

# CESI LINEACT

**Titre :** *Sélection de vues pour de l'estimation de pose d'objets industriels pour l'opérateur augmenté*

**Domaines scientifiques :** *industrie 5.0, deep et machine learning, digitalisation de l'industrie, perception*

**Mots clés :** *sélection de vues, reconnaissance d'objets industriels, estimation de pose*

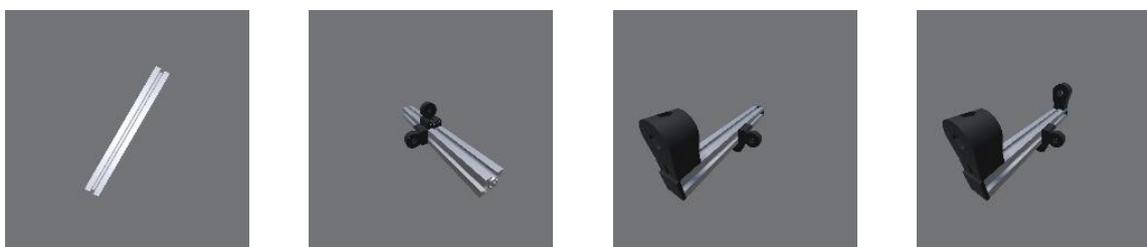
## Encadrement

Prénom Nom	Fonction, Titre	@
Nicolas Ragot	Enseignant – Chercheur Spécialisé en perception pour la robotique	<a href="mailto:nragot@cesi.fr">nragot@cesi.fr</a>
Vincent Havard	Enseignant – Chercheur Spécialisé en réalité virtuelle, réalité augmentée et apprentissage profond	<a href="mailto:vhavard@cesi.fr">vhavard@cesi.fr</a>

## Travaux de recherche proposés

### Projet de stage

Cette proposition de stage prolonge des travaux d'ores et déjà engagés au laboratoire CESI LINEACT sur de la **sélection de vues pour de la reconnaissance d'étapes d'assemblage d'un produit manufacturier** (Xu, Ragot and Dupuis, 2022). Dans sa version actuelle, notre approche, utilisant des algorithmes de clustering et d'apprentissage supervisé, vise à déterminer le nombre optimal de vues et les poses des caméras (échantillonnées selon une demi-sphère de vue) les plus informatives et discriminantes pour reconnaître un objet industriel (dans le cas des étapes d'assemblage, chaque étape est considérée comme un objet unique). Notre cas d'usage portera sur l'**assemblage d'un vélo pour enfant** (type tricycle) constitué d'un profilé aluminium et de systèmes de fixation (cf. Figure 1).



**Figure 1 :** Exemples de pièces d'assemblage du vélo d'enfant

# CESI LINEACT

Trois contributions sont ainsi envisagées :

- La première portera sur **l'estimation de pose d'un objet industriel** en cours de montage pour des applications en réalité augmentée. Nous souhaitons **évaluer et comparer** des **méthodes basées gradient** (HOG, SIFT, SURF) avec celles plus récentes utilisant des **architectures d'apprentissage profond** (CNN, DPOD (Zakharov, Shugurov and Ilic, 2019), DSC-PoseNet (Yang, Yu and Yang, 2021), etc.) dans un cadre industriel où les jeux de données sont très limités.
- Une **deuxième contribution** portera sur une **comparaison entre une estimation de pose calculée à partir d'une approche de sélection de vues et d'images sélectionnées aléatoirement**. A travers cette étude, nous souhaitons évaluer la pertinence d'une approche de sélection de vues autrement que par le prisme de la reconnaissance de forme et de ses métriques.
- Une **troisième et dernière contribution** s'intéressera à **repenser l'approche de sélection de vues** actuellement développée qui repose essentiellement sur une métrique de clustering pour identifier les vues les plus discriminantes. Cette approche se heurte à des **problèmes de confusion** puisqu'en fonction des points de vue suivant lesquels la pièce en cours de montage est observée, des étapes d'assemblage peuvent fortement se ressembler. Nous souhaiterions envisager l'utilisation d'une **autre métrique pour identifier les vues les plus discriminantes en reconnaissance d'objets**, de celles qui le sont moins.

