

Sujet de thèse CIFRE

Recalage dynamique de modèles 3D de denture sur un flux vidéo dans un contexte de réalité augmentée pour l'odontologie

Entreprise Adservio, Periosystem (Paris)

Contexte général :

Un cabinet dentaire est souvent considéré comme un lieu où un patient se rend pour recevoir des soins à ses problèmes dentaires. Or, depuis quelques années, le rôle de la dentisterie n'a cessé d'évoluer, elle n'est plus considérée comme un simple traitement médical, mais désormais, elle représente une discipline plus vaste où l'esthétique joue un rôle primordiale. Par ailleurs, la communication du praticien avec son patient sur le traitement envisagé est essentielle, elle doit être claire et accompagnée si possible de représentations visuelles permettant au patient de voir le résultat attendu.

Durant ce processus, le patient peut donner son avis, voire modifier certains éléments du traitement. Visualiser à priori le résultat d'une opération dentaire permet aussi de conforter le patient dans sa prise de décision, à savoir accepter ou refuser les soins proposés. Ce retour visuel représente aussi un outil important pour le dentiste qui lui permet de valider les soins post-traitement et ainsi de les contractualiser. Un tel travail est actuellement réalisé en grande partie de façon manuelle, nécessitant plusieurs visites du patient et plus de travail pour le praticien.

L'objectif de cette thèse est de développer de nouvelles approches de visualisation et d'interaction permettant d'automatiser le travail dans un cabinet dentaire et améliorer la communication docteur-patient. Pour cela nous proposons d'utiliser les technologies de la vision par ordinateur et les techniques de l'interaction homme-machine.

Etat de l'art :

Le domaine de la réalité augmentée est un domaine de recherche riche où plusieurs travaux autour de la détection et du suivi des objets dans un flux vidéo ont été développés ces dernières années. Parmi les algorithmes les plus connus, nous pouvons citer Linemod et Line-2D développé par Hinterstoisser et al., PWP3D par Prisacariu et al.

La majorité de ces méthodes utilisent des algorithmes de détection de contours et de recalage rigide ou mou selon le contexte et la nature du modèle 3D disponible. Disney research a récemment développé une approche permettant d'une part de détecter les dents dans une vidéo, et d'autre part de générer puis positionner un modèle 3D des dents sur ce même flux vidéo.

D'autres travaux sont proposés pour améliorer les performances des algorithmes basés sur Line-2D en utilisant des bases de données d'objets génériques texturés et/ou non texturés..

Ces dernières années, de plus en plus de travaux dans le domaine la vision artificielle et la réalité augmentée se basent sur des méthodes d'apprentissage non supervisé telles que les réseaux de neurones profonds. Ainsi, Xiang et al. utilise un modèle de réseaux de neurones pour estimer le positionnement des objets dans une scène. DeTone et al. concurrence les descripteurs connus comme ORB et SIFT. Enfin, des réseaux génératifs GAN sont utilisés pour générer des modèles et réduire la taille nécessaire pour la base d'apprentissage.



Figure 1. Algorithme de détection et génération du modèle des dents à partir d'un flux vidéo image par Wu et al.

Verrous scientifiques :

L'application directe des méthodes citées ci-dessus dans le domaine de l'odontologie se confronte à plusieurs problèmes, notamment la réflexion lumineuse, le changement de couleur des éléments à suivre à cause des caries, les gencives rétractées (récession), les dents manquantes et les implants sans prothèse (enlevé ou défectueux). Il est clair que les algorithmes de l'état de l'art doivent être adaptées à ce cas d'usage.

A notre connaissance, il n'existe pas de travaux appliquant les modèles de réseaux de neurones profonds au recalage de modèles 3D de dent sur une séquence d'images. Etant donné que les bases de données disponibles ne contiennent pas beaucoup d'images, il sera nécessaire de mettre en place une approche de modélisation 3D, nous préconisons dans le cadre de cette thèse d'utiliser des modèles génératifs (Generative adversarial networks GAN) s'il y a un besoin. Des bases de données pour les expressions faciales pourront aussi être utilisées.

Sujet de thèse :

L'objectif de cette thèse est de proposer et implémenter de nouveaux algorithmes permettant de détecter les dents dans une image ou un flux vidéo, recalculer le modèle 3D sur les dents réelles et optimiser le temps de calcul pour permettre un fonctionnement en temps réel. Le recalage de l'objet 3D doit tenir compte des occultations des lèvres, optimiser la précision et la robustesse et permettre un rendu assez réaliste.

Dans un premier temps, des approches de détection et de suivi en réalité augmentée seront appliquées sur des objets tests, ensuite sur des objets issus du contexte d'odontologie.

Un système interactif sera développé permettant au patient de sélectionner, à partir de plusieurs modèles de réparation de dents, celui qui serait recalé sur son visage

Programme de travail préliminaire :

- Réaliser un état de l'art sur les méthodes de recalage d'objets 3D sur une séquence d'images ou une vidéo RGB. + Rédaction continue de rapports sur les différents techniques existantes
- Développement et test d'une partie des méthodes existantes dans l'état de l'art incluant les méthodes classiques, à base d'apprentissage, ou mixte.
- Adaptation des techniques génériques dans le contexte spécifique de l'odontologie
 - Développement des algorithmes pour détecter les dents dans une image
- Développement et optimisation des algorithmes pour une exécution temps-réel
- Test et évaluation des approches proposées

Profil attendu :

Le domaine de la réalité augmentée nécessite un candidat avec des compétences diversifiées et solides en informatique.

