

Table des matières

1	Introduction	11
1.1	Contexte	12
1.2	Problématiques et objectifs	13
1.3	Modèle d'entité mixte	14
1.4	Structure du manuscrit	14
2	Etat de l'art	17
2.1	Réalité mixte, interfaces utilisateur naturelles et internet des objets	17
2.1.1	Préambule	17
2.1.2	Réalité mixte	20
2.1.3	Interfaces utilisateur naturelles	21
2.1.4	Internet des objets	25
2.2	Convergence de la réalité mixte, des interfaces utilisateur intuitives et de l'internet des objets	29
2.2.1	Kits de développement logiciel	30
2.2.2	Cas d'usage et prototypes	30
2.2.3	NUI interconnectées et naturalité de l'interaction utilisateur	33
2.3	Modèles et frameworks de conception ou de développement	40
2.4	Analyse	42
3	Modèle interne mixte orienté conception	47
3.1	Entités mixtes	48
3.1.1	Interface de l'entité mixte	49
3.1.2	Modèle interne mixte orienté conception d'une entité mixte	50
3.1.3	L'utilisateur en tant qu'entité mixte	56
3.1.4	Synthèse	58
3.2	Contrôle des entités mixtes	59
3.2.1	Contrôle d'une partie virtuelle d'une entité mixte	59
3.2.2	Contrôle des entités mixtes non-vivantes	61
3.2.3	Contrôle des entités mixtes vivantes	61
3.2.4	Synthèse	62
3.3	Interactions entre entités mixtes	63
3.3.1	Synthèse	67
3.4	Interactions utilisateur hybrides	68
3.4.1	Hybridation des modalités	68
3.4.2	Interactions multi-sensorielles	70
3.4.3	Interactions mono-modales à interfaces multiples	72
3.5	Conclusion	73

4 De la conception à l'application	75
4.1 Cas d'usage du ventilateur mixte	75
4.2 Vue d'ensemble du framework	76
4.3 Conception	79
4.4 Développement	81
4.5 Implémentation	84
4.5.1 Réseau de Petri DOMIM	84
4.5.2 Composants	86
4.6 Conclusion	90
5 Cas d'usage	93
5.1 Contrôle gestuel des environnements domotiques	94
5.2 Contrôle tangible des environnements domotiques	98
5.3 Téléphone mixte, interactions tangibles et rendu distribué	101
5.4 Téléphone mixte et interactions hybrides mêlant interactions gestuelles, tangibles et tactiles	104
5.5 Interactions hybrides avec un bâtiment connecté	106
5.6 Retour multisensoriel de la localisation des mains d'un utilisateur	108
5.7 Rendu stéréoscopique multicouche d'un contenu virtuel	111
5.8 Conclusion	117
6 Conclusion et travaux futurs	119
6.1 Contributions	119
6.2 Limitations et recherches à court terme	120
6.3 Recherches à moyen terme	121
6.4 Recherches à long terme	122
A Bibliographie de l'auteur	123
A.1 Publications	123
A.2 Brevets	123

Chapitre 1

Introduction

Nous assistons à l'émergence de technologies qui associent réel et virtuel. Ces technologies s'intègrent progressivement à notre quotidien. Parmi ces technologies, la réalité mixte, les interfaces utilisateur naturelles et l'internet des objets nous permettent d'interagir avec le virtuel. De nouvelles formes de systèmes interactifs en découlent, que l'on trouve dans les villes, les bâtiments, les usines connectées, etc. L'expansion de l'internet des objets, décrite par la Figure 1.1 tirée du rapport Business Insider Intelligence 2020¹, permet de nombreuses combinaisons technologiques, produisant des systèmes interactifs très variés, mais peu adaptables aux évolutions technologiques, ayant par conséquent une durée de vie limitée. Ces systèmes sont potentiellement applicables à de multiples secteurs, comme la domotique, l'industrie, le secteur énergétique ou encore la sécurité et la santé. Par exemple, un visiocasque de réalité mixte peut augmenter un téléphone intelligent intégrant un assistant vocal. Un utilisateur interagit par la voix ou le geste avec l'assistant vocal incarné qui permet, par le réseau, de visualiser et commander les appareils connectés qui l'entourent. Cette connexion grandissante des objets du quotidien, comme les montres, les téléphones ou les téléviseurs, entraîne la virtualisation de nombreux objets. En effet, les agendas, carnets d'adresses, réveils, chronomètres, lecteurs multimédia et autres voient leurs fonctionnalités être dorénavant disponibles sur des téléphones intelligents dont la pluralité de l'usage ne cesse de s'étendre.

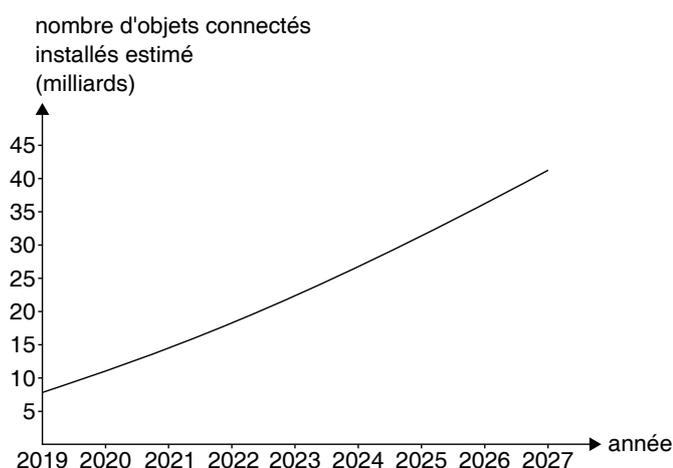


FIGURE 1.1 – Prévisions 2020 d'expansion de l'internet des objets par Business Insider Intelligence

1. <https://www.businessinsider.com/internet-of-things-report?IR=T>

Néanmoins, l'association entre réalité mixte, interfaces utilisateur naturelles et internet des objets est difficile à mettre en œuvre. Parmi les facteurs expliquant cette complexité, l'hétérogénéité des technologies employées, leur grand nombre, leur caractère fermé car sécurisé et leur manque de standardisation limitent fortement leur interconnectivité et leur extensibilité.

1.1 Contexte

Le développement de la réalité mixte, des interfaces utilisateur naturelles et de l'internet des objets offre aux utilisateurs de nouvelles façons d'interagir avec leur environnement.

La réalité mixte combine les environnements réels et virtuels. Elle fournit à l'utilisateur la possibilité d'interagir simultanément avec ces environnements aux caractéristiques bien distinctes. Par exemple, l'utilisateur équipé d'un visiocasque de réalité mixte peut être guidé dans le cadre d'un parcours touristique et interagir avec des répliques virtuelles d'objets historiques situées dans leur contexte.

Les interfaces utilisateur naturelles permettent aux utilisateurs de percevoir et d'agir sur les environnements virtuels. Elles font appel au corps de l'utilisateur afin de simuler les interactions qu'il a naturellement avec son environnement réel. Par exemple, marcher ou parler sont des interactions naturelles. Ces interactions peuvent être captées et reconnues par des ordinateurs afin d'interagir avec les environnements virtuels. Ces interfaces utilisateur ont notamment été ajoutées à certaines consoles de jeu comme la Wii de Nintendo afin de pouvoir jouer en sautant ou dansant.

L'internet des objets, en connectant les objets de notre quotidien, intègre les technologies à notre environnement. En se fondant dans notre environnement, ces technologies deviennent des services ubiquitaires permettant à leurs utilisateurs d'atteindre un objectif [Pin07]. L'internet des objets s'appuie sur des capteurs, des ordinateurs et des effecteurs disséminés dans notre environnement et reliés par internet. L'internet des objets connaît une expansion rapide : 75 milliards d'objets devraient être connectés en 2025, soit un quintuplement en une décennie². En l'espace de 5 ans, les ventes de montres connectées d'Apple ont dépassé les ventes de montres tous types confondus de l'industrie horlogère suisse³. En parallèle, le succès des téléphones intelligents entraîne des changements sociaux et comportementaux importants grâce à leur simplicité d'usage et leur hyper-connectivité.

Combiner interfaces utilisateur naturelles et réalité mixte permet à l'utilisateur d'interagir avec des objets virtuels de façon naturelle et intuitive. Certains visiocasques de réalité mixte comme Hololens ou Magic Leap embarquent depuis quelques années ce type d'interface afin de faciliter les interactions de l'utilisateur avec le virtuel. Combiner la réalité mixte avec l'internet des objets est également pertinent, car l'internet des objets fournit des services ubiquitaires connectant le réel au virtuel tandis que la réalité mixte permet à l'utilisateur d'interagir simultanément avec le réel et le virtuel. L'utilisateur peut alors interagir avec ces services affichés sous forme de métaphores virtuelles. Par exemple, une lampe connectée peut être augmentée par un bouton virtuel qui permet de contrôler son fonctionnement par le geste ou la voix. Par ailleurs, associer interfaces utilisateur naturelles et internet des objets permet à l'utilisateur d'interagir naturellement avec des services ubiquitaires. La lampe connectée peut alors être pilotée à distance grâce à une interaction pseudo-naturelle. La réalité mixte combinée aux interactions pseudo-naturelles permet aux utilisateurs d'interagir directement avec des objets, sans utiliser une interface telle qu'un téléphone ou un ordinateur intermédiaire l'interaction indirectement. Cette réduction de la distance entre réel et

2. <https://www.statista.com/statistics/471264/iot-number-ofconnected-devices-worldwide/>

3. <https://labs.letemps.ch/interactive/2019/apple-depasse-horlogerie-suisse/>

virtuel permet d'interagir plus naturellement avec les objets ainsi que d'être plus mobile et actif.

Ces domaines sont donc bénéfiques les uns pour les autres lorsqu'ils sont regroupés au sein de systèmes interactifs mixtes qui les combinent. Nous proposons de les appeler des systèmes interactifs à entités mixtes, ou *inTera*ctive *mix*ed Entity Systems en anglais (TIES). Cette convergence entre technologies mêlant réel et virtuel a pour conséquence de produire un monde de plus en plus ubiquitaire, où la frontière entre réel et virtuel devient chaque jour un peu plus ténue. Dans ce monde mixte, nous côtoyons et interagissons plus naturellement, intuitivement et fréquemment avec le virtuel.

1.2 Problématiques et objectifs

Depuis deux décennies, des TIES créent des interactions utilisateurs combinant réalité mixte, interfaces utilisateur naturelles et internet des objets. Néanmoins, les modèles et méthodologies existantes manquent de généricité, ce qui les rend peu voire pas interopérables. Une telle solution permettrait pourtant de consulter l'état d'un véhicule ou encore de voir sa vitesse courante tout en regardant la route. Les outils actuels distinguent également de façon plus ou moins confuse réel et virtuel. Or, ce sont des environnements aux propriétés distinctes. Les utilisateurs n'interagiront pas de la même manière ni pour les mêmes raisons avec ces environnements. De plus, le virtuel peut agir sur le réel et réciproquement. Ces interactions peuvent donc être complémentaires et s'enrichir mutuellement. Au final, les systèmes créés sont essentiellement des prototypes et des preuves de concept non réutilisés. Ils produisent dès lors des expériences à la durée de vie limitée. Il n'est généralement pas possible de les déployer sans passer par un processus long et fastidieux d'industrialisation. Ils sont également bridés dans leur complexité. En effet, un nombre d'objets connectés croissant rend ces systèmes potentiellement de plus en plus complexes. Or, les systèmes produits n'exploitent généralement qu'une très faible partie des capacités d'interaction qu'ils offrent.

Notre problématique est de créer des TIES de plus en plus complexes et interopérables afin de rendre leurs interactions utilisateur plus naturelles et intuitives. Nous nous intéressons aux synergies entre réalité mixte, interfaces utilisateur naturelles et internet des objets relatives à l'intuitivité de l'interaction. En particulier, nous explorons non seulement les apports de la réalité mixte et des interfaces utilisateur naturelles comme médiateurs de l'interaction avec l'internet des objets, mais également le potentiel des interconnexions entre interfaces utilisateur naturelles rendues possibles grâce à l'internet des objets. Le contexte industriel de cette recherche implique que nos travaux soient à terme reliés à des plateformes dédiées à l'internet des objets et à l'intelligence ambiante. Pour cette raison, nous ne traitons dans ce document ni d'intelligence artificielle ni d'aspects techniques spécifiques à l'internet des objets, comme la sauvegarde de contexte, la sécurité, la découverte automatique entre objets connectés, la caractérisation automatique des services ubiquitaires, ou encore les protocoles et technologies réseau dédiés. Néanmoins, il est tout à fait envisageable en aval de cette thèse d'intégrer ces problématiques à nos travaux. Nous souhaitons répondre aux questions suivantes :

- comment concevoir et mettre en œuvre des systèmes interactifs intégrant toujours plus d'entités constitutives et offrant des interactions hybrides toujours plus évoluées ? Il est nécessaire de maîtriser la complexité de ces systèmes en les modélisant de façon unifiée. Cette unification permet aux concepteurs de découper ces systèmes en sous-systèmes plus appréhendables et faciles à concevoir ;
- comment unifier ces systèmes pour les rendre interopérables et extensibles ? Créer des outils et méthodologies dédiés à leur production permet de simplifier, d'accélérer et d'uniformiser les tâches de production à effectuer. Ils fournissent des pratiques homogènes de création des TIES par et pour les développeurs ;

- quelles possibilités d'interactions innovantes offrent ces systèmes ? Ils permettent d'explorer de nouvelles techniques d'interaction intégrées à notre environnement. Ces techniques associent toujours plus étroitement virtuel et réel en étendant le nombre d'interfaces ou de modalités impliquées.

1.3 Modèle d'entité mixte

Afin de traiter nos problématiques, nous proposons un modèle d'entité mixte unifiant les trois domaines traités. Pour se faire, nous nous appuyons sur la notion d'objet mixte développée par Coutrix et al.[CN06] en la combinant à la notion de boucle d'interaction à trois corps, ces corps étant la perception, la cognition et l'action. Cette boucle d'interaction est communément utilisée dans les domaines traitant de l'interaction homme-machine. Nous utilisons le modèle résultant afin de modéliser des systèmes interactifs mixtes. Au final, ce modèle permet de :

- distinguer les entités constitutives de ces systèmes ;
- modéliser finement ces entités mixtes en :
 - décorrélant leur partie réelle et de leur partie virtuelle ;
 - synchronisant ces parties entre elles ;
 - modélisant les capacités d'interaction de chaque partie ;
- modéliser ces systèmes en les segmentant en sous-systèmes afin de :
 - maîtriser leur complexité ;
 - étudier les interactions fines se produisant :
 - entre ces entités mixtes ;
 - au sein de chaque entité mixte.

Ce modèle nous a permis de proposer un cadre méthodologique de conception et d'implémentation permettant de simplifier et d'uniformiser ces phases. Les concepteurs, les développeurs et les utilisateurs finaux bénéficient tous de notre approche, tant lors de la création que de l'usage des TIES. Les TIES basés sur notre approche ont pour caractéristiques d'être homogènes, interopérables et extensibles, afin de répondre aux enjeux futurs d'une planète partiellement dupliquée par son jumeau virtuel.

1.4 Structure du manuscrit

Ce manuscrit est organisé de la manière suivante.

Le chapitre 2 présente l'état de l'art. Il décrit tout d'abord chacun des paradigmes de la réalité mixte, des interfaces utilisateur naturelles et de l'internet des objets. Puis il présente des travaux effectuant cette convergence entre tout ou partie de ces paradigmes ou de paradigmes connexes.

Le chapitre 3 présente notre modèle unifié destiné aux TIES. Nous avons appelé ce modèle le modèle interne mixte orienté conception, ou en anglais *design-oriented mixed reality internal model* (DOMIM). Nous décrivons notre démarche puis décrivons ses caractéristiques et usages.

Le chapitre 4 présente notre cadre de conception et d'implémentation basé sur DOMIM. Il est destiné aux concepteurs et aux développeurs de TIES. Il leur fournit des outils adaptés à la grande diversité des entités constitutives de ces systèmes, à la convergence des paradigmes couverts par les TIES, et au besoin d'uniformiser les pratiques.

Le chapitre 5 présente plusieurs cas d'usage montrant l'utilité de notre approche. Ils couvrent tant les aspects d'unification des pratiques que de factorisation des compétences requises par les TIES en fournissant des composants logiciels adaptés.

Le chapitre 6 conclut sur nos travaux relatifs à la conception, à l'implémentation et à l'exploration du potentiel des TIES, et présente les orientations futures de nos travaux portant sur ces systèmes.

Rapport-Gratuit.com

Chapitre 2

Etat de l'art

Dubois et al. appellent "systèmes interactifs mixtes" [DBAG14] les systèmes incluant la réalité augmentée, les interfaces tangibles et l'intelligence ambiante. Ces systèmes ont démontré leur capacité à rendre les interactions de leurs utilisateurs plus intuitives[DBP16]. Dans ce manuscrit, nous étudions des systèmes interactifs mixtes :

- étendus à l'ensemble des interfaces utilisateur naturelles, qu'elles soient tangibles ou non,
- ayant recours à l'internet des objets comme vecteur d'ubiquité.

Dans ce chapitre, nous présentons une étude bibliographique sur les travaux de recherche proposant des systèmes interactifs mixtes autorisant des interactions entre l'utilisateur et son environnement, grâce à la réalité mixte, aux interfaces utilisateur naturelles et à l'internet des objets. Ces domaines, pris séparément, ont donné lieu à des mises en production de systèmes. Nous présentons des usages croisés de ces domaines, ainsi que des modèles et outils permettant de les concevoir et de les développer. Cette étude vise à dégager nos problématiques de recherche menant à des contributions scientifiques usuelles et techniques.

Nous décrivons dans la section suivante la réalité mixte, les interfaces utilisateur naturelles et l'internet des objets, en abordant leurs principaux aspects ainsi qu'en discutant de leurs limitations. Puis nous décrivons des travaux traitant de leur convergence. Nous concluons ce chapitre par une analyse de ces modèles.

2.1 Réalité mixte, interfaces utilisateur naturelles et internet des objets

Les domaines de la réalité mixte, des interfaces utilisateur naturelles et de l'internet des objets traitent tous de l'interaction entre réel et virtuel, mais avec néanmoins des approches distinctes, que nous décrivons dans cette section.

