

Les techniques d'analyse d'images biométriques, au service des spécialistes du suivi des maladies neurodégénératives

Wagner de Aguiar
Instituto Federal de São Paulo – IFSP, São Paulo, Brasil
wagner.aguiar@unesp.br

Guillaume Thomann
Univ. Grenoble Alpes, CNRS, G-SCOP, 38 000 Grenoble,
France
guillaume.thomann@grenoble-inp.fr

José Celso Freire Júnior
Faculdade de Engenharia de Guaratingueta – FEG/UNESP,
Brasil
jose-celso.freire@unesp.br

Gilberto Cuarelli
Instituto Federal de São Paulo – IFSP, São Paulo, Brasil
gcuarelli@ifsp.edu.br

Abstract—Un terme générique pour définir les maladies qui attaquent le système nerveux est “maladie neurodégénérative”. Actuellement, plusieurs spécialistes suivent ces patients, l'une des principales manifestations cliniques étant le trouble du mouvement. Les évolutions des outils de mesure de l'activité humaine, en situation de réhabilitation ou de diagnostic, sont nombreuses. Mais ses applications restent marginales, trop chères, ou manquent de validité auprès des communautés d'experts. La présente proposition est d'utiliser des caméras comme capteurs de mouvement appliqués à l'identification de signaux biométriques, où le besoin de percevoir la forme et le mouvement des parties du corps est essentiel. Notre proposition consiste à développer un système pour faciliter le travail du spécialiste en utilisant l'ordinateur et sa caméra comme outil de capture de données d'image. Nous voulons utiliser ces données appliquées à un du suivi des mains et des doigts grâce à la bibliothèque MediaPipe de Google, qui nous permet d'obtenir des données de positions en temps réel. Dans un premier temps, une étude des principales caractéristiques des caméras et de leur utilisation avec des programmes de traitement d'images a été réalisée. Nous nous avons procédé à une évaluation des données de ce capteur. Cette étude porte sur le suivi des patients sous traitement pour des maladies neurodégénératives. Avec l'utilisation de ce système, il est possible de mettre un outil d'aide à la disposition des spécialistes du domaine.

Keyword: *L'analyse des données; capteurs de mouvement; maladies neurodégénératives;*

I. INTRODUCTION

Pour interagir en utilisant les technologies capables de capter l'environnement, il est nécessaire d'acquérir la description géométrique de la scène. Jusqu'à présent, il était nécessaire d'utiliser des outils proposant l'utilisation de technologies avancées et de réaliser les activités en laboratoire spécifique, dans lequel il y avait la possibilité de contrôler l'environnement et d'utiliser des caméras matricielles. Ces équipements sont des capteurs capables de récupérer des informations de profondeur et de mesurer le temps pendant lequel la lumière émise par l'appareil atteint les objets

souhaités et revient au capteur. De cette façon, les données de profondeur sont disponibles dans des scènes statiques ou dynamiques [1] [2].

Selon l'Organisation Panaméricaine de la Santé - OPS, les maladies neurodégénératives figurent parmi les trois principales causes de décès dans le monde et sont responsables de près de 2 millions de décès par an. [3]. Les maladies neurodégénératives sont caractérisées par une bradykinésie, des tremblements au repos et une rigidité musculaire. Elles touchent actuellement environ 1 à 3 % des personnes âgées de 65 ans et plus, le nombre de patients augmentant, associé à une espérance de vie plus longue [4].

Avec l'objectif d'aider à l'analyse des effets de la bradykinésie permettant de vérifier l'augmentation de la maladie dégénérative, une application a été développée. Celle-ci analyse "la capacité d'engager le corps physique dans une interaction avec la technologie, par exemple par des gestes" [5].

Ce travail propose une lecture des "processus d'interaction" à l'usage des spécialistes du suivi des troubles du mouvement.

Zannaro [6] et Mutto [7] ont utilisé des instruments de pointe disponibles uniquement dans des laboratoires spécialisés, utilisant des caméras matricielles avec des capteurs qui capturent des informations de profondeur. De cette façon, les données de profondeur deviennent disponibles dans des scènes statiques ou dynamiques.

