



HAL
open science

WIM-Drone : Combiner un Monde-En-Miniature avec des Microgestes pour Faciliter la Téléopération en Réalité Augmentée

Charles Bailly, Julien Castet

► To cite this version:

Charles Bailly, Julien Castet. WIM-Drone : Combiner un Monde-En-Miniature avec des Microgestes pour Faciliter la Téléopération en Réalité Augmentée. 2023. hal-04014967

HAL Id: hal-04014967

<https://hal.science/hal-04014967v1>

Preprint submitted on 5 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

WIM-Drone : Combiner un Monde-En-Miniature avec des Microgestes pour Faciliter la Téléopération en Réalité Augmentée

WIM-Drone: Combining a World-In-Miniature and Micro-Gestures for Teleoperation in Augmented Reality

CHARLES BAILLY, Immersion, France

JULIEN CASTET, Immersion, France

Teleoperation in industrial contexts require operators to understand and visualize in real-time the remote workspace. It is nonetheless challenging to provide independent and mobile viewpoints to remote operators. This paper explores a new approach to tackle this challenge: Augmented Reality (AR) to control a drone used as aerial camera. More precisely, WIM-Drone combines a World-In-Miniature (WIM) in AR with (micro)gestures to control a drone. We present the conception and experimental evaluation of four interaction techniques based on this approach. Results from studies 1 and 2 suggest that WIM-Drone can be a promising approach for teleoperation scenario. Besides, Study 2 suggest that participants avoided collisions between the drone and humans or moving objects, but often ignored collisions with static objects.

CCS Concepts: • **Human-centered computing** → **Mixed / augmented reality**; **Gestural input**.

Additional Key Words and Phrases: Augmented Reality, Mixed Reality, World-In-Miniature, Micro-gestures, Hand gestures

Lors d'une téléopération en milieu industriel, l'opérateur a besoin de comprendre et de visualiser en temps réel l'espace de travail distant. Il est cependant complexe de fournir à un opérateur un point de vue indépendant et contrôlable à distance. Cet article explore une nouvelle approche pour pallier cette difficulté : l'utilisation de Réalité Augmentée (RA) pour contrôler à distance un drone servant de caméra aérienne. WIM-Drone combine un Monde-En-Miniature en RA avec des (micro)gestes pour contrôler le drone. Nous présentons quatre techniques d'interaction basées sur cette approche. Les résultats des deux évaluations expérimentales suggèrent que WIM-Drone est un concept prometteur pour la téléopération. Par ailleurs, l'Étude 2 suggère que les participants cherchent à éviter toute collision entre le drone et des objets en mouvement ou des humains. En revanche, les participants ont tendance à ignorer les collisions avec les objets statiques.

Mots-clés additionnels : Réalité Augmentée, Réalité Mixte, Monde-en-Miniature, Microgestes, Gestes manuels

Reference:

This is the author's draft version of the work. It is posted here for your personal use. Not for distribution. The definitive version of record was published in IHM '23: Proceedings of the 34th Conference on l'Interaction Humain-Machine.

ISBN 978-1-4503-9825-1

<https://doi.org/10.1145/3583961.3583972>

1 INTRODUCTION

Lors d'une téléopération en milieu industriel, il est primordial pour l'opérateur de bien comprendre l'environnement distant. Par exemple, un opérateur devant inspecter à distance une installation doit pouvoir l'observer sous plusieurs angles tout en évitant les obstacles présents dans l'usine (techniciens, machines, etc). Il est néanmoins difficile de fournir à un utilisateur distant le moyen de percevoir et naviguer en temps réel au sein d'un tel environnement. Par souci de clarté, nous utiliserons tout au long de cet article les termes *opérateur* pour désigner l'utilisateur à distance et *technicien* pour désigner l'utilisateur sur site.

L'une des solutions courante dans les scénarios d'assistance à distance consiste à se servir de la caméra d'un équipement

2023. Manuscript submitted to ACM

mobile (smartphone, tablette, etc). Néanmoins, cela oblige un technicien à tenir l'appareil à la main afin de montrer l'environnement à l'opérateur [8]. Utiliser un casque de Réalité Augmentée (RA) évite ce problème et libère les mains du technicien. Cette approche réduit néanmoins la marge de manoeuvre de l'opérateur, dont le point de vue reste dépendant de la position de la tête du technicien portant le casque. Cette limitation est présente même en montant une caméra 360° sur le casque du technicien [37]. Inspirés par les systèmes tels que TeleAdvisor [11], nous explorons dans cet article un concept centré sur une caméra indépendante et contrôlable à distance lors d'une téléopération. Pour cela, nous combinons un drone servant de caméra aérienne à un système de RA. Ce concept est illustré en Figure 1. Nous nous basons sur un Monde-En-Miniature (WIM en anglais, noté MeM dans cet article), qui est une réplique réduite d'un environnement [35]. Dans notre cas, nous utilisons une copie miniature en RA d'une salle des machines. Cette représentation a l'avantage d'être en 3 dimensions et de fournir une vue aérienne globale de l'environnement où vole le drone. Via des gestes et micro-gestes (mouvements de la main sur la main [32, 41]), l'opérateur peut contrôler la réplique virtuelle du drone au sein du MeM. A la manière d'un jumeau numérique [14], tout mouvement de cette réplique virtuelle du drone est reflétée sur le drone réel, qui offre alors un nouveau point de vue de l'espace de travail à l'opérateur.

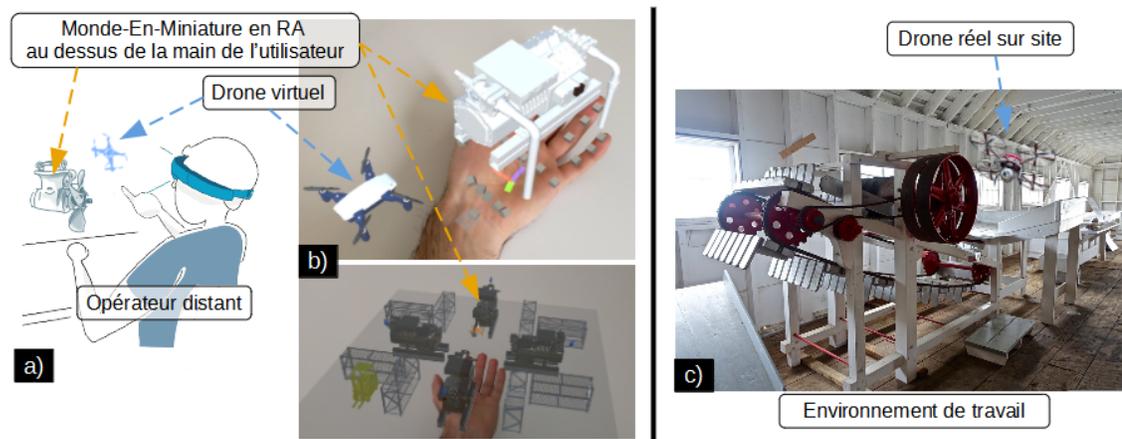


FIG. 1. Concept de WIM-Drone. a) Un opérateur distant portant un casque de RA utilise un Monde-en-Miniature (MeM) virtuel pour contrôler à distance un drone. b) Le MeM en RA, simple (haut) ou complexe (bas), suit en temps réel la main non-directrice de l'utilisateur. c) Environnement de travail distant où vole le drone réel.

WIM-Drone est directement inspiré d'un scénario industriel où un superviseur expérimenté (l'opérateur) vient aider à distance un autre technicien. L'opérateur n'est donc pas toujours dans un bureau classique, mais plutôt en situation de mobilité dans une autre partie de l'usine ou une autre usine. La RA avec un casque et des (micro)gestes permettent de préserver cette mobilité sans occuper les mains de l'opérateur. De plus, la RA permet de travailler avec un Monde-en-Miniature en 3 dimensions de la salle des machines plutôt que de travailler sur un écran 2D. Notre objectif est de tirer parti des avantages d'une visualisation en 3D stéréoscopique [1, 5] afin de faciliter le plus possible la compréhension de l'environnement et donc le pilotage du drone.

Les contributions de cet article sont les suivantes :

- Le concept de WIM-Drone ainsi que l'espace de conception correspondant. A partir de cet espace de conception, nous proposons quatre techniques d'interaction combinant Monde-En-Miniature en RA et gestes/micro-gestes pour contrôler un drone.
- L'évaluation de ces techniques au travers de deux études expérimentales. L'Étude 1 se concentre sur les performances et l'utilisabilité des techniques. L'Étude 2 se concentre sur les stratégies des participants pour éviter toute collision entre le drone et les obstacles présents dans l'espace de travail.

En tant que première exploration de WIM-Drone, nous nous focalisons dans cet article sur un drone virtuel et un environnement distant simulé. Notre but est d'évaluer le potentiel de WIM-Drone dans des conditions optimales (stabilité et précision du drone) maîtrisées de bout en bout. Après une revue de l'état de l'art pour positionner nos travaux par rapport aux études existants, nous présentons la conception de nos techniques d'interaction. Nous détaillons ensuite les deux études expérimentales et leurs résultats. Enfin, nous dressons un bilan des limitations et contributions de nos travaux.

2 REVUE DES TRAVAUX PRÉCÉDENTS

2.1 Percevoir un espace de travail distant

Young et al. ont identifié 3 exigences clés pour les systèmes de téléprésence [40]. Tout d'abord, le système doit transmettre des indices de communication tels que la voix et les gestes. Ensuite, le système doit fournir une vue partagée de l'espace de travail local en préservant les relations spatiales entre les utilisateurs et les objets. Enfin, le système doit fournir des points de vue indépendants aux différents utilisateurs [40]. Dans cet article, nous nous concentrons sur les exigences 2 et 3. Les travaux sur les indices de communication (exigence 1) tels que le regard [17] ne font pas partie du cadre de nos travaux.