2.1.1 Préambule

La réalité virtuelle tout comme la réalité augmentée sont des domaines généralement utilisés pour définir ce qu'est la réalité mixte. S'il n'existe actuellement pas de consensus sur ce qu'est la réalité virtuelle ou la réalité augmentée [SHN19], nous décrivons dans cette sous-section les définitions et approches les plus consensuelles et répandues de ces termes.

Réalité virtuelle

Arnaldi et al. [FMB06] définissent la *réalité virtuelle* de la façon suivante : « La réalité virtuelle est un domaine scientifique et technique exploitant l'informatique et des interfaces comportementales en vue de simuler dans un monde virtuel le comportement d'entités 3D, qui sont en interaction en temps réel entre elles et avec un ou des utilisateurs en immersion pseudo-naturelle par l'intermédiaire de canaux sensori-moteurs. » D'après Speicher et al. [SHN19], elle se caractérise par une vue de l'utilisateur sur un environnement entièrement virtuel. Elle nécessite un suivi de la tête de l'utilisateur, voire un périphérique d'affichage porté sur la tête. Elle immerge totalement l'utilisateur dans l'environnement virtuel.

Précurseur du domaine, Morton Heilig obtient un brevet en 1960 [Hei60] le *Telesphere Mask*. Ce casque de stéréovision et d'écoute de son spatialisé, dont les plans sont présentés par la Figure 2.1, est l'ancêtre des casques de réalité virtuelle actuels. Fuchs décrit ces casques et leur évolution dans le livre [Fuc17].

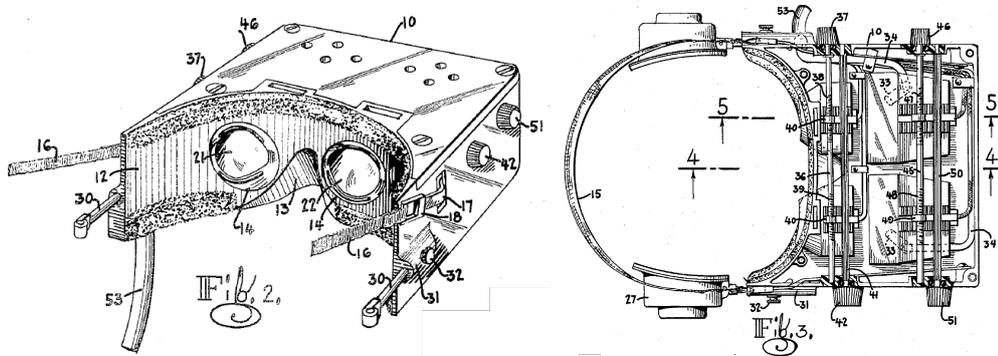


FIGURE 2.1 – Schémas extraits du brevet "Telesphere Mask" de Morton Heilig [Hei60]

La réalité virtuelle est utilisée dans de multiples secteurs d'activité, comme la formation, la visualisation de données, le divertissement, etc. Elle permet par exemple de visualiser et manipuler des modèles 3D issus d'une photogrammétrie d'objets historiques. La Figure 2.2 montre un utilisateur observer en stéréovision et manipuler par le geste un modèle du reliquaire du cœur d'Anne de Bretagne.



FIGURE 2.2 – Interaction avec un modèle 3D du reliquaire du cœur d'Anne de Bretagne. Partenariat entre Orange et le Conseil Départemental de Loire Atlantique.

Réalité augmentée

La définition la plus citée de la *réalité augmentée* est d'Azuma [Azu97]. Elle est produite par des systèmes aux caractéristiques suivantes :

- combinant réel et virtuel ;
- interactifs en temps réel ;
- localisés en trois dimensions.

D'après Speicher et al. [SHN19], la réalité augmentée est une notion très discutée et souvent confondue avec la notion de réalité mixte. Elle est perçue comme une fusion entre des graphismes rendus en 3D spatialement cohérents et l'environnement réel.

Le premier dispositif de rendu visuel en réalité augmentée est réalisé par Ivan E. Sutherland en 1968 [Sut68]. Il est nommé *Sword of Damocles*¹ en raison de son poids qui nécessitait une structure pour le soutenir, cf. Figure 2.3.

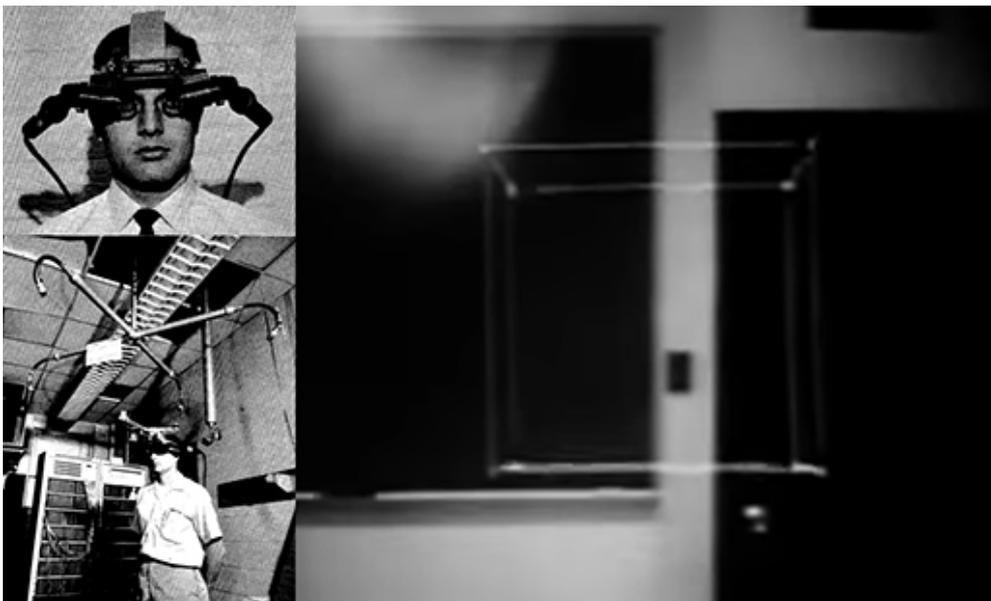


FIGURE 2.3 – The Sword of Damocles [Sut68]

Les lunettes et visiocasques de réalité augmentée nécessitent d'être mobiles et autonomes. Ils nécessitent d'être plus légers, et de couvrir un champ de vision le plus proche possible du champ de vision humain. Or, la taille et le poids des batteries, le champ de vision réduit des optiques semi-transparentes, ou encore une puissance de calcul qui doit offrir un compromis entre besoins de calcul importants et frugalité, empêchent la démocratisation de ces interfaces utilisateur [Fuc17].

Les usages actuels sont essentiellement professionnels en raison du coût du matériel. Par exemple, la filiale Testia d'Airbus développe des outils de réalité augmentée déployés dans des usines aéronautiques. Néanmoins, le potentiel de cette technologie est vaste et applicable aussi bien au divertissement, comme le jeu Pokémon Go l'a démontré, qu'au commerce, à la surveillance ou encore au tourisme.

1. <https://www.youtube.com/watch?v=NtwZXGprxag>

2.1.2 Réalité mixte

Les technologies dites de réalité mixte permettent à leurs utilisateurs d'interagir simultanément avec le réel et le virtuel. De nombreuses définitions décrivent ce qu'est la réalité mixte, sans pour autant créer de consensus [SHN19]. Les experts interrogés par Speicher et al. [SHN19] estiment que dans une dizaine d'années les notions de réalité virtuelle, de réalité augmentée et de réalité mixte se confondront, car le matériel sera alors capable de fournir tout le fonctionnel qu'elles recouvrent. Par exemple, un casque de réalité mixte pourra offusquer le réel si souhaité ou nécessaire afin de limiter la vision de l'utilisateur à l'environnement virtuel.

La définition la plus consensuelle de la *réalité mixte* est celle de Milgram et Kishino [MK94]. Ils présentent en 1994 le continuum de la virtualité, représenté dans la Figure 2.4. Ce continuum combine les environnements réels et virtuels. Selon cette définition, on peut ainsi considérer la réalité mixte comme une réalité déterminée par le niveau de virtualité de l'environnement de l'utilisateur, oscillant entre "purement virtuel" à "purement réel". Les extrêmes de ce continuum, un environnement purement réel et un environnement purement virtuel, ne sont pas considérés comme étant inclus dans ce continuum, comme l'explicitent Speicher et al. [SHN19].

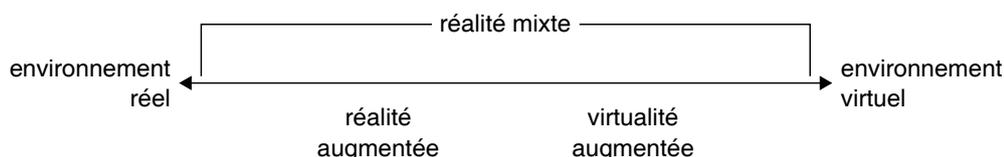


FIGURE 2.4 – Le continuum de la virtualité [MK94]

D'après Speicher et al. [SHN19], cette notion est très débattue parmi les communautés de recherche qui l'utilisent. Elle est souvent perçue comme une réalité augmentée plus évoluée, aux fonctionnalités étendues et totalement immersive. Elle est également perçue comme un terme marketing qui cherche à différencier certaines technologies. Nous nous appuyons dans notre mémoire sur la définition d'un système mixte proposée par Coutrix et al. en 2006 [CN06]. Un système mixte est un système constitué d'objets mixtes permettant à l'utilisateur d'interagir avec une réalité mixte [CNR05]. Un objet mixte est composé d'une partie réelle et d'une partie numérique. Des propriétés physiques caractérisent la partie réelle, tandis que des propriétés numériques déterminent la partie numérique, comme indiqué dans la Figure 2.5.

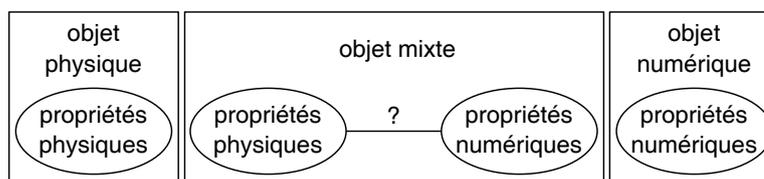


FIGURE 2.5 – Objets physiques, mixtes et numériques d'après [CNR05]. Les objets mixtes relient le réel au virtuel, une notion commune à l'internet des objets.

La synchronisation des propriétés numériques et physiques est modélisée par la boucle présentée dans la Figure 2.6. Elle inclut :

- les modalités de liaison, en l'occurrence les interfaces de captation du réel et de rendu des propriétés numériques dans le monde réel,
- les langages permettant la transmission des propriétés numériques, un langage en entrée et

un autre en sortie.

Ce modèle intègre les dispositifs de captation de propriétés réelles et de rendu de propriétés virtuelles à l'objet mixte, même si elles lui sont externes. Typiquement, l'internet des objets est également capable de fournir captation et rendu sans pour autant appartenir à l'objet mixte. Ce modèle est donc trop inclusif pour exprimer la convergence entre systèmes interactifs mixtes et internet des objets ou interfaces utilisateur naturelles.

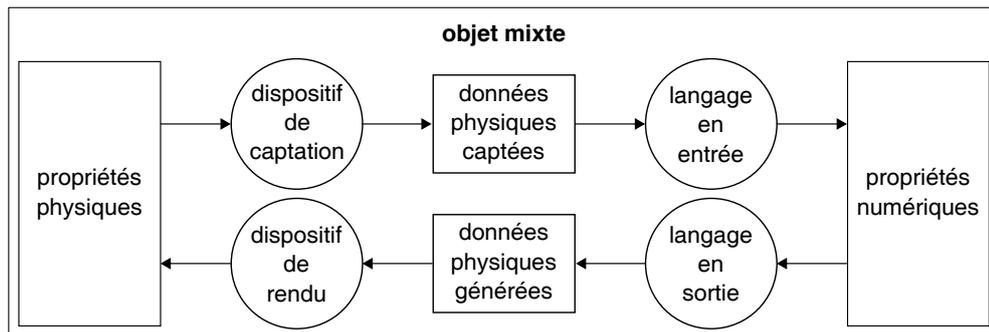


FIGURE 2.6 – Boucle d'interaction des objets mixtes d'après [CNR05]

Discussion

La réalité mixte est limitée par plusieurs aspects. En premier lieu, capturer les propriétés du monde réel pour les numériser n'est pas systématiquement possible. Certaines propriétés réelles sont difficiles voir impossibles à capturer. Les propriétés réelles capturées sont généralement la localisation d'une entité ou encore son apparence. Par exemple, il est possible de détecter des gestes effectués par les bras ou les mains d'un utilisateur suivi par une caméra. Si la capture de propriétés réelles est limitée par les caractéristiques des capteurs existants, l'actuation de propriétés numériques est quant à elle parfois difficile voire impossible. Il n'est par exemple pas possible de placer un objet sous un autre dans le monde réel sans l'intervention, par exemple, de bras robotiques capables de préhension. Dans ce cas, l'internet des objets donne à l'utilisateur d'un système interactif mixte un accès aux objets connectés. Il augmente ainsi les moyens de captation et d'actuation à la disposition de l'utilisateur.

Le matériel de réalité mixte est également un facteur de limitation de son développement. Les casques de réalité mixte sont actuellement de trois à dix fois plus chers que les casques de réalité virtuelle et l'offre d'applications dédiées est pour le moment très réduite et confidentielle. Par exemple, un casque de réalité mixte Hololens coûte 3850€ tandis qu'un casque de réalité virtuelle coûte 450€. De plus, aucun jeu commercial à gros budget n'est actuellement disponible pour les casques de réalité mixte tandis que les catalogues de jeux vidéo en réalité virtuelle ne cessent de s'étoffer.

Les interfaces utilisateur naturelles sont utilisées conjointement à la réalité mixte, afin de fournir à l'utilisateur non seulement la vision, mais également d'autres modalités d'interaction comme la voix ou le geste.

2.1.3 Interfaces utilisateur naturelles

Si la réalité mixte permet de percevoir des objets virtuels, interagir avec le virtuel nécessite des interfaces dédiées. Ces interfaces exploitent certains canaux sensori-moteurs utilisés par l'humain pour interagir avec son environnement physique.

Nous décrivons dans un premier paragraphe le paradigme d'interface utilisateur. Le lecteur connaissant cette notion pourra passer au paragraphe suivant, qui traite des interfaces utilisateur naturelles.

Interface utilisateur

Une interface utilisateur est la partie d'un système informatique grâce à laquelle un utilisateur communique et interagit avec l'ordinateur [WG93]. Elle joue le rôle de médiateur de l'interaction. Elle conditionne l'utilisabilité, la qualité, le temps d'apprentissage, les fonctionnalités et les performances d'un système interactif. Elle a donc un rôle majeur dans son succès. Par exemple, l'écran tactile a révolutionné en l'espace de 10 ans l'usage des téléphones mobiles, car sa simplicité d'utilisation et son intuitivité l'ont rendu incontournable. Une interface utilisateur utilise des capteurs afin de capter des propriétés du monde réel ou des effecteurs afin de traduire des propriétés virtuelles dans un environnement réel. Par exemple, les informations de la part de l'utilisateur sont généralement perçues par un ordinateur via des périphériques, comme une souris ou un clavier. L'interface fournit également un retour perceptible des conséquences des actions d'un utilisateur sur le système interactif. Pour cela, elle utilise des interfaces de rendu afin de rendre perceptibles certaines propriétés numériques du système. Par exemple, un écran relié à l'ordinateur permet à l'utilisateur d'obtenir un retour visuel sur son usage de la souris. La boucle d'interaction utilisateur, interface et ordinateur est résumée dans le schéma de la Figure 2.7.

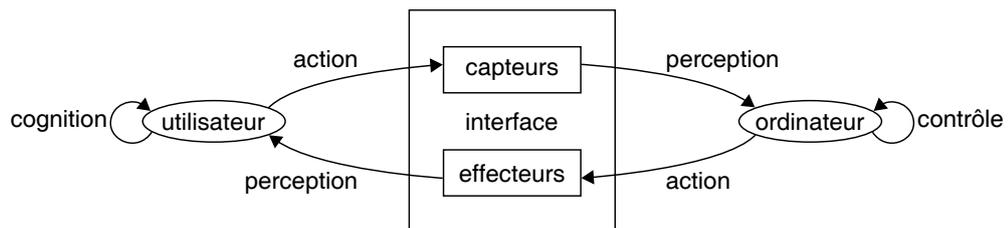


FIGURE 2.7 – La boucle d'interaction entre un utilisateur, une interface et un ordinateur

L'affordance est une notion prépondérante dans l'usage d'interfaces utilisateur. Gibsons définit [Gib77] les affordances de l'environnement comme étant ce que l'environnement a à offrir, à fournir aux êtres vivants. Norman les définit comme étant les propriétés actionnables perçues d'un objet qui définissent ses usages potentiels [Nor99]. Les affordances réelles sont émises par un objet physique. Elles donnent sens à ce que nous percevons et sont nécessaires à notre comportement. Les affordances perçues sont quant à elles produites par des mécanismes de rendu. Elles fournissent des signaux significatifs aux utilisateurs potentiels [Nor99]. Elles permettent de percevoir comment interagir avec des environnements virtuels.

Pour être naturelles et intuitives, les interfaces utilisateur doivent permettre aux utilisateurs de comprendre rapidement et facilement comment les utiliser, idéalement sans temps d'apprentissage. Les contraintes imposées par les périphériques nécessitent le recours à des métaphores, vecteurs d'affordance. Ces métaphores sont des représentations qui, associées au contexte de l'utilisateur, lui permettent de deviner comment interagir avec l'interface utilisateur.

