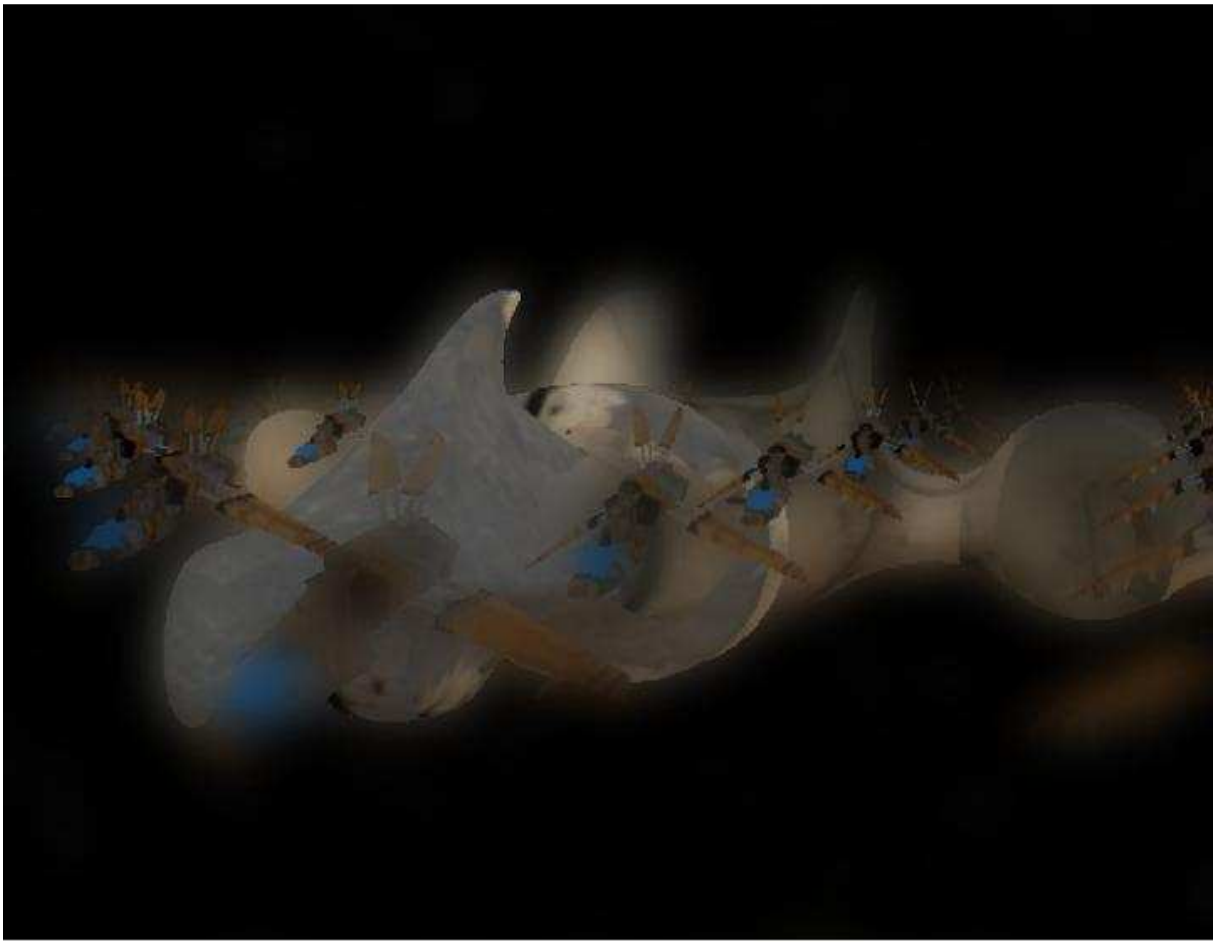


Figure 179 : traiter toute l'image

8.2.2.1. Avantages

Un des avantages est la certitude de supprimer toutes les hautes fréquences d'un seul coup. En effet, nous voyons sur la Figure 179 qu'il ne reste plus de haute fréquence ni au loin ni sur l'avant plan, peu importe le nombre d'objets ou de matériaux à traiter.

Le gros avantage pour les utilisateurs est que ce traitement peut être ajouté très simplement à une application existante. Le programmateur n'est pas obligé de prendre le programme en compte dès le départ. De plus, ce traitement n'est appliqué qu'une seule fois par monde virtuel et ne traite pas les objets cachés. Cette solution est donc beaucoup plus rapide à installer et en fonctionnement, surtout s'il y a un grand nombre d'objets.

8.2.2.2. Inconvénients

Nous ne pouvons plus faire d'exception pour un objet important ou esthétique, puisque le traitement n'est plus appliqué objet par objet. Cela signifie aussi que si un opérateur doit effectuer une tâche de précision sur une partie du monde virtuel, elle doit absolument se situer sur l'écran pour ne pas être floue.

Les objets qui sont dans le fond ne sont plus du tout discernables. Nous voyons sur la Figure 179 qu'après le sixième petit vaisseau nous ne les distinguons plus alors que les files comptent huit vaisseaux. Cela signifie que si la tâche contient un dénombrement d'objets, elle est totalement impossible.

D'autre part, sur un aspect esthétique, les utilisateurs remarquent beaucoup plus vite la différence, du fait que les objets se fondent les uns dans les autres. Par exemple, la limite entre le gros vaisseau et le ciel est très ténue à certains endroits.

8.3. Conclusion

Nos expériences ont montré que notre traitement diminue la fatigue visuelle, mais détériore quelque peu l'aspect esthétique du monde virtuel. Nous avons également discuté de la méthode à appliquer pour supprimer les hautes fréquences, matériau par matériau ou toute l'image d'un coup. Nous avons abordé les avantages et inconvénients des deux méthodes.

Conclusion et perspectives

Conclusion

Les travaux présentés dans ce mémoire traitent de la vision stéréoscopique, plus précisément de la perception des formes en immersion et de la diminution de la fatigue visuelle en vision stéréoscopique. Cette thèse est donc séparée en deux parties.

Dans le premier chapitre, nous avons évoqué les généralités de la vision humaine. Le système visuel a été décrit ainsi que son fonctionnement. Nous avons expliqué les différents moyens de perception du relief : indices monoculaires statiques et cinétiques, indices proprioceptifs et indices binoculaires. Nous avons également comparé la vision binoculaire naturelle à sa reproduction artificielle. Les avantages et inconvénients de la vision stéréoscopique en immersion ont été abordés et nous avons proposé des solutions à certaines imperfections.

Le chapitre 2 a été consacré à l'état de l'art de la perception des formes. Nous avons expliqué le fonctionnement de la perception des plans inclinés et obliques ainsi que des courbures. Elle est basée sur trois indices principaux : les dérivées première et seconde des disparités horizontales, les dérivées première et seconde des parallaxes de mouvement et pour finir les indices monoculaires (gradient de lumière ou de texture). Nous avons également abordé brièvement la reconnaissance d'objets.

Le chapitre 3 est le chapitre central de nos expérimentations sur la perception des formes en immersion. Nous y explicitons d'abord tout ce qui nous sert dans la suite du chapitre : une introduction à la psychophysique ainsi qu'à la comparaison de données statistiques. Nous présentons ensuite nos six hypothèses sur la perception des formes :

- Il est préférable d'avoir le point de vue traqué en vision monoscopique.
- Il est préférable d'avoir le point de vue traqué en vision ortho-stéréoscopique.
- Il est préférable d'avoir le point de vue traqué en vision non-ortho-stéréoscopique.
- Il est préférable d'être en vision ortho-stéréoscopique qu'en vision non-ortho-stéréoscopique lorsque le point de vue est traqué.
- Il est préférable d'être en vision ortho-stéréoscopique qu'en vision non-ortho-stéréoscopique lorsque le point de vue n'est pas traqué.
- Il est préférable d'être en vision monoscopique avec point de vue traqué qu'en vision ortho-stéréoscopique sans point de vue traqué.

Nous avons vu que le point de vue traqué est très important pour la perception des formes en profondeur, quelque soit la projection utilisée (vision monoscopique, ortho-stéréoscopique ou non). Nous avons également découvert que la vision non-ortho-stéréoscopique pouvait augmenter la détection de modifications des formes en profondeur et fronto-parallèles. Nous avons établi qu'il vaut mieux immerger des utilisateurs en vision monoscopique avec point de vue traqué qu'en vision ortho-stéréoscopique sans point de vue traqué pour la perception des formes en profondeur.

Dans la seconde partie, nous avons réalisé un état de l'art de la fatigue visuelle en vision stéréoscopique. Nous y exposons les problèmes liés à la stéréopsie, les problèmes liés à la technique et bien sûr, les problèmes liés aux hautes fréquences aux endroits de grandes disparités. Nous exposons ensuite les trois algorithmes que nous avons étudiés pour limiter les hautes fréquences aux endroits de grandes disparités. Nous commençons par expliquer la transformée en ondelettes, le Box filter et l'algorithme de flou par moyenne et moyenne glissante. Nous effectuons ensuite une comparaison

mathématique et technique entre tous ces algorithmes. L'algorithme de flou par moyenne glissante sur carte graphique a été retenu pour sa rapidité et sa modularité. Nous avons également exposé les techniques pour mesurer cette fatigue. Notre protocole de test est expliqué et les résultats sont ensuite exposés.

Les études précédentes montraient que la suppression des hautes fréquences gênantes pouvait s'avérer utile sur des images fixes, mais il n'était pas certain que sur des images suivant le point de vue de l'utilisateur ce traitement soit efficace. Grâces à des mesures objectives et subjectives, nous avons montré que notre méthode diminue de façon significative la fatigue visuelle en temps réel. Cependant, il détériore quelque peu l'aspect esthétique du monde virtuel dans lequel l'observateur évolue.

Le chapitre 8 est une discussion de nos résultats sur la fatigue visuelle. Nous y résumons nos résultats et nous discutons des avantages et inconvénients à traiter matériaux par matériaux ou bien globalement l'image en entier.

Implications industrielles

Dans ce paragraphe, nous présentons quelques résultats pratiques tirés de nos travaux au niveau de l'interfaçage visuel sur des exemples d'applications professionnelles. Cela permet d'indiquer à tout professionnel travaillant avec les outils de réalité virtuelle comment exploiter nos résultats. Nous formulons l'hypothèse suivante :

Nos expérimentations psychophysiques permettent de déterminer quand l'observateur perçoit la différence de forme entre un objet réel et sa modélisation virtuelle déformée dans une des 3 directions de l'espace. Le résultat est donné par le pourcentage de déformation permettant de voir cette différence. Nous faisons l'hypothèse que ce pourcentage correspond aussi à la perception de différences de forme entre deux pièces virtuelles. Les résultats sont une moyenne pour un certain nombre de testeurs, mais certains sont plus performants que d'autres. Des observateurs professionnels entraînés (les stylistes, les ingénieurs de conception, les ergonomes, etc.) peuvent avoir de meilleurs résultats. Mais dans tous les cas, cela donne un ordre de grandeur de la qualité perçue et permet aussi de comparer quantitativement les différentes manières de visualiser en immersion virtuelle (mono, stéréo, tête traquée ou non, etc.). Une autre hypothèse concerne les indices de perception visuelle en profondeur : nous avons vu que les indices monoscopiques et stéréoscopiques se complètent. Il ne faut donc comparer que des cas similaires. Dans nos expérimentations, les deux objets observés sont sans texture. Il ne faut donc exploiter nos résultats que pour des objets sans (ou très peu de texture). Pour « exploiter » les résultats expérimentaux, nous allons donner quelques résultats applicatifs sur des cas industriels provenant d'un constructeur automobile avec qui nous collaborons sur ce sujet.

Perception des formes

Exemple 1 : Vision ortho-stéréoscopique avec ou sans point de vue traqué

Lorsqu'un ingénieur de conception ou un styliste visualise la courbure d'un rétroviseur (en supposant sa profondeur de 20cm environ - Figure 180) en immersion stéréoscopique, il souhaite connaître à peu près l'erreur qu'il fait sur l'estimation de la perception de sa forme.

Figure 180 : rétroviseur latéral d'une voiture

Dans le cas où la porte virtuelle est vue au niveau de l'écran et est vue de face, la perception de la forme dans le sens de la profondeur est limitée.

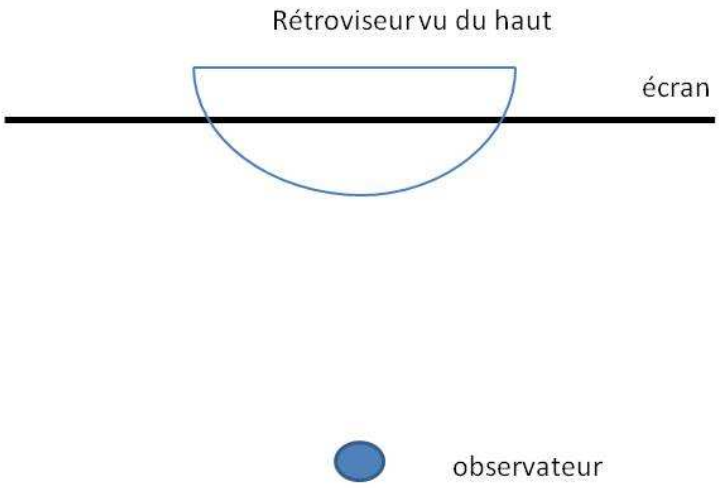


Figure 181 : courbure du rétroviseur vue de face

Nous avons vu que pour une forme virtuelle quelconque située sur l'écran, lorsque le point de vue est traqué, nous détectons une forme différente si la variation en profondeur est au moins de 14% en moyenne, donc de 3 cm. Nous pouvons donc estimer qu'avec le point de vue traqué, lorsque la variation en profondeur de la courbure du rétroviseur de voiture fait moins de **3 cm**, l'opérateur ne la percevra pas lorsque celui ci est vue de face (Figure 181). Et plus problématique, lorsque le point de vue n'est pas traqué, il faudra 62%, donc 12 cm, de variation en profondeur sur la courbure du rétroviseur avant que l'opérateur ne s'en aperçoive.

La courbure est vue de face (la portière est vue de profil)

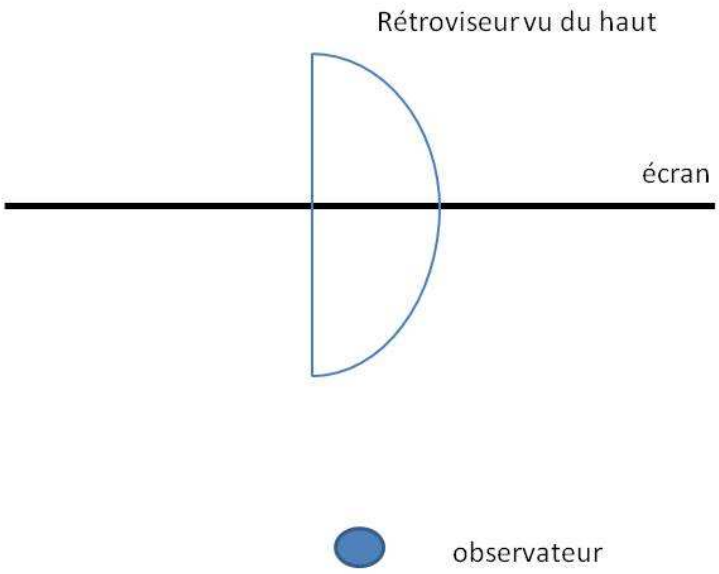


Figure 182 : courbure du rétroviseur vue de profil

Si l'opérateur tourne le rétroviseur pour avoir son profil face à lui (Figure 182), la précision qu'il aura sur la courbure est beaucoup plus importante. La précision ainsi obtenue est de 4%, soit 0.8cm sur la courbure.

Il vaut donc mieux travailler sur un rétroviseur de profil pour contrôler sa courbure. Si ce n'est pas possible, il faut savoir qu'une erreur de 3 cm au minimum est possible.

Exemple 2 : Vision monoscopique ou ortho-stéréoscopique

Dans l'exemple précédant, nous évoquions le cas d'un opérateur travaillant seul sur une courbure de rétroviseur. Dans cet exemple-ci, nous gardons le rétroviseur, mais l'opérateur est accompagné, les deux observateurs portent des lunettes stéréoscopiques traquées mais à chaque instant, un seul observateur a le bon point de vue car l'affichage ne se fait que sur une seule paire d'images stéréoscopiques. Lorsqu'il doit modifier la courbure du rétroviseur, il donne le bon point de vue à son collaborateur pour qu'il puisse observer correctement la modification. Ce dernier valide ou non la modification et redonne le bon point de vue au premier opérateur. Si la modification n'est pas tout à fait validée, il doit remodifier la courbure, redonner le bon point de vue à son collège et ainsi de suite jusqu'à ce qu'ils soient d'accord sur la nouvelle courbure. Nous le voyons, cette procédure est assez laborieuse, mais nécessaire.

Une solution serait d'ajouter un autre projecteur stéréoscopique et éventuellement de renforcer le système informatique pour qu'il puisse fournir les 4 images en temps réel. Mais c'est techniquement assez complexe actuellement et surtout d'un coût très important, surtout pour un « cave ». Une autre solution sur les installations existantes est de ne projeter qu'en monoscopique, mais de traquer le point de vue des deux utilisateurs. Cette solution est bien sûr moins onéreuse que la double stéréoscopie et envisageable car la perte de précision sur la perception de la courbure en profondeur sans stéréoscopie mais avec tête traquée est relativement faible. Nous l'avons quantifiée :

Nous avons vu au paragraphe précédent que la précision de la courbure en profondeur sur un objet de 20 cm sur l'écran en vision ortho-stéréoscopique avec le point de vue traqué est 14% soit 3 cm. En vision monoscopique avec le point de vue traqué, une modification de 25% peut être perçue sur le même objet, soit 5 cm. Nous voyons donc qu'il existe une perte de précision en vision monoscopique par rapport à la précision en vision ortho-stéréoscopique. Mais elle n'est pas très élevée,

c'est à l'industriel de savoir si cette perte d'information est envisageable par rapport au gain d'avoir deux utilisateurs ayant chacun le bon point de vue simultanément.

Fatigue visuelle en vision stéréoscopique

Nous avons vu que l'acuité stéréoscopique diminue après 30 minutes d'immersion lorsque notre traitement contre la fatigue visuelle n'est pas appliqué. Nous perdons alors en moyenne 10 secondes d'arc de disparité horizontale. Si notre observateur se situe à une certaine distance de l'écran, nous pouvons écrire les équations suivantes à partir des relations trigonométriques dans les triangles rectangles (voir Figure 183) :

$$tan\frac{10''arc}{2} = \frac{parallaxe/2}{distance\ observateur}$$

où

$$parallaxe = 2 * distance\ observateur * \tan(5''arc)$$

or, avec la Figure 167, nous avions écrit (Voir équation 7-3)

$$\frac{Distance\ objet}{parallaxe/2} = \frac{distance\ observateur - distance\ objet}{DIO/2}$$

donc

$$Distance\ objet = Distance\ utilisateur\ \frac{parallaxe}{DIO - parallaxe}$$

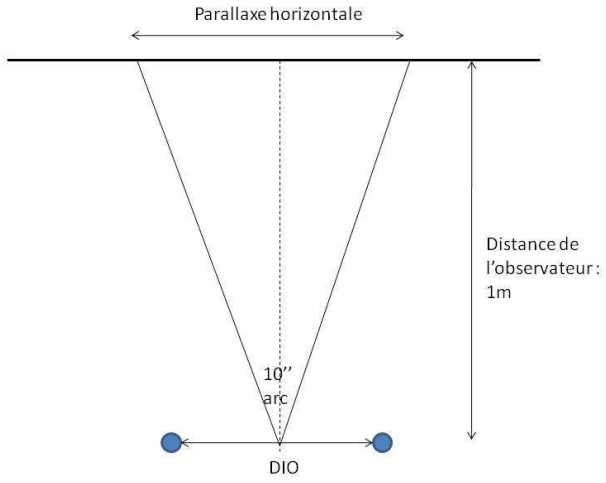


Figure 183 : perte de perception de profondeur

Pour une DIO normale (6,5 cm) et une distance d'un mètre entre l'utilisateur et l'écran, nous obtenons une «perte profondeur» en moyenne de près d'0,5 cm en moins d'une demi-heure. Certains observateurs auront une perte de précision plus importante : pour un testeur de nos expérimentations la perte de précision est de 2.4 cm.

Indépendamment de ce résultat objectif, il ne faut pas omettre que tout observateur d'images stéréoscopiques fatigue à les observer, certains étant nettement plus sensibles que d'autres. Il serait judicieux de faire une « étude de terrain » auprès de professionnels exploitant souvent les images en relief pour connaître leur fatigue, déterminée subjectivement par un question-

naire et mesurée objectivement par la perte d'acuité stéréoscopique après une séance de travail relativement longue (30 minutes ou une heure).

Perspectives

Notre étude a soulevé certains points qu'il serait intéressant de développer.

La première expérience que j'aimerais mener est de quantifier l'influence de notre traitement d'image sur la perception des formes. En effet, nous avons proposé un traitement qui diminue la fatigue visuelle. Cependant, ce traitement enlève les détails de certaines textures. Ceci pourrait provoquer une diminution de la perception des formes en relief. Nous aimerions donc refaire nos expériences sur la perception des formes en relief avec un fond et une forme contenant beaucoup de hautes fréquences. Ceci nous permettrait d'étudier l'influence de notre traitement sur la perception des formes.

Il serait également judicieux de faire la comparaison des modifications de fatigue visuelle en traitant toute l'image ou matériau par matériau. Il serait aussi intéressant de mesurer les effets de ces deux méthodes sur l'efficacité et sur la perception des formes, des tailles et de la profondeur. En effet, dans le cas du traitement sur toute l'image, nous masquons les disparités des limites des objets. Cela peut influer sur la perception de la profondeur, des tailles et des formes.

D'autre part, nous aimerions faire des tests de perception de la taille pour pouvoir quantifier la répercussion de la vision non-ortho-stéréoscopique sur la perception de la taille des éléments. Nous obtiendrions ainsi l'influence de la vision non-ortho-stéréoscopique sur la perception des formes et de la taille.

Nous aimerions également étudier plus profondément l'influence de la pixellisation sur la perception des formes en profondeur. En effet, nous avons montré que lorsque les objets se trouvent devant l'écran, la pixellisation très fine permet d'avoir des disparités horizontales très précises qui favorisent la perception des formes en profondeur. Nous aimerions donc savoir si, lorsque l'objet se situe derrière l'écran avec une pixellisation comparable à celle que nous avons avec notre système lorsque l'objet est devant l'écran, la perception des formes est améliorée dans les mêmes proportions.

Bibliographie

Albert, C., & Gosselin, L. (2002). *Fatigue visuelle, document de référence.* Paris: Direction de la santé publique.

Allison, R. S., Rogers, B. J., & Bradshaw, M. F. (2003). Geometric and induced effects in binocular stereopsis and motion parallax. *Vision Research , 43* (1879-1893).

Amzallag, F., Picciolli, N., & Bry, F. (1978). *Introduction à la statistique.* Paris: Herman.

Arden, G. B. (1988). Le standard de mesure de l'acuité visuelle. *Journal français d'ophthalmologie , 11* (11), 779-792.

ART-tracking. (s.d.). *http://www.ar-tracking.de/*. Consulté le juin 08, 2009

Bajery, R., & leaberman, L. (1976). Texture gradient as depth cue. *SGIP , 5*, 52-67.

Bay, H., Tuytelaars, T., & Van Gool, L. (2006). SURF : Speeded Up Robust Features. *Computer Vision and Image Understanding (CVIU) , 110* (3), 346-359.

Benjamin, W. J. (2007). *Borish's Clinical refraction 2nd edition.* Birmingham: Elsevier.

Benoit, C., & Dusson, A. (1999)*Tipe sur la compression de données informatiques.* MPSI.

Bergholm, F., & Francisco, A. (1996). On the horopter and Hering-Hillebrand deviation. *Cybernetic Vision, 1996. Proceedings., Second Workshop on, IEEE* .

Bonnet, C. (1986). *Manuel pratique de psychophysique.* Paris: Armand Collin.

Boyer, J. (2000). *Méthodes statistiques : médecine-biologie.* Paris: Estem.

Bradshaw, M. F., & Rogers, B. J. (1999). Sensitivity to horizontal and vertical corrugations defined by binocular disparity. *Vision research , 39*, 3049-3056.