Plusieurs approches basées sur la Réalité Mixte (RM, [34]) ont déjà été explorées pour permettre à un opérateur de voir et de naviguer de manière autonome dans un espace de travail distant. Gauglitz et al. ont proposé l'utilisation d'un smartphone pour offrir une vue à la première personne de l'espace de travail local en temps réel [8]. Un technicien peut enregistrer plusieurs points de vue auxquels l'opérateur pourra accéder via une interface graphique 2D. Mohr et al. ont exploré un concept similaire pour l'assistance à distance [22]. Les auteurs proposent notamment une procédure collaborative pour capturer l'espace de travail via des images spatialement enregistrées (*light fields*). Cependant, de telles approches mobilisent un technicien sur site pour tenir et déplacer le smartphone. Cela n'est pas toujours possible dans le cadre d'une téléopération, ni souhaitable lorsque le technicien doit déjà effectuer d'autres tâches manuelles comme une réparation ou une tâche d'assemblage [26]. L'utilisation d'un casque de RM permet de libérer les mains du technicien. Par exemple, le système proposé par Teo et al. permet à l'opérateur de faire pivoter le point de vue grâce à une vidéo 360° captée en temps réel par le technicien [37]. L'opérateur contrôle donc la rotation de point de vue mais pas sa position, qui reste liée au casque du technicien. Cette limitation représente un frein important pour l'opérateur (caméra instable, mouvements imprévisibles, etc). Les appareils portés à l'épaule [15, 29] souffrent également de ce problème. Pour donner un point de vue indépendant à l'opérateur, Gurevich et al. ont proposé le système TeleAdvisor [11]. Monter une caméra sur un bras robotisé et télécommandé permet à l'opérateur d'explorer l'espace de travail de manière indépendante. De plus, ce bras robotique est lui-même monté sur une plateforme à roues, ce qui permet de compenser le manque de mobilité des systèmes de projection traditionnels. En utilisant une configuration similaire, Lanir et al. ont suggéré que laisser le contrôle du point de vue à l'opérateur pourrait être plus bénéfique pour des tâches de construction [16].

Positionnement de nos travaux : Un système de RA qui reprend les avantages de TeleAdvisor [11], mais basé sur la caméra aérienne d'un drone.

2.2 Combiner drones et Réalité Mixte

Contrairement au système au sol comme TeleAdvisor [11], les drones peuvent fournir des vues aériennes de leur environnement. Plusieurs études tirent parti de cet avantage pour combiner drone et Réalité Mixte pour la téléopération [12, 39] ou l'entraînement au pilotage [9]. Hitchcock et al. ont proposé un système de RM où l'utilisateur pouvait basculer entre une vue au sol à la première personne et une vue aérienne à la troisième personne pour naviguer dans un labyrinthe [13]. Liu et Shen ont présenté un système auquel un drone participe en temps réel à la cartographie de l'environnement 3D. La carte 3D est reconstruite à partir de la vidéo du drone et affichée dans une interface de RA [20]. Piumsoboon et al. ont combiné un drone et de la RM pour étudier la téléprésence pour une collaboration à distance multi-échelle [28].

Dans cet article, nous nous concentrons sur le contrôle à distance du point de vue. Piloter un drone reste une opération délicate, en particulier en milieu professionnel ou dans un environnement complexe [21]. Les travaux de Hitchcock et al. sont axés sur un contrôleur classique [13]. Le système Squadrone proposé par Sabet et al. est conçu pour une collaboration à distance synchrone entre plusieurs utilisateurs en Réalité Virtuelle (RV), mais s'appuie également sur des contrôleurs [30]. Cauchard et al. ont combiné un menu projeté en RA avec des gestes de la main, mais ce système nécessite que l'utilisateur soit co-localisé avec le drone [3]. Pittman et LaViola ont étudié la rotation de la tête et les gestes de la tête comme techniques d'entrée pour contrôler un drone. Les auteurs ont néanmoins observé que la plupart des participants préféraient utiliser un contrôleur traditionnel [27]. Enfin, Chen et al. ont proposé des techniques d'interaction en RA sur smartphone pour contrôler un drone via son ombre projetée [4].

Positionnement de nos travaux : Contrairement à plusieurs travaux antérieurs [3, 4], le cas d'usage que nous visons implique que l'opérateur n'est pas co-localisé avec le drone. Nous nous concentrons également sur des techniques d'interaction en RA permettant de déplacer à distance le drone en utilisant un MeM et des microgestes.

2.3 Utiliser un Monde-En-Miniature et des microgestes

La métaphore du MeM [35] offre de nombreux avantages pour le domaine de la téléopération. En effet, un MeM peut donner une vue aérienne globale de l'espace de travail distant afin d'en faciliter la compréhension et la navigation [6]. Les Mondes-En-Miniature ont notamment été étudiés pour naviguer au sein d'un environnement à grande échelle en RV [38]. Mulloni et al. ont proposé des vues d'un MeM affichées sur un téléphone mobile pour aider un utilisateur à naviguer dans un bâtiment [23].

Les micro-gestes, c'est à dire les gestes de la main sur la main [18, 32, 41], sont une alternative prometteuse aux gestes traditionnellement utilisés en RM. Ils ont le potentiel de réduire la fatigue et la distraction tout en augmentant la confidentialité de l'utilisateur lors de l'interaction [18]. La reconnaissance précise des micro-gestes soulève encore des problématiques techniques [24, 31]. Néanmoins, de nombreuses études sont centrées sur ces défis techniques et permettent d'envisager une reconnaissance fiable de micro-gestes [19, 33]. Par exemple, le système *Soli* de Lien et al. permet de détecter le type de mouvement, l'amplitude et la vitesse d'un geste manuel avec une précision de l'ordre du millimètre [19].

Positionnement de nos travaux : Selon l'espace de conception de Danyluk et al. [6], le MeM que nous avons conçu peut être classé comme étant de *Petite* taille (20 cm), à l'échelle d'un *Objet* ou d'une *Salle* et avec des *Simplifications géométriques*. Par ailleurs, comme pour [2], nous voulions éviter tout dispositif mobile pour afficher le MeM. Ce dernier

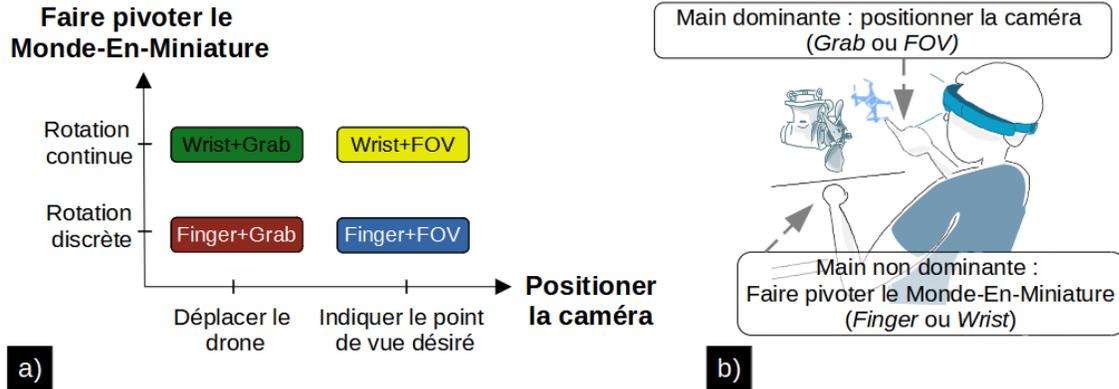


FIG. 2. a) Espace de conception lié à WIM-Drone et positionnement des quatre techniques d'interaction qui en découlent. b) Chaque main de l'opérateur est liée à un axe de l'espace de conception.

est donc affiché via un casque de RA et a pour référence une *Partie du corps de l'utilisateur* [6]. En effet, notre MeM est affiché au dessus de la main non-directrice de l'opérateur. Par ailleurs, nous nous intéressons à l'interaction bi-manuelle via des gestes de la main et micro-gestes. A notre connaissance, il n'existe actuellement pas de technique d'interaction alliant de tels gestes à un MeM en RA.

3 CONCEPTION DES TECHNIQUES D'INTERACTION

Notre exploration du concept de WIM-Drone est basée sur deux besoins de l'opérateur distant : le besoin de comprendre l'espace de travail (organisation dans l'espace, position des obstacles, etc) et le besoin de changer de point de vue au sein de cet espace. Ces deux besoins se reflètent dans les éléments clé de WIM-Drone : le MeM en RA et les gestes pour positionner la caméra du drone. Nous proposons ainsi un espace de conception basé sur deux axes : 1) le contrôle de l'orientation du MeM et 2) le contrôle de la caméra du drone (position et rotation). Cet espace de conception est illustré à la Figure 2a.

L'action correspondant à chaque axe de cet espace de conception est liée à l'une des deux mains de l'utilisateur (Figure 2b). La nouveauté des techniques d'interaction que nous proposons vient de la combinaison bi-manuelle de gestes et micro-gestes autour d'un MeM en RA.

3.1 Faire pivoter le Monde-En-Miniature

Le premier axe de notre espace de conception concerne la rotation du MeM en RA. Pour faciliter la compréhension de l'environnement 3D distant, nous avons conçu deux gestes de la main : un microgeste discret appelé *Finger* et un geste continu du poignet appelé *Wrist*. En effet, nous voulions comparer un geste rapide et ponctuel (*Finger*) à un geste continu permettant un contrôle potentiellement plus fin (*Wrist*). Comme expliqué ci-dessous, nous avons associé le geste ponctuel à des rotations discrètes du MeM et le geste continu à des rotations continues du MeM.

Avec *Finger*, l'opérateur peut déclencher des rotations discrètes du MeM affiché au dessus de sa main non-dominante (Figure 3). Pour cela, l'utilisateur peut plier l'index (seconde et troisième phalanges pliées à plus de 135°) pour déclencher une rotation de -90° du MeM. A l'inverse, plier l'annulaire déclenche une rotation de +90°. A chaque fois, la rotation est animée pendant 0.5s pour faciliter la compréhension de la rotation et du microgeste. De plus, l'opérateur peut aussi

utiliser deux doigts pour déclencher ces rotations (index+majeur ou annulaire+auriculaire) si cela est plus confortable physiquement.