Interfaces utilisateur naturelles

Les interfaces utilisateur naturelles sont un type d'interface utilisateur. Elles utilisent le corps humain comme interface. Elles font appel aux compétences que l'humain a développé depuis sa naissance, issues de son interaction permanente avec son environnement [Bla12]. A contrario une interface utilisateur non-naturelle demandera l'usage d'un outil spécifique, requérant un apprentissage long et non-intuitif. Typiquement, on trouvera dans cette catégorie des interfaces complexes

et spécialisées pour un métier et une tâche donnés. Le fait d'expérimenter la réalité façonne les connaissances de l'homme et lui permet par la suite d'interagir avec son environnement naturellement en utilisant ces mêmes connaissances [OHM⁺13]. Les interfaces utilisateur naturelles imitent les interactions que les humains ont avec leur environnement réel, si possible sans faire appel à des métaphores d'interaction. Elles produisent dès lors des interactions pseudo-naturelles. Les interfaces utilisateur s'appuient sur le système sensori-moteur de l'homme, et concernent la vue, l'haptique, la parole, le mouvement, la proprioception, mais également le goût ou encore l'odorat. Typiquement, parler à un assistant vocal imite les conversations orales entre humains. Une modalité d'interaction correspond à un canal sensori-moteur qui lui est propre [Bou07]. L'homme interagit dans le monde réel de façon multimodale, ce qui implique qu'une interaction multimodale fait partie d'une expérience naturelle [JLW11]. Les interfaces utilisateur naturelles sont donc d'autant plus naturelles à expérimenter qu'elles impliquent le maximum de canaux sensori-moteurs de l'utilisateur lors de l'interaction.

De plus, Price et al. [PDG17] ont montré que des interactions multi-sensorielles permettent d'améliorer l'efficacité de l'utilisateur dans la réalisation de tâches. Les interfaces utilisateur naturelles rendent donc plus intuitive et naturelle l'interaction d'un utilisateur avec un système interactif [Bho13]. Leur objectif est de demander très peu de temps d'apprentissage avant d'être utilisées. L'intuitivité d'une interface est un facteur clé de son succès. Blackler et al. [BH07] présentent le continuum des interactions intuitives, décrit par la Figure 2.8. Ce continuum présente comment l'affordance physique des objets anciens rend ces objets simples et intuitifs à utiliser, tandis que de nouvelles technologies s'appuient sur des fonctionnalités familières aux utilisateurs pour produire une affordance perçue et ainsi créer cette intuitivité. Dans le cas des technologies les plus innovantes, il est même nécessaire de faire appel à des métaphores issues d'autres domaines. Par exemple, lorsque les interfaces graphiques apparaissent sur les ordinateurs, on utilise la métaphore des dossiers et des fichiers pour organiser l'information, car c'était jusqu'alors le système d'organisation usuel.

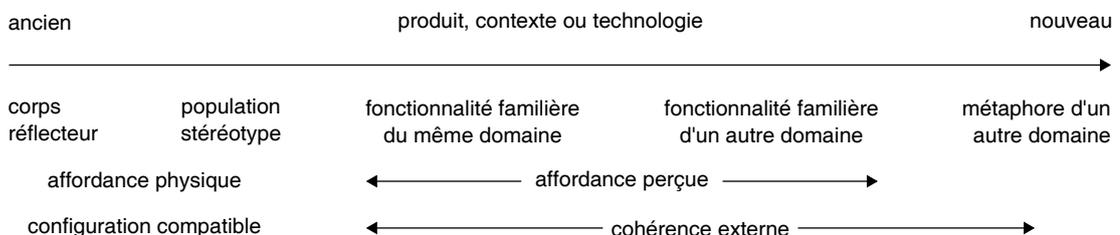


FIGURE 2.8 – Le continuum des interactions intuitives de Blackler et al. [BH07]

Parmi l'ensemble des modalités possibles, un type d'interface naturelle se distingue par sa plus grande naturalité [DBP16]. Les interfaces utilisateur tangibles offrent une forme physique à des données numériques [IU97]. Elles font appel à la perception cutanée de l'utilisateur. Une interface utilisateur tangible peut être complétée par d'autres modalités d'interaction non-tangible [Ish08]. Elle peut également s'adapter à différents objets virtuels. Aguerreche et al. présentent une interface utilisateur tangible reconfigurable permettant d'interagir avec divers objets virtuels [ADL10] présentés par la Figure 2.9.

Plusieurs objets tangibles peuvent également composer une interface tangible plus complexe [NCLN11]. Baraldi et al. présentent TANGerINE, un environnement tangible composé d'une table et de plusieurs objets tangibles intelligents [BDBL⁺07] qui permettent d'interagir avec le contenu numérique affiché par la table.

Si les interfaces utilisateur tangibles sont naturelles à l'usage, elles font appel à une perception haptique de l'utilisateur limitée, restreinte. Les interfaces utilisateur haptiques fournissent quant à

elles des sensations telles que le toucher ou la saisie d'objets virtuels, ou encore la sensation de leur poids ou des forces inertielles qui les animent. Par exemple, Murayama et al. [MBA⁺04] emploient pour leur système SPIDAR deux interfaces utilisateur haptiques identiques, un bras à retour d'effort par main afin de fournir des sensations haptiques lors de l'interaction avec des objets virtuels, comme présenté par la Figure 2.10.

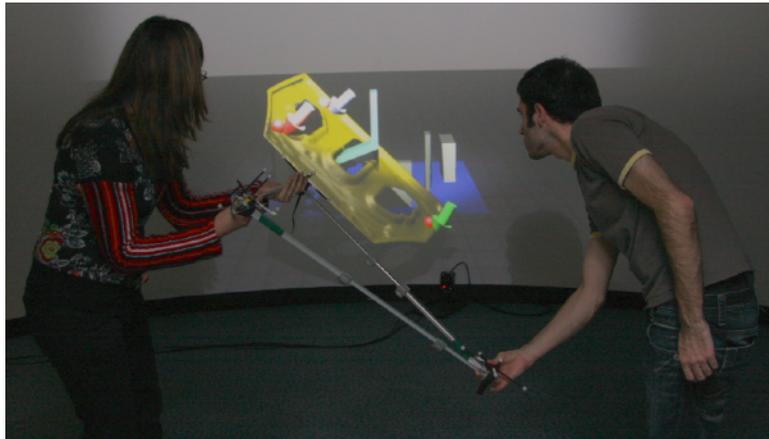


FIGURE 2.9 – Manipulation collaborative d'un capot de voiture virtuelle de Aguerreche et al. [ADL10]

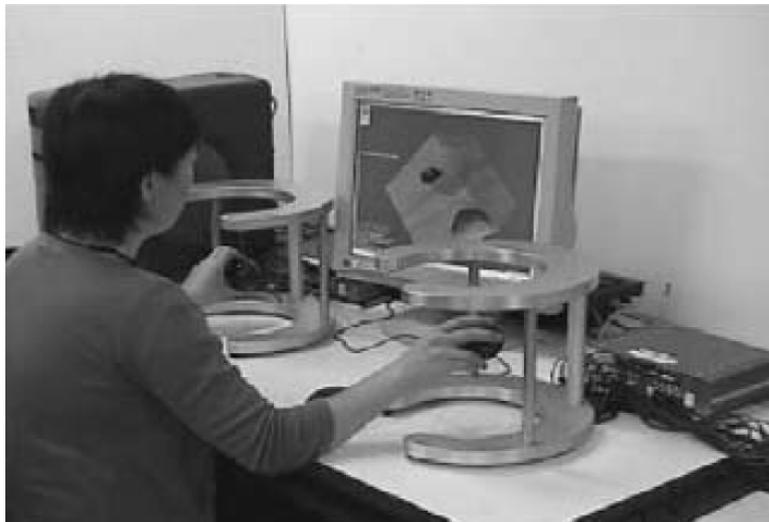


FIGURE 2.10 – Vue d'ensemble du système SPIDAR de Murayama et al. [MBA⁺04]

Il existe donc de multiples types d'interfaces utilisateur naturelles, chacune exploitant un canal sensori-moteur distinct. Interagir avec des systèmes combinant plusieurs interfaces utilisateur naturelles, plus complexes à réaliser, semble dès lors être un moyen efficace d'améliorer la naturalité d'une interaction.

Discussion

Le paradigme des interfaces utilisateur naturelles est critiqué pour plusieurs raisons, liées aux limitations des interactions pseudo-naturelles. Norman [Nor10] pointe le fait que ces interactions

nécessitent une phase d'apprentissage afin de savoir les utiliser. Or, le temps d'apprentissage requis est souvent éloigné de la quasi-instantanéité attendue. Par exemple, parler avec un assistant vocal nécessite de parler bien distinctement dans la direction du micro de l'assistant. La parole doit alors suivre un rythme et un volume sonore qui facilitent le travail de reconnaissance vocale. De plus, les interactions pseudo-naturelles requièrent généralement d'équiper l'utilisateur de capteurs et d'effecteurs. Par ailleurs, d'un point de vue sémantique, le terme "naturel" est si générique qu'il occulte la variété des technologies et des compétences concernées [HD15]. En outre, Bowman et al. [BMR12] remarquent que reproduire une interaction naturelle avec précision dans un environnement virtuel est généralement difficile voire impossible à cause des limitations techniques des systèmes de capture. Par exemple, ces systèmes souffrent d'une précision limitée de leurs capteurs, mais également des occultations lorsqu'ils font appel à des caméras. Néanmoins, cette précision s'améliore avec le temps, grâce à des capteurs plus précis et des systèmes interactifs plus performants. Cette volonté d'anthropomorpher nos interactions avec les machines se heurte à la vallée de l'étrange, ce que Mori [Mas70] appelle *uncanny valley*. La Figure 2.11 présente la courbe de l'uncanny valley. En effet, lorsqu'un robot ou un avatar humanoïde ressemble fortement à un humain en bonne santé, il n'est plus réaliste. Les interfaces utilisateur haptiques, quant à elles, se heurtent à des limitations en matière de mécatronique ou encore d'ergonomie. Par exemple, elles offrent une amplitude de mouvement limitée ou encore recouvrent partiellement certaines parties du corps, ce qui réduit les possibilités d'interaction naturelle de l'utilisateur. Elles rencontrent également des difficultés à reproduire des comportements physiques réalistes. En effet, les moteurs physiques nécessaires à la simulation des lois de la physique peinent à les reproduire de façon réaliste. Les forces rendues à l'utilisateur sont généralement peu réalistes et mal appliquées sur le corps de l'utilisateur, tout comme la friction et l'inertie des mécanismes pénalisent la crédibilité de l'expérience utilisateur [HPD17].

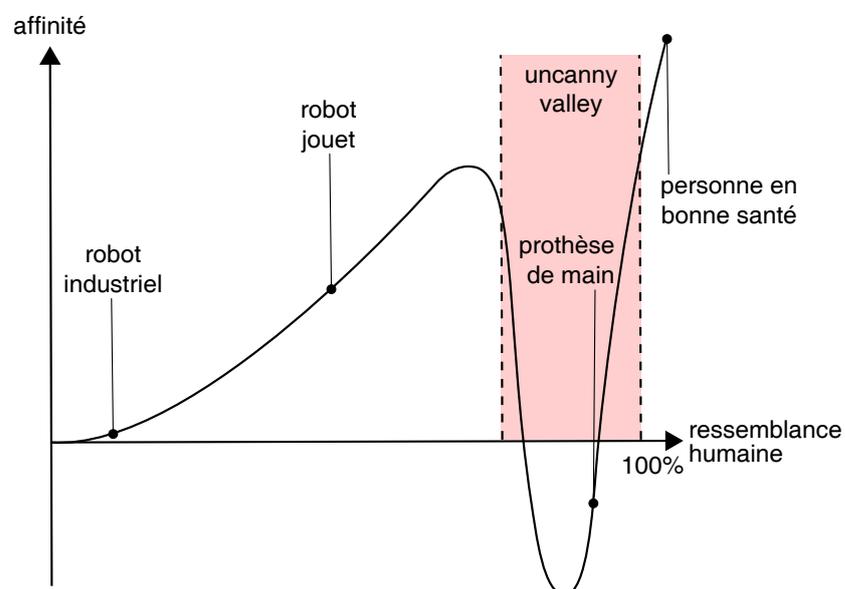


FIGURE 2.11 – The uncanny valley [MMK12].

2.1.4 Internet des objets

L'internet des objets permet de créer un réseau d'objets virtuels, eux-mêmes associés à des objets réels. La réalité mixte intègre de manière cohérente ces objets virtuels au monde réel, tandis que les interfaces utilisateur naturelles permettent à l'utilisateur d'interagir avec eux. L'internet

des objets a donc besoin d'être associé à ces technologies permettant d'interagir avec le virtuel de façon contextuelle pour prendre corps.

L'expression originelle de l'*internet des objets* est en langue anglaise *internet of things*. Une traduction littérale de cette expression en *internet des choses* n'est néanmoins pas usuelle en langue française. Pour Heidegger [Hei67], un *objet* est matériel, c'est un point de masse en mouvement dans l'ordre spatio-temporel, ou une combinaison de ces points. Pour Turner [Tur05], un *objet* est déterminé par sa pratique tandis que sa pratique est induite par son affordance. Ce processus est décrit par la Figure 2.12. Les interactions entre un utilisateur et un *objet* créent par la pratique des invariants qui permettent à la cognition de l'utilisateur d'identifier, d'utiliser et de concevoir des objets.

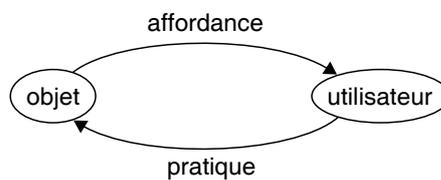


FIGURE 2.12 – Genèse d'un *objet*, d'après Turner [Tur05].

Dans cette sous-section, nous introduisons l'informatique ubiquitaire, dont fait partie l'internet des objets. Puis, nous définissons l'internet des objets ainsi que la notion d'objet intelligent, avant de donner une vue d'ensemble des problématiques et technologies réseau liées. Une discussion sur les limitations de ce paradigme conclut cette sous-section.

Mark Weiser introduit en 1988 le terme et le paradigme d'*informatique ubiquitaire* [Wei99] en réponse à l'intégration grandissante, organique et transparente des ordinateurs dans notre environnement. L'informatique ubiquitaire consiste à créer entre l'utilisateur et les ordinateurs invisibles qui l'entourent une interaction non-intentionnelle, implicite et transparente. Pour Weiser, ce paradigme est aux antipodes de la réalité virtuelle. Il nomme même virtualité incarnée les expériences utilisateur produites par l'informatique ubiquitaire. Pour lui, la réalité virtuelle crée un monde virtuel grâce aux technologies numériques, tandis que l'informatique ubiquitaire intègre ces mêmes technologies dans notre environnement réel. La Figure 2.13 décrit sa vision opposée de ces paradigmes.



FIGURE 2.13 – Informatique ubiquitaire et réalité virtuelle sont des paradigmes diamétralement opposés d'après Weiser [Wei99].

Poslad [Pos11] décrit les principaux composants de l'informatique ubiquitaire comme suit :

- les ordinateurs doivent être connectés par un réseau, distribués et accessibles de façon transparente ;
- les interactions homme-machine ont besoin d'être masquées ;
- les ordinateurs ont besoin de connaître leur environnement afin d'interagir avec lui.

Le terme *internet des objets* est introduit en 1999 par Ashton [Ash11]. Il répond au besoin de maîtriser la connexion d'un nombre grandissant d'objets à internet. L'internet des objets permet aux humains d'accéder à des services ubiquitaires. Ces services s'appuient essentiellement sur des objets capables de percevoir et d'agir dans leur environnement réel par l'intermédiaire d'une infrastructure réseau globale. L'internet des objets fournit alors le lien entre objets physiques et objets virtuels [AIM17]. Il s'étend à de multiples secteurs d'activité, comme l'industrie, le commerce, la santé, la formation, la surveillance et le contrôle de la population, la sécurité, le tourisme ou encore les loisirs.

Atzori et al. présentent un triptyque des concepts essentiels de l'internet des objets, qui sont les objets, internet et la sémantique [AIM10]. Dans ce triptyque, la sémantique a pour rôle de fournir un langage commun aux objets afin de leur permettre technologies réseau les plus de communiquer par le réseau internet (cf. Figure 2.14).

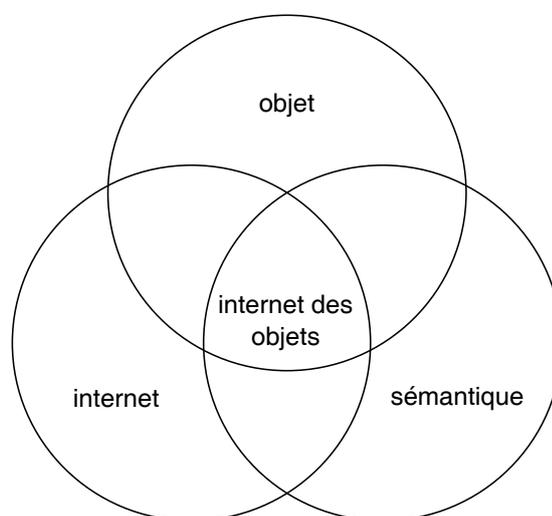


FIGURE 2.14 – Le paradigme de l'internet des objets sous forme de triptyque de ses composants essentiels qui sont les objets, internet et la sémantique, d'après [AIM10]

Ce paradigme évolue rapidement, comme le soulignent Atzori et al. [AIM17]. Ils le définissent comme étant "un framework conceptuel qui démultiplie la disponibilité de dispositifs hétérogènes et de solutions d'interconnexion, tels que des objets physiques augmentés fournissant une base d'information partagée à une échelle globale, afin de permettre la conception d'applications impliquant au même niveau de virtualité aussi bien les gens que les représentations d'objets."

D'après Greer et al. [GBWG19], l'internet des objets connecte le réel au virtuel en transduisant leurs propriétés afin d'interagir avec l'homme. Cette description n'est pas sans rappeler le paradigme de la réalité mixte. Toujours d'après Greer et al., ce paradigme converge avec le paradigme de système cyber-physique. Les systèmes cyber-physiques permettent à un utilisateur d'interagir simultanément sur les parties réelles et virtuelles d'un objet.