Butts, D. R., & McAllister, D. F. (1988). Implementation of true 3-D cursors in computer graphics. *Three-Dimensional Imaging and remote Sensing Imaging , 902*, 74-83.

Cagenello, R., & Rogers, B. J. (1993). Anisotropy in perception of stereoscopic surfaces : the role of orientation disparity. *Vision research , 33*, 2189-2201.

Cahen, O. (1990). *L'image en relief : de la photographie stéréoscopique à la vidéo 3D.* Paris: Masson.

Campbell, F., & Robson, J. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology* , 551-566.

Costello, P. J., & Howarth, P. A. (1996). *The visual effects of immersion in four virtual environments.* Sophia-Antipolis: Rapport de recherche VISERG.

Dagnelie, P. (1998). *Statistique théorique et appliquée : inférence statistique à une et à deux dimensions.* Bruxelles: De Boek Université.

DeValois, R., & DeValois, K. (1988). *Spatial Vision.* New-York: Oxford University.

Devisme, C. (2008). *Etude de l'influence des disparités horizontales et verticales sur la perception de la profondeur en champ visuel périphérique.* Paris: Thèse de l'Université de Paris VI.

Droesbeke, J.-J. (1992). *Elements de statistique.* Bruxelles: Edition de l'université libre de Bruxelles.

Durgin, F. H., Proffitt, D. R., Olson, T. J., & Reike, K. S. (1995). Comparing depth from binocular disparity to depth from motion. *Journal of experimental Psychology : Human Perception and performance , 21* (3), 679-699 .

Eggleston, R. G., Janson, W. P., & Aldrich, K. A. (1996). Virtual reality system effects on size-distance judgments in a virtual environment. *Proceedings of the IEEE virtual reality* .

Espinasse-Berrod, M.-A., & Péchereau, A. (2008). *Strabologie: Approches diagnostique et thérapeutique.* Paris: Elsevier Masson.

Fechner, G. (1860). *Elements de psychophysique.* Leipzig: Breitkopf and härtel.

Fender, D., & Julesz, B. (1967). Extension of Panum's area. *the optical Society of america , 57*, 819-830.

Fuchs, P. (2006). *La restitution visuelle stéréoscopique - traité de la réalité virtutelle.* Paris: Presse de l'Ecole des mines de Paris.

Fuchs, P., & Papin, J.-P. (2003). *Le sens et les réponses motrices de l'homme - traité de la réalité virtuelle.* Paris: Presse de l'Ecole des mines de Paris.

Fuchs, P., Ernadotte, D., Maman, D., Laurgeau, C., & Bordas, J. (1995). Téléprésence virtuelle stéréoscopique. *Interface des Mondes réels et virtuels* , 77–91.

Gabor, D. (1946). Theory of communication. *Journal of IEEE , 93* (3), 429-457.

Gibson, J. J. (1950). *The perception of the visual word .* Boston: Houghton Mifflin.

Gillam, B., & Lawergren, B. (1983). The induced effects vertical disparity and steroscopic theory. *Perception and psychophysics , 34*, 121-130.

Gillam, B., Chambers, D., & Lawergren, B. (1988). Postfusional latency in stereoscopic slant perception and the primitives of stereopsis. *Journal of experimental psychology : human perception and performance , 14*, 163-176.

Glennerster, A., Rogers, J., & Bradshaw, M. F. (1994). The effect of (i) differents cues and (ii) the observer's task in stereoscopic depth constancy ,. *145*, 159-161.

Goldstein, E. B. (2003). *Sensation & perception.* Wadsworth Publishing Company Bk&Cdr.

Gower, D. W., & Fowlkes, J. E. (1989). *Simulator Sickness in the UH-60 (Black Hawk) Flight Simulator.* U.S. Army Aeromedical Research Laboratory.

Grenez, F. (2003). *Théorie des signaux.* Bruxelles: Presse universitaire de Bruxelles.

Guillaume, P. (1979). *La psychologie de la forme.* Paris: Flammarion.

Heckmann, T., & Schor, C. M. (1989). Panum's fusional area estimated with a criterion-free technique. *Perception and psychophysics , 45*, 297-306.

Hillaire, S., Lecuyer, A., & Cozot, R. (2008). Using an Eye-Tracking System to improve Depth-of-field Blur Effect and Camera Motion in Virtual Environments . *Proceedings of IEEE Virtual Reality* , 47-51.

Hillaire, S., Lecuyer, A., Cozot, R., & Casier, G. (2008). Depth of field blur effect for Fisrt Person Navigation in virtual environments. *IEEE Computer Graphics and Application , 28* (6), 47-55.

Hillaire, S., Lecuyer, A., Cozot, R., & Casier, G. (2007). Depth of field blur effects for first person navigation in virtual Environments. *Proceedings of ACM Symposium on virtual reality Software and Technology* , 204-207.

Holway, A. H., & Boring, E. G. (1941). Determinants of apparent visual size with distance variant. *American Journal of Psychology , 54* (1), 21-37.

Howard, I. P., & Rogers, B. J. (1995). *Binocular Vision and Stereopsis.* Oxford: Oxford psychology series.

Howarth, P. A. (1999). Oculomotor changes within virtual environments. *Applied Ergonomics , 30*, 59 -67.

http://www.photofriday.com/calibrate.php. (s.d.). Consulté le juin 24, 2009

Hubbard, B. B. (1995). *Ondes et ondelettes la saga d'un outil mathématique.* Paris: Belin pour la science.

Hugonnier, S., & Hugonnier, R. (1965). *Strabismes, hétérophories, paralysie oculo-motrices : les déséquilibres oculo-moteurs en clinique.* Maillet: Masson et Cie Saint-Ouen.

Johnston, E. B. (1991). Systematic distortions of shape from stereopsis . *Vision research , 31,* 1351-1360.

Jones, G., Lee, D., Holliman, N., & Erza, D. (2001). Controlling perceived depth in stereoscopic image. *Proceedings of IEEE , 4297*, 42-53.

Julesz, J. (1971). *Fondation of cyclopean perception.* Chicago: University of Chicago Press.

Kellog, R. S., Castore, C., & Coward, R. (1980). Psychophysiological Effects of Training in a Full Vision Simulator. *Annual Scientific Meeting of the Aerospace Medical Association.*

Koenderink, J. J., van Doorn, A. J., & Koppers, A. L. (1996). Pictorial surface attitude and local depth comparisons. *Perception & psychophysics , 58* (2), 162-173.

Koenderink, J. J., van Doorn, A. J., & Koppers, A. L. (1992). Surface perception in picture. *Perception & psychophysics , 52* (5), 482-496.

Kooi, F. L., & Toet, A. (2001). Visual comfort of binocular and 3D displays. *Proceedings of SPIE , 4299*, 586-592.

Lambooij, M., Fortuin, M., Ijseelsetin, W. A., & Heynderickx, I. (2009). Measuring visual discomfort associated with 3D dispalys. *Proceedings of SPIE , 7237*.

LaViola, J. J. (2000). A Discussion of Cybersickness in Virtual. *SIGCHI Bulletin , 32* (1), 47-56.

Lemaire, P. (2006). *Abrégé de psychophysique cognitive.* Paris: De Boek, ouverture psychologique.

Levelt, W. M. (1965). *On binocular rivalry.* Soesterberg: Institute for perception.

Levelt, W. M. (1966). The alternation process in binocular rivalry. *British Journal Psychology , 58*, 143-145.

Luo, X., Kenyon, R., Kamper, D., Sandin, D., & DeFanti, T. (2007). The effect of scene complexity, stereovision, and motion parallax on size constancy in virtual environment. *IEEE Virtual Reality*, 59-66.

MacKenzie, I. S., & Ware, C. (1993). Lag as a determinant of human performance in interactive systems. *Proceedings of the ACM Conference on human Factors in Computing system* , 488-493.

Mallat, S. (2000). *Une exporation des signaux en ondelettes.* Paris: Edition de l'Ecole Polytechnique.

Manan, F. A., Jenkins, T. C., & Collinge, A. J. (2002). The effect of clinical visual stress on stereoacuity measured with the TNO test. *Malaysian Journal of medical sciences , 8*, 25-31.

Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information.* New-York: Freeman.

Massot, C. (2006). *Texture et Perception 3D dans les Scènes Naturelles : Modèles d'Inspiration.* Doctorat de l'université Joseph Fourier de Grenoble.

McKee, S. P. (1993). The spatial requirements for fine stereoacuity. *Vision Research , 23* (191-198).

Meyer, I. (1992). *Les ondelettes algorithmes et applications.* Paris: Armand Collin.

Mitchell, D. E. (1966). Retinal disparity and diplopia . *Vision research , 6*, 441-451.

Money, K. E. (1970). Motion Sickness. *Psychological Reviews , 50* (1), 1-39.

Mon-Williams, M., Warm, J. P., & Rushton, S. (1993). Binocular vision in a virtual world : visual deficits following the wearing of a head mounted display. *Ophthalmic and physiological optics , 13*, 387-391.

Nakayama, K. (1983). Kinematics of Normal and Strabismic Eyes. Dans C. Schor, & K. Ciuffreda, *Vergence Eye Movements: Basics and Clinical Aspects.* London-Boston: Butterworth.

Neveu, P. (2008). Retentissement du sur-écartement inter-pupillaire sur l'équilibre oculomoteur. *rapport de Master 2ème année* , Paris.

Nielsen, K. K., & Poggio, T. (1984). Vertical image registration in stereopsis 5. *Vision research , 24*, 1133-1140.

Ogle, K. N. (1950). *Research in Binocular vision.* Philadelphia and London: Press of W.B. Saunders Compagny.

Olivia, A., Torralba, A., & Schyns, P. (2006). Hybrib Image. *ACM Transaction on Graphics, ACM Sigraph , 25* (3), 527-530.

Paljic, A. (2004). *Interaction en environnements immersifs et retours d'effort passif.* Paris: Université Paris 6.

Paljic, A., Coquillart, S., & Burkhardt, J.-M. (2002). A study of distance of manipulation on the responsive workbench. *Immersive Projection technology Symposium* .

Palmer, D. A. (1961). Measurement of the horizontal extent of Panum's area by a method of constant stimuli. *Opical Acta , 8*, 151-159.

Palmer, S. E. (1999). *Vision science : photons to phenomenology.* Oxford: Bradford book, The MIT Press.

Perrin, J. (1998). *Profondeur et binocularité : algorithmie, étude psychophysique et intérêt pour l'ergonomie des interfaces stéréoscopiques.* Paris: Ecole des Mines de Paris.

precision-vision. (s.d.). *http://precision-vision.com*.

Prince, S. D., & Rogers, B. J. (1998). Sensitivity to disparity corrugations in peripheral vision. *Vision research , 38*, 2533-2537.

Ramousse, R., Le Berre, M., & Le Guelte, L. (1996). *Introduction aux statistiques.* Paris: Conservation et développement durable.

Ranchin, T. (2001). *Analyse Multirésolution et transformée en ondelettes.* Paris: Ecole des Mines de Paris.

Reason, J. T., & Brand, J. J. *Motion Sickness* (1975). London: Academic press.

Richards, W. (1971). Independence of Panum's near and far limits. *American Journal of optometry and archives of American academy of optometry , 48*, 103-109.

Rogers, B. J., & Graham, M. E. (1983). Anisotropy in the perception on three-dimensional surfaces. *Science , 221*, 1409-1411.

Rogers, B., & Cagenello, R. (1989). Disparity curvature and the perception of three-dimensional surfaces. *Nature , 339*, 261-270.

Roumes, C. (1995). *Contenu fréquentiel de l'image et vision binoculaire - Etude psychophysique chez l'homme.* Paris: Université de Paris.

Roumes, C., Plantier, J., & Menu, J.-P. (1992). Local contrast : a quantification tool for stereo-image. *Proceedings of IEEE , 5*, 2098-2099.

Rushton, S. K., & Mc Riddell, P. (1999). Developing visual system and exposure to virtual reality and stereo display : some concerns and speculation about the demands on accomodation and vergence. *Applied Ergonomics , 30*, 69-78.

Sakai, K., & Finkel, H. (1995). Characterisation of spatial frequency in the perception of shape form texture. *Journal ophtalmology society American , A* (12), 1208-1224.

Scarfe, P., & Hibbard, P. (2006). Disparity-defined objects moving in depth do not elicit three-dimensional shape constancy. *Vision research , 46* (10), 1599-1610.

Schor, C. M. (1986). Adaptive regulation of accommodative vergence and vergence accommodation. *American journal of Optometry & Physioloicaly Optics , 65*, 168-173.

Schor, C. M., & Tyler, C. W. (1981). Spatio temporal properties of Panum's fusional area. *Vision research , 21*, 683-692.

Schor, C. M., Heckmann, T., & Tyler, C. W. (1989). Binocular fusion limits are independant of contrast, luminance gradient and component phases. *Vision research , 29*, 821-835.

Schor, C. M., Roberston, K. M., & Wesson, M. (1986). Binocular sensory vision is limited by spatial resolution. *Vision research , 24*, 573-578.

Stelmach, L., Tam, W. J., & Meegan, D. (1999). Perceptual basis of stereoscopic video. *Proceedings of SPIE , 3639*, 260-265.

Stelmach, L., Tam, W. J., Meegan, D., Vincent, A., & Corriveau, P. (2000). Human perception of mismatched stereoscopic 3D inputs. *International Conference on Image Processing IEEE , 1*, 5-8.

Stevens, K. S. (1979). *Surface perception from local analysis of texture and contour.* Cambridge: Ph.D. Dissertation MIT.

Tam, W. J., Stelmach, L. B., & Corriveau, P. (1998). Psychovisual aspects of viewing stereoscopic video sequences. *Proceeding od SPIE , 3295*, 226-235.

Tiberghien, G. (1984). *Initiation à la psychophysique.* Paris: Presses Univeritaires de France.

upvector. (s.d.). *http://www.upvector.com*. Consulté le juin 20, 2009

Valyus, N. (1962). *Stereoscopy.* Londres: Focal Press.

Viola, P., & Jones, M. (2001). Rapid Object Detection using boosted Cascade of simple Features. *IEEE Computer society conference on computer vision and pattern recognition , 1*, 511.

Wann, J. P., Rushton, S., & Williams, M. (1995). Natural problems for stereoscopic depth perception in virtual environments. *Vision research , 35* (6), 2731-2736.

Watson, A., Barlow, H., & Robson, J. (1983). What do the eye see the best? *Nature , 302*, 419-422.

Wilson, H., & Bergen, J. (1979). A four mechanism model for spatial vision. *Vision research* , 19-32.

Wilson, H., McFarlane, D., & Phillips, G. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision research , 19* (9), 873-882.

Woo, G. C. (1974). The effect of exposure time on the foveal size of Panum's area. *Vision research , 14*, 473-480.

Wopking, M. (1995). Viewing comfort with stereoscopic pixure : an experiment study on the subjective effects of disparity magnitude and depth of focus. *Journal of SID , 3*, 101-103.

Yano, S., Ide, S., Mituhashi, T., & Thwaites, H. (2002). A study of visual fatigue and visual comfort for 3d HDTV/HDTV images. St*ereoscopic displays , 213*, 191-201.

Young, S. D., Adelstein, B. D., & Ellis, S. R. (2006). demand characterisic of a questionnaire used to assess motion sickness in a virtual environment. *Proceedings of the IEEE virtual reality conference* , 97- 102.

zeiss. (s.d.). *www.zeiss.com*.

Table des illustrations

Figure 1 : projections rétiniennes vers le corps genouillé latéral, puis le cortex strié (Fuchs, 2006) 12
Figure 2 : les principales composantes de l'œil (Fuchs, 2006) 13
Figure 3 : l'accommodation (Fuchs, 2006) 14
Figure 4 : la convergence 14
Figure 5 : la structure anatomique de la rétine (Fuchs, 2006) 15
Figure 6 : définition d'une fréquence spatiale (Fuchs, 2006) 16
Figure 7 : perspective industrielle ou cavalière 18
Figure 8 : perspective artistique ou photographique 18
Figure 9 : le relief est perçu grâce à la variation de luminosité 18
Figure 10 : intuitivement, les fleurs sont perçues comme étant de plus en plus éloignées 19
Figure 11 : l'ellipse est perçue devant le rectangle 19
Figure 12 : les détails de la texture sont moins visibles au loin 20
Figure 13 : au loin, le paysage devient flou 20
Figure 14 : aire de Panum (Fuchs, 2006) 22
Figure 15 : horoptère théorique dit géométrique (Espinasse-Berrod, et al., 2008) 23
Figure 16 : horoptère empirique, dit longitudinal Hering-Hillebrand (Espinasse-Berrod, et al., 2008)(Bergholm, et al., 1996) 23
Figure 17 : stéréogramme de points aléatoires de Julesz (Julesz, 1971) 24
Figure 18 : les yeux observent le point F, il y a des disparités verticales au point P 24
Figure 19 : représentation vectorielle des disparités d'un plan fronto-parallèle (Nakayama, 1983) 25
Figure 20 : système pour suivre la position et l'orientation du sujet (ART-tracking) 26
Figure 21 : définition de la parallaxe horizontale angulaire (Fuchs, 2006) 26
Figure 22 : vision avec parallaxe nulle 27
Figure 23 : vision avec parallaxe positive 27
Figure 24 : vision avec parallaxe négative 28
Figure 25 : différence entre vision naturelle et restitution stéréoscopique (Fuchs, 2006) 28
Figure 26 : mouvements pseudoscopiques parallèles à l'écran 30
Figure 27 : mouvements pseudoscopiques perpendiculaires à l'écran 30
Figure 28 : positionnement des deux caméras fictives ou réelles dont les axes convergent vers un point (Fuchs, 2006) 31
Figure 29 : les projections droite et gauche de la face avant d'un cube centré 32
Figure 30 : un plan oblique produit un gradient horizontal de disparités horizontales 35
Figure 31 : un plan incliné produit un gradient vertical de disparités horizontales 36
Figure 32 : effet induit et géométrique pour la perception d'une surface oblique (Devisme, 2008) 37
Figure 33 : le gradient de la texture nous incite à percevoir un plan incliné 38
Figure 34: le gradient de la texture nous incite à percevoir un plan oblique 39
Figure 35 : la fréquence de variation des disparités optimale dépend de l'excentricité (vue du dessus) 40
Figure 36 : le cylindre est vu déformé selon la distance à laquelle il se trouve (vue du dessus) 41
Figure 37 : variation de la perception de forme en fonction de la distance (Johnston, 1991) 42
Figure 38 : Certaines surfaces complexes peuvent être reproduite de façon réaliste à l'aide de gradients de texture 43

Figure 39 : Texture d'une surface courbe et traduction de la variation de celle-ci dans le domaine fréquentiel (Massot, 2006) 43
Figure 40 : Textures crées avec soit une variation des pics de fréquences (à gauche) soit une variation de la fréquence moyenne (à droite) (Sakai, et al., 1995) 44
Figure 41 : l'ombre sur une sphère (upvector) 45
Figure 42 : orientation des surfaces de la peau en fonction des ombres (Koenderink, et al., 1996) 45
Figure 43 : différentes vues d'un cube 46
Figure 44 : nous voyons toujours la même chaise, bien que l'angle du point de vue change (Palmer, 1999) 47
Figure 45 : nous distinguons les modifications de matière et de luminosité (Palmer, 1999) 47
Figure 46 : les sujets de Luo devaient redimensionner une bouteille de Coca-Cola®, en fonction du fond d'écran et de la table sur laquelle elle est posée (Luo, et al., 2007) 48
47 : déformation d'un rectangle due à une DIO du sujet plus grande (vue du dessus) 49
Figure 48 : schéma du dispositif réel 60
Figure 49 : sphère réelle 60
Figure 50 : sphère virtuelle 61
Figure 51 : forme inconnue réelle vue de près 61
Figure 52 : dispositif réel complet vue de loin 62
Figure 53 : forme inconnue virtuelle 62
Figure 54 : disposition des objets virtuels et réel ainsi que de l'observateur lors du déplacement de l'objet virtuel 63
Figure 55 : boîtier de réponse ©Nintendo 64
Figure 56 : capture d'écran lors du test : demande de confirmation d'une réponse 65
Figure 57 : sujet en train de réaliser le test 66
Figure 58 : exemple de collecte de données en vision ortho-stéréoscopique, point de vue traqué avec la sphère sur l'écran 67
Figure 59 : transformation en écarts réduits des probabilités de réponse "plus grand" dans le cas de la vision ortho-stéréoscopique, point de vue traqué avec la sphère sur l'écran 68
Figure 60 : probabilité de réponse "plus grand" pour la vision ortho-stéréoscopique, point de vue traqué avec la sphère sur l'écran et son lissage 69
Figure 61 : Influence du point de vue traqué sur la perception des formes en profondeur en vision monoscopique pour une sphère sur l'écran 70
Figure 62 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique avec une sphère sur l'écran 71
Figure 63 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique pour une sphère sur l'écran, détail 72
Figure 64 : influence du point de vue traqué sur la perception des formes en profondeur en vision monoscopique pour une sphère derrière l'écran 73
Figure 65 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique pour une sphère derrière l'écran 73
Figure 66 : influence du point de vue traqué sur la perception des formes en profondeur en vision monoscopique pour une sphère devant l'écran 74
Figure 67 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique pour une sphère devant l'écran 75
Figure 68 : influence du point de vue traqué sur la perception des formes en profondeur en vision monoscopique avec une forme non-connue sur l'écran 75
Figure 69 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique avec une forme non-connue sur l'écran 76
Figure 70 : influence du point de vue traqué sur la perception des formes en profondeur en vision monoscopique avec une forme non-connue derrière l'écran 76

Figure 71 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique avec une forme non-connue derrière l'écran 77
Figure 72 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique avec une forme non-connue derrière l'écran, détails 78
Figure 73 : influence du point de vue traqué sur la perception des formes en profondeur en vision monoscopique avec une forme non-connue devant l'écran 78
Figure 74 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision monoscopique avec une forme non-connue devant l'écran 79
Figure 75 : influence du point de vue traqué sur la perception des formes en profondeur en vision ortho-stéréoscopique pour une sphère sur l'écran 80
Figure 76 : influence du point de vue traqué sur la perception des formes en profondeur en vision stéréoscopique pour une sphère derrière l'écran 81
Figure 77 : influence du point de vue traqué sur la perception des formes en profondeur en vision stéréoscopique pour une sphère devant l'écran 82
Figure 78 : influence du point de vue traqué sur la perception des formes en profondeur en vision ortho-stéréoscopique avec une forme non-connue sur l'écran 83
Figure 79 : influence du point de vue traqué sur la perception des formes en profondeur en vision ortho-stéréoscopique avec une forme non-connue derrière l'écran 83
Figure 80 : influence du point de vue traqué sur la perception des formes en profondeur en vision ortho-stéréoscopqiue avec une forme non-connue devant l'écran 84
Figure 81 : influence du point de vue traqué sur la perception des formes en profondeur en vision non-ortho-stéréoscopique pour une sphère sur l'écran 86
Figure 82 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique pour une sphère sur l'écran 87
Figure 83 : influence du point de vue traqué sur la perception des formes en profondeur en vision non-ortho-stéréoscopique pour une sphère derrière l'écran 87
Figure 84 : influence du point de vue traqué sur la perception des formes en profondeur en vison non-ortho-stéréoscopique pour une sphère devant l'écran 88
Figure 85 : influence du point de vue traqué sur la perception de formes en profondeur en vision non-ortho-stéréoscopique avec une forme non-connue sur l'écran 89
Figure 86 : influence du point de vue traqué sur la perception des formes en profondeur en vision non-ortho-stéréoscopique avec une forme non-connue derrière l'écran 90
Figure 87 : influence du point de vue traqué sur la perception des formes en profondeur en vision non-ortho-stéréoscopique avec une forme non-connue devant l'écran 91
Figure 88 : Comparaison de la perception des formes en profondeur avec une sphère sur l'écran avec point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique 92
Figure 89 : Comparaison de la perception des formes en profondeur avec une sphère derrière l'écran avec point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique 93
Figure 90 : Comparaison de la perception des formes en profondeur avec une sphère devant l'écran avec point de vue traqué en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique 93
Figure 91 : Comparaison de la perception des formes en profondeur avec une forme inconnue sur l'écran avec point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique 94
Figure 92 : Comparaison entre la perception des formes en profondeur avec une forme inconnue derrière l'écran avec point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique 95
Figure 93 : comparaison de la perception des formes en profondeur avec une forme inconnue devant l'écran avec point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique 96
Figure 94 : Comparaison entre la perception des formes en profondeur avec une sphère sur l'écran sans point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique 97