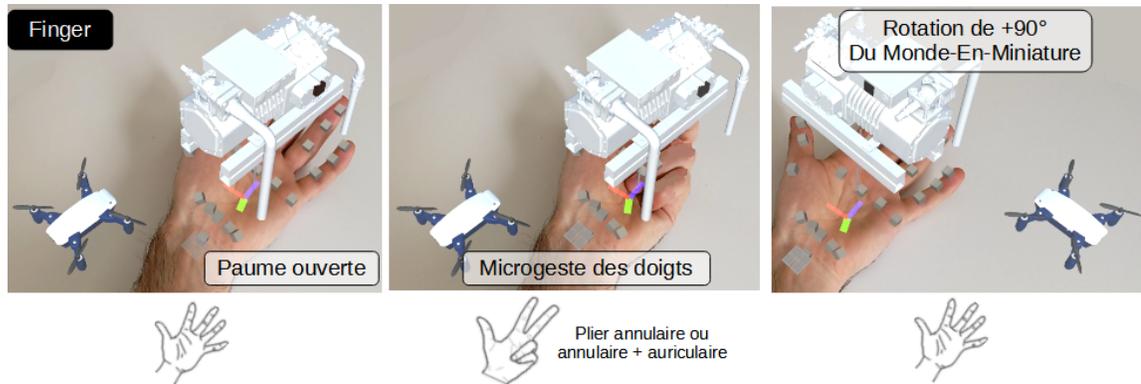


FIG. 3. *Finger* : Un microgeste avec un ou deux doigts pliés pour déclencher une rotation discrète du Monde-En-Miniature. Ici, une rotation de $+90^\circ$ est déclenchée. Une rotation de -90° peut être déclenchée en pliant l'index (ou index + majeur).

Au contraire, *Wrist* repose sur un geste continu du poignet (Figure 4). Après avoir fermé le poing, l'utilisateur contrôle en temps réel l'orientation du MeM en faisant tourner le poignet. Ainsi, tourner le poignet (gauche) vers la droite fait pivoter sur lui-même le MeM dans le sens horaire, et inversement. Cette approche reprend ainsi la métaphore *Object-in-hand* de Sundin et Fjeld [36], mais avec un MeM et un geste main libre. La position de départ (poignet aligné avec l'avant-bras) correspond à l'orientation de base (rotation de 0°). Pour pouvoir effectuer une rotation à 360° tout en restant dans une zone confortable pour le poignet (120° d'amplitude du geste max), toute rotation du poignet est multipliée par trois au niveau du MeM. L'un de nos objectifs était de comparer des rotations discrètes et continues. Ainsi, *Wrist* se veut aussi continu que possible. Seules les très petites rotations du poignet ($\leq 3^\circ$) ont été filtrées pour compenser les tremblements de la main et améliorer la stabilité du MeM.

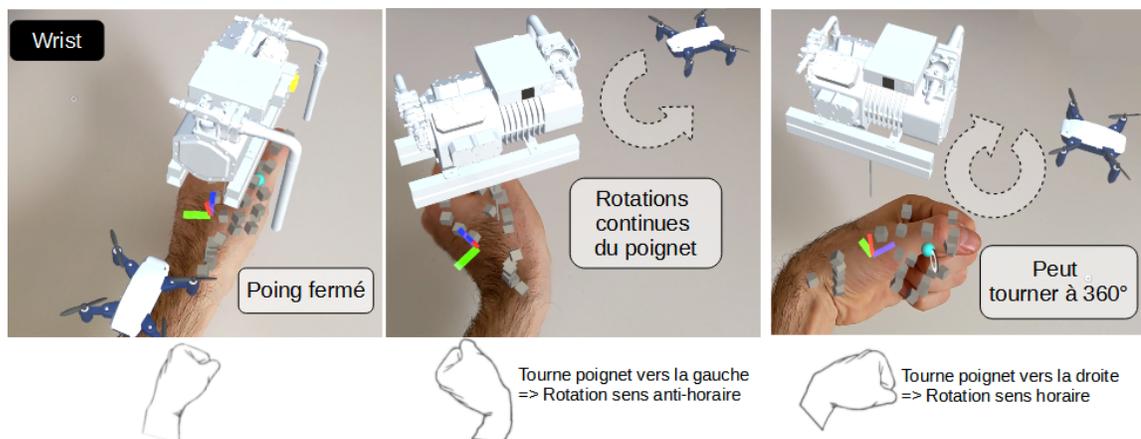


FIG. 4. *Wrist* : Un geste basé sur une rotation continue du poignet. Le Monde-En-Miniature "suit" le sens de rotation du poignet.

3.2 Positionner la caméra

Le second axe de l'espace de conception se concentre sur le contrôle du positionnement de la caméra. Inspirés par le concept de jumeau numérique, nous voulions que l'opérateur puisse interagir avec le drone virtuel affiché au sein du MeM afin de déplacer le drone réel et par conséquent changer de point de vue. Nous avons conçus deux gestes à effectuer avec la main dominante : *Grab* et *FOV*. Le premier, *Grab*, s'inspire de la réplique virtuelle proposée par Oda et al. [25]. Avec cette méthode, l'opérateur saisit une copie fantôme du drone virtuel dans le MeM grâce à un geste de pincement (*pinch gesture*). Ce geste est illustré en Figure 5). Un geste de glisser-déposer permet ensuite d'amener la copie fantôme vers l'emplacement désiré (position + orientation). Lorsque la copie fantôme est relâchée, le drone virtuel vient rejoindre cet emplacement. *Grab* permet un geste continu et donne l'impression de définir la trajectoire du drone.

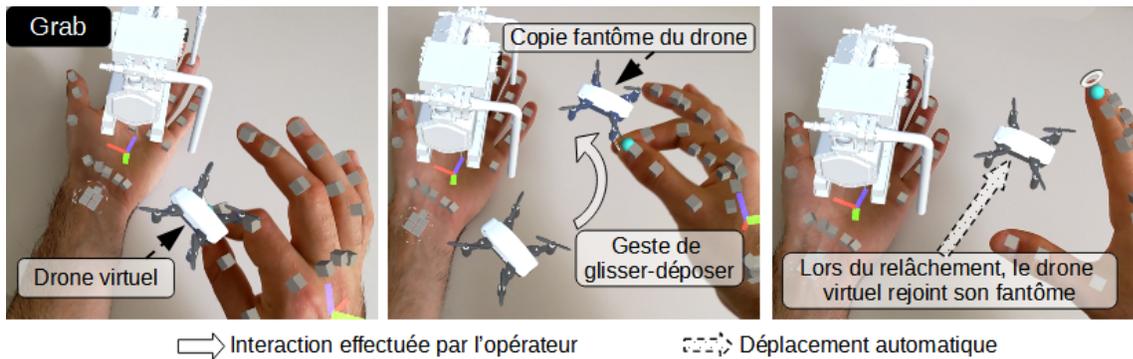


FIG. 5. *Grab* : Glisser-déposer une copie fantôme du drone pour changer de point de vue.

Nous voulions également explorer une autre métaphore où l'utilisateur pourrait directement indiquer le point de vue souhaité. Avec *FOV*, une forme pyramidale est attachée au pouce de la main dominante de l'opérateur (Figure 6). Cette forme représente le champ de vision (*Field of View*, ou *FOV*) de la caméra du drone. L'utilisateur peut pointer à un endroit donné avec le pouce et effectuer un geste de pincement pour déplacer automatiquement le drone virtuel vers cette position (position + orientation du pouce). Nous avons choisi d'utiliser le pouce comme point d'ancrage car il est beaucoup plus statique que l'index lors d'un geste de pincement, ce qui améliore la précision de l'emplacement indiqué.

Globalement, *Grab* peut se résumer à "j'amène la caméra à cette position" tandis que *FOV* correspond à "je veux voir la scène de ce point de vue". Dans les deux cas, lorsque l'opérateur déclenche un mouvement de drone, une animation de 2s s'affiche au sein du MeM pour simuler un mouvement de drone réel.

3.3 Techniques d'interaction et implémentation

Comme illustré en Figure 2a, nous avons conçu une technique d'interaction par croisement possible de nos deux axes de conception. Ainsi, nous proposons 4 techniques : *Finger+Grab*, *Finger+FOV*, *Wrist+Grab* et *Wrist+FOV*. Toutes ces techniques utilisent le même geste de paume ouverte pour faire apparaître le MeM au dessus de la main non-dominante. Nos techniques d'interaction sont des techniques bimanuelles mais il n'y a pas d'ordre imposé entre le contrôle du drone virtuel et le changement d'orientation du MeM. Les deux interactions peuvent être effectuées en simultanément ou séquentiellement. Il est important de noter que cela impacte *Grab*. En effet, tant que la copie fantôme du drone est

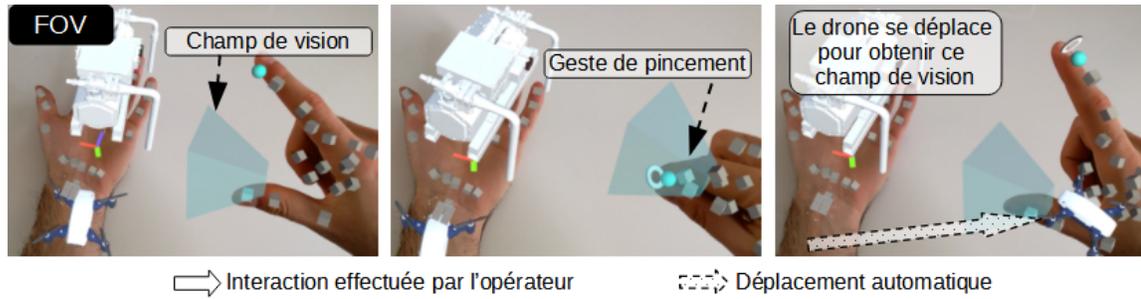


FIG. 6. *FOV* : l'opérateur définit directement un nouveau point de vue. Un geste de pincement amène automatiquement le drone vers la nouvelle position et orientation du pouce de l'opérateur.

attrapée par l'opérateur, elle ne sera pas impactée par les rotations du MeM. L'objectif est d'encourager l'opérateur à faire pivoter le MeM plutôt que de faire pivoter la copie fantôme via des positions de poignet inconfortables.

Toutes les techniques ont été implémentées en C# avec Unity 2020 (version 3.5). Nous avons utilisé le casque Hololens 2 de Microsoft pour la RA et des scripts personnalisés basés sur le plugin MRTK (version 2) pour détecter les gestes et microgestes de la main. Comme indiqué en Section 1, nous nous sommes concentrés sur un drone virtuel en simulant l'espace de travail distant.