Compétences requises :

- Ingénieur informatique ou équivalent
- Bonnes connaissances en mathématiques, imagerie, et machine learning de base
- Compétences requises en développement : C++ avec la bibliothèque OpenCV
- Capacité continue à apprendre, s'adapter aux nouveautés et évoluer au fil du temps
- Fort intérêt pour la technologie et la lecture scientifique et technique (livres, articles de recherche)

Compétences souhaités :

- Connaissances minimale des technologies 3D (OpenGL / WebGL)
- Python avec Keras, TensorFlow ou équivalent
- Avoir des démos de travaux déjà réalisés dans le domaine

Comment candidater :

Envoi du CV à :

mohamed-ikbel.boulabiar@periosystem.fr ou mohamed-ikbel.boulabiar@adservio.fr

Vous pouvez envoyer des démos et des liens vers des projets dont vous êtes fier ainsi qu'un document ou un rapport que vous avez écrit. Ils sont plus valorisant pour la sélection qu'une lettre de motivation générique.

Vous pouvez envoyer des adresses de contacts pour une référence ou des recommandations.

Encadrement de la thèse :

Direction de la thèse : Pr. Fakhreddine Ababsa

Fakhreddine.ABABSA@ensam.eu

Encadrement en entreprise: Dr. Mohamed Ikbel Boulabiar

mohamed-ikbel.boulabiar@adservio.fr

Références bibliographiques :

- [Disney] Wu, C., Bradley, D., Garrido, P., Zollhöfer, M., Theobalt, C., Gross, M., & Beeler, T. (2016). Model-based teeth reconstruction. *ACM Transactions on Graphics (TOG)*, 35(6), 220.
- [LineMod] Hinterstoisser, S., Lepetit, V., Ilic, S., Holzer, S., Bradski, G., Konolige, K., & Navab, N. (2012, November). Model based training, detection and pose estimation of texture-less 3d objects in heavily cluttered scenes. In *Asian conference on computer vision* (pp. 548-562). Springer, Berlin, Heidelberg.
- [Line2D] Hinterstoisser, S., Cagniart, C., Ilic, S., Sturm, P., Navab, N., Fua, P., & Lepetit, V. (2012). Gradient response maps for real-time detection of textureless objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(5), 876-888.
- [OLS] Seo, B. K., Park, H., Park, J. I., Hinterstoisser, S., & Ilic, S. (2014). Optimal local searching for fast and robust textureless 3d object tracking in highly cluttered backgrounds. *IEEE transactions on visualization and computer graphics*, 20(1), 99-110
- [RandomTrees] Lepetit, V., Lagger, P., & Fua, P. (2005, June). Randomized trees for real-time keypoint recognition. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* (Vol. 2, pp. 775-781). IEEE.
- [PWP3D] Prisacariu, V. A., & Reid, I. D. (2012). PWP3D: Real-time segmentation and tracking of 3D objects. *International journal of computer vision*, 98(3), 335-354.
- [TCLCHisto] Tjaden, H., Schwanecke, U., & Schömer, E. (2017). Real-Time Monocular Pose Estimation of 3D Objects using Temporally Consistent Local Color Histograms. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 124-132).
- [PoseCNN] Xiang, Y., Schmidt, T., Narayanan, V., & Fox, D. (2017). PoseCNN: A Convolutional Neural Network for 6D Object Pose Estimation in Cluttered Scenes. *arXiv preprint arXiv:1711.00199*.
- [Dtrans] Ou, X., Liu, S., Cao, X., & Ling, H. (2016, October). Beauty eMakeup: A Deep Makeup Transfer System. In *Proceedings of the 2016 ACM on Multimedia Conference* (pp. 701-702). ACM.
- [GANHand] Wan, C., Probst, T., Van Gool, L., & Yao, A. (2017, July). Crossing nets: Combining gans and vaes with a shared latent space for hand pose estimation. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit*
- [FeaturesDnn] DeTone, D., Malisiewicz, T., & Rabinovich, A. (2017). SuperPoint: Self-Supervised Interest Point Detection and Description. *arXiv preprint arXiv:1712.07629*.
- [] Hinterstoisser, S., Holzer, S., Cagniart, C., Ilic, S., Konolige, K., Navab, N., & Lepetit, V. (2011, November). Multimodal templates for real-time detection of texture-less objects in heavily cluttered scenes. In *Computer Vision (ICCV), 2011 IEEE International Conference on* (pp. 858-865). IEEE.
- [] Ferrari, V., Jurie, F., & Schmid, C. (2010). From images to shape models for object detection. *International journal of computer vision*, 87(3), 284-303
- [] Wang, B., Zhong, F., & Qin, X. (2017, June). Pose optimization in edge distance field for textureless 3D object tracking. In *Proceedings of the Computer Graphics International Conference*(p. 32). ACM.