Figure 95 : Comparaison entre la perception des formes en profondeur avec une sphère derrière l'écran sans point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique 98
Figure 96 : Comparaison entre la perception des formes fronto-parallèles avec une sphère derrière l'écran sans point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique 98
Figure 97 : Comparaison entre la perception des formes en profondeur avec une sphère devant l'écran sans point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique 99
Figure 98 : Comparaison entre la perception des formes en profondeur avec une forme non connue l'écran sans point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique 100
Figure 99 : Comparaison entre la perception des formes en profondeur avec une forme non connue derrière l'écran sans point de vue traqué en vision ortho-stéréoscopique et en vision non-ortho-stéréoscopique 101
Figure 100 : Comparaison entre la perception des formes en profondeur avec une forme non connue devant l'écran avec point de vue traqué en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique 102
Figure 101 : comparaison entre la perception des formes en profondeur avec une sphère sur l'écran sans point de vue traqué en vision monoscopique et avec point de vue traqué en vision ortho-stéréoscopique 103
Figure 102 : comparaison entre la perception des formes en profondeur avec une sphère derrière l'écran en vision monoscopique avec point de vue traqué et en vison ortho-stéréoscopique sans point de vue traqué .. 104
Figure 103 : Comparaison entre la perception des formes fronto-parallèles avec une sphère derrière l'écran en vision monoscopique avec point de vue traqué et en ortho-stéréoscopique sans point de vue traqué..... 104
Figure 104 : comparaison entre la perception des formes en profondeur avec une sphère devant l'écran en vision monoscopique avec point de vue traqué et en ortho-stéréoscopique sans point de vue traqué pour ... 105
Figure 105 : comparaison entre la perception des formes en profondeur avec une forme non connue sur l'écran en vision monoscopique avec point de vue traqué et en ortho-stéréoscopique sans point de vue traqué..... 106
Figure 106 : comparaison entre la perception des formes en profondeur avec une forme non connue derrière l'écran en vision monoscopique avec point de vue traqué et en ortho-stéréoscopique sans point de vue traqué 107
Figure 107 : est-il préférable d'être en vision monoscopique avec point de vue traqué ou en ortho-stéréoscopique sans point de vue traqué pour la perception des formes en profondeur avec une forme non connue devant l'écran ? 108
Figure 108 : moyenne de temps passé sur la tâche 109
Figure 109 : différence de taille de projection en fonction de la position de l'objet virtuel..... 114
Figure 110 : différence d'angle de point de vue pour une forme devant ou derrière l'écran 115
Figure 111 : interaction entre l'accommodation et la vergence, adaptation du modèle de Schor (Neveu, 2008) 122
Figure 112 : relation accommodation - vergence. La droite de demande est la convergence qui devrait être mise en jeu pour une quantité d'accommodation donnée et réciproquement (Neveu, 2008) 123
Figure 113 : deux lignes verticales ondulantes en opposition de phase..... 124
Figure 114 : image pour faciliter le réglage du contraste d'un écran ou d'un projeteur (htt092)..... 126
Figure 115 : relation entre les fréquences spatiales et l'aire de Panum (Schor, et al., 1986)..... 128
Figure 116 : fonction de confort (Perrin, 1998) 130
Figure 117 : décomposition paire-impaire..... 133
Figure 118 : fin de la FFT 134
Figure 119 : plan temps-fréquence en temporel et en Fourier 135
Figure 120 : plan temps-fréquence pour Fourier sur fenêtre glissante 136
Figure 121 : plan temps-fréquence pour les ondelettes..... 136
Figure 122 : décomposition en ondelettes continues d'un signal 1D 137
Figure 123 : figure à décomposer 138
Figure 124 : décomposition en ondelettes d'une image..... 139
Figure 125 : décomposition d'un signal 1D par l'algorithme de Mallat 140
Figure 126 : la pyramide de l'algorithme de Mallat 140

Figure 127 : transformée de Mallat pour un signal 1D 140
Figure 128 : algorithme de reconstruction de Mallat à une dimension 141
Figure 129 : décomposition en ondelettes rapide d'une image 142
Figure 130 : recomposition en ondelettes rapide d'une image 143
Figure 131 : comparaison théorique entre le nombre d'opérations pour un produit de FFT et un produit de convolution 144
Figure 132 : temps de décomposition et recomposition du produit de convolution 145
Figure 133 : temps de décomposition et recomposition avec la transformée de Fourier 146
Figure 134 : image de départ pour exemple de calcul BOX FILTER 147
Figure 135 : image intégrale pour exemple de calcul BOX FILTER 147
Figure 136 : représentation des points à sommer pour BOX FILTER 148
Figure 137 : représentation du processus de somme dans BOX FILTER 1 149
Figure 138 : représentation du processus de somme dans BOX FILTER 2 149
Figure 139 : explicitation des moyennes dans l'exemple BOX FILTER 150
Figure 140 : hautes fréquences dans l'exemple de calcul de BOX FILTER 150
Figure 141 : décomposition en hautes et basses fréquences avec BOX FILTER sur une base de calcul carrée de 5 pixels de côté 152
Figure 142 : décomposition en hautes et basses fréquences avec BOX FILTER sur une base de calcul carrée de 8 pixels de côté 152
Figure 143 : exemple de calcul de transformée en ondelettes de Haar 153
Figure 144 : décomposition avec une moyenne sur des carrés de 2 pixels de côté 155
Figure 145 : décomposition avec une moyenne sur des carrés de 4 pixels de côté 156
Figure 146 : matrice de départ pour la différence de calcul des moyennes 157
Figure 147 : calcul des moyennes sur des carrés de pixels 157
Figure 148 : calcul des moyennes glissantes 157
Figure 149 : moyennes glissantes sur des carrés de 2 pixels de côtés 158
Figure 150 : complexité de BOX FILTER et des moyennes en fonction de la taille du carré de calcul pour une image de 1024*1024 159
Figure 151 : complexité de BOX FILTER glissant et des moyennes glissantes en fonction de la taille du carré de calcul pour une image de 1024*1024 160
Figure 152 : moyenne quadratique glissante réalisée avec une base de calcul de 4 pixels de côté 163
Figure 153 : moyenne gaussienne glissante réalisée avec une base de calcul de 5 pixels de côté 165
Figure 154 : cercles de Landlot 168
Figure 155 : optotypes lettrés et chiffrés 169
Figure 156 : flipper Lens Test 170
Figure 157 : le Polatest de Zeiss (zeiss) 170
Figure 158 : projection de ligne avec le Polatest 171
Figure 159 : test de la mouche et test de Wirt (precision-vision) 171
Figure 160 : points de Wirt 172
Figure 161 : test de Lang 172
Figure 162 : test de Randot (precision-vision) 173
Figure 163 : test TNO 173
Figure 164 : le monde virtuel 175
Figure 165 : monde virtuel sans texture vu de face et de coté 175
Figure 166 : sujet en train de réaliser la tâche non floue 177
Figure 167 : calcul des parallaxes horizontales 178
Figure 168 : calcul du nombre de pixel par degré d'angle visuel 179
Figure 169 : la mire est fixée à 40 cm des yeux du sujet grace à un pupitre fixé au siège 180
Figure 170 : modification de la fonction proximum d'accommodation 181

Figure 171 : modification de la facilité accommodative 182
Figure 172 : modification de l'acuité stéréoscopique 182
Figure 173 : efficacité de la tâche en fonction du cylindre pointé 183
Figure 174 : quel cas est plus esthétique? 184
Figure 175 : dans quel cas la tâche est plus fatigante pour les yeux? 184
Figure 176 : dans quel cas la tâche est plus aisée? 185
Figure 177 : image de départ 188
Figure 178 : traiter objet par objet 189
Figure 179 : traiter toute l'image 190
Figure 180 : rétroviseur latéral d'une voiture 195
Figure 181 : courbure du rétroviseur vue de face 195
Figure 182 : courbure du rétroviseur vue de profil 196
Figure 183 : perte de perception de profondeur 197
Figure 184 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision ortho-stéréoscopique pour une sphère sur l'écran 215
Figure 185: influence du point de vue traqué sur la perception des formes fronto-parallèles en vision ortho-stéréoscopique pour une sphère sur l'écran, détails 216
Figure 186 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision stéréoscopique pour une sphère derrière l'écran 216
Figure 187 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision ortho-stéréoscopique pour une sphère devant l'écran 217
Figure 188 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision ortho-stéréoscopique avec une forme non-connue sur l'écran 218
Figure 189 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision ortho-stéréoscopique avec une forme non-connue sur l'écran 218
Figure 190 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision ortho-stéréoscopique avec une forme non-connue derrière l'écran 219
Figure 191 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision ortho-stéréoscopique avec une forme non-connue devant l'écran 220
Figure 192 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique pour une sphère derrière l'écran 221
Figure 193 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique pour une sphère devant l'écran 221
Figure 194 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique avec un forme non-connue sur l'écran 222
Figure 195 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique avec une forme non-connue derrière l'écran 222
Figure 196 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique avec une forme non-connue devant l'écran 223
Figure 197 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes fronto-parallèles avec une sphère sur l'écran? 224
Figure 198 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes fronto-parallèles avec une sphère derrière l'écran? 225
Figure 199 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes en profondeur avec une sphère devant l'écran? 225
Figure 200 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes fronto-parallèles avec une forme inconnue sur l'écran? .. 226

Figure 201 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes fronto-parallèles avec une forme inconnue derrière l'écran? 227
Figure 202 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes fronto-parallèles avec une forme inconnue devant l'écran? 227
Figure 203 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes fronto-parallèles avec une forme inconnue devant l'écran? Détails. 228
Figure 204 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une sphère sur l'écran?..... 229
Figure 205 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une sphère devant l'écran? 230
Figure 206 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une forme non connue sur l'écran? 230
Figure 207 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une forme non connue derrière l'écran?..... 231
Figure 208 : Est- il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes fronto-parallèles avec une forme non connue devant l'écran?..... 231
Figure 209 : est-il préférable d'être en vision monoscopique avec point de vue traqué ou en ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une sphère sur l'écran ?..... 232
Figure 210 : est-il préférable d'être en vision monoscopique avec point de vue traqué ou en ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une sphère sur l'écran ? détails 233
Figure 211 : est-il préférable d'être en vision monoscopique avec point de vue traqué ou en ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une sphère devant l'écran ?..... 233
Figure 212 : est-il préférable d'être en vision monoscopique avec point de vue traqué ou en ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une forme non connue sur l'écran ?..... 234
Figure 213 : est-il préférable d'être en vision monoscopique avec point de vue traqué ou en ortho-stéréoscopique sans point de vue traqué pour la perception des formes en fronto-parallèles avec une forme non connue derrière l'écran ?..... 234
Figure 214 : est-il préférable d'être en vision monoscopique avec point de vue traqué ou en ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une forme non connue devant l'écran ?..... 235
Figure 215 : influence de la projection sur la perception des formes en profondeur sur une sphère sur l'écran 236
Figure 216 : influence de la projection sur la perception des formes en profondeur sur une sphère derrière l'écran..... 237
Figure 217 : influence de la projection sur la perception des formes en profondeur sur une sphère devant l'écran..... 238
Figure 218 : influence de la projection sur la perception des formes fronto-parallèles pour une sphère sur l'écran..... 239

Figure 219 : influence de la projection sur la perception des formes fronto-parallèles pour une sphère derrière l'écran 240
Figure 220 : influence de la projection sur la perception des formes fronto-parallèles pour une sphère devant l'écran 241
Figure 221 : influence de la projection sur la perception des formes en profondeur pour une forme non connue sur l'écran 242
Figure 222 : influence de la projection sur la perception des formes en profondeur pour une forme non connue derrière l'écran 243
Figure 223 : influence de la projection sur la perception des formes en profondeur pour une forme non connue devant l'écran 244
Figure 224 : influence de la projection sur la perception des formes fronto-parallèles pour une forme non connue sur l'écran 245
Figure 225 : influence de la projection sur la perception des formes fronto-parallèls pour une forme non connue derrière l'écran 246
Figure 226 : influence de la projection sur la perception des formes fronto-parallèls pour une forme non connue devant l'écran 247
Figure 227 : seuillage de la transformée de Fourier de la figure de départ 254
Figure 228 : seuillage de la transformée de Fourier de la moyenne glissante sur des carrés de 2 pixels de côté 255
Figure 229 : seuillage de la décomposition de Fourier pour les moyennes glissantes sur 4, 8, 16 et 32 pixels (de droite à gauche et de haut en bas) 255

ANNEXES

Annexe A : graphiques supplémentaires pour la perception des formes en immersion

Hypothèse n°2 : Il est préférable d'avoir la tête t raquée en vision ortho-stéréoscopique

Lorsque la sphère est sur l'écran

Les résultats de la perception des formes fronto-parallèles sont exposés sur la Figure 184. Les probabilités de significativité des écarts de PSE, de JND et des écarts entre les perceptions des courbures verticales et horizontales sont respectivement 50%, 76% et 98%. Il y a donc une différence significative entre la perception des courbures horizontales et verticales.

tracking	oui/non
projection	ortho
forme	Sphère
position	sur

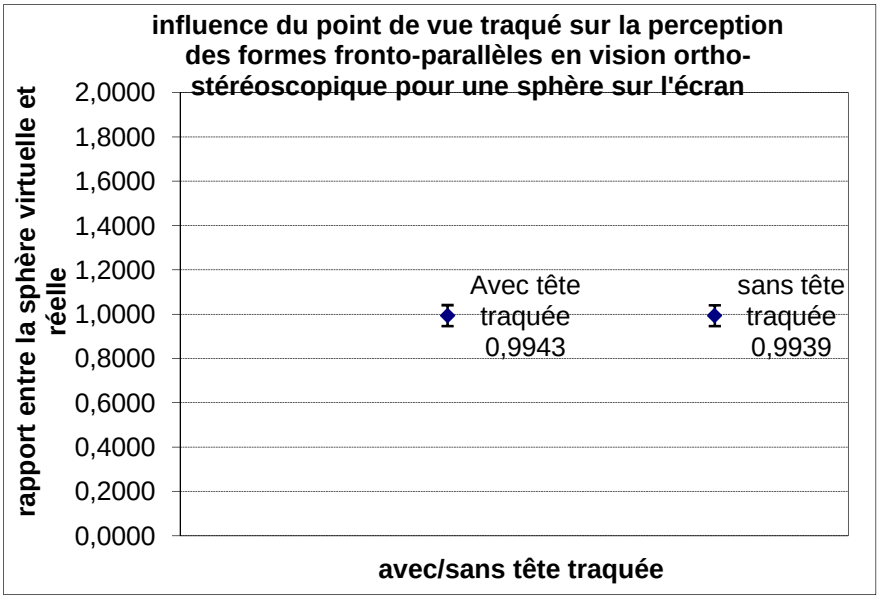


Figure 184 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision ortho-stéréoscopique pour une sphère sur l'écran

Comme la différence entre la perception en hauteur et largeur est significativement différente, nous exposons les différences entre ces perceptions sur la Figure 185. Cependant, aucun des écarts de JND et de PSE n'est significatif.

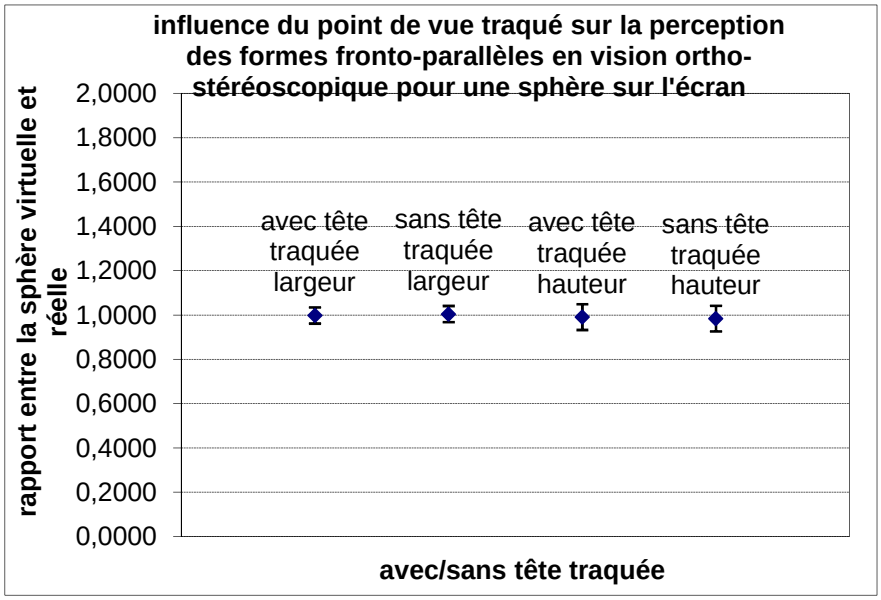


Figure 185: influence du point de vue traqué sur la perception des formes fronto-parallèles en vision ortho-stéréoscopique pour une sphère sur l'écran, détails

Lorsque la sphère est derrière l'écran

Nous pouvons réaliser le graphique pour la perception des formes fronto-parallèles. Celui-ci se trouve sur la Figure 186. Les probabilités de significativité pour l'écart de PSE, de JND et entre la largeur et la hauteur sont respectivement 44%, 87% et 92%. Aucune n'est donc significative.

tracking	oui/non
projection	ortho
forme	Sphère
position	derrière

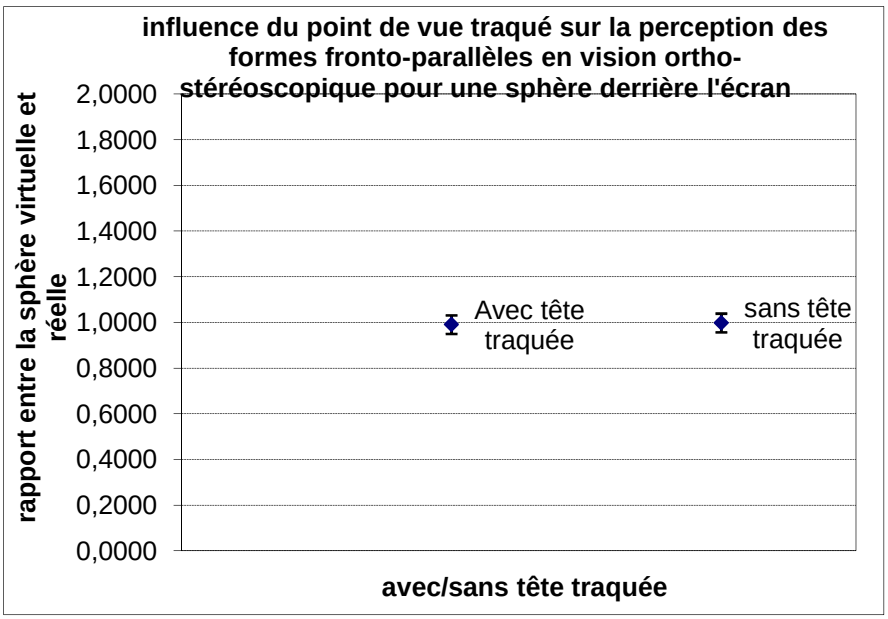


Figure 186 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision stéréoscopique pour une sphère derrière l'écran

Lorsque la sphère est devant l'écran

Les résultats de la perception des formes fronto-parallèles sont exposés sur la Figure 187. Les probabilités de significativité de différences entre les PSE, le JND et de la différence entre la largeur et la hauteur sont respectivement 55%, 87% et 66% ; aucune n'est donc significative.

tracking	oui/non
projection	ortho
forme	Sphère
position	devant

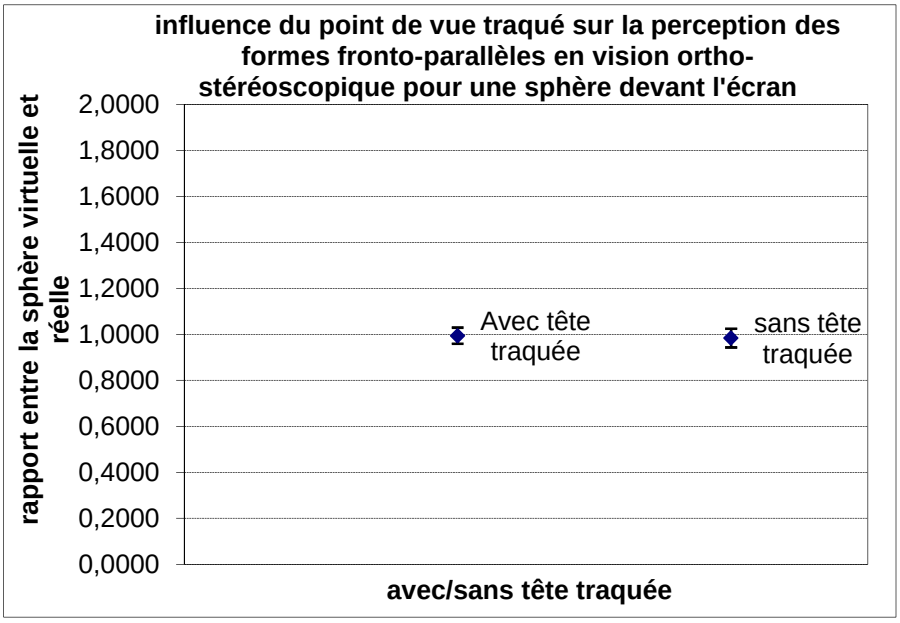


Figure 187 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision ortho-stéréoscopique pour une sphère devant l'écran

Lorsque la forme non connue est sur l'écran

Les résultats de la perception des formes fronto-parallèles sont exposés sur la Figure 188. Les PSE et JND avec point de vue traqué sont égaux à 100% et 3% des dimensions de la forme étalon, sans point de vue traqué ils sont égaux à 100% et 4%. Les probabilités de significativité de l'écart des PSE et des JND sont 51% et 93%, ils ne sont donc pas significatifs. Par contre, la probabilité d'une différence entre la largeur et la hauteur est 95%, donc cette différence est significative.

tracking	oui/non
projection	ortho
forme	Forme
position	sur

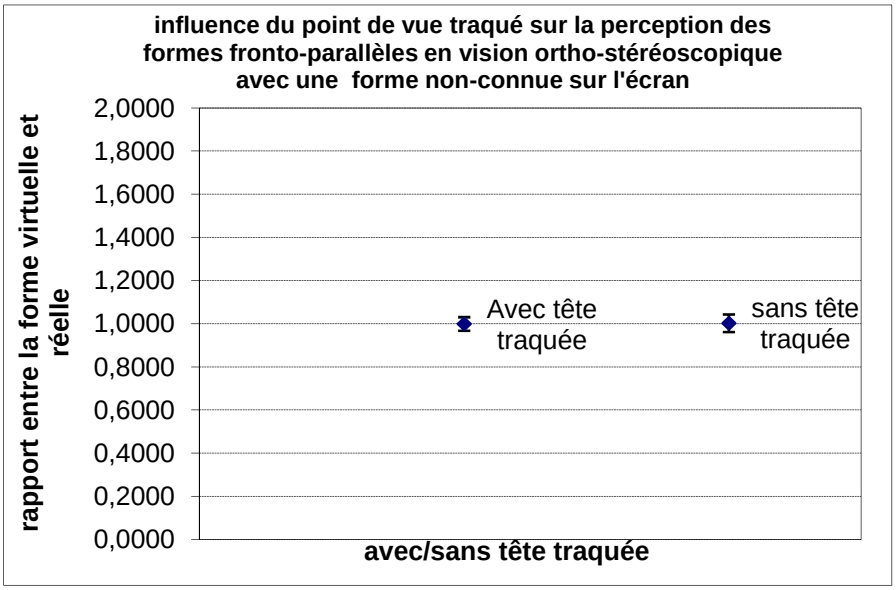


Figure 188 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision ortho-stéréoscopique avec une forme non-connue sur l'écran

Comme la différence entre la perception de la largeur et de la hauteur est significative, nous détaillons les résultats sur la Figure 189. Les PSE et JND en largeur avec point de vue traqué sont égaux à 100% et 3%, les résultats sont les mêmes lorsque le point de vue n'est pas traqué. Pour la hauteur, les PSE et JND avec point de vue traqué sont respectivement 100% et 3% et sans point de vue traqué le PSE reste le même mais le JND devient 5%. Encore cette fois, nous voyons que la différence de perception se fait en hauteur. Les probabilités de significativité en largeur pour les écarts de PSE et de JND sont, comme nous le pensions, négligeables. Par contre, les probabilités de significativité de l'écart de JND en hauteur est 98%, cet écart est donc significatif.