4 ÉTUDE 1 : UTILISABILITÉ DES TECHNIQUES

Nous avons mené deux études expérimentales autour du concept de WIM-Drone et des techniques détaillées ci-dessus. L'Étude 1 se concentre sur l'utilisabilité de nos techniques d'interaction. En effet, nous voulions vérifier que ces techniques étaient adaptées à des utilisateurs non-experts dans le pilotage de drone. Nos hypothèses de départ étaient les suivantes :

- **H1** : Effectuer des rotations continues du MeM (*Wrist*) est plus rapide et plus apprécié que de déclencher des rotations discrètes.
- **H2** : Glisser-déposer le drone (*Grab*) est plus précis mais plus lent que directement définir un nouveau point de vue pour le drone (*FOV*).

Par conséquent, nous nous attendions à ce que *Wrist+FOV* soit la technique avec les meilleurs retours quantitatifs et qualitatifs. Les participants jouaient le rôle d'un opérateur à distance essayant d'inspecter une grande machine industrielle grâce à un drone. 12 participants (6 femmes, 6 hommes, 10/12 droitiers) âgés de 21 à 55 ans (sd = 9.1) ont été recrutés pour cette étude (voir Figure 7 pour les détails démographiques).

4.1 Installation expérimentale

L'étude s'est déroulée dans une pièce calme et bien éclairée. Les participants étaient assis sur une chaise de bureau avec accoudoirs devant une table (75 cm de haut), comme le montre la Figure 8. Devant les participants (à 1.5m devant leur tête), un écran virtuel montrait l'espace de travail simulé (une salle des machines). Lorsqu'il était déclenché, le Monde-En-Miniature apparaissait à 10 cm au-dessus de la paume de l'utilisateur. Ce MeM était composé d'un modèle miniature de machine (un générateur diesel), d'un modèle miniature du drone et d'un menu en RA avec trois boutons (*Suivant*, *Correct* et *Invalide*). Trois points d'intérêt (POI) étaient positionnés sur la machine (un sur le côté gauche, un à

Démographie des participants



FIG. 7. Démographie des participants de l'Étude 1. Aucun participant n'était familier avec la notion de microgeste.

l'arrière et un sur le côté droit). Sur le MeM, les positions des POI étaient marquées par des carrés noirs représentant un écran. Sur le modèle de la salle des machines, les POI avaient les mêmes emplacements mais affichaient également un nombre représentant une température comprise entre 40°C et 61°C.

4.2 Tâche

Une tâche d'inspection de machine était divisée en 4 étapes : 1) Démarrer la tâche en appuyant sur le bouton *Suivant* et faire apparaître le MeM en effectuant le geste de paume ouverte (main non dominante). 2) Identifier quel POI est la cible actuelle. Pour ce faire, les participants devaient trouver le POI surligné en jaune dans le MeM. 3) Déplacer le drone virtuel (et continuer à faire tourner le WIM si besoin) pour pouvoir vérifier la température du POI ciblé. 4) Déterminer si la température était correcte (entre 40 et 50°) ou invalide (au delà de 50°). Ensuite, terminer la tâche en appuyant sur le bouton correspondant (bouton *Correct* ou *Invalide*).

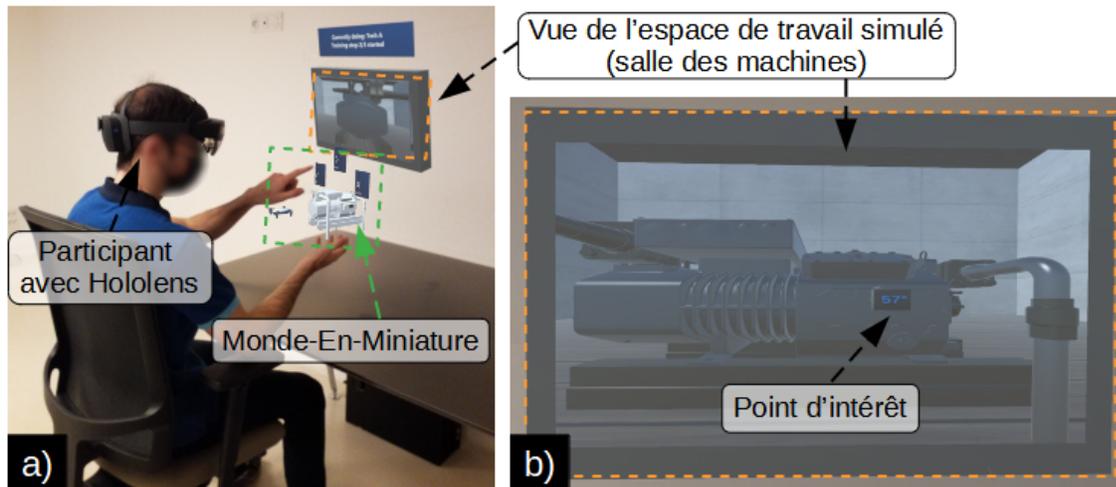


FIG. 8. a) Aperçu de l'installation expérimentale. b) Vue égo-centrée de l'écran virtuel montrant la salle des machines simulée.

4.3 Conception de l'étude

Après les explications et le consentement initial du participant, l'expérimentateur montrait l'installation expérimentale et le MeM au participant. L'objectif de cette étape était de permettre aux participants de se familiariser avec le modèle de machine et l'emplacement des trois POIs. Ensuite, pour chacune des quatre techniques, les participants ont effectué une phase d'entraînement (3 tâches d'inspection, non enregistrées) suivie d'une phase réelle (9 tâches d'inspection enregistrées). Après chaque phase réelle, les participants pouvaient faire une pause avant de répondre à quelques questions sur la technique qu'ils venaient d'essayer. Dans l'ensemble, les participants ont effectué un total de 36 (9x4) tâches d'inspection enregistrées. L'ordre des POI ciblés était aléatoire mais chaque POI était visité le même nombre de fois (3 fois par phase d'essai, d'où les 9 tâches d'inspection par phase d'essai). Les températures étaient aléatoires (probabilité de 0.5 d'être correctes) pour chaque tâche d'inspection. De plus, l'ordre des techniques était équilibré entre les participants grâce à un carré latin.

4.4 Mesures

Nous avons enregistré plusieurs mesures de temps, des données sur les rotations du MeM (nombre de rotations avec *Finger* et rotation angulaire totale avec *Wrist*) ainsi que les réponses des utilisateurs sur les températures des POI. Afin de recueillir des retours qualitatifs, la version française du questionnaire SUS [10] a été remise aux participants après chaque phase enregistrée. A la fin de l'étude, il était demandé aux participants de classer les 4 techniques selon plusieurs critères (facilité d'utilisation, rapidité, précision, appréciation globale). Un entretien semi-structuré était ensuite mené pour obtenir des commentaires supplémentaires sur le concept de WIM-Drone et sur les techniques.

4.5 Résultats

Pour mener à bien l'analyse statistique des résultats, nous avons utilisé des ANOVA à mesures répétées (avec $\alpha = 0.05$) suivies de tests post hoc (*t-tests* pairés) avec des corrections de Bonferroni. Lorsqu'un test non paramétrique était plus approprié, nous avons utilisé un test de Friedman suivi d'un test post hoc de Wilcoxon avec corrections de Bonferroni. Comme suggéré par Dragicevic [7], nous mettons l'accent sur les intervalles de confiance (IC) et les tailles d'effet dans la présentation des résultats. Nous faisons également le choix d'un langage nuancé pour limiter toute interprétation binaire. Dans toutes les figures, les barres d'erreur sont des intervalles de confiance à 95%.

4.6 Résultats quantitatifs

Comme l'illustre la Figure 9a, nous avons observé de forts indices de différence de performance entre les techniques ($F(3, 33) = 8, 6, p = 0, 002$). Il semble clair que les participants aient été plus rapides avec *Wrist+FOV* qu'avec *Finger+Grab* ($p = 0, 01$) ou *Wrist+Grab* ($p = 0, 045$). De même, les participants semblent avoir été plus rapides avec *Finger+FOV* qu'avec *Finger+Grab* ($p = 0, 025$). Nous n'avons trouvé que de faibles indices de différence entre *Finger+FOV* et *Wrist+Grab* ($p = 0, 14$) et aucune preuve de différence dans les autres cas. Globalement, les tâches d'inspection réalisées avec des techniques basées sur *FOV* ont été réalisées en environ 15,5 secondes, contre 22,5 secondes en moyenne avec des techniques basées sur *Grab*.

Pour comprendre cette différence de temps de complétion, nous avons examiné de plus près le nombre de fois où les participants ont déplacé le drone virtuel. Nous avons observé qu'en moyenne, il y avait une bonne preuve de différence entre les techniques ($X^2 = 14, 7, p = 0, 002$). Les participants semblent avoir déplacé moins de fois le drone avec *Finger+FOV* (1,32 fois en moyenne) qu'avec *Finger+Grab* (2,16 mouvements en moyenne, $p = 0, 025$) ou *Wrist+Grab*

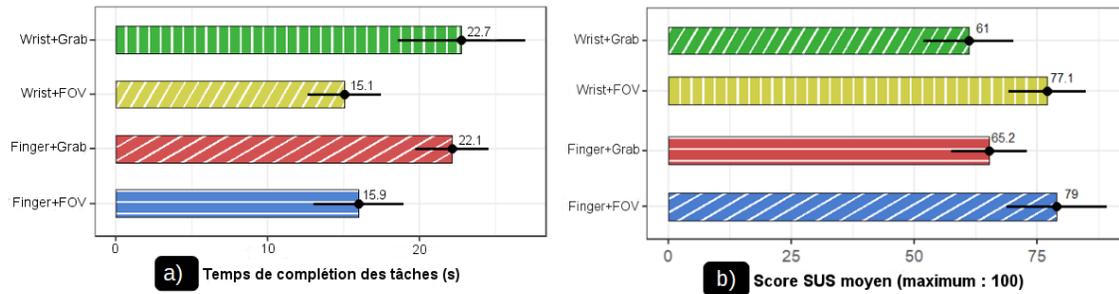


FIG. 9. a) Temps de complétion moyen d'une tâche d'inspection pour chaque technique. b) Score SUS moyen pour chaque technique. Les barres d'erreurs sont les IC à 95%.

(2,37 coups en moyenne, $p = 0,041$). Nous n'avons observé que de faibles indices de différence entre *Finger+FOV* et *Wrist+FOV* (1,64 mouvements en moyenne pour ce dernier, $p = 0,14$). Aucune preuve claire de différence n'a été trouvée dans les autres cas (grand chevauchement des IC).