L'industrie connectée est un champ d'application majeur de l'internet des objets. Grieves introduit en 2002 le concept de *jumeau numérique* dans le contexte des systèmes de production intelligents et connectés [Gri19]. Un jumeau numérique est un clone numérique d'un produit réel. Le produit réel est alors le jumeau physique de son jumeau numérique. Ces deux jumeaux sont interconnectés afin de permettre la supervision et le contrôle du produit par le fabricant et les

utilisateurs. L'aérospatiale et l'aéronautique sont des exemples d'industries qui font appel à ce paradigme [GS12]. Kritzler et al. [KFMR17] étendent ce concept et présentent le concept de *jumeau virtuel*, permettant le contrôle d'une usine connectée en interagissant avec sa maquette numérique grâce à un Hololens. La Figure 2.15 montre un exemple de maquette d'usine connectée et son jumeau virtuel interactif.

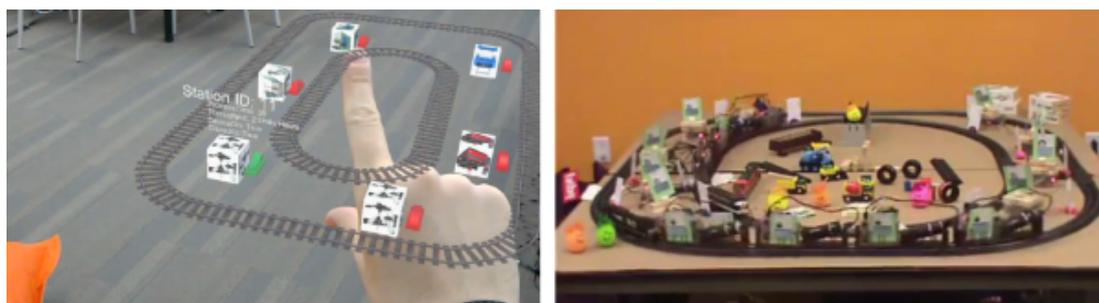


FIGURE 2.15 – Une maquette d'usine connectée et son jumeau virtuel [KFMR17]. A gauche, le jumeau virtuel de l'usine. A droite, une maquette d'usine connectée.

Les *objets intelligents* sont les éléments de construction de l'internet des objets. Ils sont définis par Kortuem et al. [KKSF10] comme étant des objets physiques et numériques autonomes augmentés de capacités de communication réseau, et de captation et d'analyse de leur environnement. Sanchez Lopez et al. [SLRHM12] présentent une architecture permettant à ces objets d'être regroupés au sein d'un même système. Les assistants personnels sont un exemple d'objet intelligent constitutif de l'internet des objets. Par exemple, ils peuvent transmettre une commande réseau à un four connecté suite à une demande vocale de l'utilisateur grâce à un protocole réseau adapté, étudié par Santos et al. [SRC⁺18]. L'objectif des concepteurs d'objets intelligents est de leur permettre de soutenir l'activité humaine de façon autonome à l'aide d'architectures dédiées, étudiées par Perera et al. [PZCG14]. Le concept d'objet intelligent est étendu par Möller et al. [MRK11] au soutien de l'activité non seulement humaine, mais également robotique, et renommé objet cognitif.

De multiples technologies réseau peuvent être utilisées par l'internet des objets. Elles sont classifiées en fonction de leur portée radio par les organismes de normalisation (IEEE/ETSI). Les technologies réseau les plus utilisées sont :

- Body Area Networks (BAN) : localisées sur, autour ou dans le corps, les entités constitutives de ces réseaux sont interconnectées (IEEE 802.15.4, ZigBee²,...);
- les réseaux personnels (PAN) : d'une portée inférieure à 100m, ils incluent RFID³, NFC, Bluetooth, ZigBee⁴, Z-Wave⁵...;
- les réseaux locaux (LAN) : d'une portée inférieure à 1km, ils incluent Wi-Fi⁶ notamment;
- les réseaux métropolitains (MAN) : d'une portée inférieure à 50km (WiMAX⁷,...);
- les réseaux étendus (WAN) : d'une portée de plus de 100km, ils incluent SigFox⁸ et LoRa⁹.

2. <https://www.zigbee.org>

3. <https://rainrfid.org/>

4. <https://www.zigbee.org>

5. <https://z-wavealliance.org/>

6. <https://www.wi-fi.org/>

7. <http://wimaxforum.org/>

8. <https://www.sigfox.com/en>

9. <https://lora-alliance.org/>

La transmission des données par les ondes radio n'est pas le seul aspect des technologies réseau. Elles font également appel à des protocoles qui permettent de normaliser le formatage des données échangées. ZWave et Wi-Fi définissent des protocoles réseau normalisés afin d'assurer leur interopérabilité. MQTT¹⁰ est un autre protocole communément utilisé par l'internet des objets en raison de sa frugalité.

Discussion

La connexion toujours grandissante des objets ne concerne pas seulement notre quotidien. Elle s'étend également des infrastructures industrielles, commerciales, médicales, administratives afin d'offrir de nouveaux services ubiquitaires. Elle simplifie et accélère un nombre grandissant de tâches que l'humain effectue. Cette évolution technologique permet par exemple de mieux mesurer puis contrôler la pollution de notre environnement grâce à des capteurs disséminés, mais néanmoins reliés. Elle est également utile dans le domaine médical, lorsque le contrôle de paramètres vitaux d'une personne est nécessaire.

Néanmoins, l'internet des objets est confronté à de multiples défis qui donnent lieu à des tendances dans les travaux concernant ce paradigme. Parmi plusieurs tendances et défis relevés par Saha et al. [SMS17], nous notons l'intérêt des grandes entreprises pour ce marché en pleine croissance. Cet engouement a pour conséquence de fragmenter et de limiter l'interopérabilité de l'internet des objets. La standardisation est nécessaire pour simplifier cette interopérabilité, mais est limitée par les problématiques de captation du marché par les entreprises. Au final, la quantité très importante d'objets connectés limite fortement l'extensibilité de l'internet des objets. Pour Nitti et al. [NPCA16], l'hétérogénéité de l'internet des objets est produite par la diversité des objets et des technologies réseau employés. Un autre problème majeur posé par l'internet des objets est son coût écologique très élevé. En effet, notre planète n'offre pas les ressources nécessaires à l'aboutissement de ce paradigme. L'internet des objets nécessite donc des solutions permettant de défragmenter les parcs d'objets connectés, et d'améliorer l'interopérabilité et l'extensibilité des objets connectés en uniformisant les pratiques de conception et d'implémentation. Ce paradigme nécessite des technologies telles que la réalité mixte et les interfaces utilisateur naturelles afin de pouvoir s'intégrer à notre environnement de façon naturelle, transparente et accessible.

2.2 Convergence de la réalité mixte, des interfaces utilisateur intuitives et de l'internet des objets

Ces paradigmes convergent depuis une vingtaine d'années dans la littérature scientifique. Cette convergence permet d'interagir en réalité mixte avec des objets connectés[LCL⁺17] à l'aide de techniques d'interaction s'intégrant à un environnement réel ou s'appuyant sur une maquette numérique [ARE⁺17]. Ils convergent également dans les usages. En effet, cette convergence permet aux utilisateurs d'interagir naturellement et ubiquitairement avec leur environnement. Elle permet également aux professionnels de concevoir des produits virtuels et de simuler leur usage avant leur production. Divers systèmes interactifs mettent en œuvre cette convergence, que nous présentons dans cette section. L'objectif de ces systèmes est de mieux intégrer les systèmes interactifs dans l'environnement de l'utilisateur. Ces systèmes visent également à améliorer l'intuitivité des interactions entre utilisateurs et environnements connectés. Néanmoins, si ces systèmes permettent de créer de nouvelles formes d'interaction avancées, les développer est rendu difficile par l'hétérogénéité et le nombre d'objets connectés disponibles [GCB⁺10].

Nous abordons dans un premier temps l'existence de kits de développement logiciel, avant de présenter des cas d'usage et des prototypes montrant l'intérêt de croiser et combiner ces domaines.

10. <http://mqtt.org/>

Dans un troisième temps, nous nous intéressons à l'interconnexion des NUI rendue possible grâce à l'IoT.

2.2.1 Kits de développement logiciel

De multiples kits de développement logiciel sont disponibles et applicables aux domaines de la réalité mixte, des interfaces utilisateur naturelles et de l'internet des objets. Ils offrent aux développeurs des implémentations de techniques évoluées centrales à ces domaines. Ils permettent d'uniformiser la mise en œuvre de ces techniques tout en les libérant d'une implémentation complexe. Par exemple, ARCore¹¹, ARKit¹² ou MixedRealityToolkit¹³ sont des briques logicielles utiles pour accéder à une mise en œuvre d'algorithmes de type SLAM [MUS16]. Ces briques logicielles sont utiles pour développer des systèmes interactifs mixtes, elles offrent des implémentations performantes et sont compatibles avec de multiples périphériques d'une plateforme donnée. Nous souhaitons les combiner afin de fournir des services et techniques d'interaction de plus haut niveau. Également, nous souhaitons créer des briques logicielles complétant les besoins de mise en œuvre des TIES.

2.2.2 Cas d'usage et prototypes

De multiples prototypes proposent d'utiliser des NUI afin d'interagir avec l'IoT. Nous présentons dans un premier temps des preuves de concept qui visent à rendre plus intuitive l'interaction d'un utilisateur avec son environnement. Dans un second temps, nous décrivons des prototypes de bâtiment intelligents faisant appel à la RM, aux NUI et à l'IoT.

NUI et objets du quotidien

Une partie des travaux de la littérature datent d'une dizaine d'années. Nous citons les travaux les plus proches de nos problématiques.

Petersen et al. [PS09] proposent une interface utilisateur naturelle basée sur la réalité augmentée permettant de corréler les propriétés physiques d'un objet réel, une fiche d'information portant sur une pompe industrielle, aux propriétés numériques de son jumeau virtuel. Ils combinent donc RM et NUI afin de compléter l'affordance de l'objet industriel dans le cadre d'une usine connectée. En effet, les informations fournies par la fiche d'information sont potentiellement fournies par le réseau local de l'usine connectée, de même que la détection et le suivi de la pompe industrielle peuvent alimenter le jumeau virtuel de l'usine connectée. Leur étude utilisateur démontre l'intérêt d'une interaction gestuelle, plus performante que l'usage d'une souris. Le dispositif utilisé n'est néanmoins pas mobile.

Mistry et al. [MM09] présentent le projet Sixthsense, qui augmente l'environnement de l'utilisateur par un dispositif projetant des informations sur des surfaces réelles comme le bras d'un utilisateur ou un journal. Cette technique d'augmentation du réel est appelée réalité augmentée spatialisée (SAR). Ce dispositif est porté par l'utilisateur sous la forme d'un médaillon. L'utilisateur interagit alors grâce à ses mains par le geste (cf. Figure 2.16). Nous sommes également dans le cas d'une augmentation du réel par le virtuel combinée à une interface utilisateur naturelle. Tout comme l'exemple précédent, la détection des objets est potentiellement transmissible à des bases de données connectées et le contenu virtuel affiché est potentiellement fourni par internet. Le dispositif porté semble peu robuste aux mouvements du corps de l'utilisateur.

11. <https://developers.google.com/ar/>

12. <https://developer.apple.com/augmented-reality/arkit/>

13. <https://github.com/Microsoft/MixedRealityToolkit-Unity>

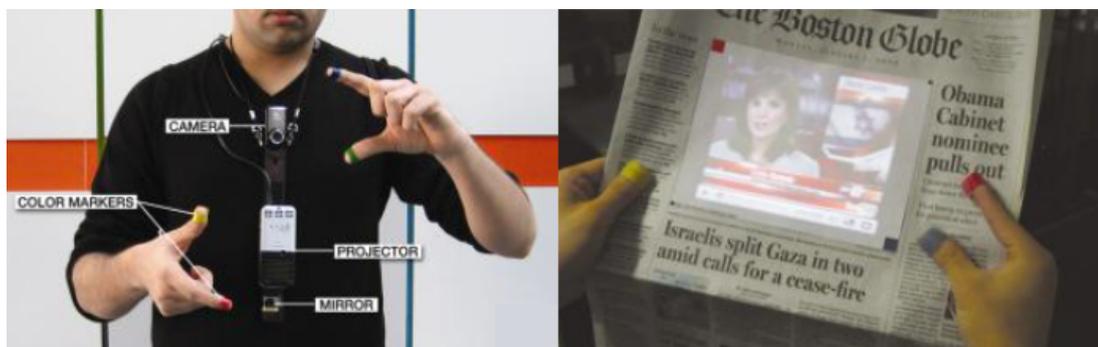


FIGURE 2.16 – Le système SixthSense et un cas d'usage de journal augmentée [MM09].

Sulisz et al. [SS12] augmentent un objet réel avec de l'information numérique. Pour ce faire, ils combinent des lunettes stéréoscopiques de réalité augmentée à un téléphone intelligent équipé d'un lecteur RFID. Le téléphone identifie un objet réel équipé d'une étiquette RFID et fournit un contenu 2D relatif à cet objet, affiché par les lunettes. Ce système affiche du texte ou de l'image et nécessite que le téléphone intelligent soit accolé à l'objet à augmenter. Là encore ce prototype associe réalité mixte à une interface utilisateur naturelle, le téléphone intelligent étant utilisé comme une interface tangible. Les puces RFID sont une composante de l'internet des objets qui permet d'identifier les objets et dès lors de transmettre de l'information relative à ces objets. L'usage d'un téléphone intelligent comme lecteur RFID tangible est très fortement bridé par la faible distance nécessaire à la communication entre lecteur et puce RFID, de l'ordre de quelques centimètres. Ce système ne semble pas estimer par ailleurs la pose de l'objet, ce qui limite très fortement les possibilités d'augmentation de l'objet par les lunettes.

Ces preuves de concept montrent le potentiel de l'usage des NUI pour interagir avec des objets du quotidien, avec toutefois une naturalité limitée par les contraintes techniques et l'encombrement des dispositifs.

Bâtiments intelligents

Lin et al. [LCL⁺17] présentent le système Ubii. Ce système permet à un utilisateur de commander des appareils intelligents comme une imprimante connectée à l'aide de lunettes de réalité augmentée monoculaires, les Google Glass, qui ne se localisent pas dans leur environnement. L'interaction consiste à pointer son regard sur un objet puis à le commander via une interface utilisateur graphique. Cette interaction s'effectue par le geste. Elle est moins naturelle qu'une interaction avec un environnement virtuel perçu en stéréovision. Néanmoins, cette technique d'interaction permet d'interagir à distance avec des objets intelligents de son voisinage (cf. Figure 2.17). L'association de l'augmentation d'objets connectés et de leur interaction basée sur le geste combine très clairement RM, NUI et IoT. Mais l'usage de lunettes monoculaires limite la naturalité de la perception du virtuel. Par ailleurs, la visée par un faisceau infrarouge du récepteur infrarouge associé à chaque objet intelligent connecté est rendu difficile par la distance entre l'utilisateur et l'objet, ce qui rend cette technique peu applicable au quotidien.

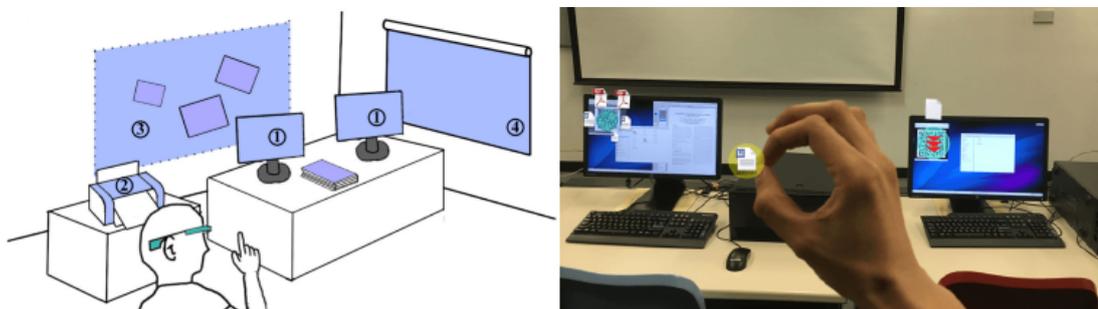


FIGURE 2.17 – Le système Ubii et un cas d’usage de glisser/déposer de document entre ordinateurs connectés [LCL⁺17].

Alice et al. [ARE⁺17] ont également étudié comment la réalité augmentée peut servir d’interface utilisateur pour l’internet des objets. Ils comparent trois modèles d’interaction avec un bâtiment connecté. Le premier positionne des icônes flottantes à proximité des objets connectés. Le second permet à l’utilisateur d’interagir avec une maquette virtuelle bâtiment. Le troisième consiste en un menu contextuel flottant (cf. Figure 2.18). Les études utilisateurs montrent qu’interagir avec le jumeau virtuel permet de mieux gérer un grand nombre d’objets connectés, mais laissent entendre que ces trois modèles d’interaction sont complémentaires.



FIGURE 2.18 – Modèles d’interaction en réalité augmentée avec un bâtiment connecté [ARE⁺17]. A gauche, l’utilisateur interagit avec le jumeau virtuel du bâtiment. Au centre, l’utilisateur interagit avec une icône flottante. A droite, l’utilisateur interagit avec un menu déroulant.

Kim et al. [KBH⁺18] étudient l’interaction entre un utilisateur et un assistant vocal incarné par un agent virtuel affiché par le visiocasque. Cette interaction multimodale a pour effet d’améliorer la confiance de l’utilisateur dans le système interactif mixte (cf. Figure 2.19). Là encore, la réalité augmentée combinée à des interactions pseudo-naturelles permet d’interagir avec des objets connectés. Néanmoins, les visiocasques de réalité mixte ne permettent pas de percevoir intégralement un agent autonome proche de l’utilisateur à taille humaine, ce qui réduit la naturalité de l’interaction avec l’agent. Norouzi et al. [NBB⁺19] étudient cette tendance anthropomorphe à associer les agents autonomes, l’internet des objets et la réalité augmentée. Ils constatent une augmentation des publications relatives à cette convergence et un potentiel de recherche important, en particulier pour des équipes de recherche interdisciplinaires.



FIGURE 2.19 – Un agent incarné piloté par la voix commande une lampe connectée [KBH⁺18].