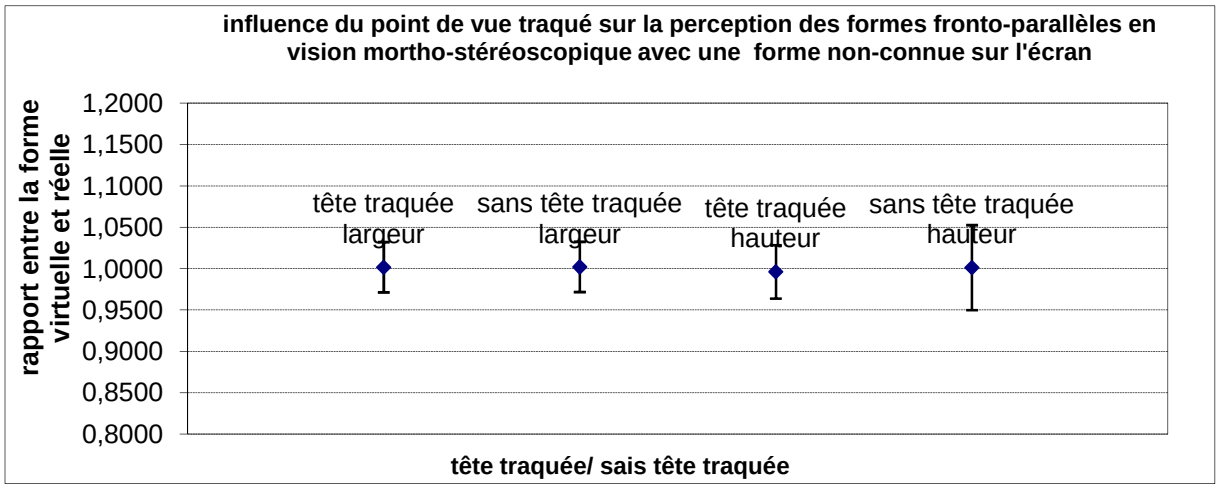


Figure 189 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision ortho-stéréoscopique avec une forme non-connue sur l'écran

Lorsque la forme non-connue est derrière l'écran

Les résultats de l'expérience sur la perception des formes fronto-parallèles sont présentés sur la Figure 190. Les PSE et JND avec point de vue traqué sont égaux à 101% et 4% des dimensions de la forme réelle. Ce sont exactement les mêmes lorsque le point de vue n'est pas traqué. Cela n'a donc aucune influence sur la perception des formes fronto-parallèles. Les probabilités de significativité des écarts de PSE et de JND sont respectivement 54% et 82%. De même la probabilité de significativité des différences entre la perception en largeur et en hauteur n'est pas significative non plus : 88%.

tracking	oui/non
projection	ortho
forme	Forme
position	derrière

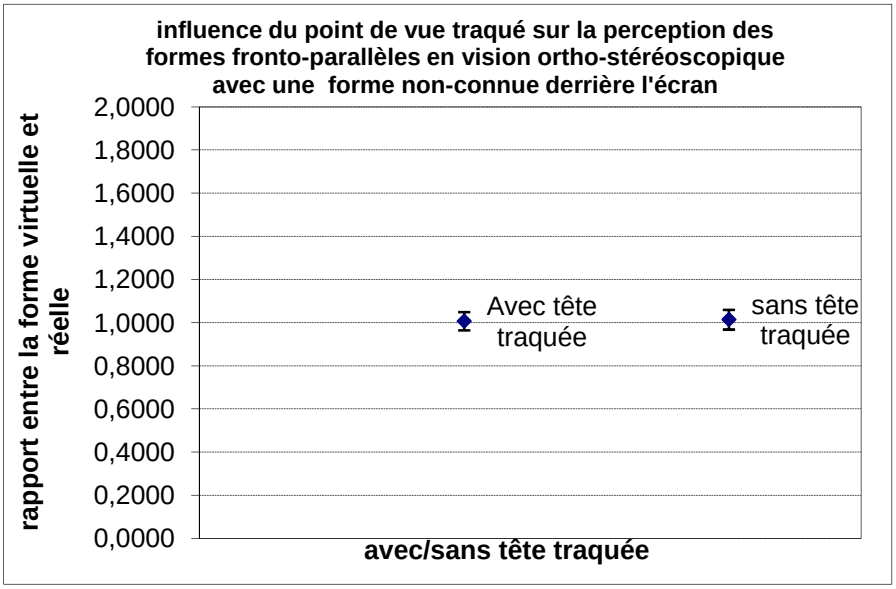


Figure 190 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision ortho-stéréoscopique avec une forme non-connue derrière l'écran

Lorsque la forme non connue est devant l'écran

Nous présentons les résultats des expériences sur les formes fronto-parallèles sur la Figure 191. Les PSE et JND avec le point de vue traqué sont respectivement 100% et 4% des dimensions de la forme. Sans point de vue traqué ils sont égaux à 100% et 5%. Les probabilités de significativité des écarts de JND et de PSE sont 85% et 49%, donc aucun des écarts n'est significatif. La différence entre la perception des courbures en largeur et en hauteur n'est pas non plus significative : 90%.

tracking	oui/non
projection	ortho
forme	Forme
position	devant

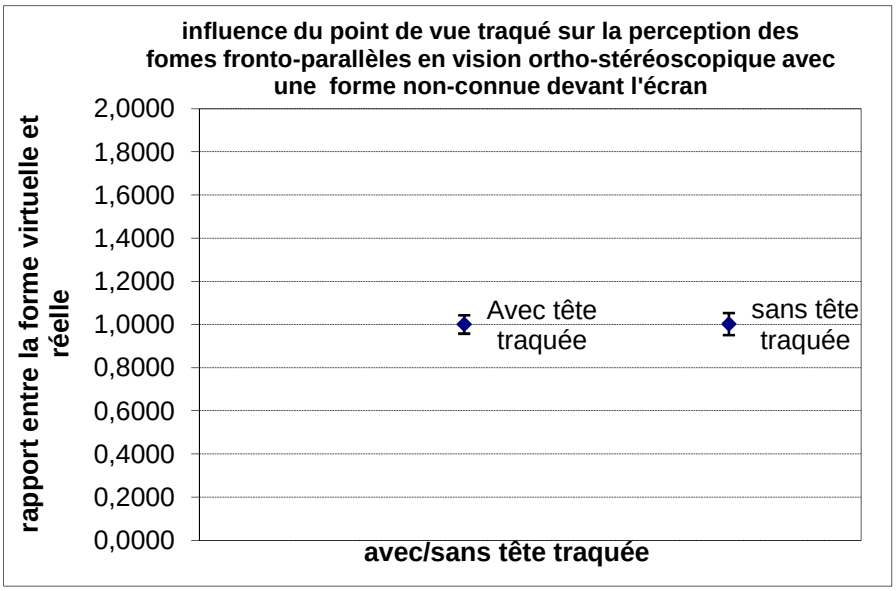


Figure 191 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision ortho-stéréoscopique avec une forme non-connue devant l'écran

Hypothèse n°3 : Il est préférable d'avoir la tête t raquée en vision non-ortho-stéréoscopique

Lorsque la sphère est derrière l'écran

L'influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique pour une sphère derrière l'écran est représenté sur la Figure 192.

tracking	oui/non
projection	Non-ortho
forme	Sphère
position	derrière

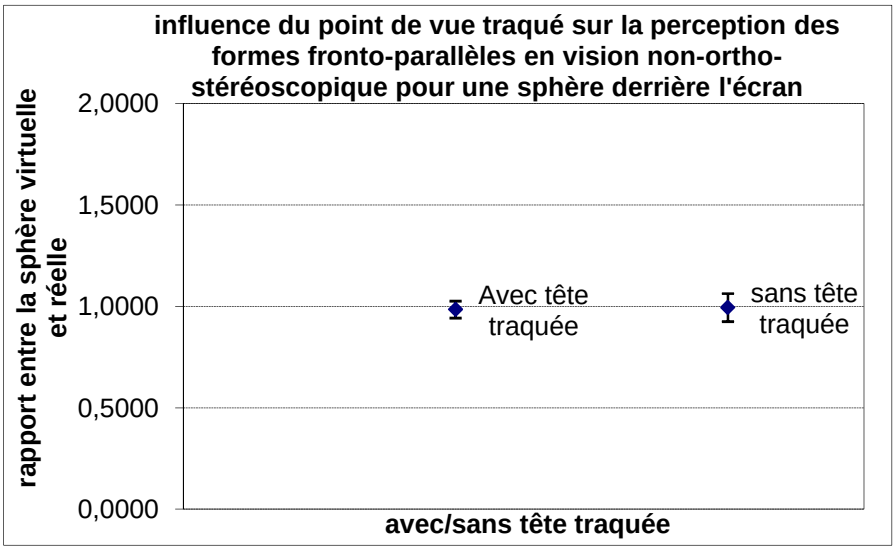


Figure 192 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique pour une sphère derrière l'écran

Lorsque la sphère est devant l'écran

L'influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique pour une sphère devant l'écran est représenté sur la Figure 193.

tracking	oui/non
projection	Non-ortho
forme	Sphère
position	Devant

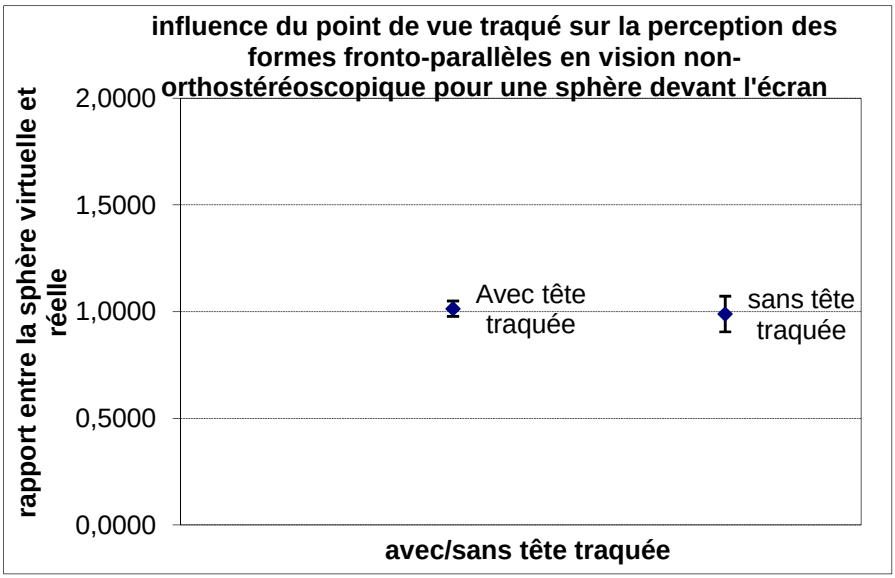


Figure 193 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique pour une sphère devant l'écran

Lorsque la forme non-connue a priori est sur l'écran

L'influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique pour une forme sur l'écran est représenté sur la Figure 194.

tracking	oui/non
projection	Non-ortho
forme	Forme
position	sur

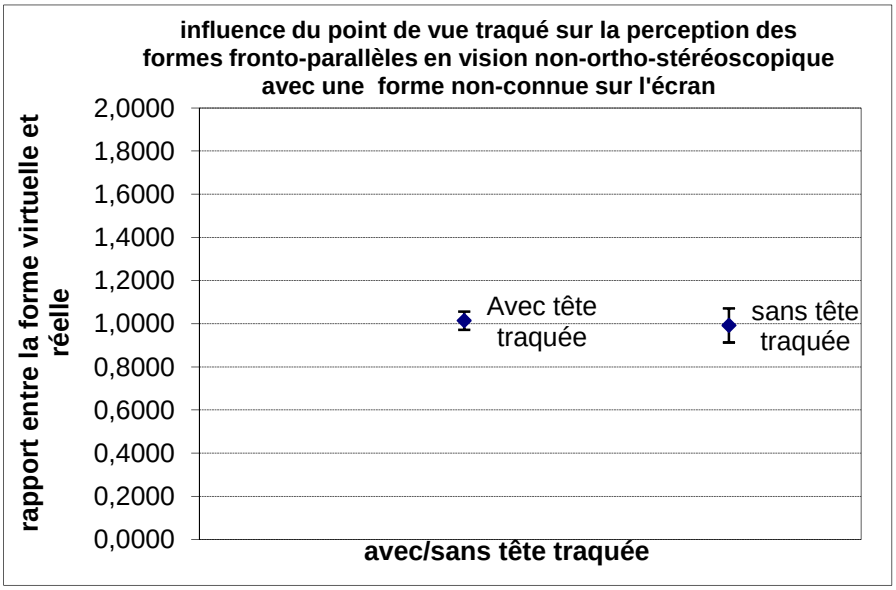


Figure 194 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique avec un forme non-connue sur l'écran

Lorsque la forme non connue est derrière l'écran

L'influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique pour une forme derrière l'écran est représenté sur la Figure 195.

tracking	oui/non
projection	Non-ortho
forme	Forme
position	derrière

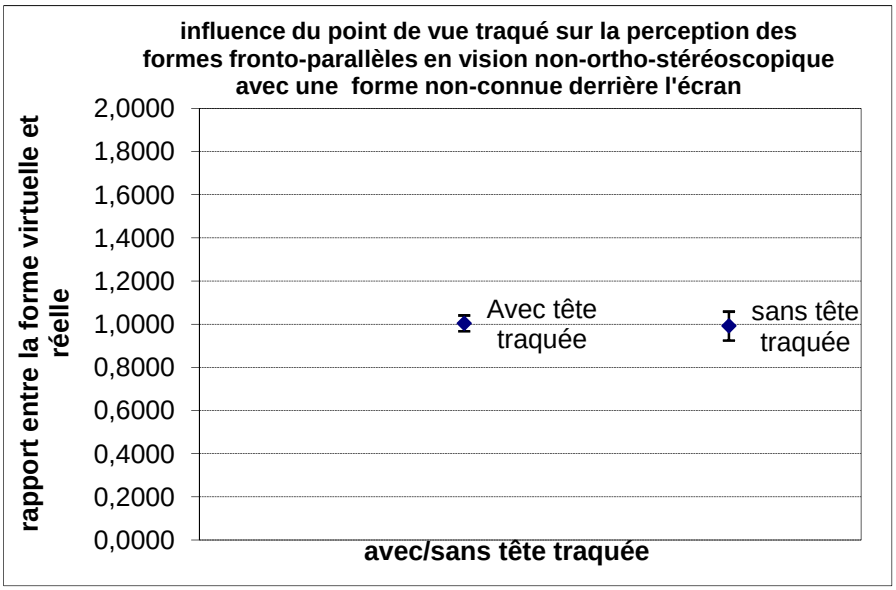


Figure 195 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique avec une forme non-connue derrière l'écran

Lorsque la forme non connue est devant l'écran

L'influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique pour une forme devant l'écran est représenté sur la Figure 196.

tracking	oui/non
projection	Non-ortho
forme	Forme
position	devant

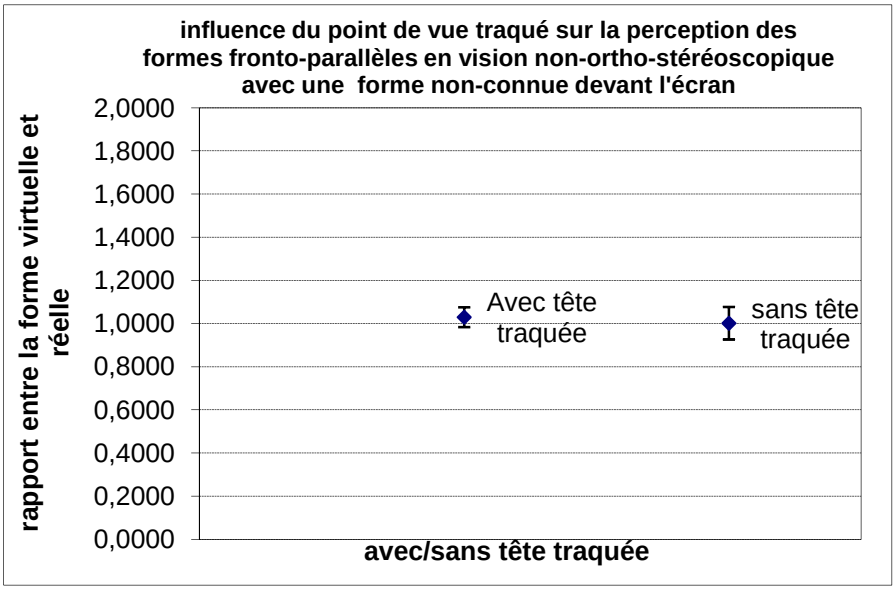


Figure 196 : influence du point de vue traqué sur la perception des formes fronto-parallèles en vision non-ortho-stéréoscopique avec une forme non-connue devant l'écran

Hypothèse n°4 : Il est préférable d'avoir la vision ortho-stéréoscopique que la vision non-ortho-stéréoscopique avec le point de vue traqué

Lorsque la sphère est sur l'écran

Les résultats de la perception des formes fronto-parallèles sont présentés en Figure 197. Les PSE et JND en vision ortho-stéréoscopique sont 99% et 5% du diamètre de la sphère réelle. En vision non-ortho-stéréoscopique, ils sont égaux à 99% et 4%. Aucun écart n'est significatif, même si la différence entre la perception des courbures horizontales et verticales obtient une probabilité de significativité de 94%.

tracking	Oui
projection	Ortho/Non-ortho
forme	Sphère
position	sur

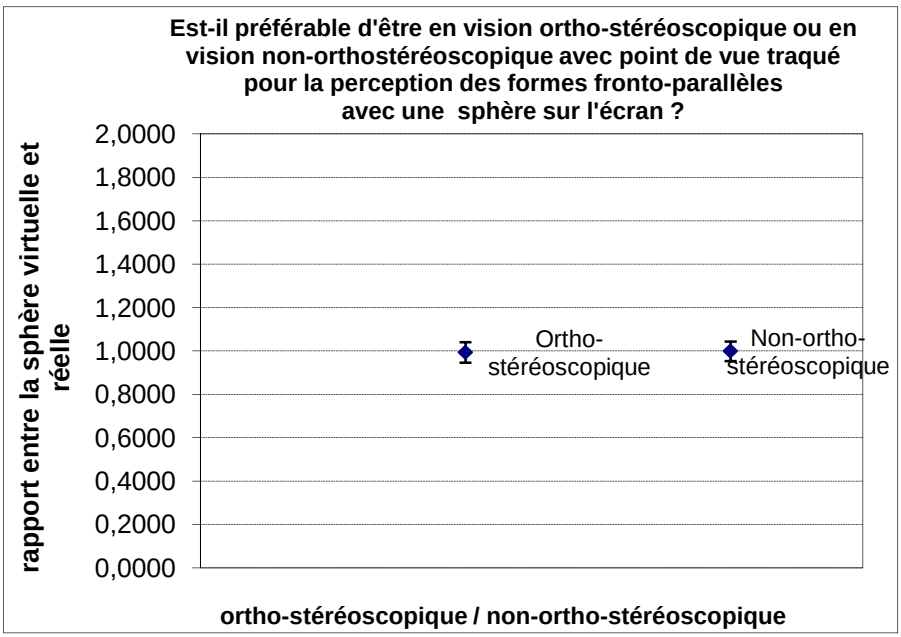


Figure 197 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes fronto-parallèles avec une sphère sur l'écran?

Lorsque la sphère est derrière l'écran

Les résultats des expériences réalisées sur les formes fronto-parallèles sont exposés sur la Figure 198. Les PSE et JND en vision ortho-stéréoscopique sont égaux à 99% et 4% du diamètre de la sphère étalon, tandis que ceux en vision non-ortho-stéréoscopique sont 98% et 4%. Aucun écart n'est significatif, les probabilités de significativité des écarts de PSE et de JND sont égales à 53% et 78% tandis que celle liée à la différence de perception de courbures horizontales et verticales est égale à 83%.

tracking	Oui
projection	Ortho/Non-ortho
forme	Sphère
position	derrière

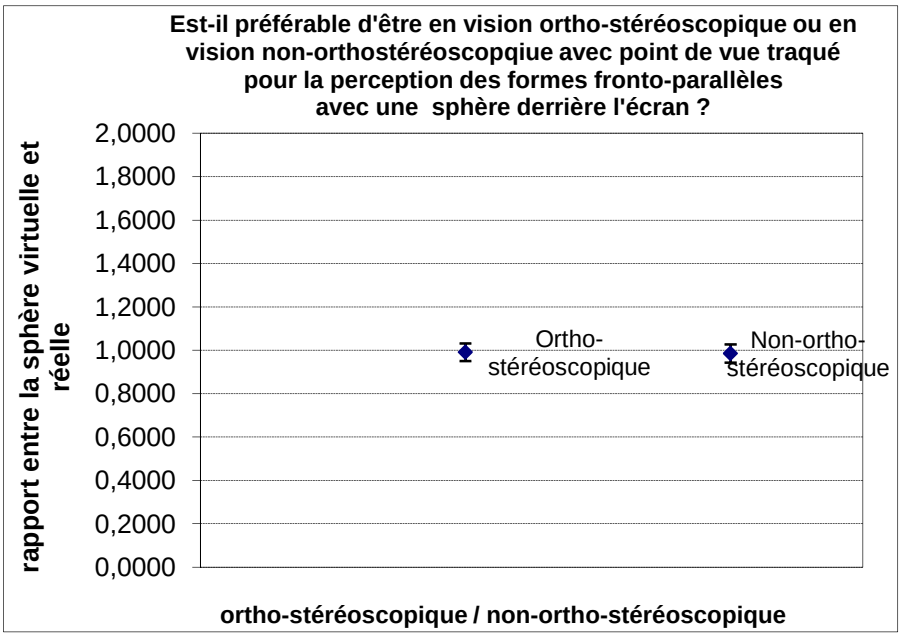


Figure 198 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes fronto-parallèles avec une sphère derrière l'écran?

Lorsque la sphère est devant l'écran

Les résultats de l'expérience sur les formes fronto-parallèles sont présentés sur la Figure 199. Les PSE et JND en vision ortho-stéréoscopique sont 99% et 3% du diamètre de la sphère initiale. En vision non-ortho-stéréoscopique, ils sont égaux à 101% et 3%. Aucun des écarts n'est significatif : 78% pour l'écart des JND, 60% pour celui des PSE et 76% pour la différence entre la perception des courbures en hauteur et en largeur.

tracking	Oui
projection	Ortho/Non-ortho
forme	Sphère
position	Devant

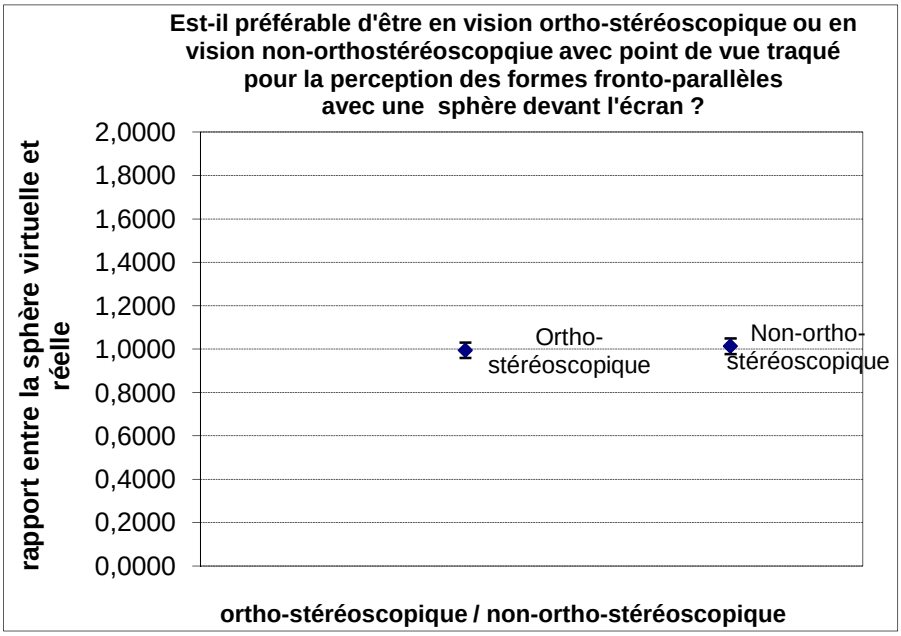


Figure 199 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes en profondeur avec une sphère devant l'écran?