La détection de la fréquence cardiaque, la détection de l'équilibre d'un individu, l'identification des crises, l'analyse des troubles résultant de maladies neurodégénératives, entre autres, utilisent l'identification biométrique. Pour mesurer et enregistrer la position et l'orientation du corps humain en mouvement, une technologie appelée Motion Capture (MoCap) est utilisée. Elle a déjà été utilisée dans plusieurs applications [8]. Parmi les techniques utilisées, on peut citer l'analyse d'images obtenues avec des capteurs de mouvement tels que Kinect [4] [9].

De nouveaux appareils de mesures tels que Microsoft Azure Connecta®, Intel RealSense® et ASUS Xtion® peuvent prendre en charge une large gamme de systèmes d'exploitation et de langages de programmation, permettant une extraction plus détaillée des données pour une analyse et une interprétation plus faciles des ces données. Ces appareils sont connus sous le nom de caméras de profondeur 3D (RGBD).

Le framework MediaPipe répond aux défis des développeurs avec un framework peu coûteux et de bonne précision. Son intégration peut se faire par programmation dans plusieurs langages tels que Python, Java et C/C++. Il s'agit d'un cadre permettant de créer des pipelines pour effectuer des inférences sur des données arbitraires, notamment en permettant la perception avec des composants graphiques modulaires, l'inférence de modèle, des algorithmes de traitement multimédia et la transformation de données [9] [10]. L'identification biométrique à l'aide de la bibliothèque MediaPipe est considérée comme une tâche difficile. La vision par ordinateur est une solution haute-fidélité et son application pour détecter les mouvements de la main devrait être utilisée [11].

II. ÉTAT DE L'ART

Les technologies de détection d'images stéréo utilisent des capteurs avec des caméras pour calculer la profondeur et permettent ainsi à ces appareils de visualiser, comprendre, interagir et apprendre avec leur environnement de manière 3D. Construire une application pour percevoir l'environnement qui vous entoure est un défi. Plusieurs étapes doivent être franchies et le framework MediaPipe répond à ces défis. En 2019, Google a lancé un projet appelé MediaPipe. À partir de juin 2021, il a ouvertement fourni ce code, qui a attiré l'attention des développeurs d'appareils qui utilisent des caméras de profondeur dans leurs applications. MediaPipe utilise l'apprentissage automatique (Machine Learning ou ML) pour déduire des jalons dans les lectures de données. Chaque main détectée/suivie affiche les 21 points de référence et chaque point est composé des coordonnées x, y et z (Figure1). Les valeurs de x et y sont normalisées [0.0, 1.0] respectivement pour la largeur et la hauteur de l'image et z représente sa profondeur. Plus la distance z, est proche de la caméra, plus sa valeur est faible. Dans la figure 1, les points peuvent être observés à partir d'un seul cadre, facilitant l'analyse de grands blocs de données [11] [12].

Les capteurs de ce type sont présents dans les appareils mobiles et les progrès des technologies appliquées aux caméras permettent une augmentation de leur intégration dans les plates-formes portables, nous avons donc un grand potentiel à explorer [10][13].

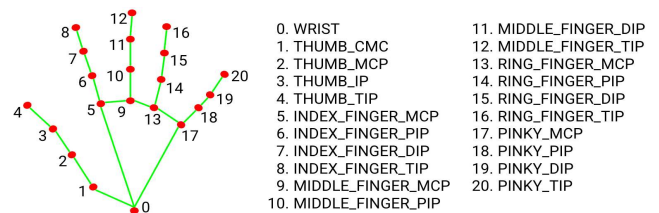


Fig 1 - Repères de la main [11]

a. Obtenir des waypoints avec Mediapipe

MediaPipe permet la création de pipelines avec une machine learning appliqué au multimodal (audio, vidéo) et dispose d'une vaste collection de modèles de détection et de suivi du corps humain, où les points sont normalisés en trois dimensions. Les développeurs de Google ont créé des modèles à l'aide de Tensorflow lite, facilement adaptables et ajustables via des graphiques [13] [14].