Uimonen et al. [Uim18] présentent un système de réalité augmentée mobile permettant de consulter et d'interagir avec la maquette numérique d'un bâtiment commercial connecté. Ce système permet de contrôler la température ou l'air conditionné du bâtiment. Il montre l'utilité de ce type de système.

Les techniques employées réduisent l'écart qui sépare virtuel et réel dans le cas des bâtiments connectés. Elles permettent d'inclure le virtuel au réel de façon transparente et de rendre plus intuitives les interactions que l'utilisateur peut avoir avec les objets connectés. Le bénéfice pour l'utilisateur est d'interagir plus efficacement avec le virtuel, mais également d'améliorer l'accès à des services ubiquitaires fournis par un bâtiment connecté. L'ensemble de ces cas d'usage montre la pertinence des systèmes interactifs mixtes combinant réalité mixte, interfaces utilisateur naturelles et internet des objets, mais se heurte à des limitations et contraintes techniques des NUI employées, qui peuvent être contournées grâce à leur interconnexion.

2.2.3 NUI interconnectées et naturalité de l'interaction utilisateur

Nous nous intéressons dans cette section à l'interconnexion des NUI rendue possible par l'IoT. Cette interconnexion permet de rendre plus naturelles les interactions de l'utilisateur avec la partie numérique des objets connectés. La vision est la principale modalité d'interaction pseudo-naturelle entre un utilisateur et son environnement mixte. La définition la plus consensuelle de la réalité mixte de Milgram et al. [MK94] est d'ailleurs centrée sur cette modalité. L'amélioration du rendu visuel des systèmes interactifs mixtes est un axe majeur de l'élévation de leur naturalité et de leur intuitivité. Dans ces systèmes, les périphériques d'affichage connectés sont les médiateurs de la perception par un utilisateur de son environnement virtuel mixte au réel. Ils sont la clé de voûte du lien entre virtuel et réel, et sont pour cette raison essentiels tant à l'IoT dont ils font partie en tant qu'objets connectés qu'à la RM qu'ils génèrent. Ces périphériques peuvent d'ailleurs devenir eux-mêmes objets de l'interaction. Ils bénéficient alors de leur connectivité pour échanger des informations avec d'autres objets connectés. Ils fournissent ainsi des interactions homme-machine plus riches et évoluées. Néanmoins, comme nous l'avons décrit dans la section précédente, les NUI basées vision ont leurs propres limitations techniques. Par exemple, leur champ de vue est restreint, ou encore elles gèrent partiellement voire aucunement les occultations. Elles peuvent être associées à d'autres interfaces utilisateur naturelles afin de combiner leurs qualités et de contourner leurs défauts. L'objet de l'interaction est alors l'interface utilisateur elle-même, qui conserve néanmoins son rôle de médiateur de l'interaction, tout en collaborant avec une autre interface utilisateur naturelle grâce à l'IoT.

En particulier, la combinaison de plusieurs affichages permet de créer des systèmes interactifs mixtes à rendu multi-couches. Kang et al. [KD99] décrivent la possibilité de composer plusieurs rendus en un seul afin de créer une vue composite d'un environnement virtuel. Ce procédé est étendu à une composition du rendu d'un environnement mixte à l'aide de plusieurs dispositifs d'affichage

communicants. Plusieurs types de combinaisons sont décrits par la littérature.

L'association de *deux affichages statiques* a entre autres la capacité d'étendre le champ de vision d'un dispositif. Jones et al. [JBOW15] présentent un dispositif qui associe un projecteur et une télévision reliée à une console de jeux. Le projecteur étend le champ de vision de la télévision en affichant autour d'elle un rendu étendu de sa vue (cf. Figure 2.20). Ce dispositif est monoscopique, et ne prend pas en compte l'occultation des affichages. Ces deux facteurs limitent l'immersion de l'utilisateur et donc la naturalité de l'expérience utilisateur.

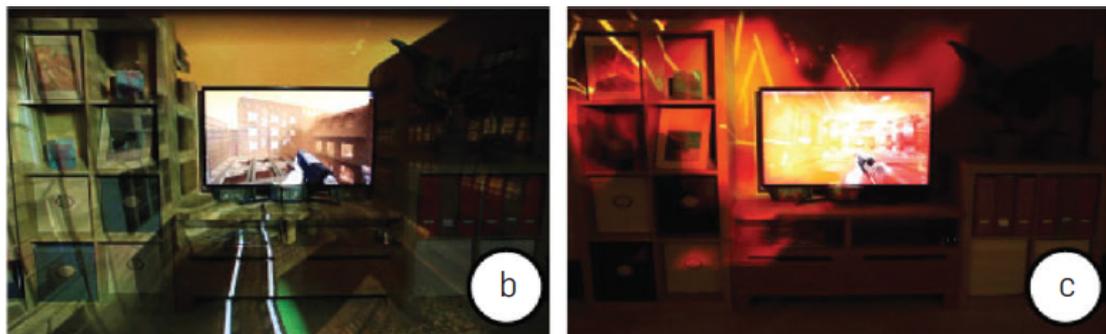


FIGURE 2.20 – Le système Illumiroom augmente le champ de vue d'un téléviseur relié à une console de jeu [JBOW15].

L'association d'*un affichage statique et d'une interface utilisateur de réalité mixte*, la plus commune, combine un affichage statique, par exemple un écran ou un projecteur, avec une interface utilisateur de réalité mixte, comme une tablette ou un casque Hololens/Magic Leap. La réalité augmentée spatialisée (SAR) utilise des projecteurs afin d'interagir intuitivement avec un environnement mixte. Elle ne requiert aucun équipement pour être perçue par l'utilisateur lorsqu'elle est monoscopique, ce qui simplifie son usage, assez fréquent notamment dans les musées. Pour cela, elle projette sur une surface réelle une texture alignée représentant un environnement virtuel. Les deux affichages sont colocalisés. Par exemple, Roo et al. [RH17] projettent une texture de volcan sur une maquette de volcan réalisée à partir de sable. Ils combinent également le rendu d'un casque de réalité mixte semi-transparent avec une maquette de moteur spatialement augmentée (cf. Figure 2.21). Ils indiquent par ailleurs que l'usage de projecteurs pose des problèmes dans la gestion des occultations.

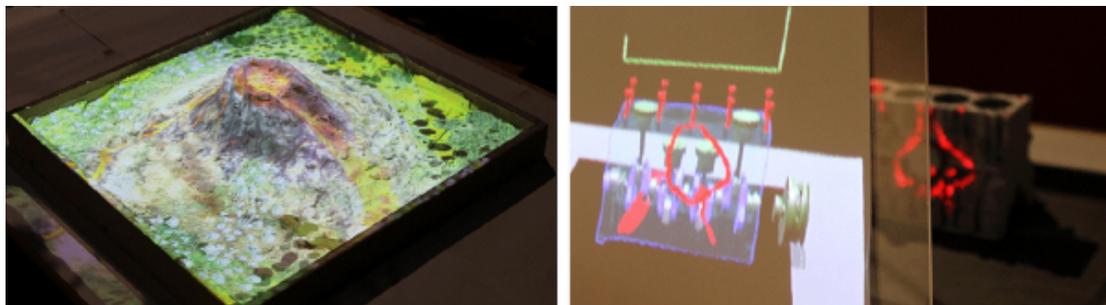


FIGURE 2.21 – Exemples de SAR par Roo et al.[RH17]. A gauche, maquette de volcan augmentée par projection. A droite, maquette de moteur augmentée par projection et combinée à un visio-casque de réalité mixte.

Autre exemple, Sandor et al. [SWM⁺02] présentent le système SHEEP. Ce système projette sur une table un jeu vidéo. Les utilisateurs jouent à ce jeu en interagissant avec plusieurs interfaces utilisateur intuitives, comme un écran tactile ou un affichage porté sur la tête (cf. Figure 2.22). Les affichages utilisés sont localisés grâce à un système ART tracking. Les occultations ne sont pas gérées par ce dispositif, ce qui nuit à la naturalité de l'interaction et produit chez l'utilisateur des ruptures de stéréopsis, entraînant de la fatigue oculaire et une rupture de l'immersion.

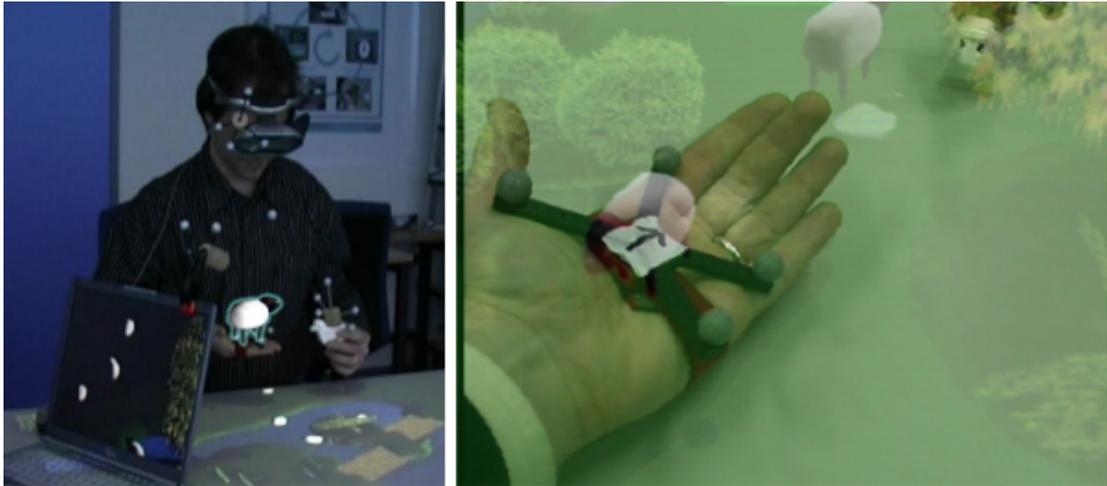


FIGURE 2.22 – Le système SHEEP combine un visiocasque de réalité augmentée à une projection monoscopique sur une table [MSW⁺03].

Benko et al. [BOZW15] présentent le système FoveAR, qui associe un casque de réalité mixte semi-transparent à une projection spatialisée augmentant un mur (cf. Figure 2.23). Son objectif est d'augmenter le champ de vision du casque par la projection. Ce système fournit une couche de rendu stéréoscopique combinée à une couche de rendu monoscopique contraint par la géométrie de la surface augmentée. Les téléviseurs sont également utilisés à la place de projecteurs dans ce type d'association.

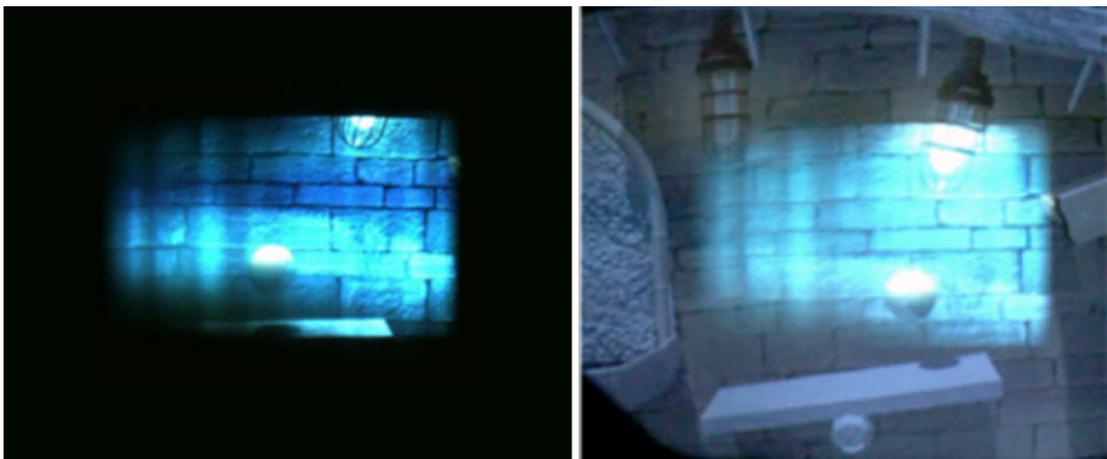


FIGURE 2.23 – Le système FoveAR combine un visiocasque de réalité augmentée à une projection monoscopique sur un mur [BOZW15]. A gauche, seul le visiocasque affiche la scène. A droite, la même scène est affichée simultanément par le visiocasque et le projecteur.

Kawakita et al. [KN14] complètent un programme télévisuel à l'aide d'un téléphone intelligent. Ce téléphone intègre une technique d'estimation de la pose de marqueurs fiduciels entourant le téléviseur afin de le colocaliser (cf. Figure 2.24). Cette technique est moins précise que les mobiles compatibles ARCore/ARKit ou les casques de réalité mixte comme Hololens ou Magic Leap, qui s'auto-localisent en reconstruisant leur environnement. De plus, elle ne résiste pas aux occultations des marqueurs et est très consommatrice de ressources. Néanmoins, les interactions utilisateur sont distribuées entre les interfaces utilisateur, ce qui apporte au système une multi-modalité utile pour simplifier et rendre plus intuitif son contrôle.

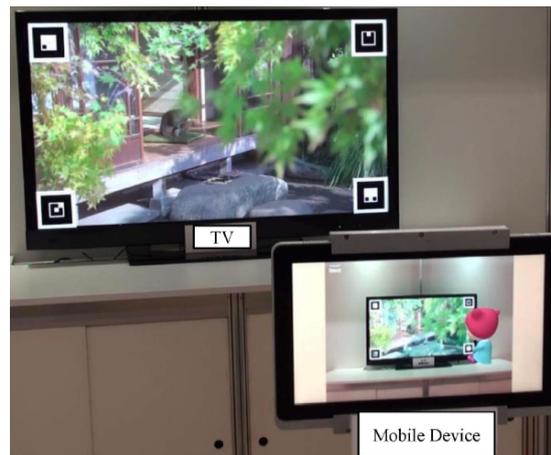


FIGURE 2.24 – Le système Augmented TV de Kawakita et al. augmente un téléviseur grâce à une tablette [KN14].

Baillard et al. [BFA⁺17] reprennent ce dispositif en utilisant un casque de réalité mixte semi-transparent et une tablette à la place d'un téléphone intelligent (cf Figure 2.25). Le casque bénéficie de sa capacité à s'auto-colocaliser avec le téléviseur pour que leur colocalisation soit plus robuste aux occultations. De plus, les occultations sont présentées comme étant gérées, même si cette gestion n'est pas décrite.



FIGURE 2.25 – Exemple d'association d'un Hololens et d'un téléviseur par Baillard et al. [BFA⁺17].

Enfin, Nishimoto et al. [NJ19] combinent un casque de réalité mixte semi-transparent, un Hololens, à une salle immersive CAVE2 afin d'étendre le champ de regard de la salle immersive au-dessus et au-dessous des écrans stéréoscopiques de la salle (cf. Figure 2.26). Ce système

combine deux rendus stéréoscopiques, l'un fourni par le casque et l'autre par la salle. Aucun mécanisme de gestion des occultations n'est présenté.

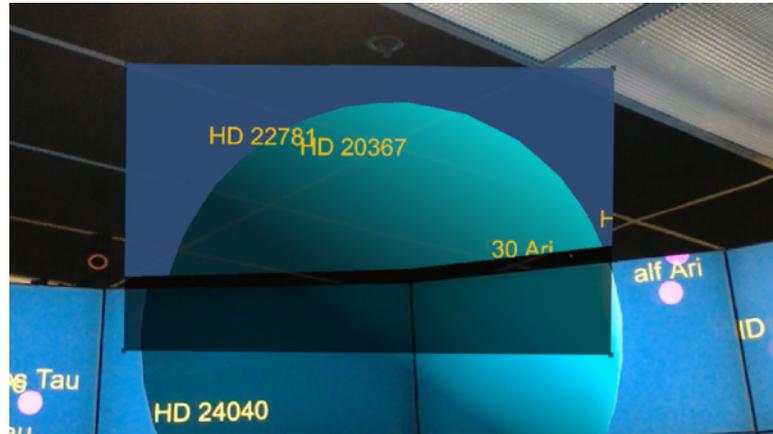


FIGURE 2.26 – Exemple d'association d'un Hololens à une salle immersive CAVE2 par Nishimoto et al. [NJ19].

Associer une interface utilisateur de réalité mixte à un affichage permet donc d'augmenter le champ de vision de l'affichage statique, mais également de composer une information publique fournie par l'affichage statique à une information personnelle fournie par l'interface utilisateur de réalité mixte. Néanmoins, la gestion des occultations n'est pas étudiée par les travaux présentés, alors qu'elles peuvent nuire à la qualité de l'expérience utilisateur.

L'association de deux interfaces utilisateur de réalité mixte combine deux interfaces utilisateur mixtes portées par l'utilisateur. Normand et al. [NM18] présentent un système interactif mixte de ce type. La première interface est un casque de réalité mixte soit semi-transparent soit vidéo, tandis que la seconde interface est un téléphone intelligent, capable de reconstruire son environnement et de s'y localiser. Son objectif est d'étendre le champ de vision du téléphone grâce au casque. Il permet également d'interagir avec le contenu affiché par le téléphone intelligent (cf. Figure 2.27). Les occultations de la main sont gérées dans le cas du casque de réalité mixte semi-transparent. Les interactions sont partagées entre les deux interfaces.



FIGURE 2.27 – Le système VESAD par Normand et al. [NM18] permet d'augmenter le champ de vision d'un téléphone intelligent à l'aide d'un visiocasque.

Ce type de système est utile pour améliorer les capacités d'interaction d'un objet connecté. Il implique néanmoins une communication et une synchronisation des comportements des interfaces utilisateur utilisées. Le système final est donc plus difficile à concevoir et implémenter qu'un simple objet connecté du fait de sa complexité.

