

# Outil de saisie par accélération en réalité virtuelle

Damien Trouchda  
Hajbi El Hassan  
El Acham Ahmad  
Eleftheriou Theophanis

January 18, 2024

## 1 Contexte et Objectifs (centré humain)

À mesure que les technologies de réalité virtuelle (RV) et augmentée (RA) continuent d'évoluer, le développement de méthodes de saisie de texte efficaces et intuitives devient cruciale. Les techniques de saisie traditionnelles, telles que les claviers projetés avec saisie par pénétration, sont souvent confrontées à des défis liés à la perception de la profondeur et à la précision, entravant l'expérience globale de l'utilisateur. En réponse à ces défis, notre rapport explore une approche novatrice de la saisie de texte en exploitant le système OptiTrack pour la capture des mouvements du doigt.

Les claviers projetés traditionnels dans les environnements RV/RA reposent fréquemment sur l'estimation visuelle de la profondeur nécessaire pour déclencher une pression de touche. Cette méthode peut entraîner des incohérences, des frustrations utilisateur et une expérience de saisie de texte sous-optimale. En revanche, la méthodologie proposée utilise la technologie OptiTrack pour capturer les mouvements précis du doigt, ce qui permet de détecter un geste cible de saisie, offrant ainsi un moyen potentiellement plus précis et rapide d'interagir avec des claviers virtuels.

Afin de mettre en lumière son potentiel pour surmonter les limitations des méthodes traditionnelles de saisie de texte, nous explorons dans ce rapport la détection d'un geste cible avec l'OptiTrack. Ainsi, en éliminant la nécessité d'évaluer la profondeur avant toute interaction avec les touches, nous anticipons une amélioration substantielle de la rapidité de l'entrée, de la satisfaction de l'utilisateur et de l'efficacité globale de la saisie de texte en RV/RA.



Figure 1: Saisie de texte en utilisant la perception de profondeur

## 2 Solution envisagée (centré humain)

Pour résoudre le problème d'ergonomie des claviers classiques, notre solution propose donc l'utilisation d'un geste-cible, consistant en un aller-retour rapide de l'index de l'utilisateur. Ce geste est détecté par l'OptiTrack, déclenchant l'action de la touche du clavier alignée avec le doigt de l'utilisateur. Cette approche vise à rendre la saisie plus précise, rapide et ergonomique, en évitant la totalité des problèmes liés à la profondeur de champ ainsi que certains des problèmes liés au calibrage des éléments de la scène.

Au cours de ce projet, nous nous sommes concentrés sur la faisabilité technique de cette solution, sans s'occuper d'une implémentation graphique. Le travail a donc consisté en un traitement des données issu de l'Optitrack, sans utilisation de Unity ou de l'Hololens.

## 3 Problèmes rencontrés et Solutions Mises en Place

### 3.1 Difficulté d'acquisition de l'instant de "Picking"

**Problème :** La difficulté de déterminer avec précision le moment du déclenchement du picking.

**Solution :** Modification du dispositif expérimental en intégrant un clavier physique. Avant d'effectuer le mouvement de picking, les participants doivent interagir avec le clavier physique. Cela permet d'établir clairement le moment du picking, créant ainsi une référence temporelle fiable pour l'analyse des données.

### 3.2 Faible volume d'acquisition

**Problème :** Du fait du petit nombre de caméra, du manque de matériel et de l'environnement de travail, le volume de travail était restreint. Ainsi, on obtient une forte baisse de qualité sur les résultats notamment sur l'évaluation du "touché de Midas".

**Solution :** Malheureusement pour ce second problème, peu de solutions ont pu être mise en oeuvre. On a maximisé la longueur du viseur ainsi que la distance du doigt à celui ci et augmenté la volume tant que possible avec le peu de points d'accroche disponibles pour les caméras.

