

# Vers des mesures subtiles de performance en Réalité Virtuelle : l'exemple d'une tâche de mobilité

Yujie Huang<sup>1</sup>

Alexandre Bruckert<sup>1</sup>

Patrick Le Callet<sup>1,2</sup>

<sup>1</sup> Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, F-44000 Nantes

<sup>2</sup> Institut universitaire de France (IUF)

{Yujie.Huang}@etu.univ-nantes.fr

{Alexandre.Bruckert, Patrick.Lecallet}@univ-nantes.fr

## Résumé

*De nombreuses pathologies de la vision peuvent significativement affecter la qualité de vie (Quality of Life, QoL) de patients, de part leurs impacts sur la façon d'effectuer diverses tâches visuelles. Il est important alors de distinguer les notions de fonction visuelle, c'est-à-dire les capacités ophtalmiques du système visuel (acuité visuelle, vision des couleurs, sensibilité au contraste), et de vision fonctionnelle, faisant référence à la capacité d'accomplir différentes tâches visuelles (lecture, navigation, reconnaissance d'objets, etc). Pour la première, il existe de nombreux outils d'évaluation quantitatifs standardisés en milieu clinique. Cependant, les outils d'évaluation actuels de la vision fonctionnelle sont soit trop subjectifs, soit non standardisés. Plusieurs protocoles d'évaluation de la vision fonctionnelle fondés sur une évaluation de l'orientation et de la mobilité (O&M) en environnement virtuel ont cependant récemment été proposés. Dans cette étude, nous proposons une étude de données issues d'un tel test, afin de mettre en évidence les avantages et inconvénients des protocoles existants. Nous constatons notamment des niveaux de difficulté variables dans différentes configurations environnementales, menant à un biais potentiel dans l'interprétation des résultats de ces tests. Nous avons également identifié des causes potentielles d'erreurs. En combinant ces résultats à une analyse de sujets sains et de patients utilisant des métriques précédentes, nous montrons que se fier uniquement aux indicateurs statistiques est insuffisant, soulignant le besoin de nouvelles méthodologies de traitement de ces données.*

## Mots clefs

Vision fonctionnelle, Réalité virtuelle, Qualité de vie, Orientation et mobilité, Modélisation.

## 1 Introduction

Les déficiences visuelles représentent un problème de santé publique majeur qui affecte la qualité de vie d'individus de tous âges [1]. En France, on estime que le nombre de personnes souffrant de déficiences visuelles dépasse 2% de la

population totale [2], ce qui pose de nombreux défis en matière de soins médicaux et sociaux. Face à cette réalité, des investissements importants sont réalisés pour développer des dispositifs de remédiation, des prothèses, ou encore de nouvelles thérapies géniques prometteuses [3].

Cependant, un frein à l'adoption de telles thérapies est le manque d'outils pour mesurer objectivement la vision fonctionnelle. En effet, il est important de séparer les notions de fonction visuelle et de vision fonctionnelle. La fonction visuelle est liée à "la performance des yeux" [4], telle que l'acuité visuelle ou la sensibilité au contraste, et de nombreux outils de mesure standardisés existent. La vision fonctionnelle quant à elle est liée à "la performance de la personne dans les tâches visuelles" [4]. Les méthodes de mesure pour évaluer cette vision fonctionnelle présentent cependant certaines lacunes importantes : les questionnaires manquent d'objectivité [5] ; le test physique d'orientation et de mobilité (O&M) manque de reproductibilité [1], est coûteux et chronophage ; le test d'O&M basé sur la réalité virtuelle actuel nécessite un déplacement physique du patient [6, 7]. Nous avons donc proposé un test O&M basé sur la réalité virtuelle [8], qui présente trois avantages principaux par rapport aux tests précédents : la facilité de configuration de l'environnement, une sécurité accrue pour les participants et l'accès à plus de données comportementales pendant le test.