Lorsque la forme non connue est sur l'écran

Nous présentons les résultats des expériences avec les formes fronto-parallèles sur la Figure 200. Les PSE et JND en vision ortho-stéréoscopique sont 101% et 4% des dimensions de la forme initiale, tandis que ceux en vision non-ortho-stéréoscopique sont égaux à 101% et 5%. Aucun des écarts relevés (entre le JND, le PSE et entre la perception des courbures horizontales ou verticales) ne sont significatif (94%, 59% et 77%).

tracking	Oui
projection	Ortho/Non-ortho
forme	Forme
Position	sur

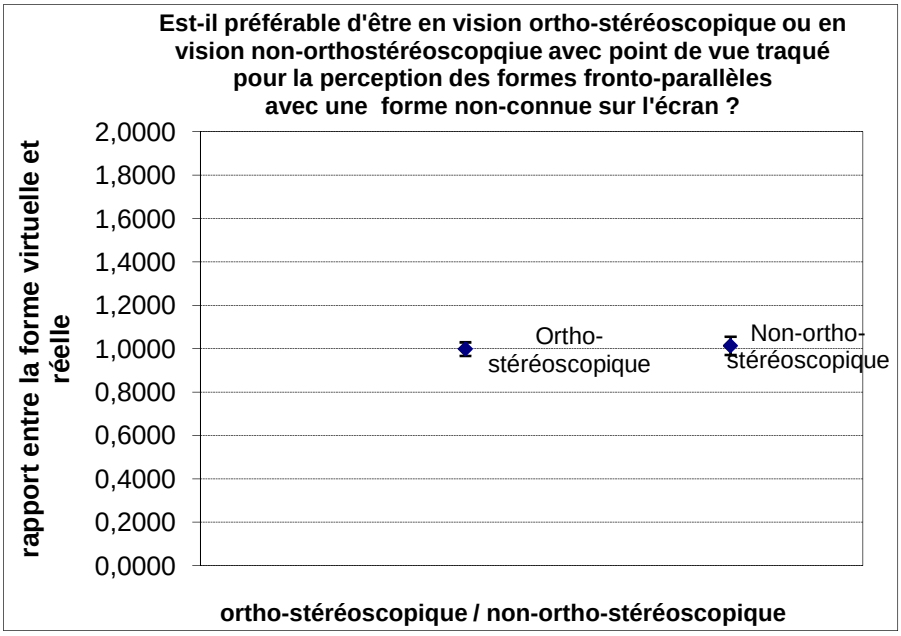


Figure 200 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes fronto-parallèles avec une forme inconnue sur l'écran?

Lorsque la forme est derrière l'écran

Les résultats des expériences pour les formes fronto-parallèles sont exposés sur la Figure 201. Les PSE et JND lorsque nous sommes en vision ortho-stéréoscopique sont 100% et 4% contre 100% et 3% des dimensions initiales lorsque nous sommes en vision non-ortho-stéréoscopique. Aucun des écarts n'est significatif : l'écart entre les JND a une probabilité de significativité de 87%, celui des PSE, une probabilité de 51% et la différence entre la perception des courbures horizontales et verticales, 80%.

tracking	Oui
projection	Ortho/Non-ortho
forme	Forme
position	Derrière

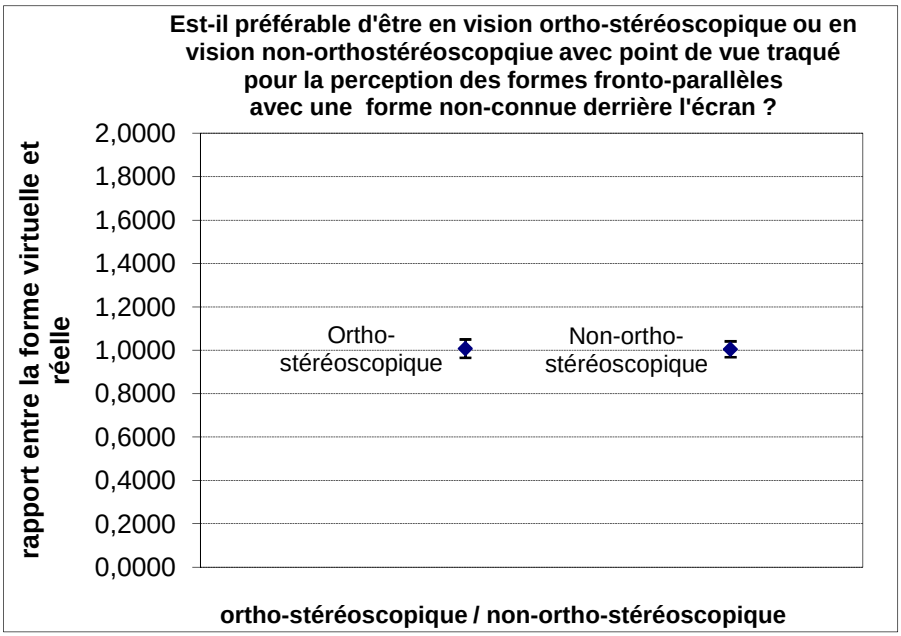


Figure 201 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes fronto-parallèles avec une forme inconnue derrière l'écran?

Lorsque la forme non connue est devant l'écran

Les résultats de l'expérience sur les formes fronto-parallèles sont présentés sur la Figure 202. Les PSE et JND en vision ortho-stéréoscopique sont 100% et 3% des dimensions de la forme initiale et sont les mêmes en vision non-ortho-stéréoscopique. Les différences entre les JND et les PSE ne sont pas significatives. Par contre, la probabilité de significativité d'une différence de perception des formes en largeur ou en hauteur est 95%, donc significative.

tracking	Oui
projection	Ortho/Non-ortho
forme	Forme
position	devant

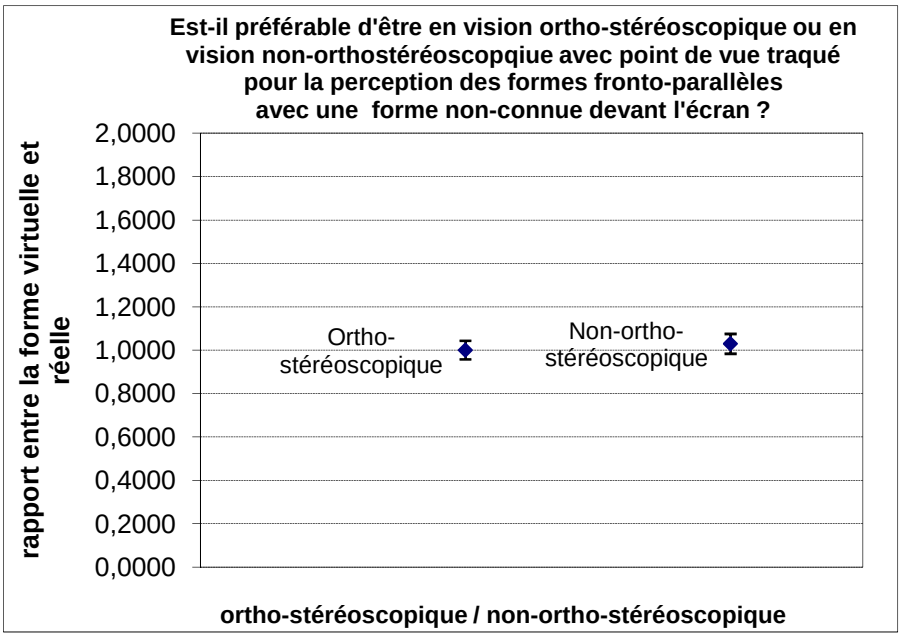


Figure 202 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes fronto-parallèles avec une forme inconnue devant l'écran?

Comme l'écart entre la perception de la largeur et de la hauteur est significatif, nous exposons les résultats pour la largeur et la hauteur en détails sur la Figure 203. Les PSE et JND en largeur en vision ortho-stéréoscopique sont 98% et 5,3% des dimensions de la forme étalon, tandis qu'en vision non-ortho-stéréoscopique, ils sont égaux à 103% et 5%. En hauteur, les PSE et JND en vision ortho-stéréoscopique sont 100% et 3%, alors qu'en vision non-ortho-stéréoscopique ils sont 102% et 4%. Mais aucun de ces écarts n'est significatif.

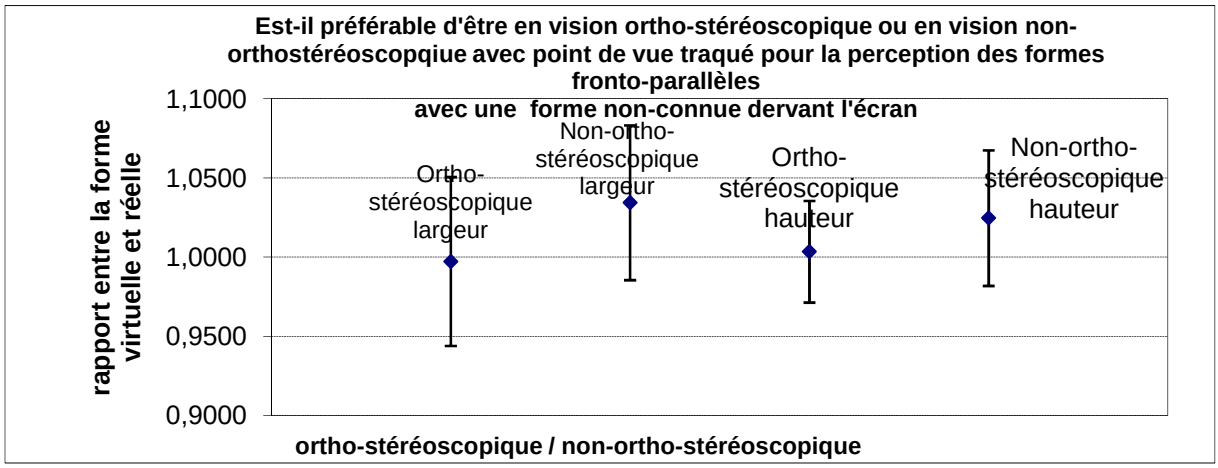


Figure 203 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes fronto-parallèles avec une forme inconnue devant l'écran? Détails.

Hypothèse n°5 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique sans point de vue traqué ?

Lorsque la sphère est sur l'écran

Les résultats de l'expérience sur les formes fronto-parallèles sont exposés sur la Figure 204. Aucun des écarts n'est significatif, même si celui entre les JND obtient une probabilité de significativité de 94%.

tracking	non
projection	Ortho/Non-ortho
forme	Sphère
position	sur

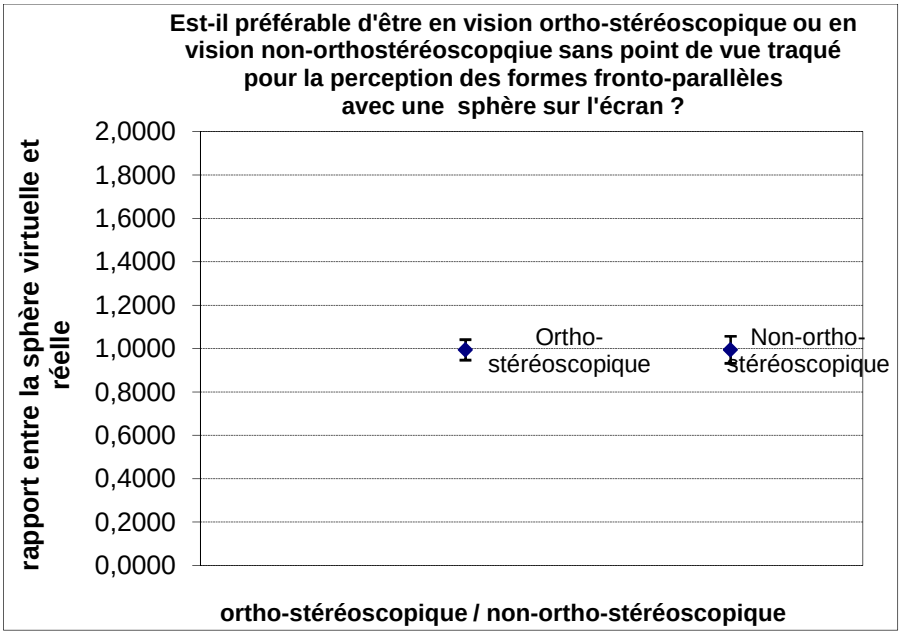


Figure 204 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une sphère sur l'écran?

Lorsque la sphère est devant l'écran

Les résultats de l'expérience sur les formes fronto-parallèles sont exposés sur la Figure 205.

tracking	non
projection	Ortho/Non-ortho
forme	Sphère
position	devant

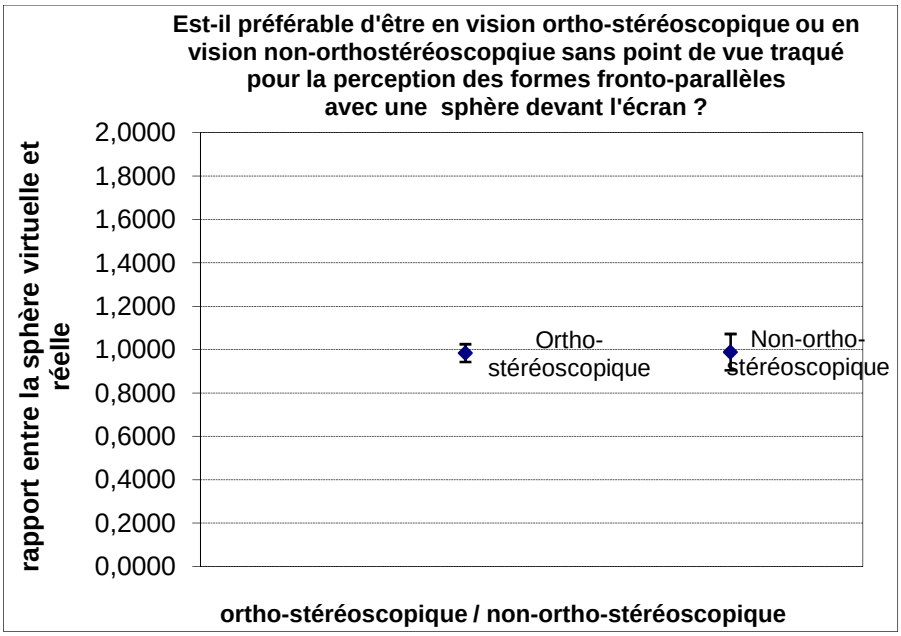


Figure 205 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une sphère devant l'écran?

Lorsque la forme non connue est sur l'écran

Les résultats de l'expérience sur les formes fronto-parallèles sont exposés sur la Figure 206.

tracking	non
projection	Ortho/Non-ortho
forme	Forme
position	sur

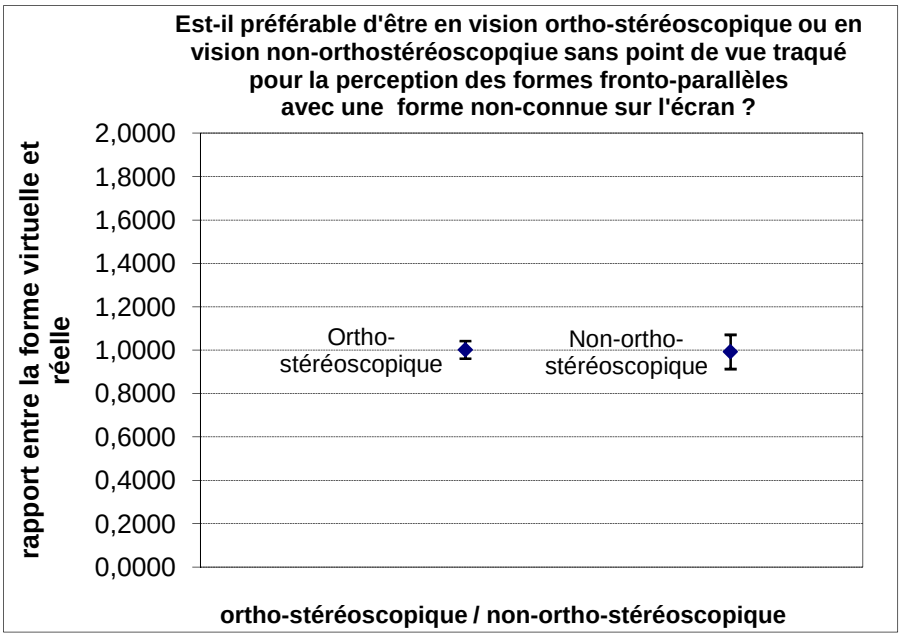


Figure 206 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une forme non connue sur l'écran?

Lorsque la forme non connue est derrière l'écran

Les résultats de l'expérience sur les formes fronto-parallèles sont exposés sur la Figure 207Figure 207.

tracking	non
projection	Ortho/Non-ortho
forme	Forme
position	derrière

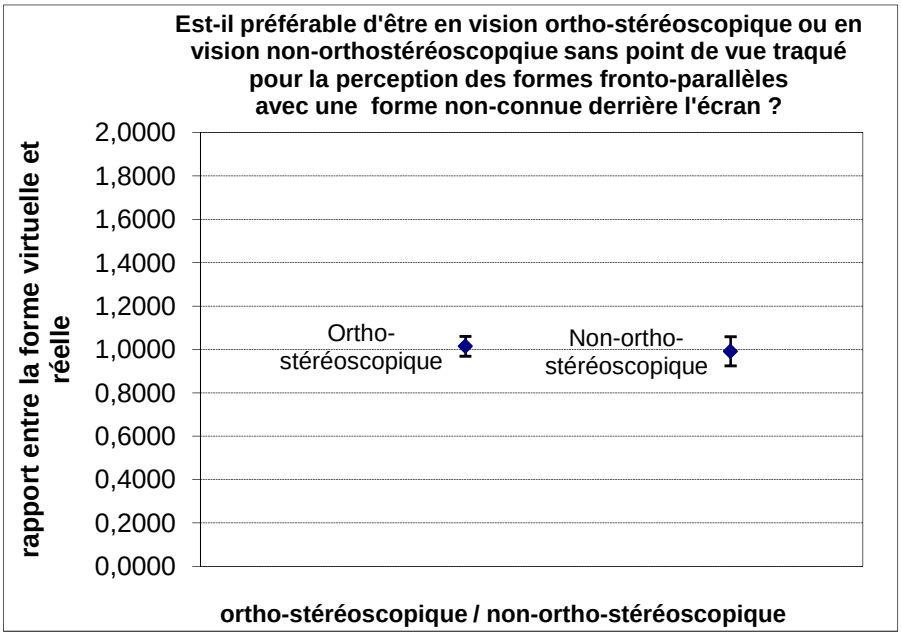


Figure 207 : Est-il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une forme non connue derrière l'écran?

Lorsque la forme non connue est devant l'écran

Les résultats de l'expérience sur les formes fronto-parallèles sont exposés sur la Figure 208.

tracking	non
projection	Ortho/Non-ortho
forme	Forme
position	devant

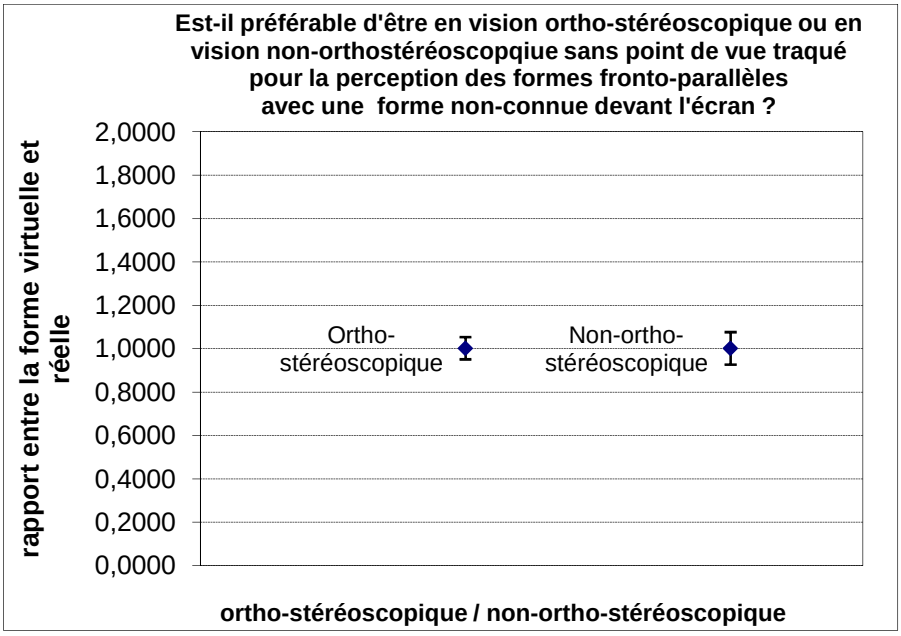


Figure 208 : Est- il préférable d'être en vision ortho-stéréoscopique ou en vision non-ortho-stéréoscopique avec point de vue traqué pour la perception des formes fronto-parallèles avec une forme non connue devant l'écran?

Hypothèse n°6 : Est-il préférable d'être en vision monoscopique avec le point de vue traqué ou en vision ortho-stéréoscopique sans point de vue traqué ?

Lorsque la sphère est sur l'écran

Les résultats de l'expérience sur les formes fronto-parallèles sont exposés sur la Figure 209. Les PSE et JND en vision monoscopique avec point de vue traqué sont égaux à 99% et 3,5% du diamètre étalon. En vision ortho-stéréoscopique sans point de vue traqué, ils sont égaux à 99% et 4,7%. Les écarts de JND et de PSE ne sont pas significatifs (94% et 5%) par contre, celui entre la perception des courbures en largeur et en hauteur l'est (98%).

tracking	Oui/non
projection	Mono/ortho
forme	Sphère
position	Sur

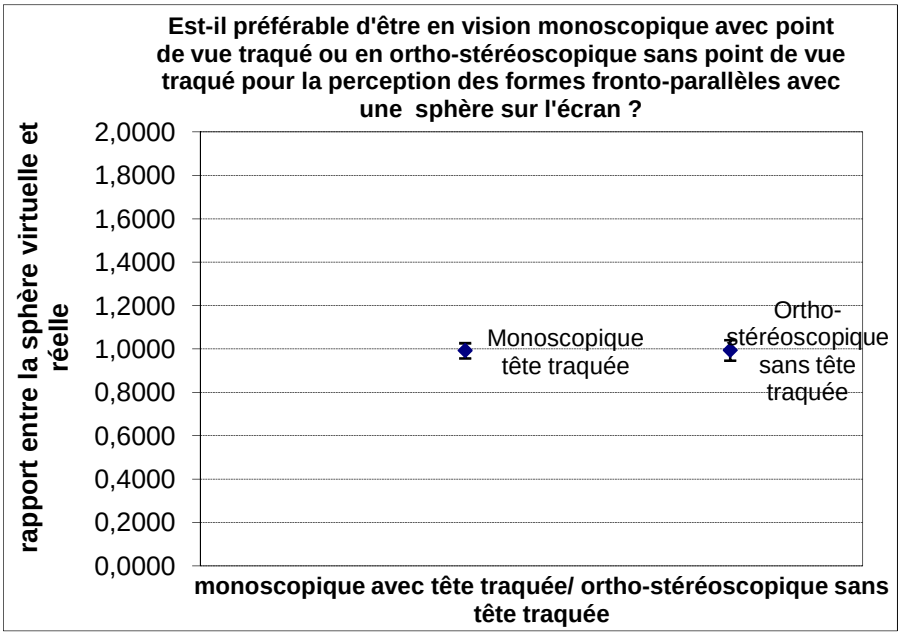


Figure 209 : est-il préférable d'être en vision monoscopique avec point de vue traqué ou en ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une sphère sur l'écran ?