Travaux	RM		NUI		IoT	
	réel	virtuel	vision	autres modalités	objet	réseau
Sandor et al. 2002	-	jeu vidéo	projection monoscopique visiocasque tablette	gestes main tactile tangibles	-	local
Petersen et al. 2009	papier	papier	écran monoscopique statique	gestes main toucher	pompe industrielle	-
Mistry et al. 2009	journal	vidéo montre carte	projection monoscopique portée	gestes main	journal	-
Sulisz et al. 2012	objet	contenu associé	visiocasque stéréoscopique	téléphone tangible	objet tagé RFID	RFID
Kawakita et al. 2014	-	programme TV	écran monoscopique statique tablette	tactile	TV monoscopique	local
Benko et al. 2015	mur	mur	projection monoscopique visiocasque stéréoscopique	locomotion	-	local
Lin et al. 2017	-	-	visiocasque monoscopique	gestes main regard	ordinateur imprimante projecteur haut-parleurs	local
Roo et al. 2017	maquette volcan maquette moteur	volcan informations	projection monoscopique visiocasque stéréoscopique	gestes main	maquettes	local
Baillard et al. 2017	- -	contenu associé	visiocasque stéréoscopique tablette	gestes main	TV monoscopique	local
Kim et al. 2018	-	assistant personnel incarné	visiocasque stéréoscopique	voix gestes main locomotion	assistant personnel intelligent	local internet
Normand et al. 2018	téléphone	téléphone	visiocasque stéréoscopique	gestes main	téléphone intelligent	-
Nishimoto et al. 2019	-	graphe	visiocasque stéréoscopique écran stéréoscopique statique	-	-	local

TABLE 2.1 – Comparaison des cas d'usage explorant la convergence entre réalité mixte (RM), interfaces utilisateur naturelles (NUI) et internet des objets (IoT).

L'ensemble des travaux présentés dans cette sous-section, regroupés dans la Table 2.1, explorent le potentiel des systèmes combinant RM, NUI et IoT. Ces travaux ont également la caractéristique commune d'avoir produit des preuves de concept et des prototypes non-réutilisables. Ils ne proposent pas de modèle unifiant les domaines de la réalité mixte, des interfaces utilisateur naturelles et de l'internet des objets. Ils ne proposent pas non plus d'outils ou de frameworks de conception et de développement. Or, la gestion de la complexité de ces systèmes est un des principaux défis qu'ils posent. Ils nécessitent des outils permettant de gérer la complexité inhérente à la quantité d'objets connectés à ces systèmes. Un modèle unifié de ces systèmes est également primordial afin de les rendre interopérables et extensibles.

2.3 Modèles et frameworks de conception ou de développement

Des modèles et frameworks sont utilisés depuis le début des systèmes interactifs afin de les produire plus efficacement. Ils permettent de mettre en place des méthodologies simplifiant et facilitant la création de ces systèmes. Nous choisissons les critères suivants afin de les comparer :

- la synchronisation entre réel et virtuel : elle est nécessaire à la fois à la réalité mixte et à l'internet des objets. En effet, ces deux paradigmes relient réel et virtuel, la réalité mixte du point de vue utilisateur, et l'internet des objets du point de vue réseau ;
- la césure entre réel et virtuel : ne pas distinguer le réel du virtuel est rédhibitoire dans la modélisation de l'ensemble des domaines que nous couvrons dans ce manuscrit. Cela a pour conséquence de produire de la confusion entre virtuel et réel lors de la conception et du développement ;
- des patrons de projet dédiés aux plateformes couvertes permettent d'accélérer et d'uniformiser le développement d'un système interactif mixte ;
- des bibliothèques de composants logiciels dédiées à la réalité mixte, aux interactions pseudo-naturelles et à l'internet des objets sont des outils fournis aux développeurs. Ils permettent de raccourcir les temps de développement, d'uniformiser les systèmes interactifs mixtes et de contourner la nécessité de faire appel à des spécialistes de ces domaines ;
- la flexibilité de l'interconnectivité des TIES produits est importante pour leur interopérabilité et leur extensibilité. Elle leur permet de s'adapter à des contextes et des environnements divers et de s'interconnecter à moindre coût à d'autres systèmes.

Modèles de conception d'interfaces graphiques

Divers modèles ont été créés afin de concevoir et mettre en œuvre des interfaces utilisateur graphiques. Ils modélisent une interface sous la forme d'une composition d'objets interactifs. Les motifs d'architecture Modèle-Vue-Contrôleur (MVC) [Wal08] et Présentation-Abstraction-Contrôle (PAC) [Cou87] datent des années 80. Chaque objet interactif présente à l'utilisateur son état interne et permet des opérations sur cet état [HC97]. MVC distingue les modèles des données à afficher, la vue qui présente ces données et le contrôleur en charge des actions de l'utilisateur. PAC, quant à lui, considère également un triptyque de concepts. Le premier, la présentation, est en charge des interactions avec l'utilisateur. Le second, l'abstraction, gère les données et les fonctions du système. Le troisième, le contrôleur, fait le lien entre présentation et abstraction. Ces motifs d'architecture sont inadéquats pour décrire de manière détaillée des interactions homme-machine en réalité mixte. En effet, ils découpent un système par leur fonctionnel, mais ne distinguent pas les entités composant ces systèmes ni ne décorrèlent les composantes réelles et virtuelles de chacune.

Plus récemment, des approches comme Lacoche et al. [LLCVF19] ont tenté d'étendre ces modèles pour les appliquer à l'IoT. Ils présentent un modèle et des outils basés sur PAC afin de simuler

et d'implémenter des environnements ubiquitaires grâce à la réalité virtuelle. Néanmoins, ce framework ne propose pas de modèle générique des entités constitutives de ces systèmes. Il permet de reproduire des interactions se produisant dans un environnement réel ou virtuel dans son environnement jumeau, sans compléation entre interactions dans un environnement réel et interactions dans un environnement virtuel. Il n'est au final pas adapté à l'interaction en RM, et donc inapplicable à la conception ou à l'implémentation de systèmes interactifs mixtes.

Frameworks de conception de systèmes interactifs mixtes

Jacob et al. [JGH⁺08] proposent un framework descriptif. Destiné à la conception de nouvelles interactions, il englobe un ensemble de recommandations thématiques visant à comprendre, comparer et décrire les interactions émergentes des systèmes interactifs mixtes. Les thèmes couverts sont la connaissance et la maîtrise des lois physiques qui régissent notre monde, de son propre corps, de son environnement et enfin de la sociabilité. Si ces recommandations sont utiles aux concepteurs, elles ne lui fournissent pas de modèle, ni d'outils pour modéliser un système ou le développer.

Bouzekri et al. [BCMP18] présentent une architecture générique des systèmes interactifs mixtes. Cette architecture tient compte de l'interaction de l'utilisateur, des interfaces utilisateur, d'un modèle comportemental de l'interface, mais ne corrèle pas le réel du virtuel ni ne décrit leurs comportements respectifs. Elle ne propose pas de modèle unifié des différentes entités constitutives de ces systèmes.

Pfeiffer et al. [PPL18] proposent un framework de conception en réalité virtuelle des systèmes interactifs mixtes. Il prend place lors de la première étape d'inspiration et d'exploration de la méthode de conception innovante dite du double diamant¹⁴. Ce système permet de simuler des prototypes combinant internet des objets et interfaces utilisateur naturelles. Ces prototypes en réalité virtuelle sont basés sur une image panoramique à 360° d'un environnement réel incluant des objets connectés. Ils permettent de confronter virtuellement un prototype à son environnement futur afin d'améliorer sa conception avant sa mise en production. Ce framework permet de décrire les interactions avec l'internet des objets ainsi qu'avec les interfaces utilisateur naturelles. Il fait appel au protocole MQTT¹⁵ de l'internet des objets. C'est un protocole de machine à machine léger, utilisable dans le cas de réseaux à faible bande passante. Ce framework utilise également Node-RED¹⁶, une application utilisée pour relier objets connectés, interfaces logicielles et services de l'internet des objets. Les prototypes obtenus offrent alors un aperçu du système interactif mixte final, qui reste à développer.

Frameworks d'implémentation de systèmes interactifs mixtes

Les systèmes interactifs mixtes ont été étudiés dans le cadre de la convergence entre réalité augmentée, interfaces utilisateur tangibles et informatique ubiquitaire. MacWilliams et al. [MSW⁺03] présentent un framework de développement dédié à cette convergence appelé DWARF. Ce framework vise un prototypage et offre un processus de développement connecté. Il propose aux développeurs d'utiliser un réseau de Petri pour implémenter le comportement de leurs systèmes interactifs mixtes en créant le graphe comportemental de chaque interface utilisateur du système. Il ne présente néanmoins pas d'outil de conception ni de modèle générique applicable aux systèmes interactifs mixtes combinant RM, NUI et IoT.

Dubois et al. [DBAG14] proposent un framework basé modèles afin de développer des systèmes interactifs mixtes. Ce framework permet de créer de multiples modèles d'un système interactif mixte dont les règles et la syntaxe sont définies par un métamodèle. Ces modèles font le lien entre la représentation d'aspects saillants du système et le système final. Ces modèles sont ensuite regroupés dans un graphe de modèles. Néanmoins, ce framework ne propose pas de modèle unifié de l'ensemble des entités qui constituent les systèmes interactifs mixtes ni ne décrit les

14. <https://www.economie.gouv.fr/apie/publications/infographies/methode-double-diamant>

15. <http://mqtt.org/>

16. <https://nodered.org/>

mécanismes précurseurs du comportement de ces entités. Il tend également à confondre réel et virtuel.

Frameworks d'implémentation orientés internet des objets

Kelaidonis et al. [KSF⁺12] proposent un framework de virtualisation des objets réels ainsi que de gestion de la cognition de ces objets virtuels. Cette gestion permet de combiner les fonctionnalités des objets virtuels afin de répondre aux demandes de service des utilisateurs. Elle permet aux objets virtuels de reconnaître un contexte et de prendre une décision en fonction de ce contexte. Ce framework offre une approche de la relation entre réel et virtuel basée sur la sémantique dans un contexte IoT. Néanmoins, l'interaction avec l'utilisateur n'est pas abordée.

Nitti et al. [NPCA16] étudient comment les objets virtuels sont devenus un élément majeur de l'internet des objets. Les objets virtuels sont essentiels à la découverte des objets connectés ainsi qu'à la composition de services. Ils jouent également un rôle crucial dans la gestion de la complexité, de l'hétérogénéité et de l'extensibilité de l'internet des objets. Cette étude décrit la relation entre objets réels et objets virtuels dans l'internet des objets, mais ne traite pas d'interaction pseudo-naturelle entre les objets connectés et l'utilisateur.

Ces travaux n'abordent pas dans sa totalité notre problématique de modélisation et de méthodologie de mise en œuvre des systèmes interactifs mixtes combinant RM, NUI et IoT. Certains proposent des méthodologies de conception ou de mise en œuvre de systèmes s'approchant de ces systèmes, mais sans prendre en compte l'ensemble de leurs spécificités. Or, la convergence des domaines de la réalité mixte, des interfaces utilisateur naturelles et de l'internet des objets doit tenir compte des spécificités communes à ces domaines tant lors de leur conception que de leur mise en œuvre.

Travaux	synchro. réel/virtuel	césure réel/virtuel	projet-type plateforme	bibliothèques RM/NUI/IoT	flexibilité
MacWilliams et al. 2003	oui	oui	non	non	statique
Kelaidonis et al. 2012	oui	oui	non	non	dynamique
Dubois et al. 2014	oui	oui	non	non	dynamique
Nitti et al. 2016	oui	oui	non	non	non
Bouzekri et al. 2018	oui	non	non	non	non
Pfeiffer et al. 2018	simulée	simulée	non	non	non
Lacoste et al. 2019	oui	oui	non	non	statique

TABLE 2.2 – Comparaison des modèles et frameworks applicables aux systèmes interactifs mixtes combinant la réalité mixte (RM), les interfaces utilisateur naturelles (NUI) et l'internet des objets (IoT).

Nous en concluons qu'aucun modèle ou framework de la littérature n'est adapté à une conception et une mise en œuvre unifiées des systèmes interactifs mixtes combinant RM, NUI et IoT. Un modèle unifiant ces domaines est donc nécessaire pour concevoir et mettre en œuvre ces systèmes à l'aide d'outils dédiés.

2.4 Analyse

L'utilisateur des systèmes interactifs mixtes est confronté à des systèmes aux propriétés spécifiques. La Figure 2.28 présente ces spécificités en combinant la synergie entre pratique et affordance des objets selon Turner [Tur05], les objets mixtes de Coutrix et al. [CNR05], et le continuum des interactions intuitives de Blackler et al. [BH07]. Cette figure décrit comment l'utilisateur perçoit l'affordance mixte d'un objet mixte. Cette affordance combine une connaissance ancienne d'un

objet réel à une affordance perçue d'une métaphore comprise et connue ou non. L'affordance physique est alors un élément facilitateur de l'apprentissage de l'affordance perçue.

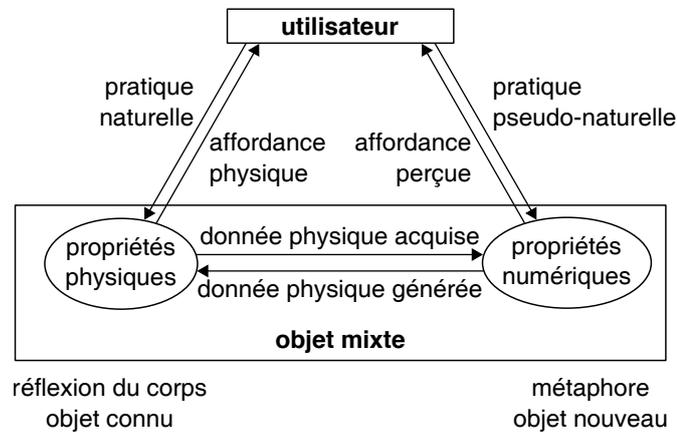


FIGURE 2.28 – L'utilisateur perçoit l'affordance mixte de l'objet mixte, composée de l'affordance physique et de l'affordance perçue de cet objet.

La réalité mixte permet de percevoir des environnements réels et virtuels simultanément. Elle permet d'interagir sur chacun de ces environnements en utilisant au mieux leurs spécificités et de combiner leurs avantages tout en contournant leurs inconvénients. Les interfaces utilisateur naturelles fournissent un moyen d'interagir avec le virtuel en simulant des interactions réelles, rendant ces interactions naturelles et intuitives. L'internet des objets connaît quant à lui un développement rapide. Il fournit des services ubiquitaires en reliant réel et virtuel grâce au réseau. Nous pensons que l'utilisateur d'un TIES (système interactif à entités mixtes, ou inTeractive mIxed Entity Systems en anglais) bénéficie d'une expérience utilisateur améliorée comparativement aux autres systèmes interactifs. En effet, l'utilisateur interagit grâce à ces systèmes avec le réel et le virtuel de façon ubiquitaire, naturelle et intuitive. La Figure 2.29 présente ce triptyque sous la forme d'un diagramme de Venn.

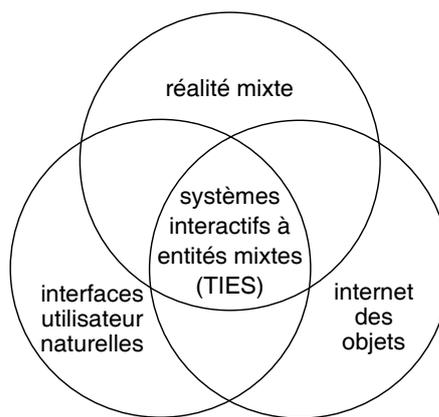


FIGURE 2.29 – Diagramme de Venn illustrant la convergence entre réalité mixte, interfaces utilisateur naturelles et internet des objets sous la forme des TIES.

La conception et la mise en œuvre des TIES présentent de multiples défis. Le nombre très élevé d'objets connectés est un frein au développement de ces systèmes. D'une part, il rend ces systèmes complexes à modéliser, à développer et à étendre. D'autre part, leur développement se

heurte à l'hétérogénéité des dispositifs, ce qui entraîne une fragmentation élevée de ces systèmes. Cette fragmentation nuit à leur interopérabilité et à l'apparition d'un modèle unifié de référence. Les approches présentées ne couvrent pas la diversité inhérente aux systèmes interactifs mixtes combinant à la fois réalité mixte, interfaces utilisateur naturelles et internet des objets. Néanmoins, ils couvrent voire combinent certains de ces domaines. La proposition d'un modèle de référence est donc nécessaire pour que ces systèmes puissent croître en nombre. Or, un modèle unifiant l'ensemble de ces domaines permettrait de créer des outils de conception et de développement nécessaires à la croissance et à l'interopérabilité des systèmes interactifs mixtes. Il permettrait également de pouvoir fractionner ces systèmes en sous-systèmes afin de maîtriser leur complexité.

De plus, les travaux étudiant de nouvelles interactions grâce à ces systèmes sont des prototypes et des preuves de concept. Ces travaux bénéficieraient d'un tel modèle et d'outils basés sur ce modèle afin de les pérenniser et de les rendre interopérables. L'interopérabilité d'un système est sa capacité à fonctionner avec d'autres systèmes de la façon la plus transparente et fluide possible. Ces nouvelles interactions visent à améliorer l'expérience des utilisateurs de ces systèmes. La naturalité de ces interactions est un autre aspect qui mérite d'être approfondi. Rendre ces systèmes complexes simples et intuitifs à utiliser est nécessaire pour que les utilisateurs les adoptent. Cette facilité d'usage doit masquer les mécanismes qui les composent. Elle requiert l'amélioration des techniques d'interaction pseudo-naturelle existantes. Nous décrivons des solutions permettant d'améliorer la naturalité et l'intuitivité de l'interaction entre utilisateurs et systèmes interactifs mixtes dans le chapitre 5. Cette amélioration peut résulter de leur extension à la multimodalité et la multi-sensorialité, qui sont des facteurs essentiels des interactions naturelles entre l'homme et son environnement. Elle peut également provenir de la combinaison des affordances d'un objet mixte.