Dans cette étude, nous utilisons les données de sujets sains et de patients malvoyants recueillies lors de ce test. Sur la base des données des sujets sains, nous avons tout d'abord analysé la similarité des différents labyrinthes de test. Nous montrons que des différences de difficultés peuvent exister entre les différentes configurations de parcours. Nous avons ensuite étudié les caractéristiques des erreurs commises par les participants et avons constaté que la plupart des objets manqués avaient les valeurs RGB les plus basses. En examinant le nombre de points de fixation oculaires sur chaque objet manqué pendant le test, nous avons déduit l'existence de trois types d'erreurs différents. En ajoutant les données des patients, nous avons démontré que l'utilisation des seules métriques de temps et de nombre d'erreurs sur le parcours n'est pas suffisante pour évaluer la

vision fonctionnelle de manière précise. Il est nécessaire de combiner ces indicateurs avec des données issues du comportement des participants pour fournir une évaluation robuste.

## 2 Travaux connexes

La vision fonctionnelle est traditionnellement mesurée à l'aide de méthodes subjectives et objectives. Les méthodes subjectives reposent largement sur des auto-évaluations par les patients à travers des questionnaires ou des échelles de notation [5]. Ces méthodes sont faciles à mettre en œuvre mais ne produisent que des scores subjectifs, qui ne sont pas toujours fiables, et peuvent manquer de détails. Les méthodes objectives sont principalement basées sur le test d'Orientation et de Mobilité (O&M) [6]. Le test O&M le plus couramment utilisé aujourd'hui est le Multi-Luminance Mobility Test (MLMT) [1]. Pour représenter la performance du sujet, la durée du test et le nombre de collisions pendant le parcours sont enregistrés. Cependant, ces parcours de mobilité physique présentent plusieurs limitations en raison des conditions environnementales et de la difficulté de reproductibilité [1].

La réalité virtuelle permet quant à elle de concevoir des environnements divers et facilement modifiables à faible coût. Ainsi, sur la base du MLMT, une version basée sur la réalité virtuelle VR-O&M a été proposée et validée en utilisant le même système de notation que le MLMT [6]. Malgré l'avantage de la reproductibilité, cette méthode nécessite encore que les participants se déplacent physiquement. De plus, les systèmes de notation des méthodes actuelles sont basés sur des métriques simples, principalement la durée du test et le nombre d'erreurs. Par conséquent, il reste nécessaire d'évaluer ces outils ainsi que leurs limites, afin de pouvoir proposer un ensemble de méthodes d'analyse de données adapté.

Pour surmonter les lacunes des méthodes existantes, nous avons conçu un nouveau test O&M basé sur la réalité virtuelle [8]. Le test comprend huit labyrinthes différents, chacun ayant la même longueur de parcours et le même nombre d'objets (deux pour l'entraînement et six pour les tests). Un exemple est illustré dans la Figure 1. Lors de ce test, les participants restent assis pour garantir leur sécurité. Il leur est donné la consigne d'atteindre la sortie du labyrinthe, tout en détruisant les objets qu'ils détectent, en les touchant pendant 2 secondes. Dans notre test, le fait de ne pas détruire un objet est considéré comme une erreur. En utilisant cet environnement, nous avons collecté des données auprès de sujets sains et de patients (42 sujets sains, 9 patients). Sur la base de cette étude, nous avons mené plusieurs analyses préliminaires.

Tous les tests précédents d'orientation et de mobilité (O&M) ont utilisé des configurations de parcours différentes pour chaque test, ce qui est également le cas dans notre étude. Pour le test MLMT, ainsi que ses équivalents en VR actuels, le système de notation est défini comme :

$$TimeScore = t_{duration} + t_{penalties} \quad (1)$$



FIGURE 1 – Structure d'un labyrinthe

$$AccuracyScore = \frac{N_{penalties}}{N_{obstacles}} \quad (2)$$

où  $t_{penalties}$  était fixé arbitrairement à 15 secondes par erreur simple, c'est-à-dire par collision, déviation du parcours et erreurs de contournement ; et à 30 secondes pour les erreurs de redirection. Les scores dans chaque configuration ont ensuite été comparés empiriquement pour s'assurer qu'aucune n'était particulièrement facile ou difficile. Cependant, dans le test VR-O&M, la difficulté de chaque configuration n'a pas été vérifiée rigoureusement, ce qui pourrait entraîner des résultats d'analyse biaisés.