## 4 État des lieux de ce qui a été réalisé (centré technique)

### 4.1 Méthodologie

Pour effectuer nos expérimentations, nous avons utilisé des données du flux optitrack avec une fréquence d'acquisition de 240Hz. Les tests d'acquisition sont faits par 3 personnes différentes. En travaillant sur les acquisitions, nous avons observé que la détection précise du picking, définie comme le moment où l'utilisateur initie un mouvement d'accélération du doigt en vue de la sélection, est cruciale pour améliorer les méthodes de saisie de texte en réalité virtuelle et augmentée. Dans cette étude, nous avons adopté une approche basée sur l'analyse de l'accélération, en particulier en appliquant un seuil sur la norme L2 de l'accélération. Cependant une autre approche basée sur l'alignement entre deux séquences d'accélération (query et référence) a été implémenté et mérite plus d'expériences, celle-ci permettrait de retrouver la séquence du début de picking à sa fin, si la séquence de référence est bien choisie. Ainsi, le processus de détection du picking repose sur une analyse de la norme 2 de l'accélération: Nous avons un seuil (relativement haut) d'accélération permettant de détecter le mouvement, puis à partir du franchissement de ce seuil on revient dans le temps jusqu'à trouver une zone où l'accélération moyenne est inférieure à un second seuil (relativement bas) pendant une durée donnée. On trouve ainsi le début du picking. Nous avons installé un setup avec 3 marqueurs dont deux sur un plan fixe afin de constituer un viseur (Points 1 et 2) et un sur le doigt dont le mouvement est capté (Point 3). On effectue ensuite plusieurs séries de picking en visant par alignement du doigt avec les deux points du viseur au début du geste.

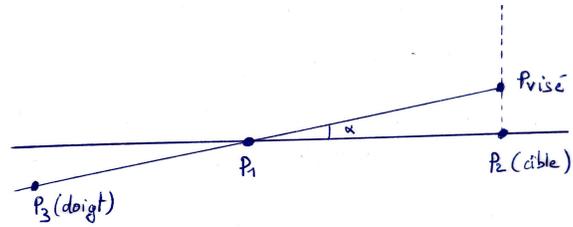


Figure 2: Schéma de l'expérience

## 4.2 Résultats

### 4.2.1 Détection temporelle du picking

La Figure 3 illustre le processus de détection du picking. Nous avons tracé la norme de l'accélération et son gradient, mettant en évidence le point où l'accélération atteint son maximum. Le seuil de détection du picking est défini à  $5 \times 10^{-6} m.s^{-2}$  pour distinguer les mouvements significatifs des fluctuations indésirables. Ce seuil a été fixé expérimentalement et dépend du  $\delta t$  choisi dans le calcul des accélérations par différence finie.

De manière significative, la ligne pointillée rouge marque le moment réel du picking, que l'on récupère en enfonçant une touche du clavier. Nous avons également introduit une zone colorée en jaune, représentant la région de faible gradient, soulignant la période précédant le picking.