Nous avons vu dans ce chapitre qu'il n'existe pas de modèles ni d'outils de conception et de développement des systèmes interactifs mixtes unifiant la réalité mixte, les interfaces utilisateur naturelles et l'internet des objets. Les solutions décrites fournissent des réponses partielles aux problèmes rencontrés. D'autres solutions sont dédiées à certains aspects techniques de la mise en œuvre des TIES. Par exemple, des bibliothèques logicielles telles que ARCore, ARKit ou MixedRealityToolkit fournissent des composants logiciels intégrables à des applications. Elles permettent à des ordinateurs équipés de caméras et d'un affichage de se localiser dans leur environnement réel et de reconstruire son jumeau virtuel. Ils peuvent offrir des composants basiques d'interaction pseudo-naturelle. Mais d'autres outils de l'internet des objets sont nécessaires pour interconnecter les entités constitutives de ces systèmes. Ces outils ne sont pas conçus pour travailler de façon complémentaire et ne s'adressent pas à la même plateforme. De plus, aucun modèle n'unifie l'ensemble de ces domaines. Des approches plus évoluées proposent des modèles [CNR05], des frameworks de conception et de développement dédié aux systèmes interactifs mixtes [DBAG14]. Mais ils ne s'appuient pas sur un modèle fédérateur capable d'unifier la réalité mixte, les interfaces utilisateur naturelles et l'internet des objets. Par exemple, le modèle d'objet mixte [CNR05] décrit les propriétés ainsi que les interactions internes et externes des objets mixtes composant les systèmes interactifs mixtes. Ce modèle distingue le réel du numérique et décrit leur corrélation. De plus, il intègre à l'objet mixte les interfaces utilisateur naturelles. Or, l'absence de modèle unifié de ces entités de différentes natures ne permet pas aux concepteurs et aux développeurs de développer des outils génériques. Ce manque ne favorise pas la création des systèmes interactifs, ni leur uniformisation, ni leur interopérabilité. Il ne modélise pas non plus les comportements réels et numériques d'un objet mixte. Les concepteurs et développeurs sont par conséquent limités lors de la création de comportements réels et numériques. Ces comportements sont de fait synchrones mais pas complémentaires. Or, les concepteurs et développeurs ont besoin de modèles permettant d'abstraire ces systèmes. Ils ont aussi besoin de méthodes et d'outils pour les concevoir et les développer. Ces modèles, méthodes et outils facilitent l'uniformisation de leur création, dont découle leur interconnectivité et leur extensibilité.

Actuellement, un concepteur d'un système interactif mixte assemble des interfaces utilisateur et des objets connectés en fonction des besoins exprimés. L'architecture qui en découle relie les utilisateurs aux objets connectés grâce à des interfaces utilisateur naturelles et des technologies réseau. Or, aucun modèle générique ne permet de décrire à la fois les utilisateurs, les interfaces utilisateur naturelles et les objets connectés grâce à un seul modèle générique. En particulier, un amalgame entre réel et virtuel s'opère. Ces environnements sont souvent confondus alors qu'ils ont des caractéristiques bien distinctes. L'architecture qui en résulte est donc hétérogène, peu interopérable et faiblement extensible. Les développeurs en charge du développement du système utilisent alors les outils à leur disposition, qui ne sont pas dédiés à ces systèmes. Ils implémentent des solutions spécifiques à chacun des systèmes qu'ils créent. Les applications qui animent ces systèmes s'appuient alors sur des technologies de communication et d'interaction naturelle spécifiques à l'écosystème de leurs concepteurs et développeurs. Les temps de développement sont augmentés par ce manque de factorisation des ressources de développement. Cette approche du développement rend également l'interopérabilité de ces systèmes faible. Cela implique de reprendre le développement de ces systèmes pour les rendre interopérables et extensibles.

Les TIES souffrent donc d'une complexité importante, liée à la multiplicité et à l'hétérogénéité des plateformes impliquées. Les TIES existants confondent souvent réel et virtuel et ne s'appuient pas sur des comportements réels et virtuels complémentaires. Ils sont longs et fastidieux tant à concevoir qu'à implémenter. Ils ne bénéficient pas de pratiques les unifiant, ils sont donc peu interopérables et extensibles. De plus, ils requièrent des compétences de développement poussées dans des domaines très divers.

Chapitre 3

Modèle interne mixte orienté conception

L'hétérogénéité et la diversité des ressources employées [DBAG14] rendent la conception et la mise en œuvre des systèmes interactifs mixtes complexe [RBMN06]. Le chapitre 2 a mis en évidence l'intérêt de solutions pour implémenter ces systèmes [MSW⁺03], pour les modéliser [CNR05], et pour unifier les pratiques [DGN10]. La conception et la mise en œuvre des TIES (systèmes interactifs à entités mixtes, ou *inT*eractive *mx*ed Entity Systems en anglais) rencontrent des problématiques similaires, plus l'inclusion de l'ensemble des interfaces utilisateur naturelles et l'intégration des spécificités de l'internet des objets.

De cette nécessité de traiter la complexité des TIES, est né le besoin de les structurer. Notre approche consiste donc à proposer un modèle unifié des entités hétérogènes constitutives des TIES et de leurs interactions. Nous appelons ce modèle : modèle interne mixte orienté conception, ou en anglais *Design-Oriented Mixed Reality Internal Model* (DOMIM) (cf. Figure 3.1).

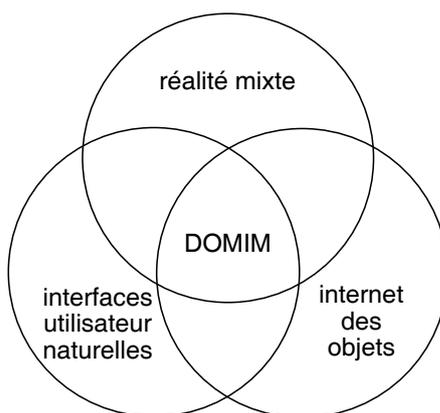


FIGURE 3.1 – Le modèle DOMIM permet d'uniformiser la réalité mixte, les interfaces utilisateur naturelles et l'internet des objets.

DOMIM permet d'unifier les TIES puisque tout ce qui les constituent est entité. Ces entités sont les utilisateurs, les interfaces utilisateur et les objets connectés. DOMIM s'appuie sur la notion d'entité mixte, une entité possédant à la fois une partie réelle et une partie virtuelle qui sont corrélées tout en ayant chacune la capacité d'interagir avec leur environnement. Les entités mixtes

résolvent ainsi le problème de la cohérence entre réel et virtuel. Chaque entité mixte possède un modèle externe, qui résout le problème des relations entre entités et de l'interaction utilisateur si cet utilisateur est une entité. Une entité mixte possède également un modèle interne, qui résout de manière opérationnelle la synchronisation bi-directionnelle entre ses parties réelles et virtuelles. Si DOMIM permet de modéliser chacun des domaines de la réalité mixte, des interfaces utilisateur naturelles et de l'internet des objets, il exprime tout son potentiel lorsque ces domaines sont associés.

DOMIM a pour objectifs :

- de faciliter la modélisation des systèmes interactifs mixtes grâce à un modèle unifié. DOMIM vise à fournir un formalisme graphique adapté à la conception des TIES ;
- d'utiliser les outils développés grâce à ce modèle. Les concepteurs et les développeurs peuvent s'appropriier ces outils afin d'uniformiser, de simplifier et d'accélérer la réalisation de leurs tâches.

Ce chapitre est structuré comme suit.

Dans un premier temps, nous décrivons DOMIM, un modèle générique des entités constituant les TIES et de leurs interactions. Ce modèle permet de décomposer ces systèmes en sous-systèmes composés d'entités mixtes, ainsi que de décrire et d'uniformiser simplement les interactions entre entités mixtes.

Dans un second temps, nous nous intéressons aux mécanismes internes des entités mixtes. DOMIM modélise les capacités comportementales de leurs parties réelles et virtuelles. Nous étudions le contrôle d'une entité mixte en fonction de sa nature. En effet, une entité mixte peut être asservie, autonome ou partiellement contrôlable.

Dans un troisième temps, nous étudions les interactions entre les entités d'un système interactif mixte. Nous analysons notamment l'interaction entre l'utilisateur et les autres entités de ce système. Nous montrons comment DOMIM permet la modélisation et le développement d'interactions utilisateurs hybrides, combinant les capacités d'interaction de plusieurs interfaces utilisateur naturelles.

3.1 Entités mixtes

DOMIM définit la relation entre réel et virtuel au sein de chaque entité d'un TIES en les décorrélant tout en les associant. Ce modèle décrit également l'ensemble des interactions se déroulant dans les TIES, qu'elles soient externes, c'est-à-dire entre entités ou sous-systèmes, ou internes à chaque entité. DOMIM s'appuie pour ce faire sur le concept d'entité mixte, que nous définissons de la façon suivante :

Une entité mixte possède une partie réelle et une partie virtuelle. Chacune de ces parties possède des propriétés qui la décrivent ainsi qu'un comportement propre. Les propriétés et comportements corrélés des parties réelles et virtuelles sont synchronisés en temps réel.

Les utilisateurs, les interfaces utilisateur naturelles et les objets connectés sont tous des entités mixtes. Le modèle d'entité mixte se décompose entre un modèle externe et un modèle interne. Le modèle externe décrit les interactions d'une entité mixte avec son environnement. Le modèle interne décrit ses interactions internes. Par exemple, un ventilateur mixte est une entité mixte. Il a une partie réelle, l'objet physique avec des pales et un bouton pour l'actionner. Il possède également une partie virtuelle, constituée d'un modèle 3D ainsi que de propriétés, qui copie ou complète sa partie réelle. Copier la partie réelle implique de reconstruire sa géométrie afin de créer le jumeau virtuel de l'objet. La compléter implique de créer une partie virtuelle additionnelle qui étend la nature de l'objet mixte. L'utilisateur peut alors interagir avec elle grâce à des interactions pseudo-naturelles. Le geste de l'entité «utilisateur» exprimant son intention d'allumer l'entité «ventilateur

connecté» est une interaction externe. L'entité «casque de réalité mixte» capte et interprète ce geste et transmet alors l'intention de l'utilisateur au ventilateur. Un exemple d'interaction interne du ventilateur connecté est sa perception du toucher de la main virtuelle de l'utilisateur qui déclenche alors sa mise en route dans son environnement virtuel (cf. Figure 3.2).

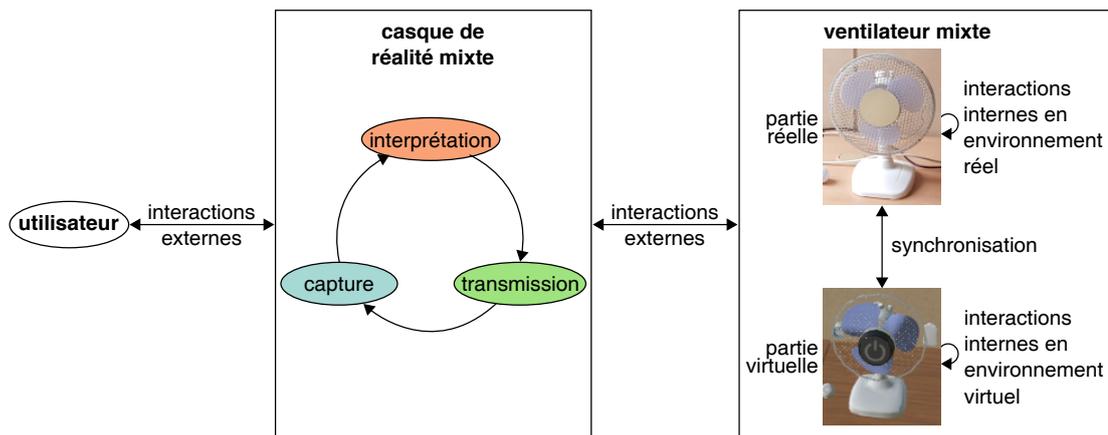


FIGURE 3.2 – Exemple d'entité mixte : un ventilateur composé d'une partie réelle et d'une partie virtuelle synchronisées. Les interactions entre le ventilateur et son environnement sont ses interactions externes. Les interactions à l'intérieur de chaque partie sont ses interactions internes réelles ou virtuelles. Les interactions internes entre les parties du ventilateur sont ses interactions internes mixtes.

Comme nous l'avons décrit précédemment, la complexité des systèmes interactifs mixtes est un frein important à leur conception et à leur développement. Décomposer un système complexe en sous-systèmes interdépendants est une solution qui permet de maîtriser cette complexité. DOMIM permet de segmenter un système interactif mixte en sous-systèmes à granularité multiple. Cette granularité s'étend de la description des relations entre systèmes au fonctionnement interne d'une entité d'un système, en passant par la description interne d'un sous-système.

Nous décrivons dans les sections suivantes notre modèle DOMIM, étape par étape. Tout d'abord, dans la section 3.1.1, nous présentons le modèle externe d'une entité mixte. Puis, dans la section 3.1.2, nous décrivons son modèle interne étape par étape. Nous étudions ensuite dans la section 3.1.3 le cas de l'utilisateur. Nous y décrivons les interactions internes au corps de l'utilisateur, ses interactions internes réelles. Nous détaillons également les interactions internes à son avatar, ses interactions internes virtuelles.

3.1.1 Interface de l'entité mixte

L'interface d'une entité mixte décrit les interactions des parties d'une entité mixte avec leur environnement (cf. Figure 3.3). Il modélise de la même manière toutes les entités d'un TIES. Il permet de créer des entités mixtes aux comportements réels et virtuels distincts. Cela a pour effet de décorréler leurs interactions qui se déroulent dans des environnements distincts. Les TIES peuvent alors améliorer l'expérience de leurs utilisateurs. En effet, chaque environnement a des caractéristiques qui lui sont propres. Ces caractéristiques conditionnent les interactions qui peuvent s'y dérouler. Les entités mixtes bénéficient ainsi des qualités du réel et du virtuel tout en contournant leurs défauts. Dès lors, il permet d'identifier et de généraliser l'usage de patrons de conception des TIES. Il améliore également la factorisation des ressources, méthodes et outils utiles à la conception et à la mise en œuvre de ces systèmes. En effet, en reproduisant des patrons d'architecture, il permet de

réutiliser les travaux antérieurs et d'uniformiser les pratiques usuelles de mise en œuvre. Il permet également de rendre ces systèmes interopérables. L'uniformisation de la mise en œuvre permet effectivement de mettre en place des mécanismes d'échange entre entités mixtes par le réseau s'appuyant sur un langage commun. En outre, il améliore leur extensibilité en facilitant l'interconnexion de plusieurs TIES. Les TIES interconnectés deviennent alors des sous-systèmes d'un TIES de plus grande dimension.

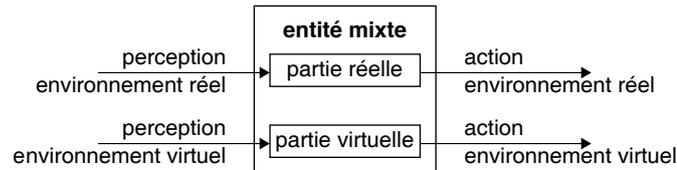


FIGURE 3.3 – L'interface d'une entité mixte. Chaque partie interagit avec son environnement.

3.1.2 Modèle interne mixte orienté conception d'une entité mixte

Notre proposition est de compléter l'interface d'une entité mixte en décrivant pour chacune de ses parties un modèle interne de ses mécanismes d'interaction et en les synchronisant. En effet, une entité mixte est constituée de deux parties distinctes, ayant chacune un comportement dans son environnement pouvant se répercuter dans l'autre environnement. Une entité mixte n'est pas obligatoirement un objet réel doté d'un jumeau virtuel. Certes, les parties réelles et virtuelles nécessitent généralement d'être colocalisées pour être perçues simultanément. Mais elles peuvent également être complémentaires et bénéficier des avantages de chaque monde. Cela permet de créer des comportements partiellement autonomes pour chaque partie d'une entité mixte. Par exemple, le robot B15 de Microsoft est un exemple d'entité mixte qui illustre notre approche. Ce robot mixte possède une partie réelle qui est la partie inférieure de son corps tandis que la partie supérieure de son corps est visualisée en réalité augmentée, comme le présente la Figure 3.4.



FIGURE 3.4 – Le robot mixte B15 de Microsoft. Ce robot est composé d'une partie réelle mobile et d'une partie virtuelle rendue par un Hololens.

Afin de pouvoir concevoir et développer une entité mixte, il est nécessaire de décrire et de synchroniser ses comportements. Une entité mixte doit tout d'abord s'informer sur son environnement. Ensuite, elle analyse cette connaissance de son environnement afin de décider comment agir. Enfin, elle agit dans son environnement. Pour modéliser son comportement, nous utilisons la boucle perception-décision-action [FMG11], couramment utilisée pour décrire les mécanismes internes d'interaction d'une entité (cf. Figure 3.5).

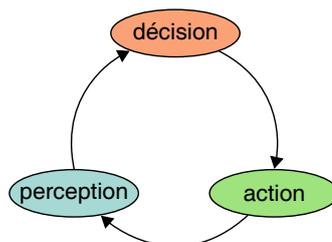


FIGURE 3.5 – La boucle perception-décision-action

La perception d'une entité est une interface entre son environnement et son intériorité[BH09]. Elle permet d'organiser, d'identifier et d'interpréter des informations sensorielles afin d'appréhender son environnement [Sch11]. Les informations sensorielles sont capturées par le système perceptuel. Ce système est composé de plusieurs sous-systèmes, également appelés canaux sensoriels. Ces canaux sont par exemple la vision ou l'audition, utilisés communément par les êtres humains, les interfaces utilisateur naturelles et les objets connectés. Cette prise d'information est active, car elle nécessite d'activer les canaux sensoriels et d'agir pour obtenir ces informations, par exemple en naviguant dans son environnement.

La décision est le résultat d'un processus décisionnel complexe, que Simon [Sim72] modélise

comme la composante de trois phases :

- la phase d'intelligence combine l'exploitation des informations sensorielles et l'identification des décisions possibles ;
- la phase de conception recherche des solutions connues et conçoit de nouvelles solutions ;
- la phase de choix détermine la solution optimale et valide ce choix.

L'action est nécessaire à la perception active. Elle permet de se mouvoir et d'émettre des informations sensorielles destinées à son environnement comme le son. Elle sollicite le système moteur de l'entité mixte afin de se mouvoir et d'activer ses effecteurs.