## 3 Analyse et résultat

Afin de garantir une comparabilité entre les différents participants et les sessions de tests, ainsi que pour éliminer l'effet de mémoire, il est important que les labyrinthes conçus soient de difficulté équivalente. Nous avons donc analysé et vérifié cette équivalence à partir des données préliminaires, en utilisant les données des sujets sains, plus nombreuses que celles des patients. Sur 479 sessions valides, chaque labyrinthe comptait en moyenne  $80 \pm 1.5$  essais.

Nous utilisons deux métriques différentes pour cette analyse. En effet, le protocole de collecte étant différent par rapport à l'approche MLMT [6], nous ne pouvons pas directement appliquer le système de score utilisé lors de cette étude 1, 2. De plus, ces scores reposent sur une attribution de pénalité de temps arbitraire, peu indiquée dans l'objectif d'une méthode répliquable d'analyse. Nous nous concentrerons donc sur les variables brutes  $t_{duration}$  et  $n_{error}$ . Comme illustré dans sur la Figure 2a, aucune différence significative de durée de parcours n'a été trouvée (Kruskal-Wallis test :  $P=0.771$ ), indiquant une similarité entre toutes les configurations de labyrinthes. Ces résultats étaient attendus, étant donné la contrainte de similarité de longueur lors de la création des différents parcours.

Une autre variable que nous utilisons est le nombre d'objets manqués. Comme le montre la Figure 2b, une différence significative a été observée en termes de nombre d'objets manqués (Kruskal-Wallis test :  $P \ll 0.001$ ). Nous avons ensuite réalisé un test par paires de Dunn, dont les résultats indiquent que les participants ont commis significativement moins d'erreurs dans les Labyrinthes A et B par rapport aux Labyrinthes C, D, E et F. Cela démontre que malgré le maintien du même nombre d'objets, certaines configurations affectent significativement la difficulté de réussir à

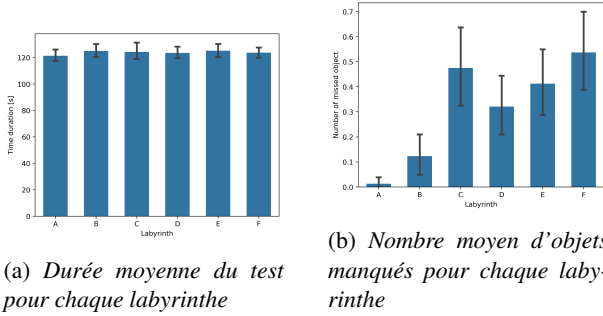


FIGURE 2 – Scores des participants sains sur les différents labyrinthes

détecter puis détruire chaque objet.

Dans les tests précédents d'orientation et de mobilité (O&M),  $n_{error}$  est une métrique utilisée pour évaluer la performance. Cependant, nous avons constaté que parmi les sujets sains, il existe une différence significative dans le nombre d'erreurs à travers différents labyrinthes, qui n'est donc a priori pas due à une baisse de la vision fonctionnelle. Il est donc important de comprendre la cause de ces erreurs. Nous avons extrait certaines caractéristiques de tous les objets manqués, y compris les valeurs RGB, les niveaux de luminosité et le nombre de regards portés sur ces objets pendant le test. Comme le montre la Figure 3a, la plupart (87%) des objets manqués ont la plus basse valeur RGB. De plus, comme indiqué dans la Figure 3b, la plupart (91%) des objets manqués ont été regardés moins de 8 fois. Il semble alors que, pour les sujets sains, les objets plus sombres sont plus susceptibles d'être manqués. Il est à noter que certains objets ont tout de même été manqués malgré un nombre élevé de points de fixations. Nous supposons que ces cas représentent un autre type d'erreur, probablement causé par une mauvaise utilisation de l'environnement de test. De plus, les objets avec des fixations autour de 10 fois pourraient indiquer ceux qui sont intrinsèquement difficiles à distinguer du fond.