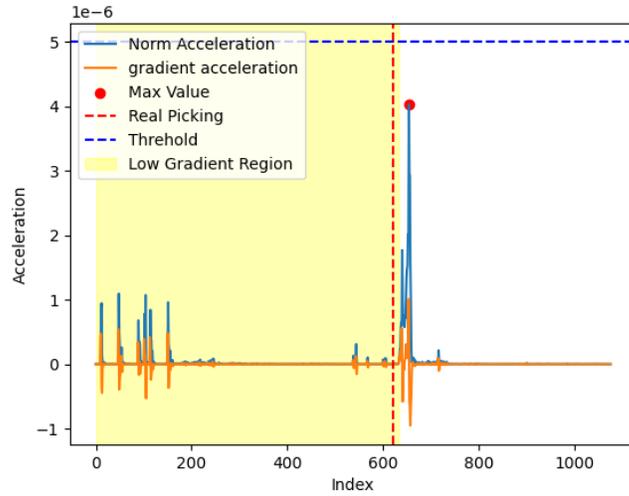


Figure 3: Analyse d'accélération

Pour retrouver le point du picking, une descente de gradient sur l'accélération en partant du maximum, nous donnerait un minimum local qui correspond exactement à la base d'un des pics voisins (en général 2 à 4 pics au moment du picking) ce qui n'est pas à souhaiter. Notre approche consiste donc plutôt à trouver l'instant à partir duquel une fenêtre d'accélération ne contient que des valeurs petites et donc inférieur à notre deuxième seuil choisi à  $5 \times 10^{-6} m.s^{-2}$  (ce qui correspond à l'instant d'immobilité avant le picking).

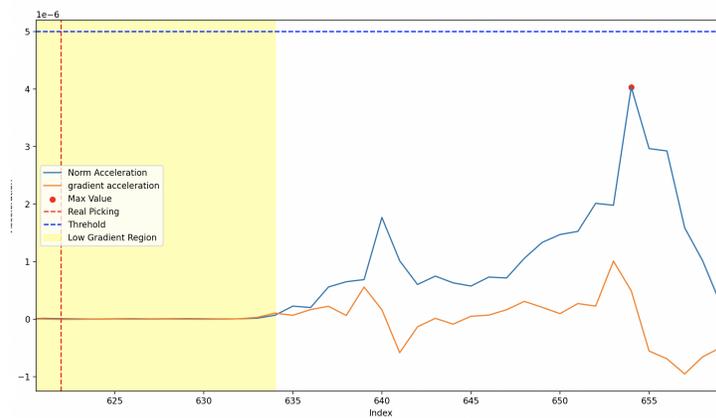


Figure 4: Zoom sur la figure 3

Remarquons que la difficultés des acquisitions (bon mouvement et appuie sur le clavier au bon instant) ainsi que le nettoyage des données (on obtient souvent une à trois valeurs fausses lorsqu'on clique sur le clavier), font qu'on dispose d'une dizaine (15) d'acquisitions exploitables.

Si on en avait plus, alors un test statistique sur la distribution des accélération peut être très intéressant. Notamment maximum (pour la valeur du pic) et moyenne de la séquence d'accélération.

On obtient une proportion de succès de notre approche égale à  $\frac{8}{15}$ . La difficulté étant dans la différence d'échelle pour la valeur du maximum (pic du picking), un test statistique avec un très grand nombre d'échantillons permettrait de trancher.

A noter que pour le rendu nous nous sommes restreins à l'analyse et détection de picking sans autres mouvements. Cependant voici une visualisation (trajectoire et norme d'accélération) d'un mouvement horizontale suivi d'un picking.

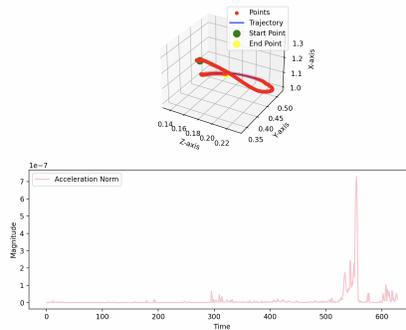


Figure 5: Mouvement horizontal suivi d'un picking

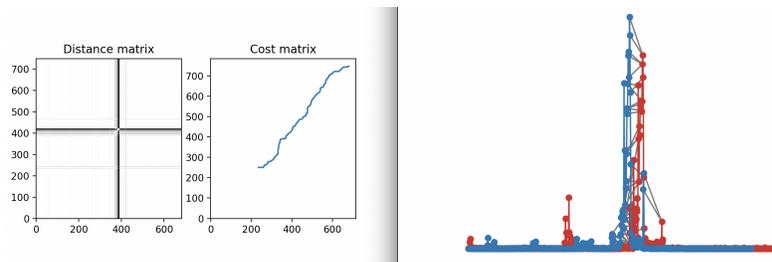


Figure 6: Alignement DTW

Cette dernière visualisation, montre l'intérêt d'utiliser une approche basée

sur l'alignement pour de futures expérimentations. À gauche : une matrice des distances qui permet de retrouver les indices des pickings par exemple et une matrice de coût qui donne l'alignement optimal des deux séquences. À droite l'alignement des deux pics d'accélération sur deux acquisitions.