La figure 2 illustre une proposition d'architecture pour le pré-traitement des images. A première étape se situent les images reçues par la caméra. À l'étape suivante, l'identification de la main et la capture des données de positionnement sont effectuées, puis les données sont exportées vers un fichier .csv et sont disponibles pour analyse.

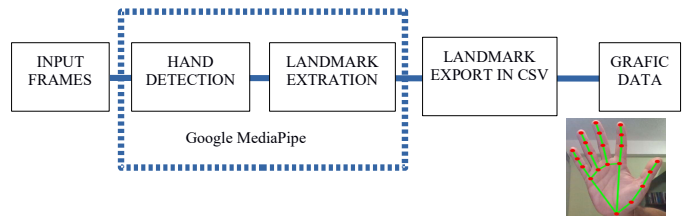


Fig. 2 – Proposition d'architecture pour le pré-traitement d'images

b. bradykinésie

Des études montrent la faisabilité d'évaluer objectivement la marche de patients atteints de maladies neurodégénératives dans des environnements contrôlés, à l'aide de Kinect et, avec la proposition de quantifier ces anomalies [15]. Dancra [16] a mené une étude pour classer les stades de la maladie de Parkinson à l'aide d'un capteur Kinect basé sur l'évaluation des troubles de la marche.

Ben [4] évoque l'échelle d'évaluation unifiée de la maladie de Parkinson (UPDRS), une échelle subjective réalisée et notée par des médecins. Dans son étude, l'outil proposé prend la position des échelles UPDRS utilisées aujourd'hui, qui sont des mesures subjectives. Compte tenu de l'importance que les maladies neuro-dégénératives ont pris dans la société aujourd'hui, cette méthode est très prometteuse pour la perception générale des conditions humaines. En outre, l'utilisation de capteurs de mouvement couplés à des applications avec de nouvelles techniques peuvent permettre de réaliser à faible coût, de manière efficace et précise, une étude d'aide au traitement des troubles résultant de maladies neuro-dégénératives.

L'analyse d'images a montré un grand potentiel dans le domaine médical, notamment dans le diagnostic de la maladie d'Alzheimer, de la maladie de Parkinson et de la sclérose en

plaques. L'application de ces méthodes peut encore améliorer la capacité à analyser des données d'images multimodales complexes et améliorer l'efficacité de ces diagnostics [17].

III. MATÉRIEL ET MÉTHODE

A. Matériel

Les données initiales ont été obtenues avec une caméra d'ordinateur portable RVB HD standard avec une résolution HD de 1280x720 à 30 ips. Deuxièmement, nous utilisons la caméra RealSense D455i et un capteur RVB et CMOS avec une résolution de 1280x720 à 90 ips.

À l'aide d'un programme développé en JavaScript, il a été possible de visualiser les données reçues par la caméra, comme le montre la figure 3. Lors de l'utilisation du programme, il faut d'abord choisir la caméra qui sera utilisée, puis déterminer le temps d'échantillonnage, qui peut être réalisé en 10, 20 ou 30 secondes. Bien entendu, il est toujours possible de déterminer le type de rendu visuel qui sera effectué.