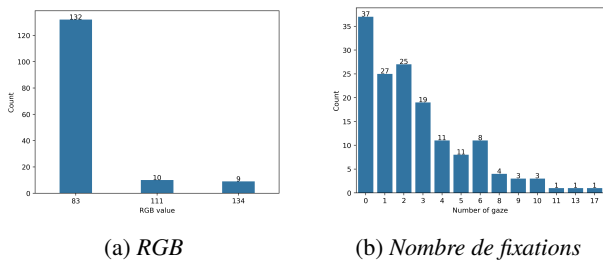
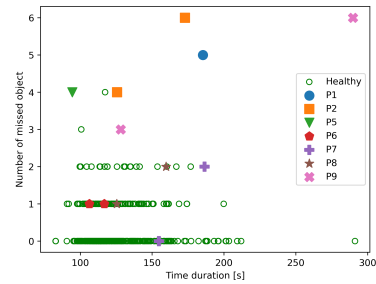


FIGURE 3 – Nombre d'erreurs selon différentes caractéristiques

Il semble alors nécessaire de distinguer les types d'erreurs. Les erreurs dues à l'inattention et à une mauvaise utilisation de l'environnement de test ne sont pas liées à la vision fonctionnelle, et doivent être exclues dans son évaluation. En revanche, les erreurs dues à la difficulté peuvent

être utilisées pour évaluer la gravité de l'état du patient. Comme mentionné précédemment, les différentes configurations de nos six labyrinthes de test entraînent une différence significative dans le nombre d'erreurs. Par conséquent, identifier les configurations qui entraînent des types spécifiques d'erreurs sera notre direction de recherche future. Une fois que nous aurons compris quelles configurations provoquent quelles erreurs, nous pourrions modifier la difficulté de notre environnement de test en manipulant ces configurations. Cela aidera à réduire l'occurrence d'erreurs non pertinentes et à améliorer le pouvoir discriminant de l'environnement de test en ajustant la difficulté du labyrinthe.



leurs RGB les plus basses. En examinant le nombre de regards portés sur chaque objet manqué pendant le test, nous avons identifié trois types différents d'erreurs : manquement d'objet par négligence, par mauvaise utilisation de l'environnement de test, et en raison de la difficulté d'observation. Nous avons illustré l'importance de distinguer les types d'erreurs afin d'éliminer le biais des erreurs non liées à la vision.

En intégrant les données des patients malvoyants, nous avons démontré que l'utilisation uniquement des métriques de temps et de précision, comme dans les tests précédents, est insuffisante pour une évaluation nuancée de la vision fonctionnelle. Des métriques d'évaluation plus robustes devraient être proposées.

Étant donné la richesse de nos données comportementales, les analyses futures se concentreront sur l'explication des différents comportements et interactions dans l'environnement virtuel. Deux approches peuvent être envisagées. Les modèles stochastiques ont le potentiel d'extraire des caractéristiques ou de découvrir des motifs sous-jacents. Par exemple, Shaily et al. [9] ont proposé l'utilisation de Modèles de Markov Cachés (HMMs) pour les données séquentielles dans la reconnaissance des activités humaines, tandis que Ben-Gal et al. [10] ont employé des chaînes de Markov pour la modélisation de la mobilité. Les méthodes basées sur les graphes sont également efficaces pour la représentation des données. Rossi et al. [11] ont utilisé une approche basée sur les graphes pour regrouper les sujets ayant des patterns de navigation similaires en réalité virtuelle. Avec ces diverses stratégies de modélisation, des méthodes d'apprentissage associées peuvent être explorées plus en profondeur.