De plus, nous avons également trouvé une bonne preuve de différence concernant les rotations du Monde-En-Miniature ($X^2 = 23,3$, $p = 3e - 5$). Nous avons observé que les participants effectuaient un total de rotations de près de 580° en moyenne avec *Wrist+Grab*. L'analyse statistique suggère qu'il existe une nette différence entre *Finger+FOV* (204° en moyenne, $p = 0,015$) et *Finger+Grab* (308° en moyenne, $p = 0,02$). Par ailleurs, nous avons également observé une bonne preuve de différence entre *Finger+FOV* et *Wrist+FOV* (454° en moyenne, $p = 0,02$).

Comme le montre la Figure 9b, les résultats du questionnaire SUS suggèrent que *Finger+FOV* et *Wrist+FOV* ont une bonne utilisabilité. Ils ont respectivement un score moyen de 79 et 77, ce qui est supérieur au palier de 71/100 [10]. Les résultats pour les deux autres techniques sont plus mitigés avec un score moyen autour de 65 pour *Finger+Grab* et 61 pour *Wrist+Grab*. Ces observations sont étayées par notre analyse statistique. En effet, nous avons observé une preuve convaincante de différence entre les techniques ($X^2 = 15,1$, $p = 2e - 3$). Lors de l'exécution de tests post hoc, nous n'avons trouvé que des preuves claires de la différence entre *Wrist+Grab* et *Wrist+FOV*.

4.7 Résultats qualitatifs

Lors d'entretiens semi-structurés, les participants ont exprimé des opinions divergentes sur *Finger* et *Wrist*. 6/12 participants ont préféré les techniques basées sur *Finger*. Par exemple, P9 a comparé le geste du doigt à : "un clic où la machine fait un tour fixe" et a trouvé cette technique plus facile à utiliser. P1 et P3 ont également mentionné le sentiment d'être plus en contrôle avec cette approche. Au contraire, 6/12 participants ont préféré utiliser des rotations continues avec *Wrist*. P7 a expliqué que : "il était évident de faire tourner le moteur" et a trouvé cette approche plus rapide. Trois de ces participants ont également signalé que les gestes des doigts étaient moins confortables physiquement à effectuer. P7 et P8 ont ajouté que *Finger* était contraignant et un peu frustrant par rapport à *Wrist*.

Nous avons observé un consensus plus clair sur la comparaison de *FOV* et *Grab*. 8/12 participants ont préféré *FOV*, 3 ont préféré *Grab* et 1 était indécis. Les participants préférant *FOV* ont expliqué que ce geste était plus simple, intuitif et plus facile à utiliser. 3 d'entre eux ont eu quelques difficultés à saisir le drone virtuel et à le repositionner rapidement, ce qui n'est pas arrivé avec *FOV*. Parmi ceux qui ont préféré *Grab*, P7 a mentionné que : "vous attrapez le drone et il vous suit. Avoir ce retour immédiat me donne un sentiment de précision". Les deux autres ont également mentionné qu'ils trouvaient la manipulation directe du drone virtuel plus agréable. Enfin, *Finger+FOV* a été la technique classée favorite le plus

Enseignements de l'Étude 1	
WIM-Drone	Combinaison de Monde-En-Miniature, de drone et microgestes bien accueillie par les participants Deux techniques proposées (<i>Finger+FOV</i> et <i>Wrist+FOV</i>) ont un score d'utilisabilité prometteur.
Faire pivoter le Monde-En-Miniature	Déclencher des rotations discrètes (<i>Finger</i>) est plus lent et plus fatigant physiquement. En revanche, cette approche est plus simple à contrôler que des rotations continues (<i>Wrist</i>).
Positionner la caméra	Directement indiquer un nouveau point de vue (<i>FOV</i>) est plus rapide et préféré par une majorité de participants

FIG. 10. Résumé des enseignements tirés de l'Étude 1.

de fois (5/12 participants l'ont classée première). *Finger+Grab* a été classé comme le moins apprécié le plus souvent (5/12 participants l'ont classé en dernier). La distinction entre le deuxième et le troisième favori n'était pas claire entre *Wrist+Grab* et *Wrist+FOV*.

4.8 Leçons apprises

Le concept général de WIM-Drone a reçu des retours positifs. Les participants ont compris l'idée d'utiliser un drone virtuel pour une tâche d'inspection à distance et l'intérêt d'un Monde-En-Miniature pour visualiser l'environnement distant. L'interaction bi-manuelle (pouvoir faire tourner le MeM et déplacer la caméra) a également été bien accueillie par l'ensemble des participants. Ces résultats sont encourageants pour la combinaison de MeM en RA et de drone pour des scénarios de téléopération. Un bilan des enseignements tirés de l'Étude 1 est disponible en Figure 10.

4.8.1 Faire pivoter le WIM. Nos hypothèses de départ ne se sont qu'en partie vérifiées. Nous nous attendions à ce qu'un contrôle continu de la rotation du Monde-En-Miniature (*Wrist*) soit plus rapide et plus apprécié que de déclencher des rotations discrètes une par une. Les résultats qualitatifs remettent en cause H1 et suggèrent que les participants ont des avis partagés. Nous pensions que les participants avec moins d'expérience en RA préféreraient utiliser les rotations discrètes de *Finger* plutôt que *Wrist*, mais nous n'avons pas observé ce phénomène. Les deux approches offrent des avantages et limitations différentes en terme de rapidité, de sensation de contrôle et de confort physique qui semblent hiérarchisées différemment selon les participants. D'autres travaux sont requis pour approfondir l'amplitude des rotations du MeM selon la disposition spatiale des points d'intérêt ainsi que la fatigue physique liée aux microgestes utilisées dans cet étude.

4.8.2 Contrôler la caméra. Les résultats expérimentaux semblent soutenir l'hypothèse H2. Indiquer directement un nouveau point de vue (*FOV*) s'est avéré plus rapide et plus apprécié des participants que glisser-déposer le drone virtuel (*Grab*). Les deux techniques basées sur (*FOV*) ont également obtenus de meilleurs scores d'utilisabilité. Ce constat est néanmoins à nuancer par rapport à la qualité du système de détection des gestes manuels. Les limitations techniques de l'Hololens ont en effet pu influencer certains participants, notamment pour attraper et glisser-déposer le drone virtuel avec *Grab*.

5 ÉTUDE 2 : GESTION DES OBSTACLES DANS L'ENVIRONNEMENT DISTANT

Dans l'Étude 1, nous nous sommes concentrés sur l'utilisabilité des techniques d'interaction à l'aide d'un MeM. Cependant, de nombreux environnements ont des caractéristiques beaucoup plus complexes, notamment des murs, des objets et des êtres humains. Un drone doit éviter de tels obstacles lors de son vol. Cependant, les utilisateurs interagissant avec le MeM en RA interagissent avec un drone virtuel non affecté par une quelconque collision. Cette différence

Technique	Illustration	Manipulation du drone	Description	Hypothèses
Contrôleur (référence)		Indirecte, synchronisée	Manette de Xbox pour contrôler le drone Joysticks pour la position, flèches pour l'orientation Pas de Monde-En-Miniature, seulement une carte 2D de l'espace de travail	Technique la plus rapide Charge mentale plus importante pour trouver le point d'intérêt, plus d'erreurs lors de l'inspection
Wrist* ou Finger* +Grab		Directe, asynchrone	Attraper le drone, puis glisser-déposer sa copie fantôme Pas d'automatisation de la trajectoire. Le drone peut traverser les obstacles, créant des collisions virtuelles et physiques	Sentiment de contrôle Les participants ont plus de chances d'ignorer les obstacles et les collisions
Wrist* ou Finger* +Grab+Traj		Directe, asynchrone	Attraper le drone, puis glisser-déposer sa copie fantôme La trajectoire prévisionnelle du drone s'affiche en temps réel. Les collisions physiques sont évitées automatiquement	Sentiment de réalisme grâce à l'adaptation de la trajectoire Des collisions virtuelles sont encore possibles
Wrist* ou Finger* +FOV+Traj		Automatisée	Geste de pincement pour que le drone vienne automatiquement se placer (position + orientation du pouce). La trajectoire prévisionnelle du drone s'affiche en temps réel. Les collisions physiques sont évitées automatiquement	Sentiment de réalisme grâce à l'adaptation de la trajectoire Aucune collision Faible charge mentale grâce à l'automatisation complète

FIG. 11. Les techniques évaluées lors de l'Étude 2.

est au cœur de l'Étude 2. En particulier, nous avons voulu approfondir deux questions principales. Premièrement, les utilisateurs franchiraient-ils des obstacles virtuels ou éviteraient-ils activement les collisions lors des interactions avec le MeM? Deuxièmement, quel serait l'impact d'une adaptation automatique de trajectoire du drone (c'est-à-dire simulant un véritable évitement d'obstacle) sur les participants? Pour explorer ces questions, il est important de distinguer trois types de collisions :

- Les **collisions physiques** ont lieu lorsque le drone réel percute un obstacle dans l'espace de travail. Cela pourrait arriver si le drone ne dispose pas de sécurité concernant les obstacles. L'adaptation automatique de la trajectoire cherche à éviter ces collisions physiques, par exemple en forçant le drone à passer par un autre chemin pour arriver à destination.
- Les **collisions simulées** sont presque identiques aux collisions physiques. Néanmoins, au lieu d'avoir lieu dans un espace de travail réel, elles sont provoquées dans un espace de travail virtuel simulé, comme ce que nous utilisons dans l'Étude 1 et 2.
- Les **collisions virtuelles** n'ont lieu qu'au sein du MeM entre des objets virtuels. Par exemple, l'opérateur utilisant *Wrist*+Grab* peut attraper la copie fantôme du drone et lui faire traverser un mur virtuel afin de gagner du temps. Sans adaptation de la trajectoire, cette collision virtuelle créerait une collision physique lorsque le vrai drone suit le mouvement indiqué. Avec une adaptation automatique de la trajectoire, le vrai drone suivrait un chemin différent de celui désigné par l'opérateur pour atteindre sa destination sans collision physique.