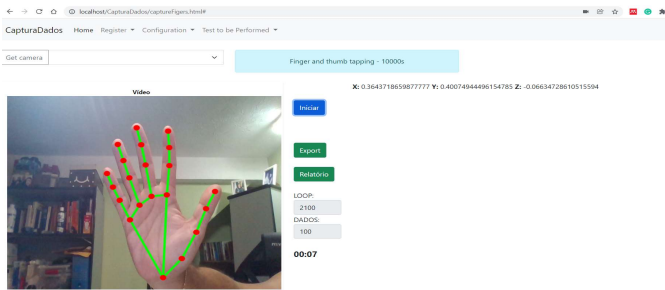


Fig 3 – Écran de capture de données

Dans l'image, nous avons les 21 points identifiés de la main. A la fin du temps de lecture, il est possible de générer un rapport ou d'effectuer une nouvelle lecture. Lorsqu'un nouveau rapport est généré, un graphique est présenté, regroupant tous les points lus. Seules les coordonnées x et y du détecteur sont prises en compte, et chaque entrée nulle d'image est vérifiée et rejetée lorsque la main n'est pas présente ou que l'image est floue. Ci-dessous, la figure 4 montre le graphique avec les 21 points de la main représentés. Lorsque le curseur est arrêté sur l'une des lignes, les valeurs x et y sont affichées, comme le montre la figure ci-dessous.

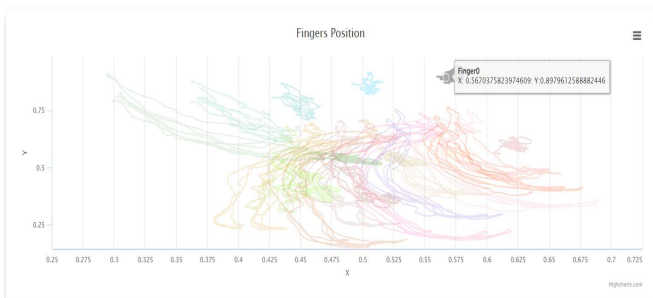


Figure. 4 - Graphique linéaire de mouvement en coordonnées x et y

A partir de cette étape du programme, il est possible de continuer le processus pour stocker plus de points comme

ceux déjà obtenus, ou de relancer le processus. Si les informations de lecture sont satisfaisantes, il est possible de générer une feuille de calcul .csv avec les données obtenues en appuyant sur le bouton d'exportation. Des informations sur la personne, telles que le nom, l'âge et le sexe, sont demandées. Ces données, ainsi que le code du type de lecture effectuée, doivent composer le nom du fichier. Exemple name_age_sex_codxx.csv

Avec ces données stockées, il est possible d'effectuer la recherche disponible dans le programme et de sélectionner l'individu et le test effectué.

La figure 5 montre cet écran dans lequel il est possible de sélectionner le graphique avec les 21 points à lire ou d'individualiser chacun et ses graphiques. Dans la figure 6, seul le point 0 sélectionné est affiché.

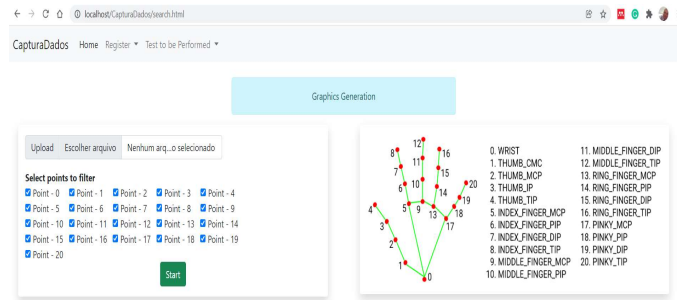


Fig 5 – Écran pour sélectionner le point à fichier

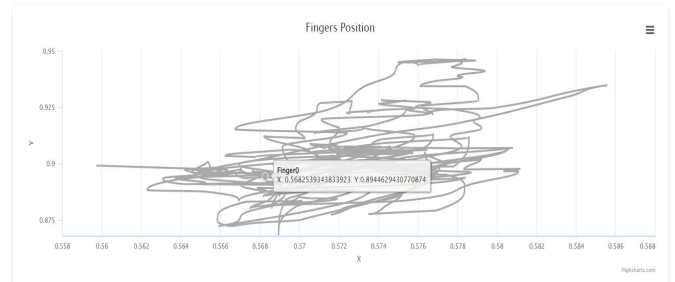


Fig 6 - Graphique du point 0

Le mouvement représenté sur le graphique de la figure 6 concerne le point 0 et sa variation est visualisée sur les axes x et y. Sur la figure 7, le même point 0 est représenté, mais cette fois par rapport à l'axe z. Dans ce graphique, la taille du cercle indique la position par rapport à la distance z. Cet axe représente la distance entre le point et la caméra.