À l'avenir, les tests d'orientation et de mobilité basés sur la réalité virtuelle pourraient ne pas se limiter aux tâches traditionnelles de "navigation". Les chercheurs pourraient explorer le potentiel d'utiliser des tâches plus immersives et liées au mode de vie, telles que la cuisine ou la conduite, pour tirer pleinement parti des capacités de la réalité virtuelle.

## Références

- [1] Daniel C Chung, Sarah McCague, Zi-Fan Yu, Satha Thill, Julie DiStefano-Pappas, Jean Bennett, Dominique Cross, Kathleen Marshall, Jennifer Wellman, et Katherine A High. Novel mobility test to assess functional vision in patients with inherited retinal dystrophies. *Clinical & experimental ophthalmology*, 46(3) :247–259, 2018.
- [2] Seth R Flaxman, Rupert RA Bourne, Serge Resnikoff, Peter Ackland, Tasanee Braithwaite, Maria V Cicinelli, Aditi Das, Jost B Jonas, Jill Keeffe, John H Kempen, et al. Global causes of blindness and distance vision impairment 1990–2020 : a systematic review and meta-analysis. *The Lancet Global Health*, 5(12) :e1221–e1234, 2017.
- [3] Albert M Maguire, Katherine A High, Alberto Auricchio, J Fraser Wright, Eric A Pierce, Francesco Testa, Federico Mingozzi, Jeannette L Bennicelli, Guishuang Ying, Settimio Rossi, et al. Age-dependent effects of rpe65 gene therapy for leber's congenital amaurosis : a phase 1 dose-escalation trial. *The Lancet*, 374(9701) :1597–1605, 2009.
- [4] August Colenbrander. Assessment of functional vision and its rehabilitation. *Acta ophthalmologica*, 88(2) :163–173, 2010.
- [5] Vijaya K Gothwal, Jan E Lovie-Kitchin, et Rishita Nutheti. The development of the lv prasad-functional vision questionnaire : a measure of functional vision performance of visually impaired children. *Investigative ophthalmology & visual science*, 44(9) :4131–4139, 2003.
- [6] Tomas S Aleman, Alexander J Miller, Katherine H Maguire, Elena M Aleman, Leona W Serrano, Keli B O'Connor, Emma C Bedoukian, Bart P Leroy, Albert M Maguire, et Jean Bennett. A virtual reality orientation and mobility test for inherited retinal degenerations : testing a proof-of-concept after gene therapy. *Clinical Ophthalmology*, pages 939–952, 2021.
- [7] Jean Bennett, Elena M Aleman, Katherine H Maguire, Jennifer Nadelmann, Mariejel L Weber, William M Maguire, Ayodele Maja, Erin C O'Neil, Albert M Maguire, Alexander J Miller, et al. Optimization and validation of a virtual reality orientation and mobility test for inherited retinal degenerations. *Translational Vision Science & Technology*, 12(1) :28–28, 2023.
- [8] Audrey Crozet, Lucas Communier, Toinon Vigier, Pierre Lebranchu, et Patrick Le Callet. A virtual mobility test to evaluate functional vision of visual impaired patients. Dans *IMXw'23 : ACM International Conference on Interactive Media Experiences Workshops*. ACM, 2023.
- [9] Shagun Shaily et Veenu Mangat. The hidden markov model and its application to human activity recognition. Dans *2015 2nd International Conference on Recent Advances in Engineering & Computational Sciences (RAECS)*, pages 1–4. IEEE, 2015.
- [10] Irad Ben-Gal, Shahar Weinstock, Gonen Singer, et Nicholas Bambos. Clustering users by their mobility behavioral patterns. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 13(4) :1–28, 2019.
- [11] Silvia Rossi, Francesca De Simone, Pascal Frossard, et Laura Toni. Spherical clustering of users navigating 360 content. Dans *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4020–4024. IEEE, 2019.