Nous présentons les résultats en détails pour la perception de largeurs et hauteurs sur la Figure 210 puisque cette différence s'était montrée significative. Le seul écart significatif est celui du JND en hauteur.

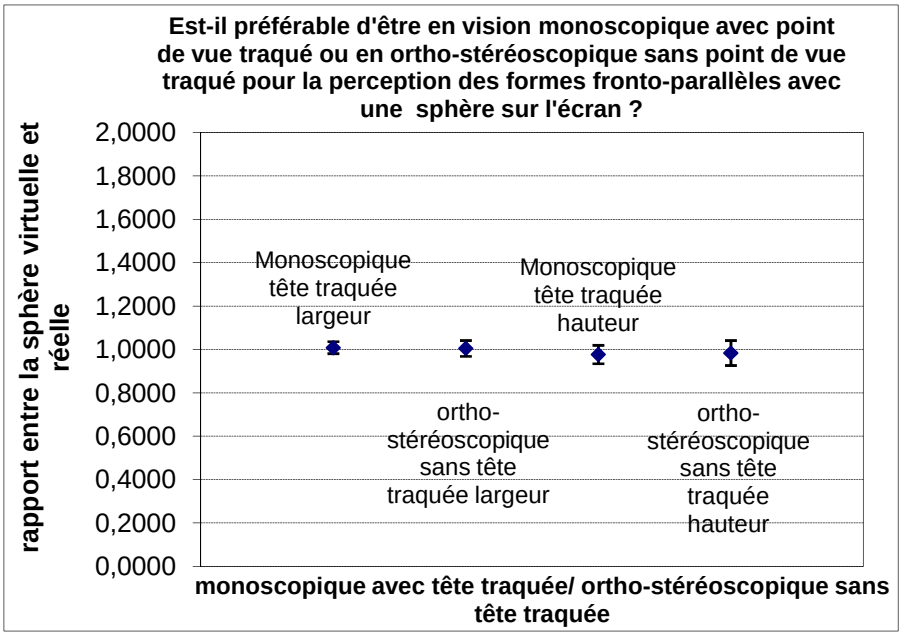


Figure 210 : est-il préférable d'être en vision monoscopique avec point de vue traqué ou en ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une sphère sur l'écran ? détails

Lorsque la sphère est devant l'écran

Les résultats de l'expérience sur les formes fronto-parallèles sont exposés sur la Figure 211.

tracking	Oui/non
projection	Mono/ortho
forme	Sphère
position	Devant

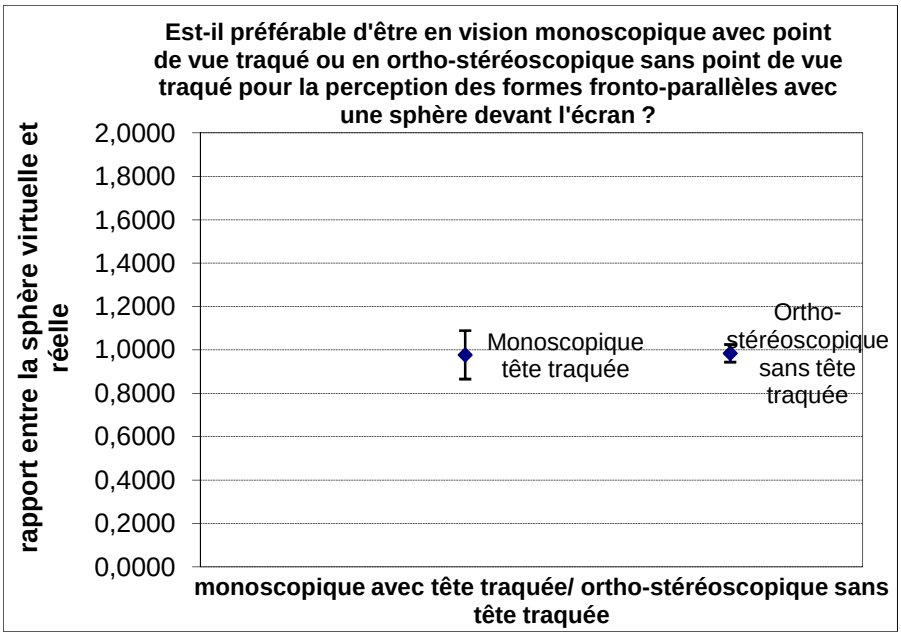


Figure 211 : est-il préférable d'être en vision monoscopique avec point de vue traqué ou en ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une sphère devant l'écran ?

Lorsque la forme non connue est sur l'écran

Les résultats de l'expérience sur les formes fronto-parallèles sont exposés sur la Figure 212.

Tracking	Oui/non
projection	Mono/ortho
Forme	Forme
Position	Sur

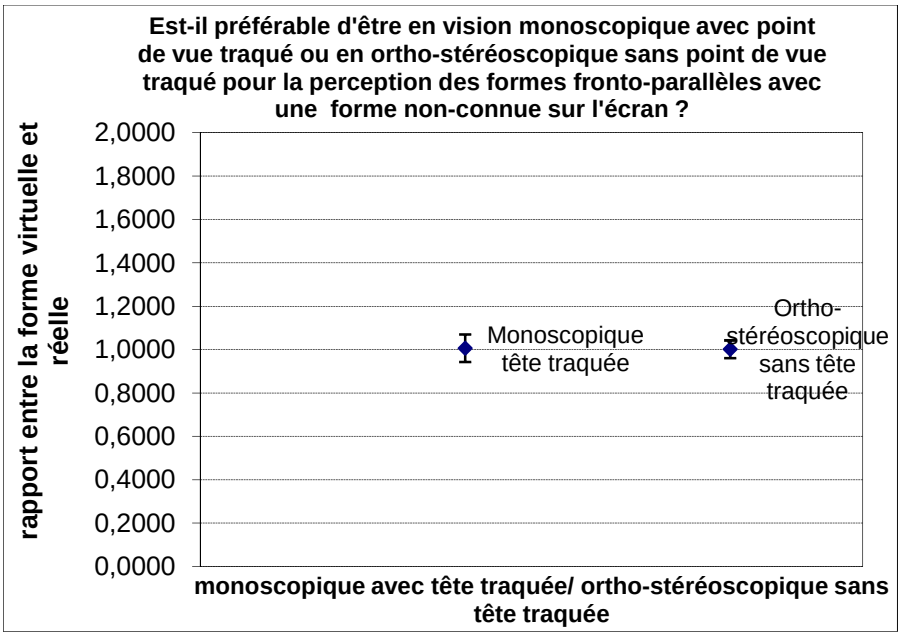


Figure 212 : est-il préférable d'être en vision monoscopique avec point de vue traqué ou en ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une forme non connue sur l'écran ?

Lorsque la forme non connue est derrière l'écran

Les résultats de l'expérience sur les formes fronto-parallèles sont exposés sur la Figure 213.

Tracking	Oui/non
projection	Mono/ortho
Forme	Forme
Position	Derrière

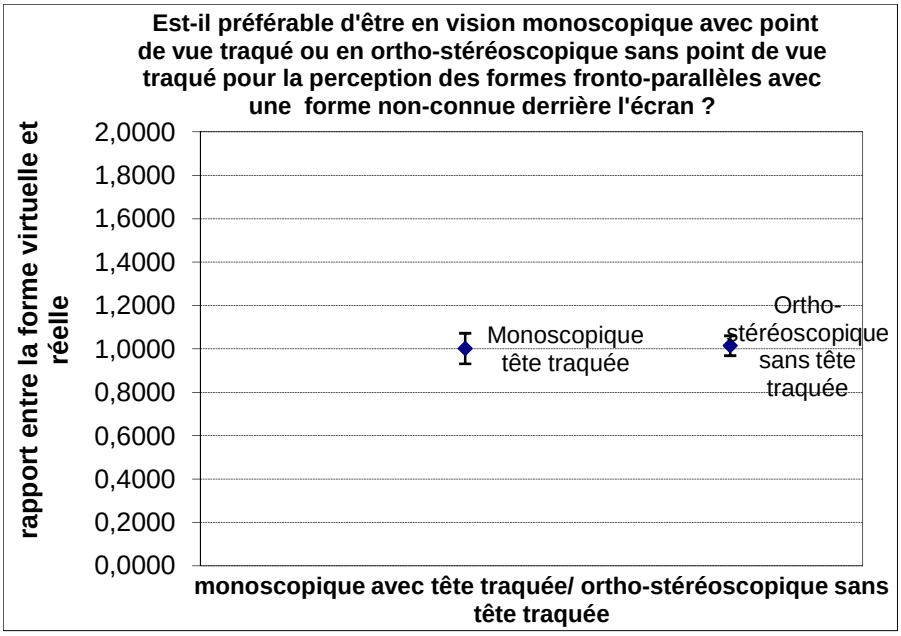


Figure 213 : est-il préférable d'être en vision monoscopique avec point de vue traqué ou en ortho-stéréoscopique sans point de vue traqué pour la perception des formes en fronto-parallèles avec une forme non connue derrière l'écran ?

Lorsque la forme non connue est devant l'écran

Les résultats de la perception des formes fronto-parallèles sont exposés sur la Figure 214. Les PSE et JND en vision monoscopique point de vue traqué sont 99% et 6%, tandis qu'en vision ortho-stéréoscopique point de vue non traqué, ils sont égaux à 100% et 5%. Aucun des écarts n'est significatif.

Tracking	Oui/non
projection	Mono/ortho
Forme	Forme
Position	Devant

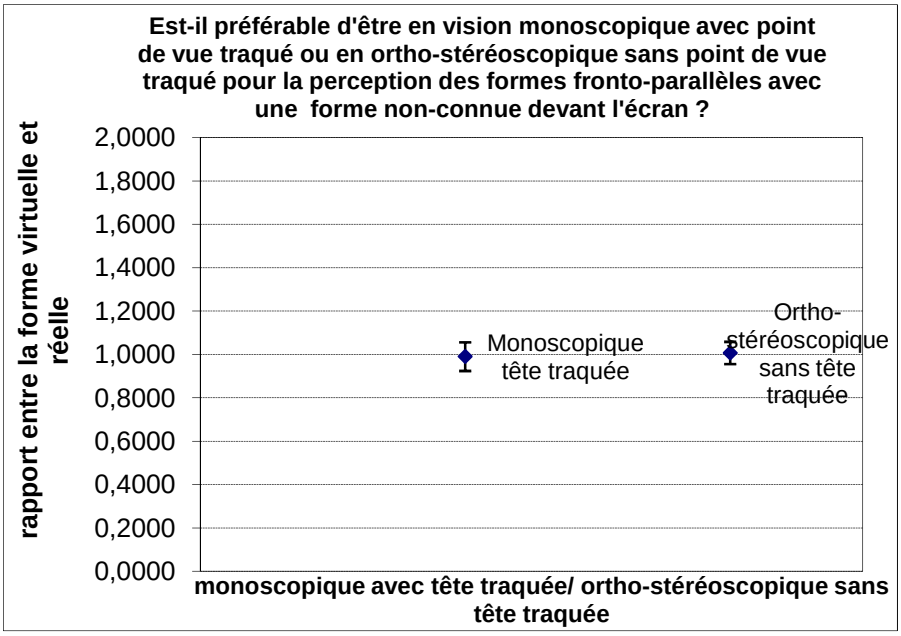


Figure 214 : est-il préférable d'être en vision monoscopique avec point de vue traqué ou en ortho-stéréoscopique sans point de vue traqué pour la perception des formes fronto-parallèles avec une forme non connue devant l'écran ?

Annexe B : tableau récapitulatif pour la perception des formes en immersion

L'influence de la projection sur la perception des formes en profondeur sur une sphère sur l'écran est reportée sur la Figure 215.

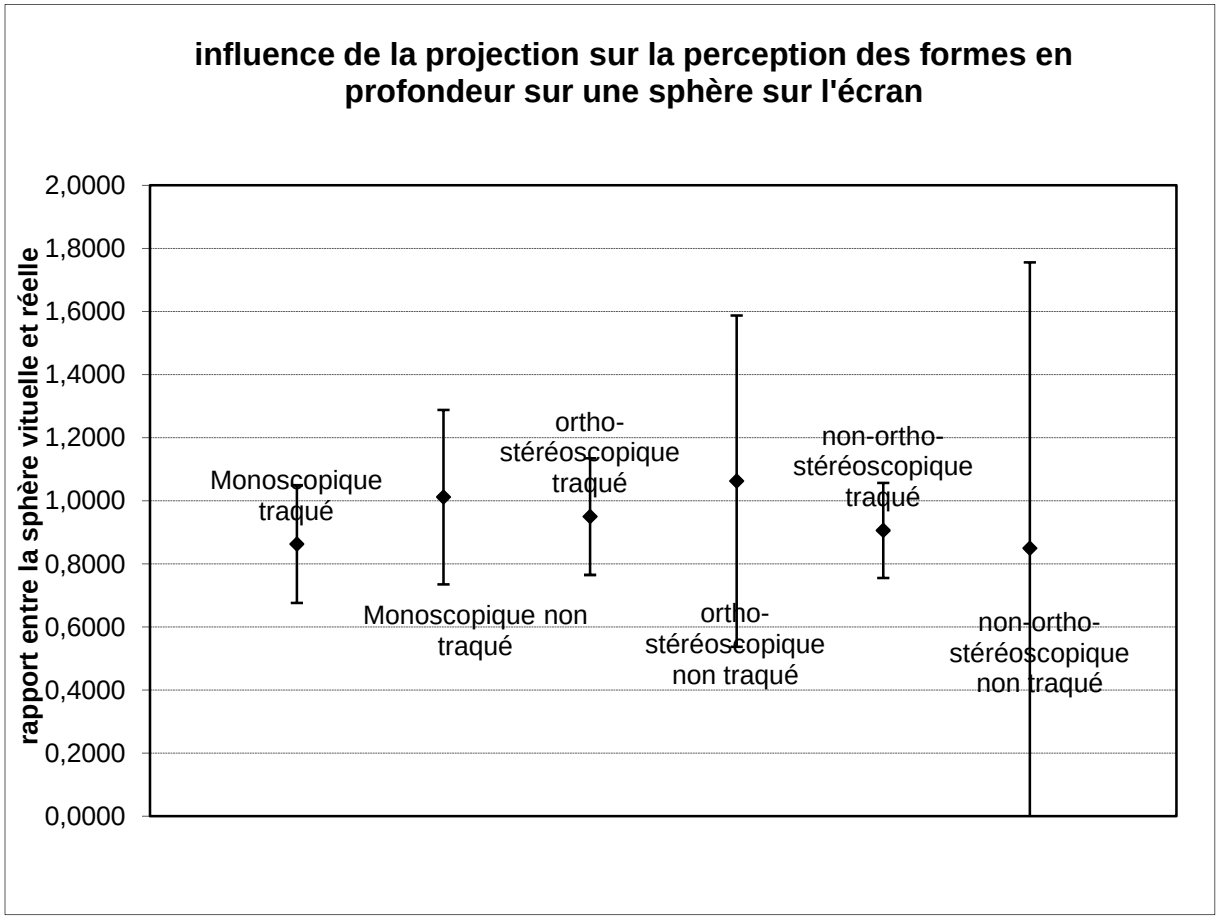


Figure 215 : influence de la projection sur la perception des formes en profondeur sur une sphère sur l'écran

L'influence de la projection sur la perception des formes en profondeur sur une sphère derrière l'écran est reportée sur la Figure 216.

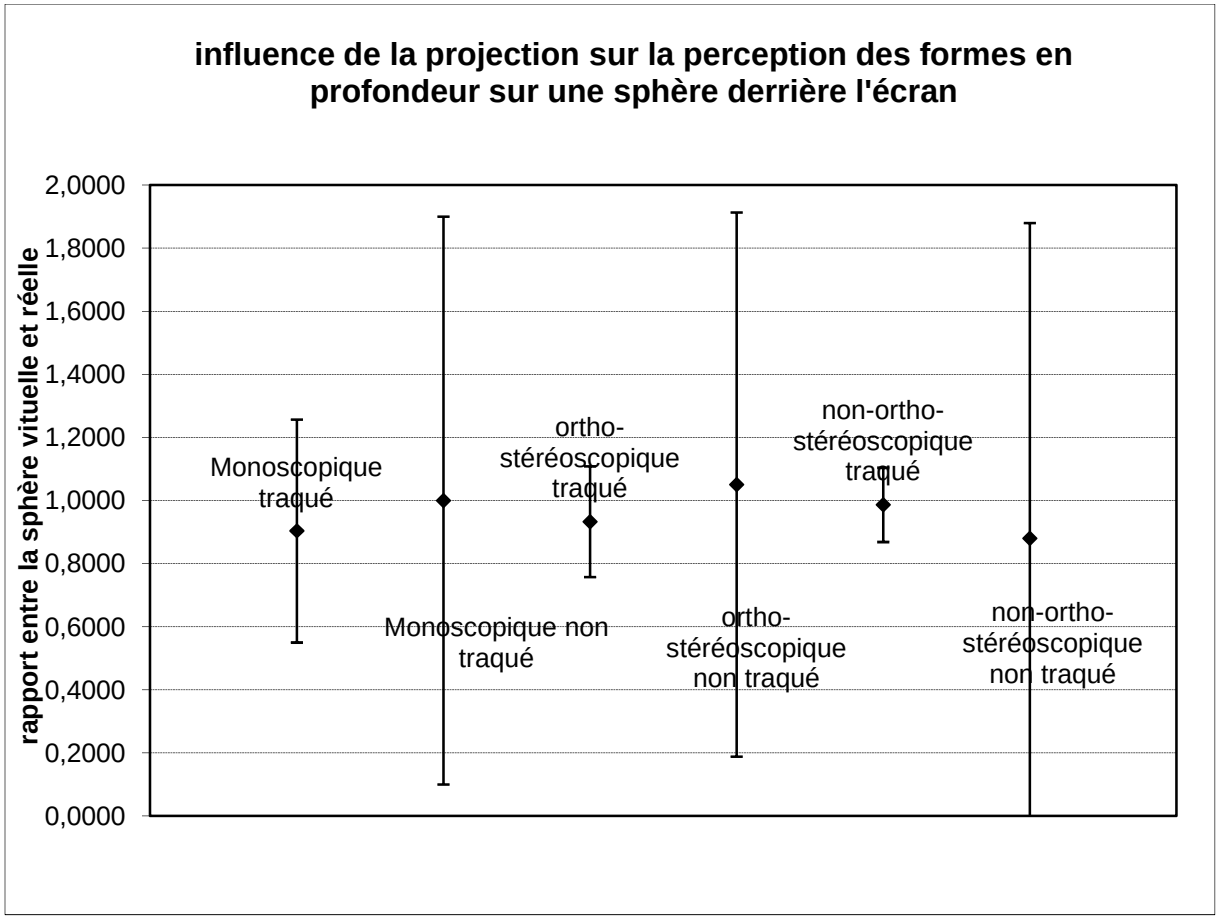


Figure 216 : influence de la projection sur la perception des formes en profondeur sur une sphère derrière l'écran

L'influence de la projection sur la perception des formes en profondeur sur une sphère devant l'écran est reportée sur la Figure 217.

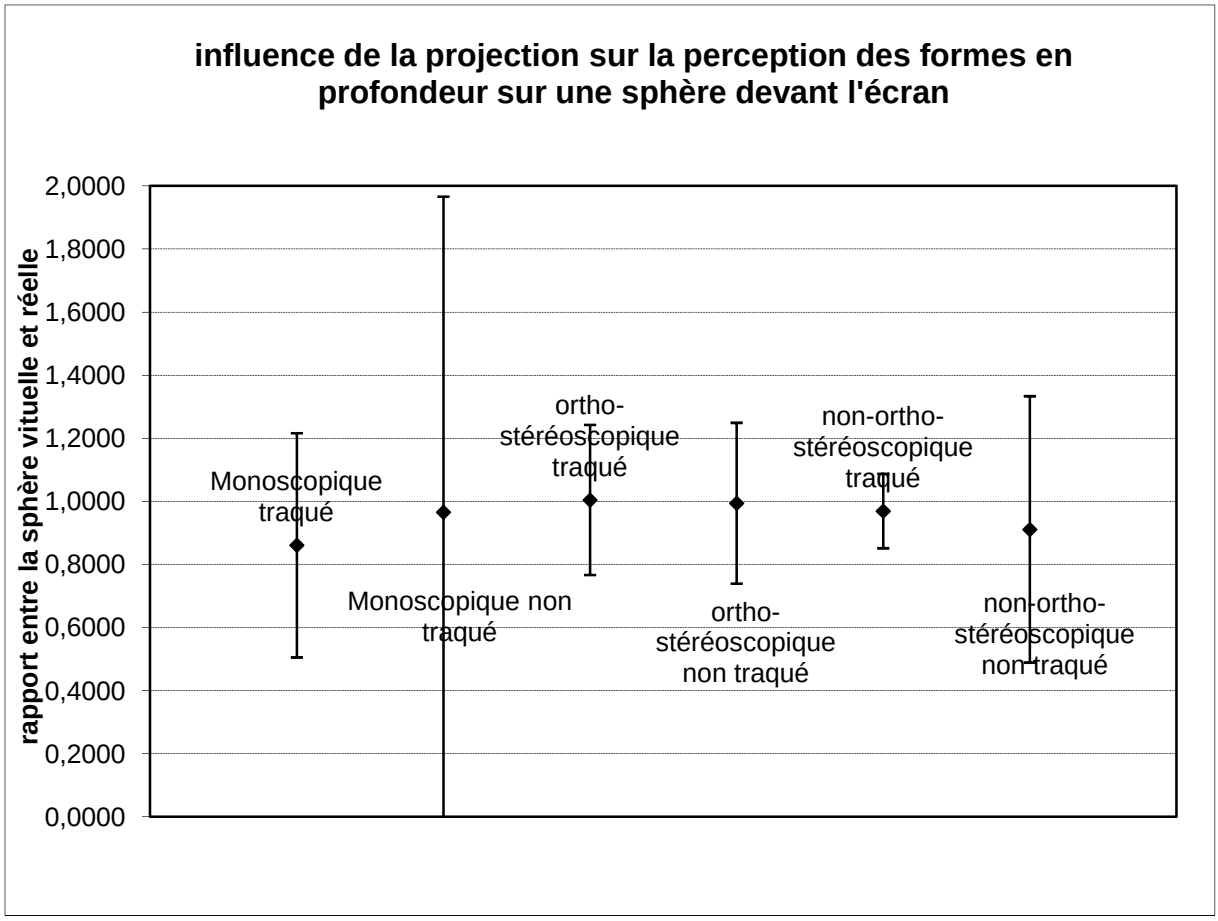


Figure 217 : influence de la projection sur la perception des formes en profondeur sur une sphère devant l'écran

L'influence de la projection sur la perception des formes fronto-parallèle pour une sphère sur l'écran est reportée sur la Figure 218.