## Contexte scientifique

Dans le contexte de l'industrie 5.0, les jumeaux numériques, la réalité virtuelle et la réalité augmentée sont des technologies clés pour concevoir, simuler et optimiser les systèmes de production cyber-physiques (CPPS). Elles permettent également d'interagir avec les CPPS, que ce soit directement, à distance ou de manière collaborative (Havard, Jeanne, *et al.*, 2019), (Sjarov *et al.*, 2020). Les avancées récentes en réalité augmentée notamment ont permis d'envisager de nouveaux modes d'interaction entre les robots et les humains : interfacés avec le jumeau numérique, les dispositifs en RA permettent à l'opérateur de visualiser des informations ou des instructions sur les opérations à réaliser, comme des tâches d'assemblage ou de maintenance (Havard, Sahnoun, *et al.*, 2019), (Ojer *et al.*, 2020), (Havard *et al.*, 2021). Ces nouveaux modes d'interaction doivent être fluides et confortables pour que la collaboration humain – robot soit la plus efficace possible.

Actuellement, la validation du passage d'une tâche d'assemblage à une autre est réalisée par l'opérateur par une opération manuelle. Un enjeu vise donc à automatiser cette procédure en reconnaissant automatiquement les étapes d'assemblage du produit.

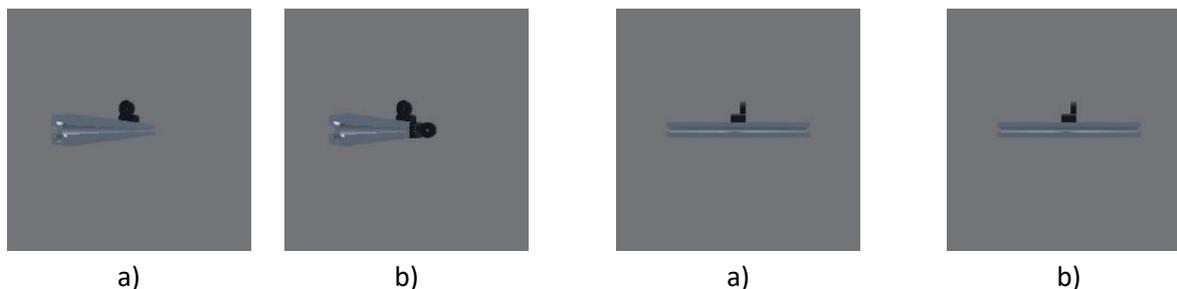
Les méthodes actuelles de la littérature en reconnaissance de formes utilisent les réseaux d'apprentissage profond. **Ces approches nécessitent de grands volumes de données**, ce qui est bien souvent **incompatible avec le secteur industriel** où secret et propriété intellectuelle empêchent la constitution et la publication de larges jeux de données.

Les **objets industriels sont très souvent peu texturés**. Les approches classiques en estimation de pose peuvent se retrouver en défaut. Dans le même temps, celles utilisant l'apprentissage profond se heurtent aux **faibles volumes de données disponibles** pour entraîner les modèles et ont donc des performances limitées.

Dans le processus d'assemblage d'un produit industriel, des **étapes successives peuvent se confondre** très fortement. En fonction du point de vue suivant lequel les images sont prises, **les pièces ajoutées d'une étape à l'autre peuvent occultées ou en masquées d'autres**. Il en résulte que **2 images prises à**

## CESI LINEACT

2 étapes différentes peuvent paraître totalement similaires, rendant particulièrement complexe le processus d'identification des vues les plus discriminantes. A titre d'exemple la Figure 2 illustre deux vues de la pièce en cours d'assemblage, respectivement aux étapes 2 et 3. Celle-ci montre clairement l'ajout d'une pièce de fixation entre l'étape 2 et 3. La Figure 3 illustre deux autres vues de cette pièce pour les mêmes étapes. Suivant ce point de vue, la pièce de fixation a « disparu » puisqu'elle est totalement masquée par la barre d'aluminium.