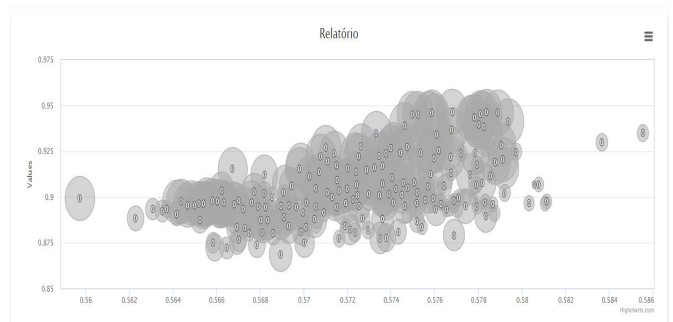


Fig 7 - Graphique du point 0 avec l'axe z

Dans l'analyse des données obtenues, la suggestion est de valider environ 95% des données dans la représentation du modèle souhaité, où les données initiales qui peuvent être corrompues ou invalides sont rejetées. Un calcul approximatif est effectué pour déterminer les variations de l'intervalle de temps sélectionné avec les valeurs obtenues. Dans l'intervalle de 10 secondes, une moyenne de 5670 mesures ont été effectuées, se référant aux 21 points.

En excluant les 210 premières mesures, nous avons environ 5460 mesures valides. Selon l'équation (1)

$$mv = md - 210 \quad (1)$$

où md sont les mesures prises et mv sont les mesures valides.

Ensuite, la moyenne des échantillons à la position 0 est déterminée, ce qui indique s'il y a eu un changement significatif dans le positionnement de la main. Dans l'exemple, cela représente 5460/21, ce qui donne 260 échantillons au point 0, et la moyenne est prise. En cas de changements importants, une nouvelle lecture est proposée ou, selon le spécialiste, ces données peuvent être conservées.

IV. Résultats

Selon les données obtenues, il a été possible de déterminer que la méthode aide à identifier la variation des points sur la main. Ces données peuvent provenir d'un seul point ou d'une comparaison de points et les valeurs obtenues entre eux peuvent être représentées graphiquement. Les points à analyser ou à comparer sont définis par le spécialiste qui aura la liberté de choix, car les valeurs nécessaires pour effectuer cette analyse ont déjà été obtenues et stockées.

V. Conclusion

Les données obtenues grâce à la caméra permettent d'interpréter le positionnement des doigts et de la main par rapport au référentiel de la caméra. Cette information peut permettre au spécialiste d'identifier plus facilement une augmentation ou une diminution du degré d'un trouble neurologique tel que la maladie d'Alzheimer.

De nombreuses autres applications peuvent être envisagées, comme dans le domaine de la kinésithérapie où ce système peut être utilisé, dans la récupération des mouvements des doigts en chirurgie correctrice, par exemple.

References

- [1] Zennaro, S., P. and Bernardi, A. and Ghidoni, S. and Menegatti, E., Performance evaluation of the 1st and 2nd generation Kinect for multimedia applications, IEEE, 2015
- [2] Mutto, C. and Zanuttigh, P. and Cortelazzo, G. M., Time of flight Cameras and Microsoft KinectTM, editora Springer-Verlag Berlin Heidelberg, 2013
- [3] Organização Pan-Americana de Saude - OPAS. OPAS Brasil, 2020 disponível em: <https://www.paho.org/pt/noticias/9-12-2020-oms-revela-principais-causas-morte-e-incapacidade-em-todo-mundo-entre-2000-e> – Acesso em 20/01/2022
- [4] Dror, Ben, Eilon Yanai, Alex Frid, Nimrod Peleg, Nadav Goldenthal, Ilana Schlesinger, Hagit Hel-Or, Shmuel Raz, Automatic Assessment of Parkinson's Disease From Natural Hands Movements Using 3D Depth