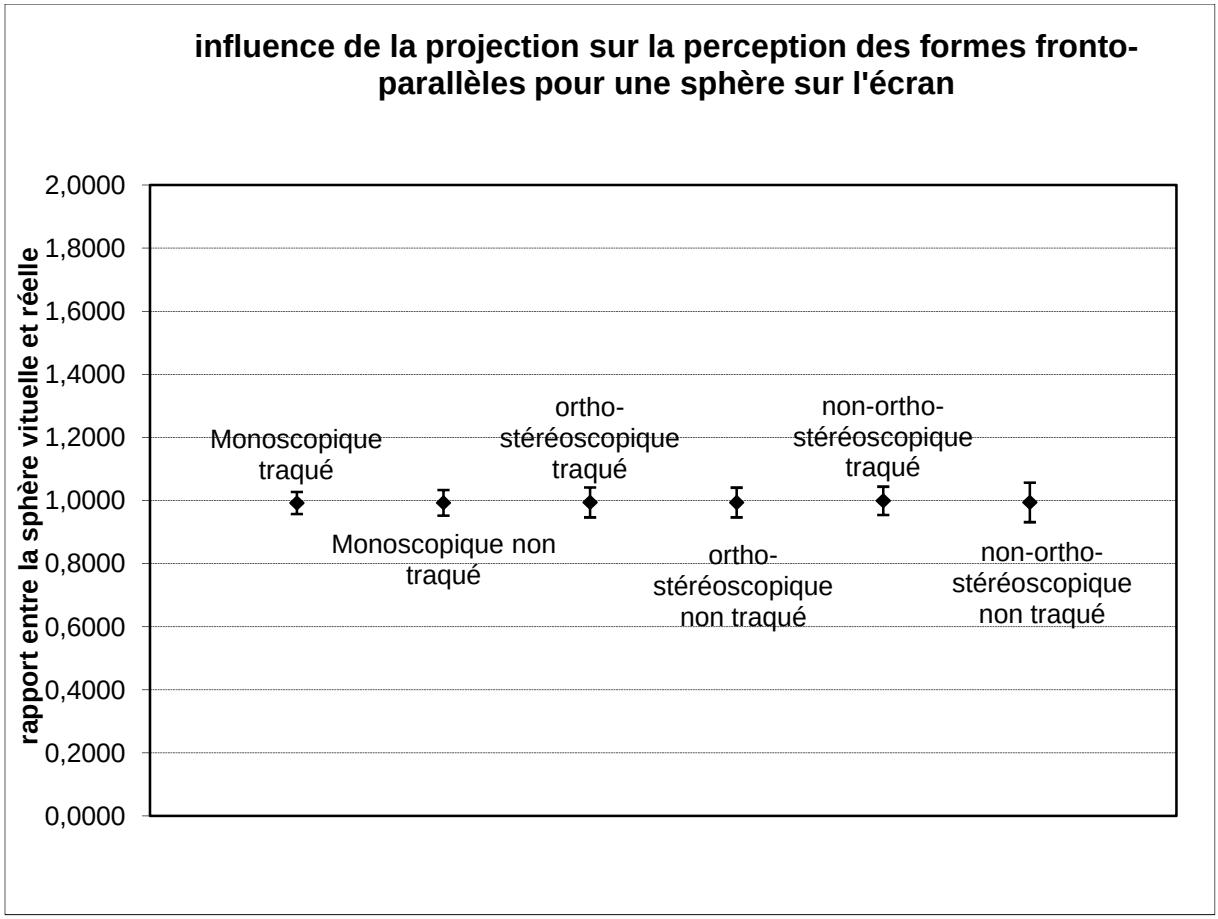


Figure 218 : influence de la projection sur la perception des formes fronto-parallèles pour une sphère sur l'écran

L'influence de la projection sur la perception des formes fronto-parallèle pour une sphère derrière l'écran est reportée sur la Figure 219.

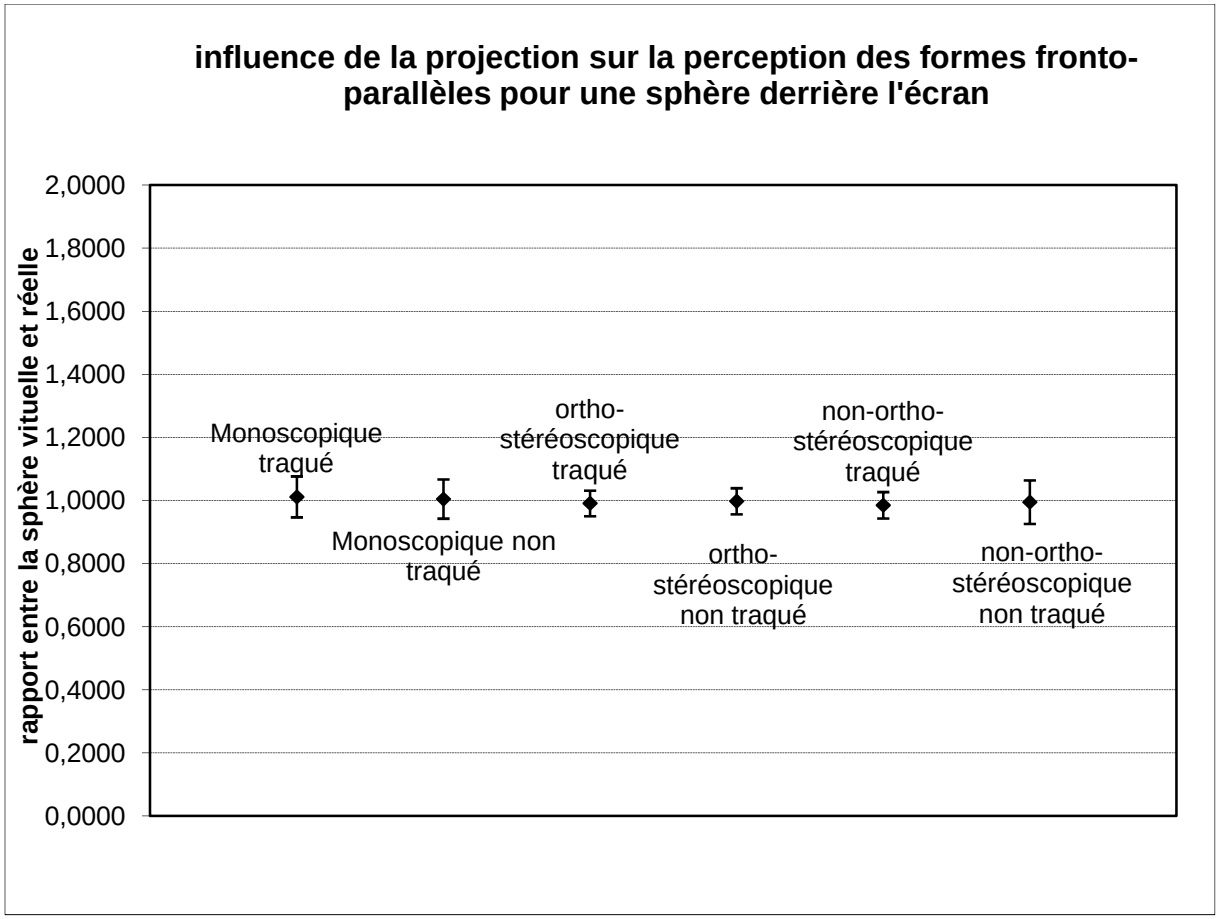


Figure 219 : influence de la projection sur la perception des formes fronto-parallèles pour une sphère derrière l'écran

L'influence de la projection sur la perception des formes fronto-parallèle pour une sphère devant l'écran est reportée sur la Figure 220.

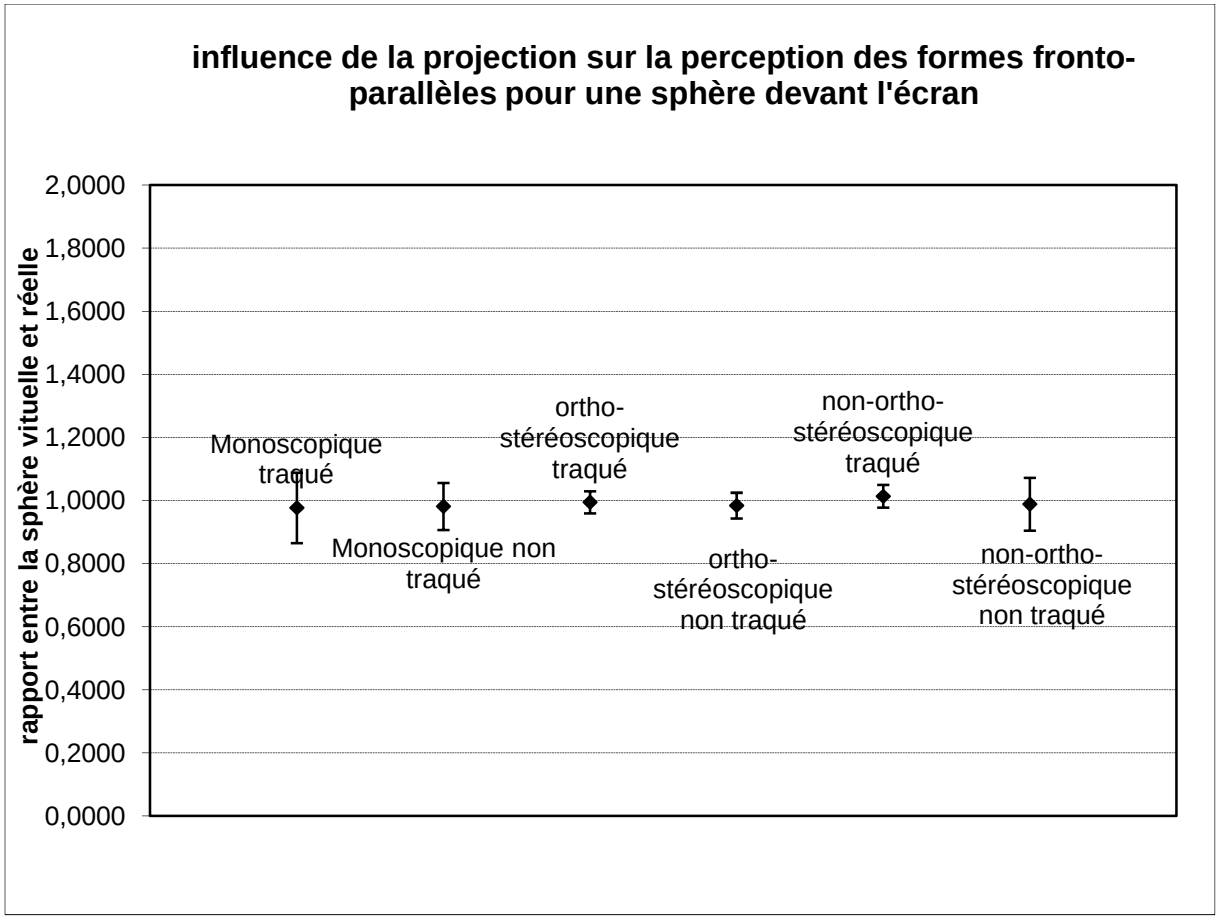


Figure 220 : influence de la projection sur la perception des formes fronto-parallèles pour une sphère devant l'écran

L'influence de la projection sur la perception des formes en relief pour une forme sur l'écran est reportée sur la Figure 221.

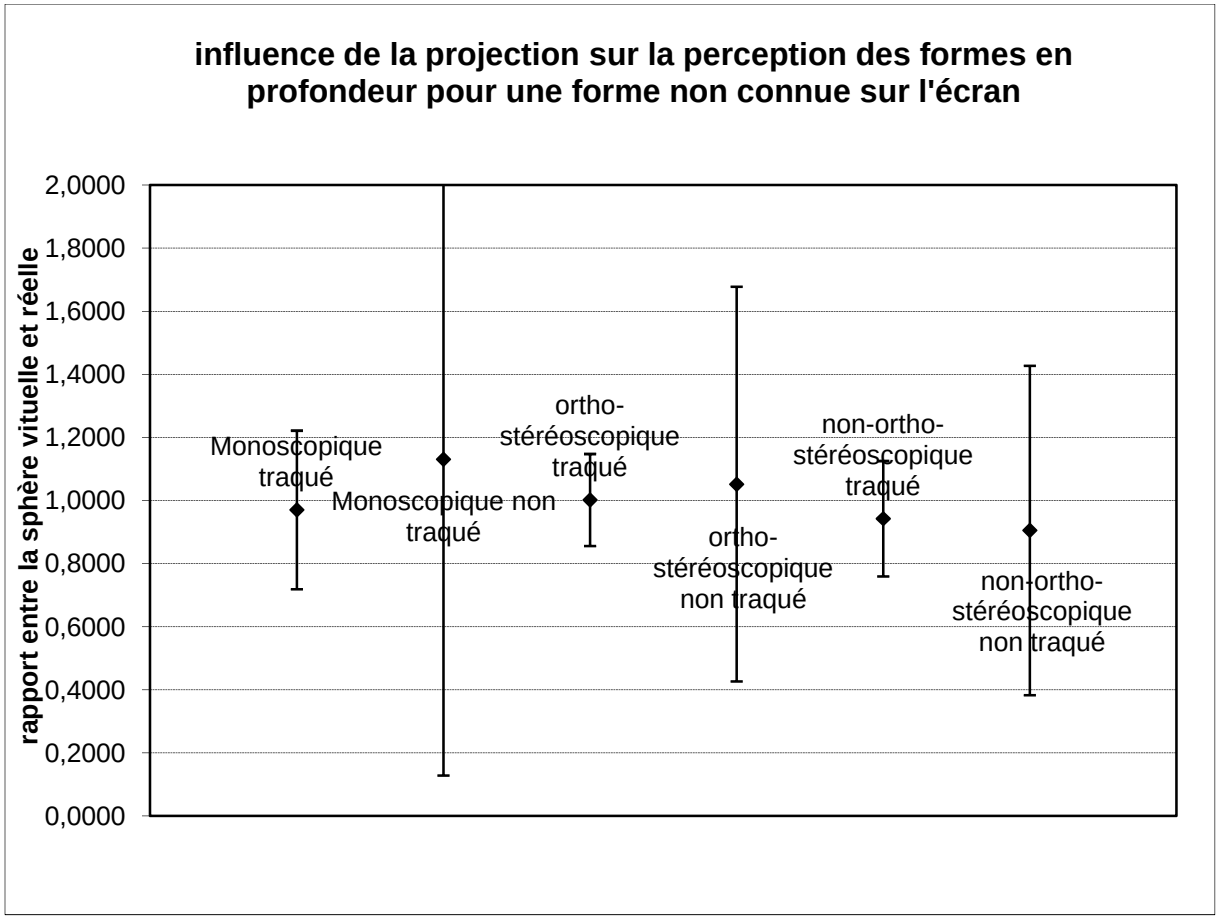


Figure 221 : influence de la projection sur la perception des formes en profondeur pour une forme non connue sur l'écran

L'influence de la projection sur la perception des formes en profondeur pour une forme non connue derrière l'écran est reportée sur la Figure 222.

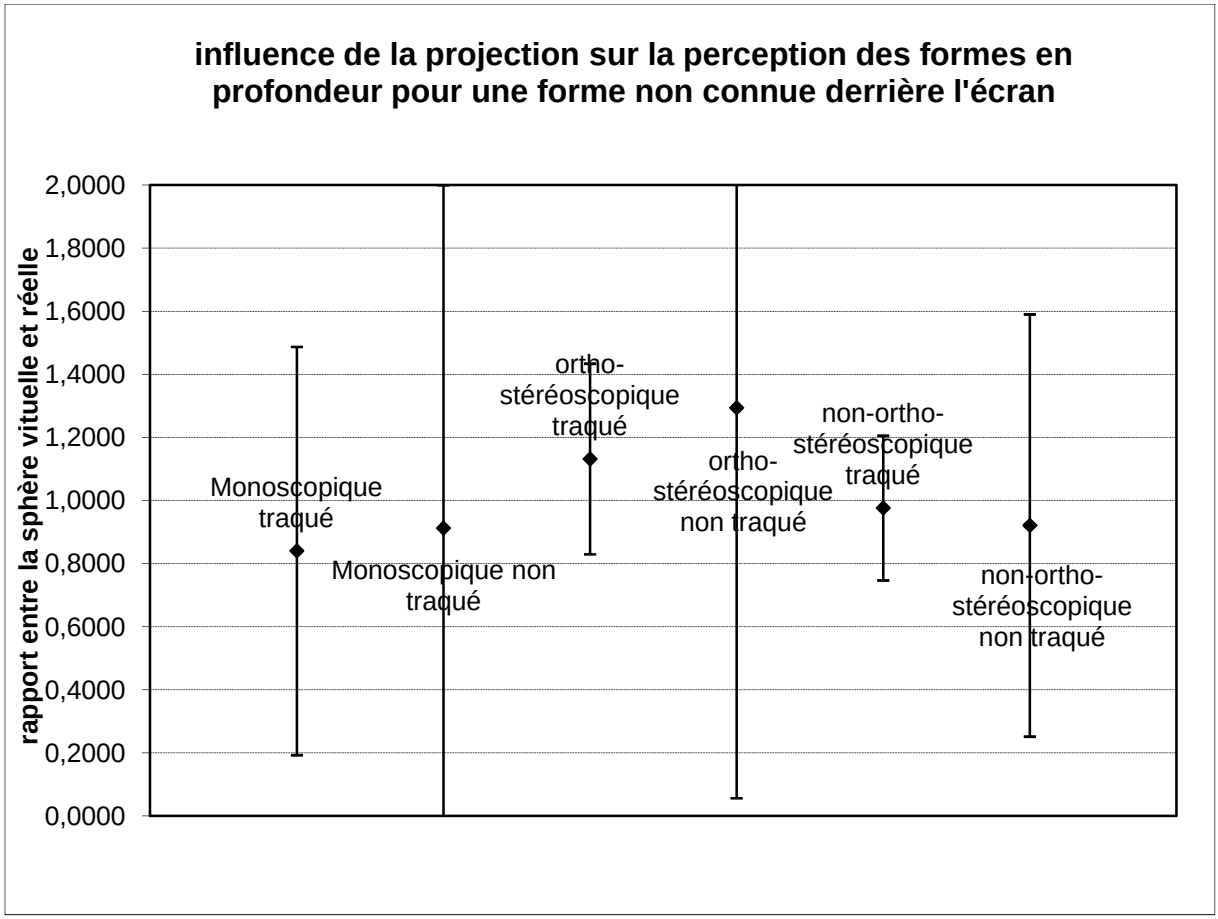


Figure 222 : influence de la projection sur la perception des formes en profondeur pour une forme non connue derrière l'écran

L'influence de la projection sur la perception des formes en profondeur pour une forme non connue devant l'écran est reportée sur la Figure 223.

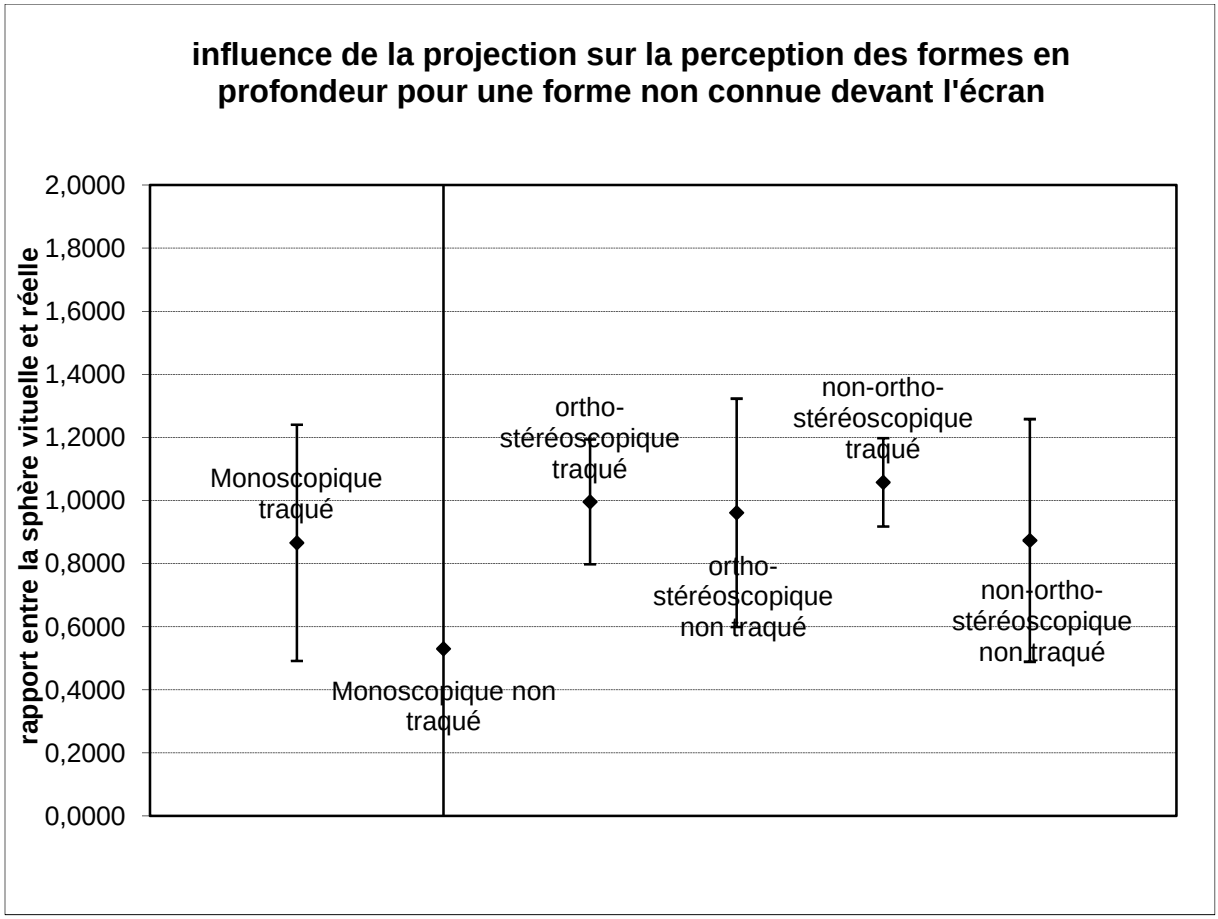


Figure 223 : influence de la projection sur la perception des formes en profondeur pour une forme non connue devant l'écran

L'influence de la projection sur la perception des formes fronto-parallèle pour une forme non connue sur l'écran est reportée sur la Figure 224.

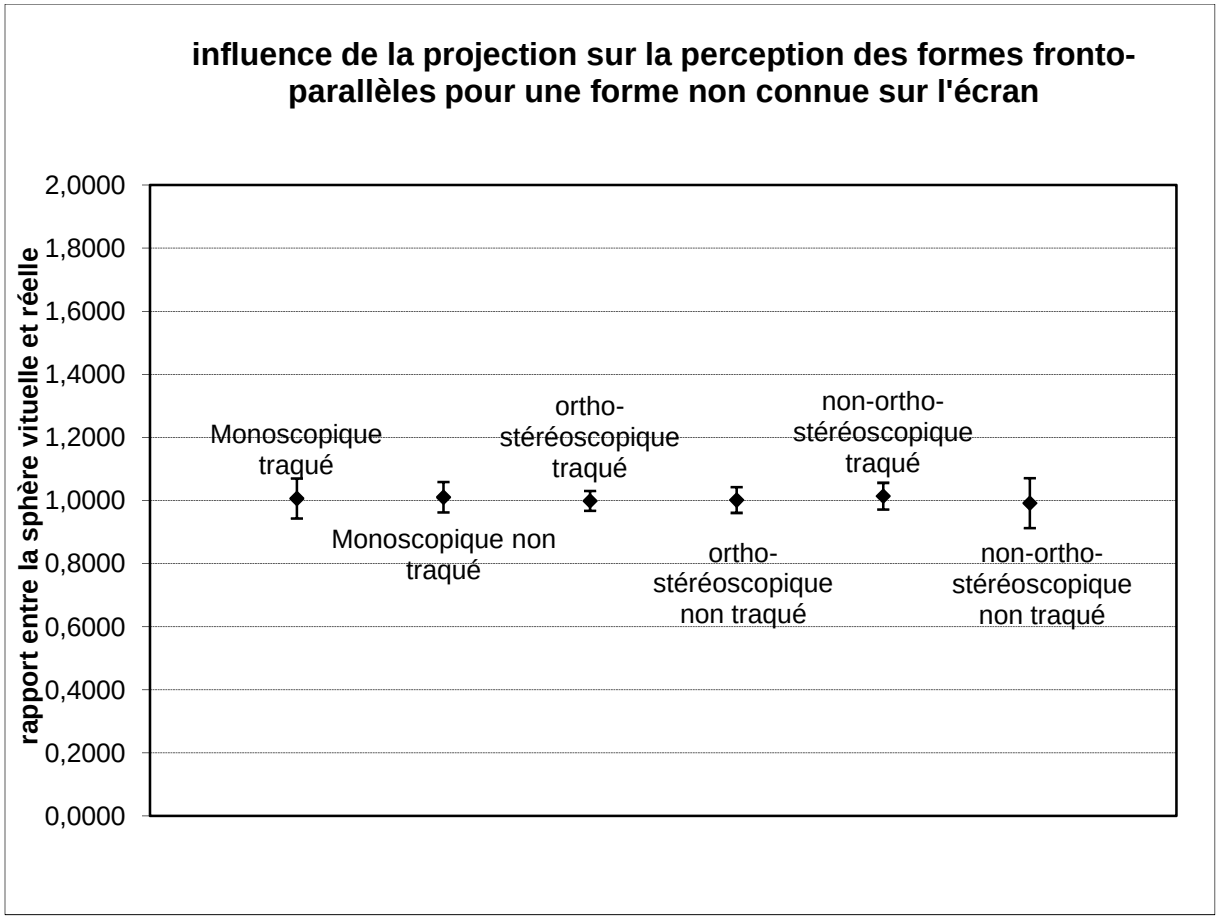


Figure 224 : influence de la projection sur la perception des formes fronto-parallèles pour une forme non connue sur l'écran

L'influence de la projection sur la perception des formes en profondeur pour une forme non connue derrière l'écran est reportée sur la Figure 225.