**Figure 2** : images du même objet prise sous un angle de vue identique : a) étape 2 ; b) étape 3. La 2<sup>e</sup> pièce de fixation apparaît clairement sur l'image de l'étape 3

**Figure 3** : images du même objet prise sous un angle de vue identique : a) étape 2 ; b) étape 3. Suivant cet angle de vue, la 2<sup>e</sup> pièce de fixation est masquée donnant l'impression que les 2 vues sont identiques

L'**objectif du stage** est donc d'automatiser le passage d'une tâche d'assemblage à une autre en **reconnaissant automatiquement chaque étape**. Pour ce faire, nous nous appuyons sur les **3 contributions** détaillées dans la section « projet de stage ».

Ces travaux de stage s'inscrivent dans le cadre du projet JENII (Jumeaux d'Enseignement Numériques, Immersifs et Interactifs). Il s'agit d'un démonstrateur de l'enseignement supérieur répondant à l'AMI « DemoES » 2021 et regroupant : ENSAM, CNAM, CESI et CEA. L'ambition du projet est de proposer une offre de formation spécifique pour l'industrie du futur, fondée sur la technologie des jumeaux numériques immersifs et collaboratifs de systèmes industriels réels avec deux objectifs majeurs : 1) mettre à disposition des environnements de formation technologique accessibles à distance ; 2) permettre le développement et la diffusion rapides et mutualisants des outils pédagogiques nécessaires pour accompagner les évolutions technologiques et les métiers de l'industrie.

# CESI LINEACT

## Programme de travail

	PERIODE DE STAGE					
Etat de l'art : sélection de vues, reconnaissance d'étapes d'assemblage, estimation de pose						
Evaluation : estimation de pose à partir de méthodes basées gradient et de celles utilisant des CNN (ex. PoseNet, DPOD)						
Evaluation : estimation de pose à partir des « Top » vues et celles prises aléatoirement sur la demi-sphère d'observation centrée sur l'objet						
Implémentation d'une méthode de sélection de vues où les images sont filtrées par leur indice de similarité (ex. PNSR)						
Evaluation de cette méthode et comparaison avec les performances de la méthode initialement développée reposant sur un calcul de score basé FMI						
Rédaction du rapport de stage & préparation de la soutenance						
Livrables				L.A1		L.A2 / L.D1

## Production scientifique/technique attendue

- L.D1 - Démonstrateur de reconnaissance automatique des étapes d'assemblage associant les informations liées au gestes/actions de l'opérateur.
- L.A1 - Production d'un article synthétisant les résultats obtenus pour l'estimation de pose basée HOG/SURF/SIFT VS apprentissage profond combiné ou non avec une sélection de vues
- L.A2 – production d'un article synthétisant la nouvelle approche de sélection de vues où les données sont filtrées par leur indice de similarité

## Présentation du laboratoire

LINEACT CESI (EA 7527), Laboratoire d'Innovation Numérique pour les Entreprises et les Apprentissages au service de la Compétitivité des Territoires, anticipe et accompagne les mutations technologiques des secteurs et des services liés à l'industrie et au BTP. La proximité historique de CESI avec les entreprises est un élément déterminant pour nos activités de recherche, et a conduit à concentrer les efforts sur une recherche appliquée proche de l'entreprise et en partenariat avec elles. Une approche centrée sur l'humain et couplée à l'utilisation des technologies, ainsi que le maillage territorial et les liens avec la formation, ont permis de construire une recherche transversale; elle met l'humain, ses besoins et ses usages, au centre de ses problématiques et aborde l'angle technologique au travers de ces apports.