#### 4.2.2 Précision spatiale du picking (Midas touch)

Après la détection temporelle, intéressons à l'aspect spatial du picking. Le but du picking étant de sélectionner une touche de clavier, il est essentiel d'avoir une bonne précision et une bonne exactitude. Pour évaluer ces caractéristiques on a effectué la mesure de la distance entre le Point cible (C) ou Point 2 et le Point visé (V). Le point visé en question est défini par l'intersection de la droite passant par les points 1 et 3 et le plan passant par le point 2 et orthogonal à la droite (Point 1-Point 2). On obtient ainsi à partir de 40 acquisitions le graphe suivant:

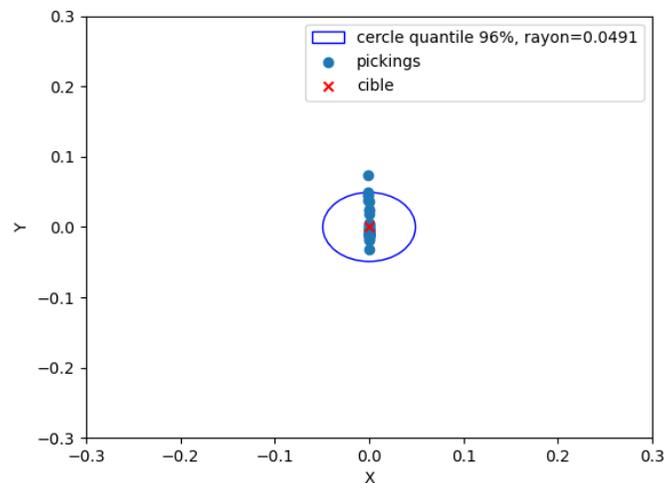


Figure 7: Répartitions des pickings

On observe qu'un cercle centré sur la cible et de rayon 4.9 cm englobe 96% des résultats. Dans le cadre de l'expérience, la distance entre  $P_1$  et  $P_2$  est de 68 cm. Ce cercle de 5cm de rayon correspond donc à un demi-angle d'erreur  $\alpha = 4.1^\circ$ .

## 5 Évaluation de l'interaction créée (centré humain)

L'évaluation de l'interaction créée a montré des résultats plutôt positifs en terme de précision en effet on observe que malgré une faible distance entre le doigt et le viseur on a une bonne précision car la touche idéale devrait faire une quinzaine de centimètres. On peut donc supposer que dans le cas d'une application finale cette taille serait bien plus faible (car la réelle distance d'intérêt est la distance oeil-doigt et que celle ci est plus grande). Cependant, l'interaction ne semble pas résoudre le problème du "gorilla arm effect" qui a été par exemple fortement ressenti lors des tests. En outre le point fort anticipé, à savoir la rapidité de la saisie n'a pas pu être évalué et comparé à la saisie classique par manque de temps. Dans de prochaines expérimentations, il serait intéressant d'effectuer les tests sur hololens pour pouvoir tester l'effet de "gorilla arm" dans différentes configurations (par exemple avec projection au sol pour reposer les bras), mais aussi pour avoir une comparaison réelle d'efficacité avec les claviers classiques.

## 6 Conclusion

En conclusion, le prototype d'outil de saisie par accélération en réalité virtuelle a démontré un beau potentiel notamment en terme de précision. Bien que des limitations liées à la dépendance à l'OptiTrack aient été identifiées, des pistes d'optimisation et d'amélioration sont envisageables pour rendre l'outil encore plus performant.