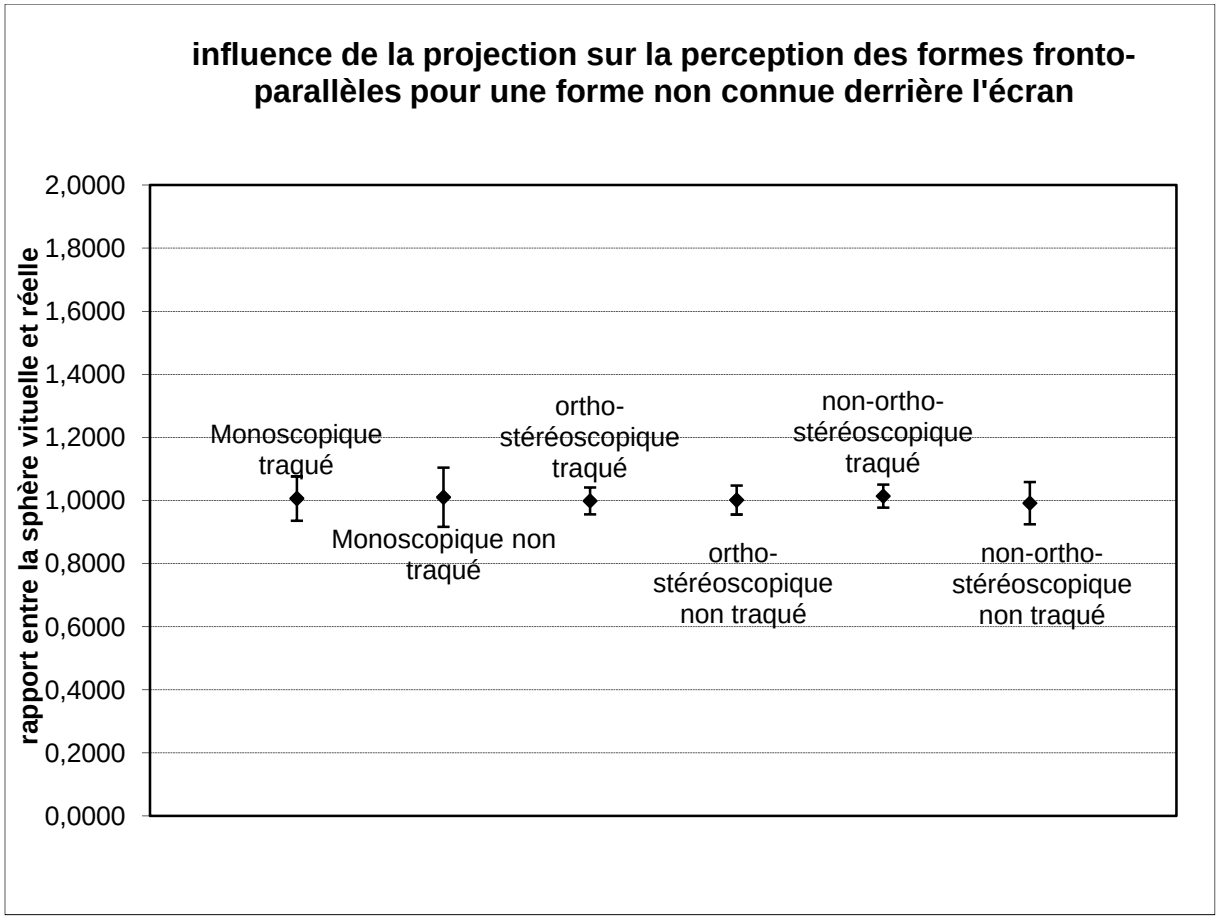


Figure 225 : influence de la projection sur la perception des formes fronto-parallèls pour une forme non connue derrière l'écran

L'influence de la projection sur la perception des formes fronto-parallèle pour une forme non connue devant l'écran est reportée sur la Figure 226.

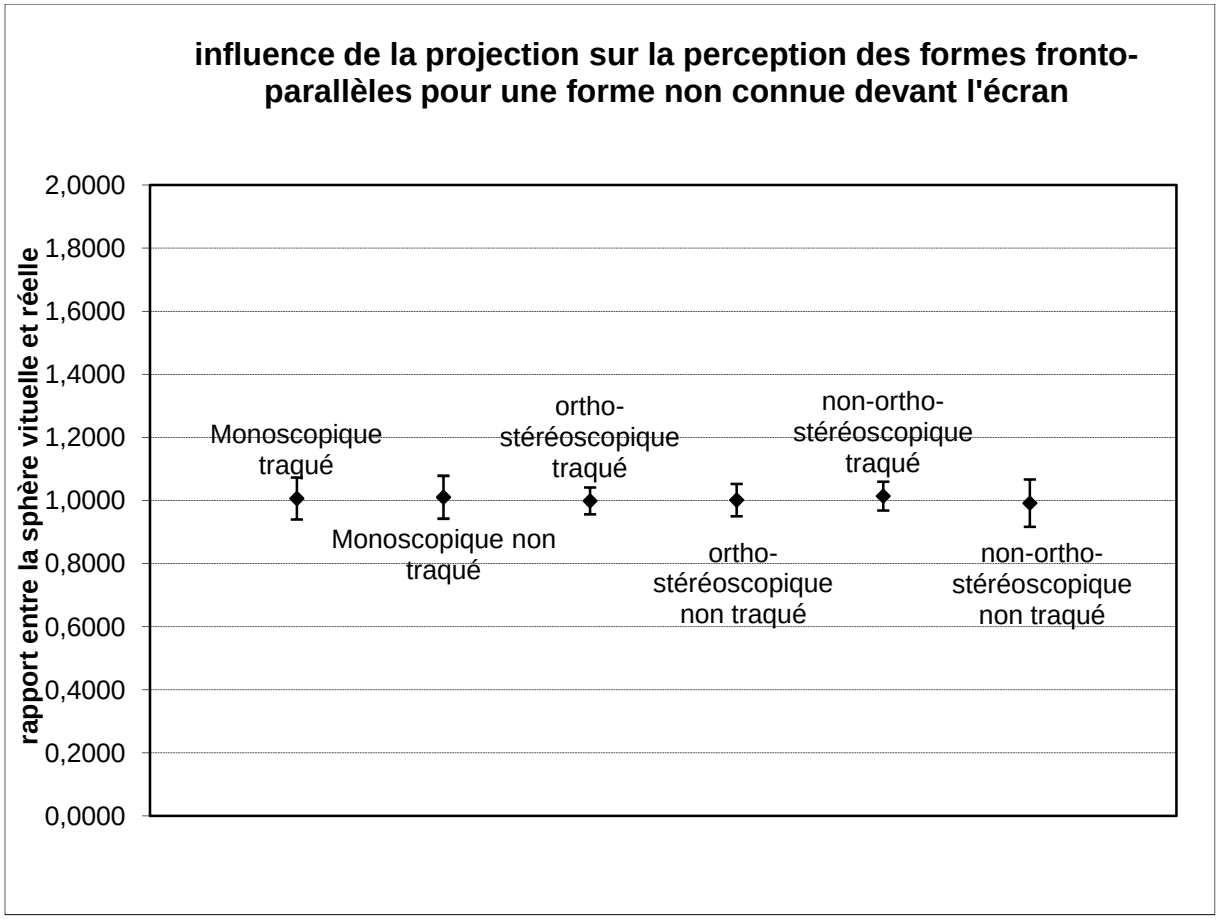


Figure 226 : influence de la projection sur la perception des formes fronto-parallèls pour une forme non connue devant l'écran

Annexe C : calcul de significativité pour l'amplitude accommodative

Nous avons donc des effectifs de populations :

$$n_1 = n_2 = 19$$

	Moyenne des différences [cm]	Ecarts-types [cm]
Sans traitement	1.210	1.127
Avec traitement	-0.07	1.159

Voyons si la différence d'écart type est significative :

$$T = \frac{\sigma_{\max}^2}{\sigma_{\min}^2} = \frac{1.159^2}{1.127^2} = 0.639$$

Ce qui nous donne, en regardant dans les tables de la loi normale, une probabilité associée de 73%, ces deux écarts-types peuvent être considéré comme égaux, puisqu'elle est inférieure à 95%

Calculons la différence entre les deux moyennes :

$$D = -1.289cm$$

Comme les échantillons suivent une loi normale (vérifié avec le test de Shapiro-Wilk), nous pouvons appliquer la formule de l'équation 3-8

$$T^* = \frac{D}{\sqrt{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}}} \approx N(0,1)$$

$$Abs(T^*)=3.1410$$

Ce qui nous donne, en regardant dans la table de la loi normale, une probabilité associée de 99.9%. Nous voyons donc que la différence de moyenne est significative puisque la probabilité qui lui est associée est supérieure à 95%.

Annexe D : Comparaison de plusieurs distributions

Comparaison de moyennes – L'ANOVA

Au paragraphe **Erreur ! Source du renvoi introuvable.**, nous avons étudié la comparaison statistique de deux populations. Nous pouvons généraliser cette étude à la comparaison de plusieurs (*K*) populations. Il faut cependant savoir qu'une ANOVA ne peut se calculer que sur des populations normales de même écart-type.

Nous souhaitons détecter des différences entre les moyennes $\mu_1, \mu_2, ..., \mu_K$ de plusieurs groupes $U_1, U_2, ..., U_K$.

Nous voulons donc tester l'hypothèse *Ho :* $\mu_1 = \mu_2 = ... = \mu_K$. Et l'hypothèse inverse : au moins deux moyennes sont différentes.

Nous définissons :

- T_k la somme des composantes $\{x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}\}$ de l'échantillon *k ;*
- n_k le nombre de composantes de l'échantillon *k ;*
- $\overline{x}_k$ la moyenne de l'échantillon *k* : $\left(\overline{x}_k = \dfrac{T_k}{n_k}\right)$;
- s_k^2 l'estimateur de la variance de l'échantillon *k* : $\left(s_k^2 = \dfrac{1}{n_k}\sum_i (x_i^{(k)} - \overline{x}_k)^2\right)$;
- T la somme de tous les T_k : $\left(T = \sum_k T_k\right)$;
- n la somme de tous les n_k : $\left(n = \sum_k n_k\right)$;
- $\overline{x}$ la moyenne des $\overline{x}_k$: $\left(\overline{x} = \dfrac{T}{n}\right)$;
- s_e^2 l'estimateur de la variance de toutes les observations : $\left(s^2 = \dfrac{1}{K}\sum_k (\overline{x}_k - \overline{x})^2\right)$.

Nous pouvons dès lors, calculer la somme des carrés totale (*SCT*) des écarts entre toutes les observations $x_i^{(k)}$ et la moyenne globale $\overline{x}$:

$$SCT = \sum_i \sum_k \left(x_i^{(k)} - \overline{x}\right)^2$$

0-1

Nous allons scinder l'expression 3-13 en deux expressions : l'une concernant la somme des carrés dans les groupes (*SCD*), l'autre concernant la somme des carrés entre les groupes (*SCE*)

$$SCT = \sum_i \sum_k \left(x_i^{(k)} - \bar{x}_k\right)^2 + \sum_k \left(\bar{x}_k - \bar{x}\right)^2$$

0-2

Nous posons :

$$\begin{cases} SCD = \sum_i \sum_k \left(x_i^{(k)} - \bar{x}_k\right)^2 \\ SCE = \sum_k \left(\bar{x}_k - \bar{x}\right)^2 \end{cases}$$

0-3

Donc :

$$\begin{cases} SCD = \sum_k n_k s_k^2 \\ SCE = \sum_k n_k s_e^2 \end{cases}$$

0-4

On sait que si Ho est exacte, σ^2 peut être approximé par

$$S_E^2 = \frac{SCE}{K-1}$$

Et

$$S_D^2 = \frac{SCD}{n-K}$$

0-5

Si *Ho* est vérifiée, leur rapport doit être proche de 1, dans le cas contraire, S_D^2 reste un estimateur de σ^2, mais ce n'est plus le cas de S_E^2. Dès lors, nous pouvons calculer :

$$\frac{S_E^2}{S_D^2} = \frac{\frac{SCE}{K-1}}{\frac{SCD}{n-K}} \sim F_{K-1,n-k}$$

0-6

Nous rejetons Ho si S_E^2 / S_D^2 est supérieur à $F_{K-1,n-k;1-\alpha}$.

Comparaison des variances

De façon courante, deux tests sont couramment employés :

- le test de Bartlett, qui nécessite des calculs plus compliqués mais peut s'appliquer à des populations qui ne contiennent pas le même nombre d'échantillons ;
- le test de Hartley, quant à lui, est plus rapide mais ne tolère pas de différence dans le nombre d'échantillons (Dagnelie, 1998)(Boyer, 2000).

L'hypothèse Ho, à rejeter ou confirmer, est :

$$\sigma_1 = \sigma_2 = ... = \sigma_K$$

Le test de Bartlett

Le test de Bartlett utilise la somme des carrés des écarts (SCE) et la variance estimée globale :

$$\hat{\sigma}^2 = \frac{SCE}{n-k}$$

0-7

Nous devons calculer :

$$\chi^2_{obs} = \frac{(n-k)\ln\hat{\sigma}^2 - \sum_{i=1}^{k}[(n_i - 1)\ln\hat{\sigma}^2]}{1 + \frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\frac{1}{n_i - 1} - \frac{1}{n-k}\right)} \sim \chi^2_{\text{k-1}}$$

0-8

Le rejet de l'hypothèse nulle (Ho) revient à dire :

$$\chi^2_{obs} > \chi^2_{k-1,1-\alpha}$$

0-9

Le test de Hartley

Le test de Hartley ne s'utilise que lorsque les distributions ont le même nombre d'échantillons. Il est basé sur la comparaison des deux variances extrêmes :

$$H_{obs} = \frac{\sigma^2_{\max}}{\sigma^2_{\min}}$$

0-10

à comparer aux valeurs critiques particulières de la table de Hartley $H_{1-\alpha}$

Annexe E : Lien entre Flou et hautes fréquences

La technique de moyenne glissante rend nos images « floues », c'est entendu, mais nous pourrions nous demander si le fait de rendre floue une image correspond réellement à la suppression des hautes fréquences. Au premier abord, nous serions tentés de répondre « oui », mais nous allons vérifier cette réponse intuitive. Pour ce faire, passons les images données par nos algorithmes des moyennes glissantes dans des transformées de Fourier.

Seuillons la transformée de Fourier de l'image de départ pour mieux mettre en évidence les fréquences spatiales présentes. Sur la Figure 227 nous voyons qu'il y a beaucoup de hautes fréquences dans notre image initiale.

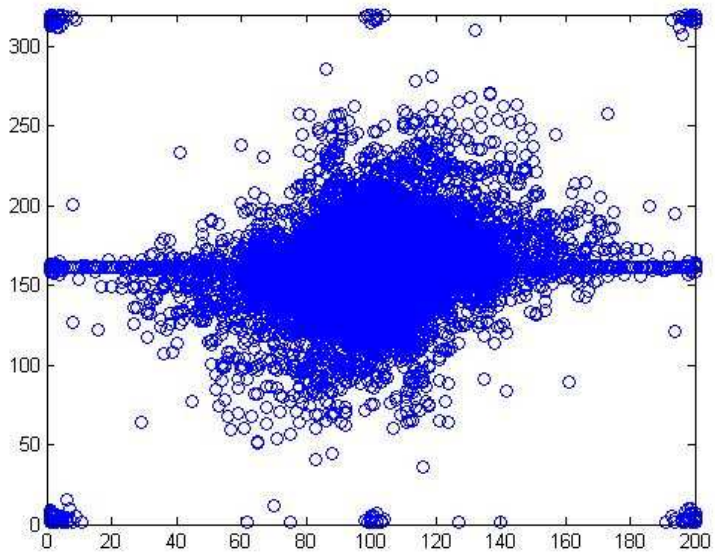


Figure 227 : seuillage de la transformée de Fourier de la figure de départ

Voyons à présent le même seuillage sur la transformée de Fourier de l'image obtenue par la moyenne glissante de notre Clown sur 2 pixels (Figure 228). Nous remarquons que rien que le fait de faire cette moyenne glissante sur de tout petits carré enlève déjà beaucoup de hautes fréquences.

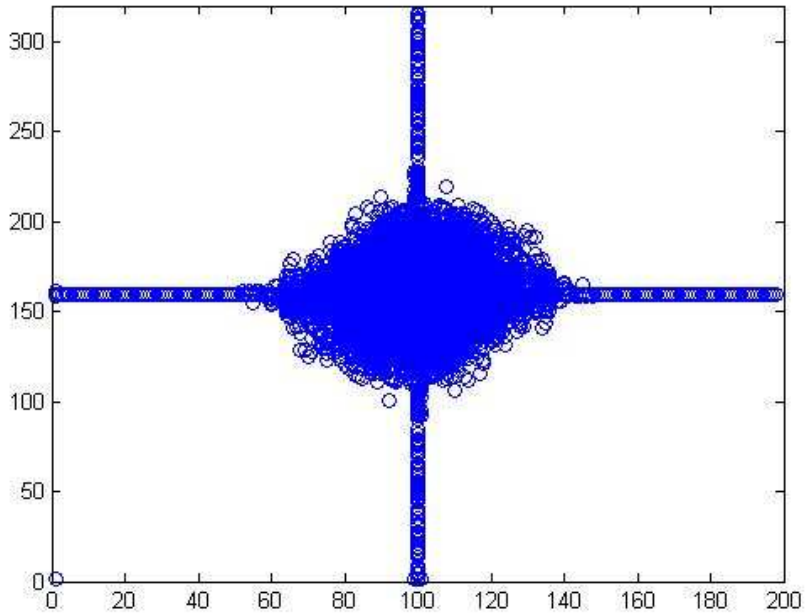


Figure 228 : seuillage de la transformée de Fourier de la moyenne glissante sur des carrés de 2 pixels de côté

Nous pouvons, à titre d'exemple montrer le même seuillage pour des carrés de 4, 8, 16 et 32 pixels de côté sur la Figure 229

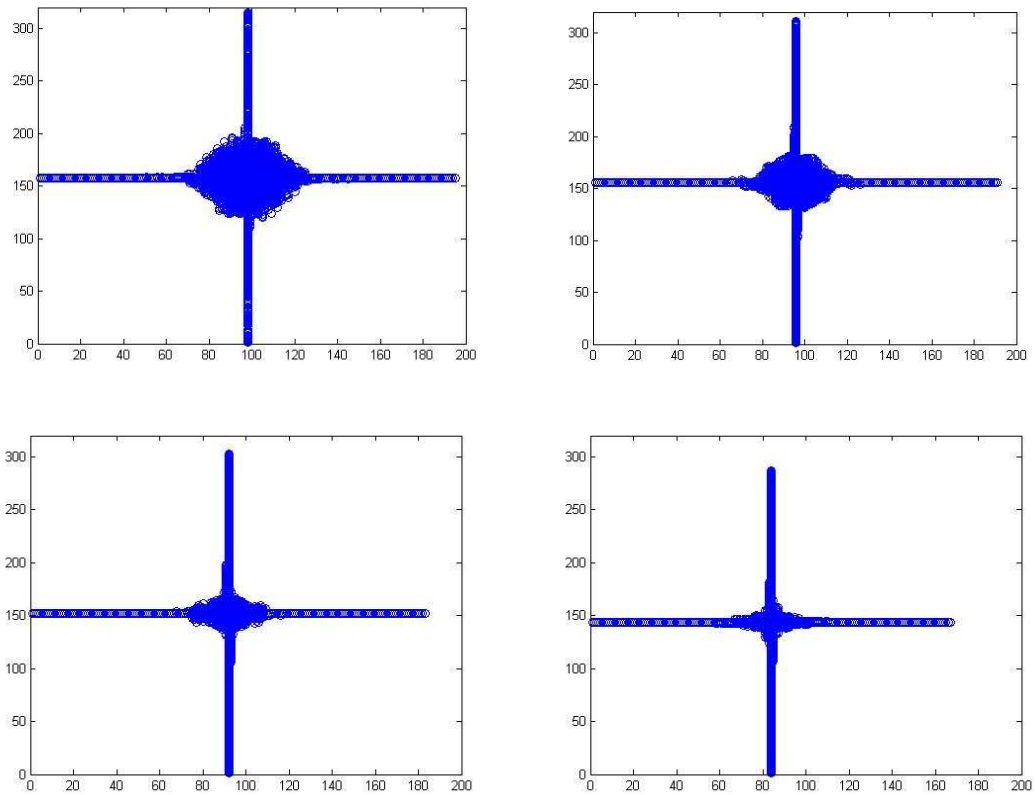


Figure 229 : seuillage de la décomposition de Fourier pour les moyennes glissantes sur 4, 8, 16 et 32 pixels (de droite à gauche et de haut en bas)

Nous voyons donc bien que plus la moyenne glissante est effectuée sur un grand nombre de pixels, plus les hautes fréquences disparaissent. Pour chacune de ces moyennes glissantes, nous pouvons chiffrer la plus haute fréquence présente. Pour caractériser cette fréquence, nous n'utiliserons pas les cpd (cycle par degré de vision oculaire) puisque nous ne savons pas à quelle distance se trouve l'œil, mais nous pouvons les comparer à la période (exprimée en pixels) de la sinusoïde correspondante.

Nombre de pixels pour la moyenne glissante	Période des oscillations restantes en pixels en vertical	Période des oscillations restantes en pixels en horizontal
2	6.5	3.5
4	7	7.5
8	14	8
16	17	14
32	24	17.5

Tableau 4 : correspondance entre le nombre de périodes des oscillations restantes et le nombre de pixels pour la moyenne glissante

Résumé :

Les dispositifs d'immersion stéréoscopique sont très répandus et tendent à l'être de plus en plus (travail en immersion, film projeté en vision stéréoscopique, écrans auto-stéréoscopiques…). Dans certains cas, les immersions peuvent être longues et la fatigue visuelle s'installe. Nous donnons une méthode pour diminuer la fatigue visuelle liée à la vision stéréoscopique. Certaines applications demandent également une perception des formes très performante. Nous déterminons donc les paramètres améliorant cette perception et nous les quantifions. Ce mémoire de thèse est composé de deux parties : l'une consacrée à la perception des formes en immersion, l'autre traite de la diminution de la fatigue visuelle en vision stéréoscopique.

Dans la première partie, nous commençons par expliquer comment nous percevons les formes en visions naturelle et artificielle. Ensuite nous exposons un protocole de test pour connaître quels paramètres peuvent influencer cette perception en immersion. Nous nous interrogeons notamment sur l'amélioration que peut apporter le fait de traquer le point de vue de l'utilisateur en vision monoscopique, ortho-stéréoscopique ou non-ortho-stéréoscopique. Nous évaluons également l'influence de la vision ortho-stéréoscopique et non-ortho-stéréoscopique sur la perception des formes Nous quantifions les différences entre tous ces modes de projections et déterminons quels paramètres sont les plus importants à respecter pour une bonne perception des formes.

Dans la seconde partie, nous exposons les causes possibles de fatigue visuelle en vision stéréoscopique. Nous nous concentrons ensuite sur une de ces causes : les hautes fréquences spatiales liées à de grandes disparités. Nous avons développé plusieurs algorithmes traitant en temps réel les images stéréo pour supprimer les hautes fréquences gênantes (transformée en ondelette, box filter et programmation sur carte graphique d'une moyenne glissante) avant d'expliquer le choix de l'algorithme. Nous exposons ensuite notre protocole de test pour montrer que notre traitement diminue de façon objective et subjective la fatigue visuelle en vision ortho-stéréoscopique. Nous quantifions ensuite les améliorations apportées par notre algorithme sur cette fatigue.

Mots clef : Perception des formes, ortho-stéréoscopie, flou adaptatif, psychophysique, fatigue visuelle, point de vue traqué, réalité virtuelle