La Figure 11 illustre les techniques incluses dans l'Étude 2 pour étudier ces deux types de collisions. Nous avons tout d'abord choisi de nous servir d'une technique de référence basée sur une télécommande, qui est le dispositif classique pour contrôler un drone. Ensuite, nous avons implémenté des variantes de nos techniques d'interaction pour répondre aux observations de l'Étude 1. *Wrist** est ainsi une variante de *Wrist* qui cherche à améliorer la sensation de contrôle du MeM. Cette approche utilise des paliers de 15° de rotation au lieu des 3° originaux. De plus, *Finger** remplace le plissement des doigts de *Finger*, microgeste jugé peu confortable par certains participants. A la place, l'utilisateur vient appuyer sur l'angle de son pouce pour déclencher une rotation de +45°, +90° ou +180° avec l'index, le majeur ou l'annulaire respectivement. Ces techniques alternatives ont été évaluées lors d'une étude pilote dédiée avec 3 participants.

Nous avons recruté 8 participants droitiers de 22 à 43 ans ($sd = 8,5$) pour participer à l'Étude 2. Les détails démographiques sont disponibles en Figure 12. Le protocole de l'Étude 2 partage de nombreuses similitudes avec le protocole de l'Étude 1 (installation expérimentale, méthodes d'analyse statistique, etc). **Ci-après, seules les différences sont rapportées.**

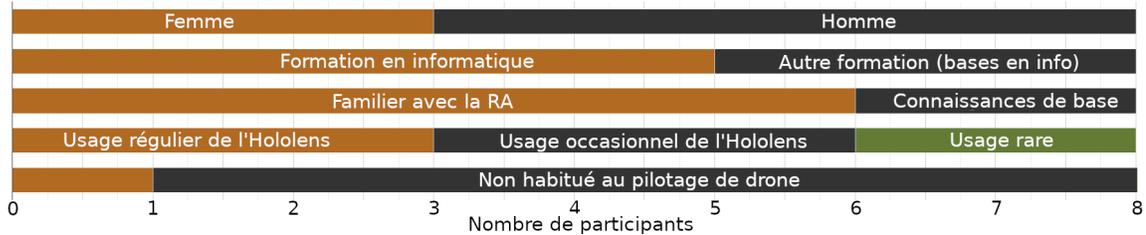


FIG. 12. Démographie des participants de l'Étude 2.

5.1 Configuration et tâche

L'une des différences entre les deux études provient de la complexité de l'environnement industriel simulé. Contrairement à l'Étude 1 où la salle des machines était simplifiée, nous avons ici utilisé un environnement de salle des machines plus complet et réaliste avec différents types d'obstacles. Comme le montre la Figure 13a, le MeM reflétant cet environnement était plus grand, plus détaillé et plus réaliste.

Les participants ont été invités à effectuer une tâche d'inspection similaire à l'Étude 1 en vérifiant la température indiquée sur l'une des quatre machines dans l'espace de travail distant. Une flèche bleue s'affichait sur la carte 2D (condition de référence) ou dans le MeM (autres conditions) pour indiquer le point d'intérêt ciblé. Chaque tâche d'inspection comprenait un type d'obstacle : objet statique, objet en mouvement, humain statique et humain en mouvement (voir Figure 13b). Dans la condition *objet statique*, seuls les objets fixes dans la pièce (machines et étagères) étaient présents. Dans les trois autres conditions d'obstacle, un obstacle supplémentaire (chariot élévateur en mouvement, groupe de travailleurs statique/en mouvement) était présent dans l'un des deux couloirs menant au point d'intérêt actuel (voir Figure 13a). L'autre couloir vers la cible était toujours libre d'obstacle. Pour implémenter la trajectoire d'évitement d'obstacles, nous avons adapté un GridGraph du projet A* Pathfinding ¹.

5.2 Conception de l'étude et mesures

Les participants avaient pour consigne d'utiliser le système comme s'ils contrôlaient un vrai drone dans un environnement distant. La moitié des participants a utilisé (*Contrôleur*, *Wrist*+Grab* sans adaptation de trajectoire, *Wrist*+Grab* avec adaptation de trajectoire et *Wrist*+FOV*). L'autre moitié des participants a testé la même chose avec *Finger** : (*Controller*, *Finger*+Grab* sans adaptation de trajectoire, *Finger*+Grab* avec adaptation de trajectoire et *Finger*+FOV*). Dans l'ensemble, les participants ont effectué 32 (8 tâches * 4 techniques testées) tâches d'inspection enregistrées. L'ordre des techniques était équilibré entre les participants. L'ordre des conditions d'obstacles pour chaque technique était aléatoire mais équilibré. En plus des mesures déjà prises lors de l'étude 1, nous avons également enregistré le nombre et la durée des collisions entre le drone et chaque type d'obstacle.

¹<https://arongranberg.com/astar/features>

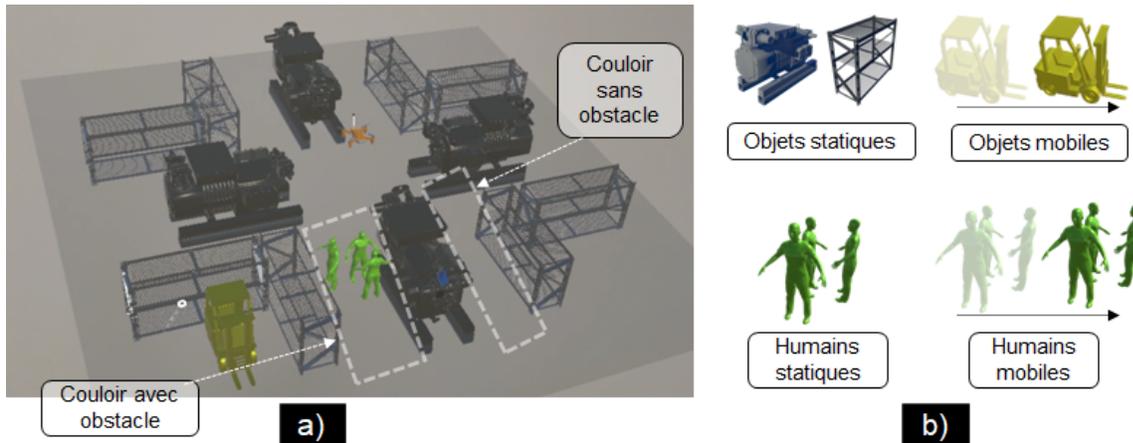


FIG. 13. a) Aperçu du monde-En-Miniature plus complexe utilisé lors de l'Étude 2. b) Les 4 types d'obstacles utilisés dans l'Étude 2.

5.3 Résultats

5.3.1 Comparaison avec la technique de référence. Les participants ont été plus rapides en utilisant *Contrôleur*, avec une moyenne de 18,3 secondes pour effectuer une tâche d'inspection. La méthode pour faire pivoter le MeM, *Wrist** ou *Finger**, ne semble pas avoir eu d'influence notable sur le temps de complétion. Les participants ont été plus rapides avec *Contrôleur* qu'avec *Wrist*+Grab* ($p = 0,004$), *Wrist*+Grab+Traj* ($p = 0,01$), *Finger*+Grab* ($p = 0,009$) ou *Finger*+Grab+Traj* ($p = 0,012$). Nous n'avons trouvé que de faibles preuves de différence entre *Contrôleur* et *Wrist*+FOV+Traj* ($p = 0,16$) et *Finger*+FOV+Traj* ($p = 0,21$). En revanche, 7/8 participants ont été désorientés au moins une fois lors d'une tâche d'inspection avec *Contrôleur* et sa carte 2D. Par exemple, P2 a mentionné : "j'étais perdu au début, et je croisais les doigts en espérant que c'était la bonne machine". Cela s'est produit même si P2 a expérimenté la condition *Contrôleur* en dernier et était donc familier avec l'espace de travail simulé. Au contraire, tous les participants ont déclaré avoir apprécié le MeM et n'ont signalé aucune désorientation lors de son utilisation.

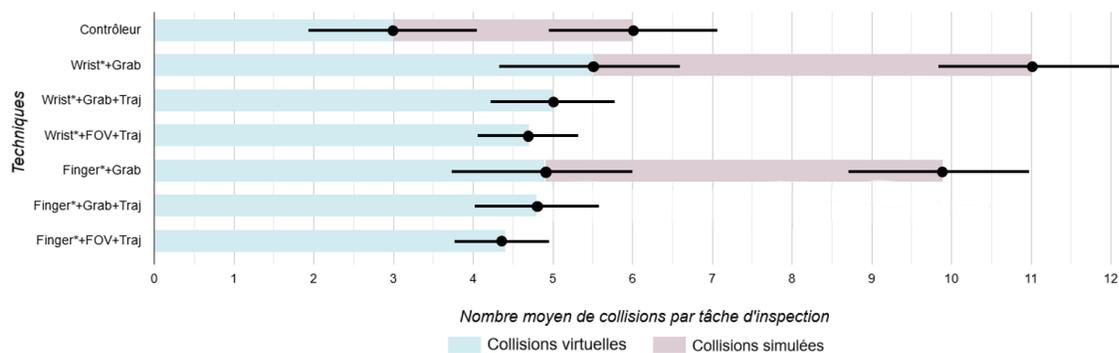


FIG. 14. Nombre moyen de collisions virtuelles et physiques lors des tâches d'inspection de l'Étude 2. Il est important de noter que pour la plupart des techniques, une collision virtuelle entraînait obligatoirement une collision simulée.

5.3.2 Gestion des obstacles et collisions. Concernant le nombre de collisions entre le drone et les obstacles (Figure 14), nous n'avons observé des preuves de différence statistique que entre *Contrôleur* et les autres techniques. Là encore, nous n'avons pas observé d'effet notable de la technique pour faire pivoter le MeM. Indépendamment des techniques, environ 97% des collisions se sont produites avec des objets statiques ou des murs. De manière surprenante, nous n'avons observé pratiquement aucune collision avec des objets en mouvement, des humains statiques ou des humains en mouvement. La hitbox du drone était volontairement agrandie pour simuler une marge de sécurité (un vrai drone ne devrait pas frôler les obstacles). De plus, les couloirs de la salle des machines (Figure 11a) étaient volontairement étroits pour dissuader les participants de tenter de se faufiler entre les obstacles. La combinaison de ces deux facteurs a entraîné un nombre important de collisions simulées dans notre étude, là où un vrai drone serait simplement passé (trop) près d'un obstacle. Enfin, les obstacles dynamiques étaient positionnés dans les couloirs comme les objets statiques, restaient dans ces couloirs lors de leurs (faibles) déplacements et nous avons revérifié a posteriori que les collisions avaient été correctement enregistrées.