- Sensor, Convention of Electrical and Electronics Engineers in Israel, IEEE (2014)
- [5] Hartson, R. and Pyla, P. S., The UX Book -- Process and Guidelines for Ensuring a Quality User Experience, Editora Morgan Kaufmann, 2012.
- [6] Zennaro, S. and Munaro, M. and Milani, S. and Zanuttigh, P. and Bernardi, A. and Ghidoni, S. and Menegatti, E., Performance evaluation of the 1st and 2nd generation Kinect for multimedia applications, IEEE International Conference on Multimedia and Expo (ICME), pages 1-6, (2015)
- [7] Mutto, C. and Zanuttigh, P. and Cortelazzo, G. M., Time of flight Cameras and Microsoft KinectTM, The Universal Access Handbook, Springer-Verlag Berlin Heidelberg pages 17-47 (2013)
- [8] Regazzoni, G. de Vecchi, and C. Rizzi, RGB cams vs RGB-D sensors: Low cost motion capture technologies performances and limitations, Journal of Manufacturing Systems, vol. 33, no. 4, pp. 719-728, 2014. doi:10.1016/j.jmsy.2014.07.011.
- [9] Qiannan Li, Yafang Wang, Andrei Sharf, Ya Cao, Changue Tu, Baoquan Chen, Shengyuan Yu, Classification of gait anomalies from Kinect, IEEE. (2016)
- [10] Castro, M. C. L., J. P. A. Xavier, P. F. F. Rosa, and J. C. de Oliveira, "Interaction by Hand-Tracking in Virtual Reality" SBCOPENLIB Anais do simposio de realidade virtual e aumentada (SVR) Novembro 2020. DOI: https://doi.org/10.5753/svr_estendido.2020.12954.
- [11] Lugaresi, C., Tang, J., Nash, H., Mcclanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C., Yong, M. G., Lee, J. Chang, W., Hua, W., Georg, M. and Grundmann, M., MediaPipe: A Framework for Building Perception Pipelines, Google Research, arXiv:1906.08172v1, Jun 2019
- [12] MEDIAPIPE, soluções ML personalizáveis e multiplataforma, recuperado 13 dezembro 2021, de <https://github.com/google/mediapipe>.
- [13] Bazarevsky, V. I. Grishchenko, K. Raveendran, T. Zhu, F. ZhangG, and M. Grundmann "Blaze Pose: On-device Real-time Body Pose tracking" arXiv:2006.10204v1 [cs.CV] 17 Jun 2020
- [14] Arpita H. and Akshit T. Real-Time Vernacular Sign Language Recognition using MediaPipe and Machine Learnig – International Journal of Research Publication and Review – Journal homepage : www.ijrpr.com ISSN 2582-7421 pg 9 – 17 may 2021
- [15] Das, P., Ahmed, T., and Ali, M. F. Static Hand Gesture Recognition for American Sign Language using Deep Convolutional Neural Network. In november 2020 IEEE Region 10 Symposium (TENSYP) 2020, June.(pp. 1762-1765). IEEE.
- [16] Farnood Gholami, Daria A. Trojan, József Kóvecses, Wassim M. Haddad, *Fellow, IEEE*, and Behnood Gholami, A Microsoft Kinect-Based Point-of-Care Gait Assessment Framework for Multiple Sclerosis Patients, IEEE Journal of biomedical and Health Informatics (2017).
- [17] Lacramioara Dranca, Lopez de Abetxuko Ruiz de Mendarozketa, Alfredo Goñi, Arantza Illaramendi, Irene Navalpotro Gomez, Manuel Delgado Alvarado and Maria Cruz Rodriguez-Oroz. In: BMC Bioinformatics (2018). DOI: s12859-018-2488-4.
- [18] Brito, E. N. D. B., Figueiredo, B. Q., Souto D. N., Nogueira J. F., Melo A. L. S. C., Silva, I. T., Oliveira, I. P. and Almeida M. G. Inteligencia Artificial no diagnostico de doenças Neurodegenerativas: Uma revisão sistematica de literatura. Research, Society and Development, V. 10 ISSN 2525-3409 – set. 2021