Sa recherche est organisée selon deux thèmes scientifiques interdisciplinaires et deux domaines applicatifs.

# CESI LINEACT

- Le thème 1 "Apprendre et Innover" relève principalement des Sciences cognitives, Sciences sociales et Sciences de gestion, Sciences et techniques de la formation et celles de l'innovation. Les principaux objectifs scientifiques visés par ce thème sont la compréhension des effets de l'environnement, et plus particulièrement des situations instrumentées par des objets techniques (plateformes, ateliers de prototypage, systèmes immersifs...) sur les processus d'apprentissage, de créativité et d'innovation.
- Le thème 2 "Ingénierie et Outils Numériques" relève principalement des Sciences du Numérique et de l'Ingénierie. Les principaux objectifs scientifiques de ce thème portent sur la modélisation, la simulation, l'optimisation et l'analyse de données de systèmes industriels ou urbains. Les travaux de recherche portent également sur les outils d'aide à la décision associés et sur l'étude des jumeaux numériques couplés à des environnements virtuels ou augmentés.

Ces deux thèmes développent et croisent leurs recherches dans les deux domaines applicatifs de l'Industrie du Futur et de la Ville du Futur, soutenues par des plateformes de recherche, principalement celle de Rouen dédiée à l'Usine du Futur et celles de Nanterre dédiée à l'Usine et au Bâtiment du Futur.

## Organisation du stage

Financement : JENII

Lieu de travail : CESI campus de Rouen

Date de démarrage : 01/03/2022

Durée : 6 mois (avec report à prévoir des jours de fermeture début Août, à reprogrammer à la date de réouverture de l'établissement)

Candidat pressenti :

Etablissement de provenance :

## Références

Havard, V., Sahnoun, M., *et al.* (2019) 'An Architecture for Data Management, Visualisation and Supervision of Cyber-Physical Production Systems', in. Available at: <https://doi.org/10.3233/ATDE190015>.

Havard, V., Jeanne, B., *et al.* (2019) 'Digital twin and virtual reality: a co-simulation environment for design and assessment of industrial workstations', *Production & Manufacturing Research*, 7(1), pp. 472–489. Available at: <https://doi.org/10.1080/21693277.2019.1660283>.

Havard, V. *et al.* (2021) 'A use case study comparing augmented reality (AR) and electronic document-based maintenance instructions considering tasks complexity and operator competency level', *Virtual Reality*, 25(4), pp. 999–1014. Available at: <https://doi.org/10.1007/s10055-020-00493-z>.

## CESI LINEACT

Ojer, M. *et al.* (2020) 'Projection-Based Augmented Reality Assistance for Manual Electronic Component Assembly Processes', *Applied Sciences*, 10(3), p. 796. Available at: <https://doi.org/10.3390/app10030796>.

Sjarov, M. *et al.* (2020) 'The Digital Twin Concept in Industry – A Review and Systematization', in *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*. 2020, pp. 1789–1796. Available at: <https://doi.org/10.1109/ETFA46521.2020.9212089>.

Xu, K., Ragot, N. and Dupuis, Y. (2022) 'View Selection for Industrial Object Recognition', in *Proceedings of the 48th Annual Conference of the Industrial Electronics Society. Annual Conference of the Industrial Electronics Society*, Brussels.

Yang, Z., Yu, X. and Yang, Y. (2021) 'DSC-PoseNet: Learning 6DoF Object Pose Estimation via Dual-scale Consistency'. arXiv. Available at: <https://doi.org/10.48550/arXiv.2104.03658>.

Zakharov, S., Shugurov, I. and Ilic, S. (2019) 'DPOD: 6D Pose Object Detector and Refiner', in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea (South): IEEE, pp. 1941–1950. Available at: <https://doi.org/10.1109/ICCV.2019.00203>.