Les participants ont rapporté des stratégies mitigées concernant les obstacles. 6/8 participants sont passés au-travers d'éléments virtuels du MeM lors de l'utilisation de *Grab*. Ils ont expliqué qu'ils n'avaient pas prêté attention aux collisions. Certains de ces participants étaient enclins à les ignorer après avoir vu une première fois le drone traverser des objets pour atteindre sa destination. P6 et P8 ont indiqué qu'ils se sentaient "*moins en sécurité*" à propos de ce comportement, qui "*diminuait la confiance et le réalisme*". Au contraire, voir la trajectoire du drone a été apprécié par tous les participants. P3 et P7 ont reporté qu'ils faisaient confiance au drone pour éviter les collisions, mais qu'ils avaient explicitement traversé des objets virtuels au sein du MeM pour gagner du temps. Seul 1/8 participant a systématiquement essayé d'éviter toute collisions, indépendamment de la technique.

5.3.3 Variantes des techniques d'interaction. Nous n'avons pas observé de différence marquante au niveau des résultats quantitatifs entre *Finger** et *Wrist**. Les retours qualitatifs suggèrent que *Finger** est plus confortable physiquement que *Finger*. Par ailleurs, aucun participant n'a reporté de sentiment de manque de contrôle avec *Wrist**, ce qui est encourageant par rapport à *Wrist*.

6 LIMITATIONS ET LEÇONS APPRISSES

Nos travaux présentent plusieurs limitations. Tout d'abord, nos techniques restent en partie limitées par la précision de la détection des gestes avec l'Hololens 2. De plus, notre Monde-En-Miniature offrait une marge de manoeuvre limitée pour le drone (couloirs étroits, espaces libres réduits), ce qui a pu augmenter artificiellement le nombre de collisions. Par ailleurs, l'Étude 2 combine de nombreux facteurs, ce qui limite la généralisation de certains résultats. Le niveau de réalisme du MeM ainsi que l'influence du positionnement des obstacles au sein de l'environnement restent ainsi à étudier de manière plus approfondie dans de futurs travaux. Enfin, il serait intéressant d'explorer d'autres correspondances entre (micro)gestes et rotation du MeM, en associant par exemple un microgeste discret (*Finger*) à des rotations continues. Nous proposons néanmoins plusieurs enseignements tirés de cette seconde étude.

6.0.1 Le Monde-En-Miniature limite la désorientation. Les résultats supérieurs de la technique *Contrôleur* ne sont pas surprenants. Tout d'abord, tous les participants avaient une expérience non négligeable dans l'utilisation d'une manette, essentiellement via des jeux vidéo. De plus, nous avons conçu *Contrôleur* comme une référence parfaite, sans les problèmes classiques du pilotage de drone (stabilité du drone, latence, inertie du freinage, vent...). Néanmoins, 7/8 participants ont été désorientés au moins une fois lors d'une tâche d'inspection avec *Contrôleur*. Après la consultation initiale de la carte 2D, les participants ont été contraints de se concentrer uniquement sur la vue à la première personne

de l'espace de travail. Cette concentration peut influencer à la fois l'orientation spatiale et la détection d'obstacles, notamment d'obstacles mobiles. Nous pensons que cette vue à la première personne a favorisé une sous-estimation du volume du drone, ce qui expliquerait qu'il y ait eu des collisions avec cette technique. Même dans un environnement simple et un contrôle parfait du drone, avoir une carte 2D était moins préféré qu'avoir un MeM 3D.

6.0.2 Automatiser la trajectoire du drone n'a éliminé que les collisions simulées. Comme attendu, l'automatisation de la trajectoire du drone (techniques avec *Traj*) a permis d'éliminer toute collision simulée. En revanche, cette automatisation n'a pas réduit le nombre de collisions virtuelles. Une majorité de participants n'ont pas pris en compte ou ont délibérément ignoré les collisions lors de l'utilisation de *Grab*, même si cela affectait leur confiance dans le système. Certains utilisateurs comme P3 et P7 ont délibérément traversé des éléments virtuels du MeM lors de l'affichage de la trajectoire du drone. Ils ont expliqué que grâce à la trajectoire, ils étaient convaincus que le drone pouvait éviter de lui-même les obstacles et s'efforçaient donc d'être aussi efficaces que possible. Concernant les techniques avec *FOV*, une explication possible est que les participants ont créé des collisions virtuelles avec des machines lors de la définition d'un nouveau point de vue. 3/8 participants ont rapporté que *FOV* n'était parfois pas assez précise, notamment à cause du mouvement du pouce lors du geste de pincement.

6.0.3 Les participants ont évité les collisions avec les humains et les objets en mouvement. Cependant, même lorsqu'ils ne faisaient pas attention ou ignoraient délibérément les collisions potentielles, les participants ont toujours évité les collisions avec les objets en mouvement et les humains. Cette observation est particulièrement intéressante pour les techniques WIM-Drone. Cela implique que, intentionnellement ou non, les participants ont toujours choisi un chemin sans obstacle (par exemple, un couloir libre), même s'il était plus long. En outre, même des participants comme P3 et P7, qui ont délibérément traversé des obstacles virtuels pour aller plus vite, ont tout de même évité les objets en mouvement et les humains statiques ou en mouvement. Cela suggère que certains participants établissent des priorités parmi les types d'obstacles. Ils ont considéré que les objets statiques au sein du MeM (machines, étagères, murs) pouvaient être ignorés pour gagner du temps, mais que les objets en mouvement et les humains devaient être évités.

7 CONCLUSION

Cet article contribue à résoudre un problème difficile en téléopération : permettre à l'utilisateur distant d'avoir un point de vue indépendant sur la scène. Dans notre étude, nous proposons plusieurs techniques d'interaction combinant un Monde-En-Miniature (MeM) en Réalité Augmentée et des micro-gestes pour contrôler ce point de vue. L'utilisateur peut faire pivoter le MeM et déplacer un drone virtuel pour inspecter à distance un objet d'intérêt. Les résultats expérimentaux de l'Étude 1 suggèrent que WIM-Drone est une approche prometteuse pour la téléopération dont le concept a été apprécié par les participants. Pointer directement pour indiquer un nouveau point de vue est plus rapide et préféré dans notre étude pour de glisser-déposer le drone virtuel. L'Étude 2 a mis en évidence que les participants adoptent des stratégies différentes pour éviter ou non les obstacles présents dans l'espace de travail distant. Les participants ont notamment évité toute collision avec des objets en mouvement et des humains, mais ont souvent ignoré les collisions avec les objets statiques.

Dans l'ensemble, nous espérons que nos travaux encourageront une exploration plus approfondie des concepts et du potentiel du WIM-Drone. Notre prochaine étape consiste à évaluer WIM-Drone dans un contexte collaboratif d'assistance à distance avec un véritable drone.

REMERCIEMENTS

This project has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No 101016608.

RÉFÉRENCES

- [1] Ikpe Justice Akpan and Murali Shanker. 2018. A comparative evaluation of the effectiveness of virtual reality, 3D visualization and 2D visual interactive simulation : an exploratory meta-analysis. *SIMULATION* (Feb. 2018), 003754971875703. <https://doi.org/10.1177/0037549718757039>
- [2] Blaine Bell, Tobias Höllerer, and Steven Feiner. 2002. An annotated situation-awareness aid for augmented reality. In *Proceedings of the 15th annual ACM symposium on User interface software and technology* (New York, NY, USA, 2002-10-27) (*UIST '02*). Association for Computing Machinery, 213–216. <https://doi.org/10.1145/571985.572017>
- [3] Jessica R. Cauchard, Alex Tamkin, Cheng Yao Wang, Luke Vink, Michelle Park, Tommy Fang, and James A. Landay. 2019. Drone.io : A Gestural and Visual Interface for Human-Drone Interaction. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Daegu, Korea (South), 2019-03). IEEE, 153–162. <https://doi.org/10.1109/HRI.2019.8673011>
- [4] Linfeng Chen, Kazuki Takashima, Kazuyuki Fujita, and Yoshifumi Kitamura. 2021. PinpointFly : An Egocentric Position-Control Drone Interface Using Mobile AR. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (*CHI '21*). Association for Computing Machinery, New York, NY, USA, Article 150, 13 pages. <https://doi.org/10.1145/3411764.3445110>
- [5] Alex Dan and Miriam Reiner. 2017. EEG-based cognitive load of processing events in 3D virtual worlds is lower than processing events in 2D displays. *International Journal of Psychophysiology* 122 (Dec. 2017), 75–84. <https://doi.org/10.1016/j.ijpsycho.2016.08.013>
- [6] Kurtis Danyluk, Barrett Ens, Bernhard Jenny, and Wesley Willett. 2019. A Design Space Exploration of Worlds in Miniature. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2021-05-06) (*CHI '21*). Association for Computing Machinery, 1–15. <https://doi.org/10.1145/3411764.3445098>
- [7] Pierre Dragicjevic. 2016. Fair statistical communication in HCI. In *Modern Statistical Methods for HCI*. Springer, 291–330. https://doi.org/10.1007/978-3-319-26633-6_13
- [8] Steffen Gauglitz, Benjamin Nuernberger, Matthew Turk, and Tobias Höllerer. 2014. In touch with the remote world : remote collaboration with augmented reality drawings and virtual navigation. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology - VRST '14* (Edinburgh, Scotland). ACM Press, 197–205. <https://doi.org/10.1145/2671015.2671016>
- [9] Yong-Guk Go, Jong-Won Lee, Ho-San Kang, and Soo-Mi Choi. 2020. Interactive Training of Drone Flight Control in Mixed Reality. In *SIGGRAPH Asia 2020 XR* (Virtual Event Republic of Korea, 2020-12-04). ACM, 1–2. <https://doi.org/10.1145/3415256.3421496>
- [10] Guillaume Gronier and Alexandre Baudet. 2021. Psychometric Evaluation of the F-SUS : Creation and Validation of the French Version of the System Usability Scale. *International Journal of Human-Computer Interaction* 37, 16 (Oct. 2021), 1571–1582. <https://doi.org/10.1080/10447318.2021.1898828>
Publisher : Taylor & Francis_eprint : <https://doi.org/10.1080/10447318.2021.1898828>
- [11] Pavel Gurevich, Joel Lanir, and Benjamin Cohen. 2015. Design and Implementation of TeleAdvisor : a Projection-Based Augmented Reality System for Remote Collaboration. 24, 6 (2015), 527–562. <https://doi.org/10.1007/s10606-015-9232-7>
- [12] Hooman Hedayati, Michael Walker, and Daniel Szafir. 2018. Improving Collocated Robot Teleoperation with Augmented Reality. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (Chicago IL USA, 2018-02-26). ACM, 78–86. <https://doi.org/10.1145/3171221.3171251>
- [13] Aaron Hitchcock and Kelvin Sung. 2018. Multi-view augmented reality with a drone. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (Tokyo Japan). ACM, 1–2. <https://doi.org/10.1145/3281505.3283397>
- [14] David Jones, Chris Snider, Aydin Nassehi, Jason Yon, and Ben Hicks. 2020. Characterising the Digital Twin : A systematic literature review. 29 (2020), 36–52. <https://doi.org/10.1016/j.cirpj.2020.02.002>
- [15] T. Kurata, N. Sakata, M. Kourogi, H. Kuzuoka, and M. Billinghurst. 2004. Remote collaboration using a shoulder-worn active camera/laser. In *Eighth International Symposium on Wearable Computers* (2004-10), Vol. 1. 62–69. <https://doi.org/10.1109/ISWC.2004.37> ISSN : 1530-0811.
- [16] Joel Lanir, Ran Stone, Benjamin Cohen, and Pavel Gurevich. 2013. Ownership and control of point of view in remote assistance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2013-04-27) (*CHI '13*). Association for Computing Machinery, 2243–2252. <https://doi.org/10.1145/2470654.2481309>
- [17] Khanh-Duy Le, Ignacio Avellino, Cédric Fleury, Morten Fjeld, and Andreas Kunz. 2019. GazeLens : Guiding Attention to Improve Gaze Interpretation in Hub-Satellite Collaboration. In *Human-Computer Interaction – INTERACT 2019*, David Lamas, Fernando Loizides, Lennart Nacke, Helen Petrie, Marco Winckler, and Panayiotis Zaphiris (Eds.). Springer International Publishing, Cham, 282–303.
- [18] Guangchuan Li, David Rempel, Yue Liu, Weitao Song, and Carisa Harris Adamson. 2021. Design of 3D Microgestures for Commands in Virtual Reality or Augmented Reality. *Applied Sciences* 11, 14 (Jan. 2021), 6375. <https://doi.org/10.3390/app11146375>
- [19] Jaime Lien, Nicholas Gillian, M. Emre Karagozler, Patrick Amihood, Carsten Schwesig, Erik Olson, Hakim Raja, and Ivan Poupyrev. 2016. Soli : ubiquitous gesture sensing with millimeter wave radar. *ACM Transactions on Graphics* 35, 4 (July 2016), 142 :1–142 :19. <https://doi.org/10.1145/2897824.2925953>

- [20] Chuhao Liu and Shaojie Shen. 2020. An Augmented Reality Interaction Interface for Autonomous Drone. (2020). arXiv:2008.02234 <http://arxiv.org/abs/2008.02234>
- [21] Sara Ljungblad, Yemao Man, Mehmet Aydin Baytaş, Mafalda Gamboa, Mohammad Obaid, and Morten Fjeld. 2021. What Matters in Professional Drone Pilots' Practice? An Interview Study to Understand the Complexity of Their Work and Inform Human-Drone Interaction Research. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 159, 16 pages. <https://doi.org/10.1145/3411764.3445737>
- [22] Peter Mohr, Shohei Mori, Tobias Langlotz, Bruce H. Thomas, Dieter Schmalstieg, and Denis Kalkofen. 2020. Mixed Reality Light Fields for Interactive Remote Assistance. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu HI USA, 2020-04-21). ACM, 1–12. <https://doi.org/10.1145/3313831.3376289>
- [23] Alessandro Mulloni, Hartmut Seichter, and Dieter Schmalstieg. 2012. Indoor navigation with mixed reality world-in-miniature views and sparse localization on mobile devices. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (New York, NY, USA, 2012-05-21) (AVI '12). Association for Computing Machinery, 212–215. <https://doi.org/10.1145/2254556.2254595>
- [24] Viet Nguyen, Siddharth Rupavatharam, Luyang Liu, Richard Howard, and Marco Gruteser. 2019. HandSense : capacitive coupling-based dynamic, micro finger gesture recognition. In *Proceedings of the 17th Conference on Embedded Networked Sensor Systems* (New York New York, 2019-11-10). ACM, 285–297. <https://doi.org/10.1145/3356250.3360040>
- [25] Ohan Oda, Carmine Elvezio, Mengu Sukan, Steven Feiner, and Barbara Tversky. 2015. Virtual Replicas for Remote Assistance in Virtual and Augmented Reality. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (New York, NY, USA, 2015-11-05) (UIST '15). Association for Computing Machinery, 405–415. <https://doi.org/10.1145/2807442.2807497>
- [26] S. K. Ong, M. L. Yuan, and A. Y. C. Nee. 2015. Augmented reality applications in manufacturing : a survey. 46, 10 (2015), 2707–2742. <https://doi.org/10.1080/00207540601064773>
- [27] Corey Pittman and Joseph J. LaViola. 2014. Exploring head tracked head mounted displays for first person robot teleoperation. In *Proceedings of the 19th international conference on Intelligent User Interfaces* (Haifa Israel, 2014-02-24). ACM, 323–328. <https://doi.org/10.1145/2557500.2557527>
- [28] Thammathip Piumsomboon, Gun A. Lee, Barrett Ens, Bruce H. Thomas, and Mark Billinghurst. 2018. Superman vs Giant : A Study on Spatial Perception for a Multi-Scale Mixed Reality Flying Telepresence Interface. 24, 11 (2018), 2974–2982. <https://doi.org/10.1109/TVCG.2018.2868594>
- [29] Thammathip Piumsomboon, Gun A. Lee, Andrew Irlitti, Barrett Ens, Bruce H. Thomas, and Mark Billinghurst. 2019. On the Shoulder of the Giant : A Multi-Scale Mixed Reality Collaboration with 360 Video Sharing and Tangible Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow Scotland Uk, 2019-05-02). ACM, 1–17. <https://doi.org/10.1145/3290605.3300458>
- [30] Mehrnaz Sabet, Mania Orand, and David W. McDonald. 2021. Designing Telepresence Drones to Support Synchronous, Mid-air Remote Collaboration : An Exploratory Study. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama Japan, 2021-05-06). ACM, 1–17. <https://doi.org/10.1145/3411764.3445041>
- [31] Yu Sang, Laixi Shi, and Yimin Liu. 2018. Micro Hand Gesture Recognition System Using Ultrasonic Active Sensing. *IEEE Access* 6 (2018), 49339–49347. <https://doi.org/10.1109/ACCESS.2018.2868268> Conference Name : IEEE Access.
- [32] Adwait Sharma, Joan Sol Roo, and Jürgen Steimle. 2019. Grasping Microgestures : Eliciting Single-Hand Microgestures for Handheld Objects. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3290605.3300632>
- [33] Toby Sharp, Cem Keskin, Duncan Robertson, Jonathan Taylor, Jamie Shotton, David Kim, Christoph Rhemann, Ido Leichter, Alon Vinnikov, Yichen Wei, Daniel Freedman, Pushmeet Kohli, Eyal Krupka, Andrew Fitzgibbon, and Shahram Izadi. 2015. Accurate, Robust, and Flexible Real-time Hand Tracking. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15). Association for Computing Machinery, New York, NY, USA, 3633–3642. <https://doi.org/10.1145/2702123.2702179>
- [34] Maximilian Speicher, Brian D. Hall, and Michael Nebeling. 2019. What is Mixed Reality?. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–15. <https://doi.org/10.1145/3290605.3300767>
- [35] Richard Stoakley, Matthew J. Conway, and Randy Pausch. 1995. Virtual reality on a WIM : interactive worlds in miniature. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '95). ACM Press/Addison-Wesley Publishing Co., USA, 265–272. <https://doi.org/10.1145/223904.223938>
- [36] Martin Sundin and Morten Fjeld. 2009. Softly Elastic 6 DOF Input. *International Journal of Human-Computer Interaction* 25, 7 (Sept. 2009), 647–691. <https://doi.org/10.1080/10447310902964124>
- [37] Theophilus Teo, Gun A. Lee, Mark Billinghurst, and Matt Adcock. 2019. Investigating the use of Different Visual Cues to Improve Social Presence within a 360 Mixed Reality Remote Collaboration*. In *The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry* (Brisbane QLD Australia, 2019-11-14). ACM, 1–9. <https://doi.org/10.1145/3359997.3365687>
- [38] Samuel Truman and Sebastian von Mammen. 2020. An Integrated Design of World-in-Miniature Navigation in Virtual Reality. In *International Conference on the Foundations of Digital Games* (New York, NY, USA, 2020-09-15) (FDG '20). Association for Computing Machinery, 1–9. <https://doi.org/10.1145/3402942.3402994>
- [39] Michael E. Walker, Hooman Hedayati, and Daniel Szafrir. 2019. Robot Teleoperation with Augmented Reality Virtual Surrogates. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Daegu, Korea (South), 2019-03). IEEE, 202–210. <https://doi.org/10.1109/HRI.2019.8673306>
- [40] Jacob Young, Tobias Langlotz, Matthew Cook, Steven Mills, and Holger Regenbrecht. 2019. Immersive Telepresence and Remote Collaboration using Mobile and Wearable Devices. 25, 5 (2019), 1908–1918. <https://doi.org/10.1109/TVCG.2019.2898737> Conference Name : IEEE Transactions on

Visualization and Computer Graphics.

- [41] Klen Čopič Pucihar, Christian Sandor, Matjaž Kljun, Wolfgang Huerst, Alexander Plopski, Takafumi Taketomi, Hirokazu Kato, and Luis A. Leiva. 2019. The Missing Interface : Micro-Gestures on Augmented Objects. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow Scotland Uk, 2019-05-02). ACM, 1–6. <https://doi.org/10.1145/3290607.3312986>