Nous souhaitons que le modèle interne d'une entité mixte soit générique à toutes ses parties afin de simplifier leur coordination. Pour cela, nous intégrons la boucle perception-décision-action à chaque partie de notre modèle d'entité mixte et les synchronisons. Nous pouvons ainsi l'appliquer aux humains, mais également aux interfaces utilisateur naturelles et aux objets connectés. Cette application est évidente à l'homme, dans la mesure où ce modèle comportemental est issu de son étude. Elle est également applicable aux objets intelligents, qui ont la capacité d'observer et d'analyser leur environnement [SLRHM12] et d'y agir de façon autonome [Pos11]. Nous considérons dans ce document les objets intelligents comme étant des objets mixtes. Typiquement, un robot-aspirateur reconstruit une représentation virtuelle de son environnement réel. Son environnement réel est généralement le domicile de son propriétaire. Le robot mixte détermine et simule à l'aide de cette reconstruction ses comportements virtuels avant de les effectuer dans le monde réel. Il peut alors aspirer le sol des pièces qu'il explore. Les interfaces utilisateur naturelles ont également pour fonction de percevoir le comportement de l'utilisateur. Elles traduisent ensuite ce comportement en propriétés numériques. Enfin, elles restituent certaines de ces propriétés numériques dans leur environnement réel. Elles peuvent interpréter les informations perçues et décider d'actions à mener d'elles-mêmes. Elles peuvent communiquer avec une autre entité par le réseau. La phase d'intelligence des objets mixtes et des interfaces utilisateur naturelles est assurée alors par les applications qui les pilotent.

Cette boucle d'interaction permet de décrire les interactions internes à une entité mixte qui s'appuient sur leur perception, leurs décisions et leurs actions. Les comportements des entités mixtes sont issus de ces interactions internes. Les interactions entre entités mixtes d'un système interactif mixte découlent de ces comportements. Nous qualifions ces interactions des interactions externes aux entités mixtes. Par exemple, un ventilateur mixte perçoit un signal réseau émis par un casque de réalité mixte. Il identifie ce signal comme un message qui lui est destiné, décide de le décoder, puis le transmet à sa partie virtuelle. Sa partie virtuelle perçoit et analyse ce message demandant l'activation de son moteur. Elle compare cette commande à l'état de la propriété virtuelle équivalente. Elle peut alors déterminer si la meilleure solution est de déclencher le fonctionnement du moteur si le ventilateur mixte n'est pas en train de fonctionner. Son choix d'action est alors d'activer le fonctionnement du ventilateur. Cette activation a pour effet de déclencher la rotation de ses pales virtuelles et, par synchronisation entre partie virtuelle et partie réelle de cet état, de ses pales réelles.

Le modèle résultant est composé de deux boucles d'interaction synchronisées. Ces boucles permettent de capturer des propriétés de leur environnement, de choisir un comportement à adopter puis à agir pour réaliser ce comportement choisi. Cette boucle est applicable à toutes les parties d'une entité mixte, qu'elles soient réelles ou virtuelles. La boucle d'interaction de chaque partie décrit ses interactions internes. La synchronisation entre ces parties est l'interaction interne mixte de l'entité (cf. Figure 3.6). Une version condensée de cette figure est présentée par la Figure 3.7. Nous l'utiliserons par la suite pour modéliser des systèmes interactifs mixtes. Nous détaillons le

fonctionnement de cette boucle dans la Figure 3.8.

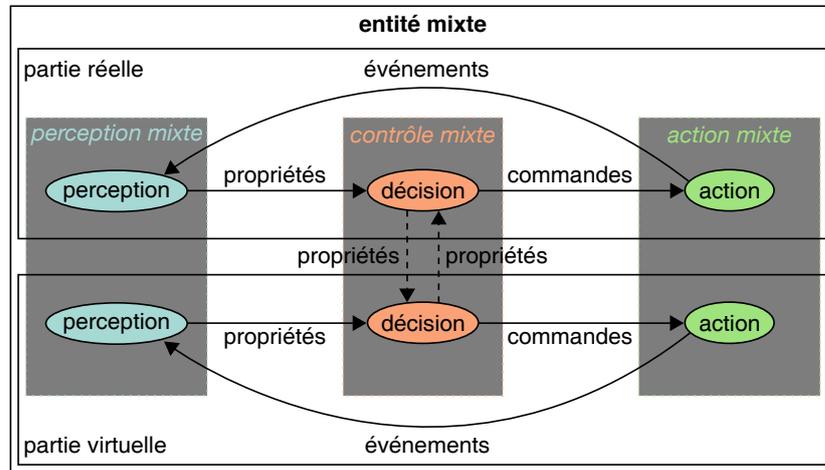


FIGURE 3.6 – Modèle interne mixte orienté conception d'une entité mixte (DOMIM).

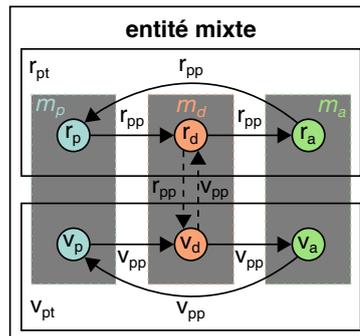


FIGURE 3.7 – Version compacte du modèle interne d'une entité mixte.

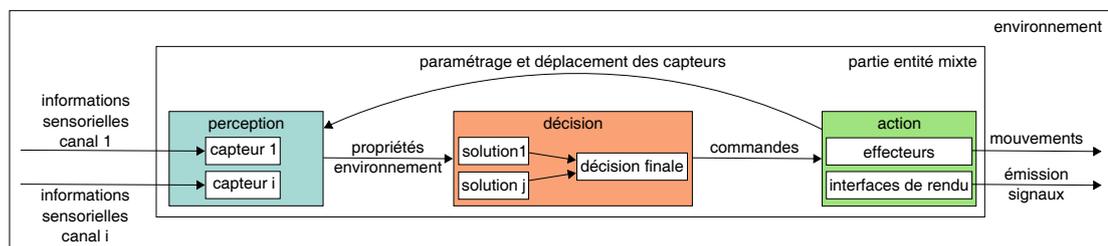


FIGURE 3.8 – Le fonctionnement de chaque partie d'une entité mixte est modélisé par la boucle perception-décision-action.

Suivant la nature de l'entité mixte, les parties et leurs composants ne sont pas systématiquement implémentées. Par exemple dans le cas d'un utilisateur, seule la partie virtuelle de l'utilisateur sera implémentée par une application. Elle n'aura que dans certaines situations exceptionnelles la possibilité de venir contrôler le corps de l'utilisateur, comme nous allons le détailler dans la

sous-section 3.1.3. Néanmoins, les composants logiciels constituant les entités mixtes implémentées sont des scripts ou des bibliothèques logicielles. Ils peuvent être réutilisés par d'autres entités mixtes ayant des besoins communs. Par exemple, un composant réseau sera réutilisé par d'autres entités mixtes avec l'avantage d'uniformiser leurs communications et d'accélérer leur développement, comme nous le verrons dans le chapitre 4.

Les interactions internes à chaque partie d'une entité mixte permettent de modéliser son comportement :

- de la perception à la décision : les informations perçues sont transmises pour analyse. Par exemple, une caméra transmet à une application les images qu'elle capture. Une caméra virtuelle simule la captation d'images de son environnement puis la transmet au processus décisionnel de l'application ;
- de la décision à l'action : une décision a été prise, elle est transmise au système moteur pour être exécutée. Par exemple, un robot aspirateur décide d'aspirer d'où il se trouve jusqu'à une porte, son moteur doit générer le mouvement tandis que l'aspirateur doit faire tourner la brosse et aspirer l'air au niveau du sol. Un agent virtuel autonome décide de marcher jusqu'à une chaise, il demande à son système moteur simulé de l'animer pour simuler la marche ;
- de l'action à la perception : une action induit la perception. Un robot se tourne vers la zone qu'il souhaite observer pour compléter sa connaissance de son environnement. Un agent virtuel autonome perçoit un flux de données qui lui indique le risque de collision avec d'autres agents s'il poursuit sa trajectoire actuelle.

Les interactions internes mixtes s'effectuent entre les parties réelles et virtuelles d'une entité mixte :

- la perception mixte permet le partage des propriétés de l'environnement réel et de l'environnement virtuel. Ces connaissances peuvent se compléter et étendre le niveau de connaissance de l'entité mixte. Par exemple, un utilisateur mixte perçoit la porte fermée de la pièce dans laquelle il se trouve tandis que sa partie virtuelle pourrait percevoir les avatars et agents qui se déplacent également derrière cette porte ;
- le contrôle mixte résulte de la coordination des décisions prises par chaque partie d'une entité mixte afin d'assurer la cohérence de leurs comportements respectifs. Comme nous le verrons dans la section 3.2, ce contrôle est très lié à la nature de l'entité mixte et est de facto porté principalement par une des parties de l'entité mixte. Par exemple, la partie virtuelle d'une voiture autonome simule son environnement réel à partir des informations environnementales captées par sa partie réelle. En cas de détection d'un danger imminent par la simulation, la partie virtuelle de la voiture contraint sa partie réelle à un freinage d'urgence afin de protéger ses passagers ;
- les actions mixtes combinent :
 - les actions dupliquées aux deux parties. Par exemple, la partie virtuelle du robot B15 est constamment localisée au-dessus de sa partie réelle et avec la même orientation. Les actions dupliquées nécessitent qu'une propriété réelle puisse être capturée et traduite en propriété virtuelle lorsque le contrôle est assuré par la partie réelle de l'entité. Réciproquement, elles nécessitent qu'une propriété virtuelle puisse être traduite en propriété réelle lorsque le contrôle est assuré par la partie virtuelle de l'entité. Ces actions dupliquées permettent à une partie de répliquer le comportement de l'autre partie. Par exemple, dans le cas d'une caméra de profondeur reconstruisant le squelette de l'utilisateur, un avatar virtuel de l'utilisateur reproduit dans son environnement virtuel ses déplacements ;
 - les actions complémentaires de chaque partie. Par exemple, la partie virtuelle du robot B15 anime son buste, ses bras et sa tête tandis que sa partie réelle se déplace dans

l'environnement réel qu'elle perçoit. Les actions dans un environnement virtuel sont programmables comme nous allons le décrire dans la section 3.2.

Toutes les entités mixtes ne sont pas forcément constituées que d'une seule partie réelle et d'une seule partie virtuelle. Une entité mixte peut également avoir plusieurs parties dans un environnement donné. Par exemple, une voiture mixte est équipée d'une multitude de capteurs, de mécanismes à actionner, d'interfaces utilisateur et autres appareils électroniques reliés à un ordinateur de bord qui les contrôle. DOMIM permet de modéliser des entités mixtes possédant plusieurs parties réelles ou virtuelles lorsqu'une ou des parties de même nature contrôle une multitude de parties de nature distincte. Dans ce cas, la modélisation et la mise en œuvre d'une partie virtuelle sont réutilisables pour créer les autres parties virtuelles. Un exemple est un joueur de jeux de rôle massivement multijoueurs (MMORPG) qui possède plusieurs personnages sur plusieurs MMORPG, comme indiqué dans la Figure 3.9.

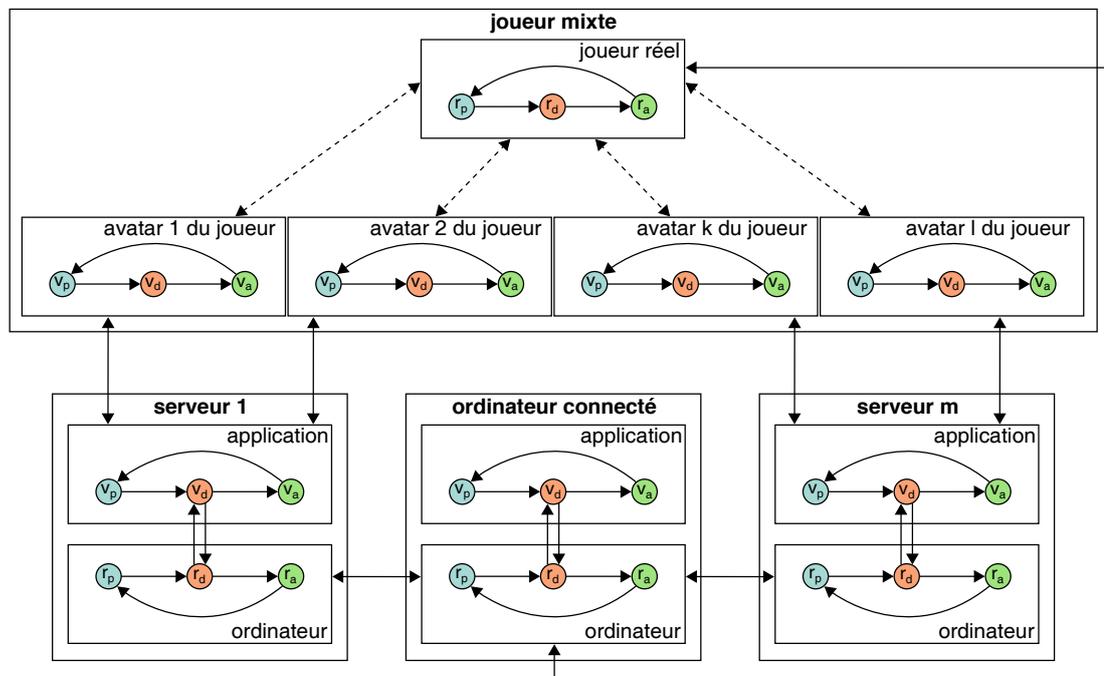


FIGURE 3.9 – Exemple d'entité mixte à parties virtuelles multiples : un joueur de jeux vidéo en ligne peut contrôler de multiples avatars virtuels. Une entité mixte peut avoir une partie réelle et de multiples parties virtuelles.

Une entité mixte peut dès lors interagir simultanément dans des environnements réels et virtuels. Ses comportements réels et virtuels doivent donc être synchronisés pour assurer la cohérence du comportement mixte de l'entité. DOMIM, en modélisant ces comportements, permet de décrire et d'étudier comment ils sont composés, par exemple en étant interpolés. En particulier, DOMIM permet de distinguer quelles propriétés dupliquées sont communes à toutes les parties de l'entité, et quelles propriétés sont spécifiques à chaque partie. Typiquement une prise connectée mixte a la capacité de recevoir une commande transmise par un réseau. Si la prise réelle est mise sous tension alors la prise virtuelle le sera également lors de la synchronisation de leurs propriétés. Réciproquement, si l'environnement virtuel met la prise virtuelle sous tension, le mécanisme de synchronisation met sous tension la prise réelle. Par exemple, un agent autonome peut mettre

sous tension la prise virtuelle. Tout comme un utilisateur réel peut mettre sous tension la prise en appuyant sur un bouton physique. La partie virtuelle peut également décider de se mettre sous tension, par exemple si une minuterie a été programmée. La Figure 3.10 décrit ce scénario. Nous considérons que le module Wi-Fi embarqué par la partie réelle de la prise mixte est un capteur réel. En effet, il perçoit et interprète des signaux Wi-Fi de son environnement réel. Ces signaux sont ensuite traduits en propriétés numériques, qui sont alors transmises à son jumeau virtuel.

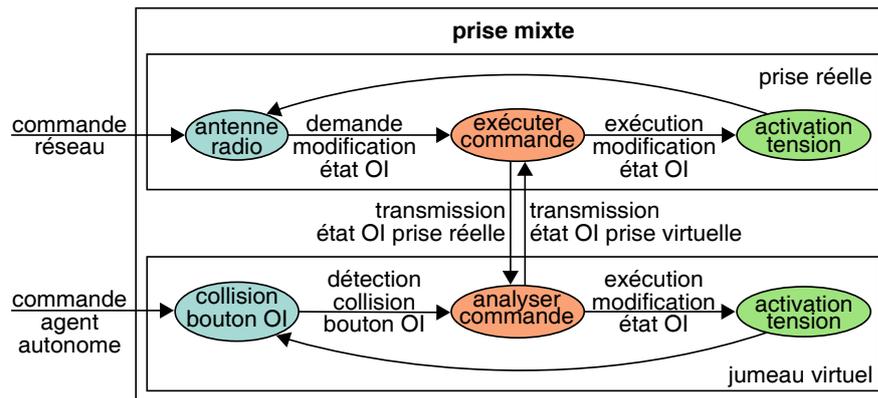


FIGURE 3.10 – Modèle interne d'une prise mixte. Cette prise peut être commandée par une commande réseau ou par un agent autonome virtuel dans son environnement virtuel.

Une entité mixte possède des interfaces internes et communique avec des interfaces externes. Une interface externe est potentiellement utilisable par plusieurs entités mixtes grâce à leurs échanges réseau. Le modèle interne d'une entité mixte est utile pour distinguer l'interface propre d'une entité mixte des interfaces externes avec lesquelles elle interagit. Elles complètent ses capacités de perception et d'action. Dans le cas d'un bâtiment intelligent, l'interface de sa partie réelle lui permet de communiquer avec des capteurs. Par exemple, un serveur supervisant le fonctionnement du bâtiment est connecté à des caméras ou des détecteurs de présence, et des effecteurs, qui peuvent être des écrans ou des hauts parleurs. Ces interfaces additionnelles lui permettent de détecter la présence et l'activité d'individus et de leur transmettre des informations visuelles ou sonores comme une sirène d'alarme. Sa partie virtuelle simule alors l'état courant du bâtiment. Cette simulation lui permet de détecter les événements présents ou à venir. Ou encore, un téléphone mobile intelligent transmet sa reconstruction de son environnement réel à un Hololens. Leurs reconstructions respectives sont alors mutualisées et permettent de créer des comportements complémentaires. Le téléphone peut alors compléter les informations affichées par les Hololens, à la manière du concept de VESAD de Normand et al. [NM18].

Le modèle interne d'une entité mixte permet donc de définir de façon détaillée comment le comportement d'une entité mixte est la résultante de la synchronisation entre ses comportements réels et virtuels. Il décrit une architecture permettant la mise en œuvre unifiée de chaque entité mixte d'un TIES.

3.1.3 L'utilisateur en tant qu'entité mixte

DOMIM prend implicitement en compte l'utilisateur, comme toutes les entités constitutives des TIES. Considérer l'utilisateur en tant qu'entité mixte permet notamment de passer plus facilement d'un humain réel à un humain virtuel autonome. DOMIM permet alors de créer des composants logiciels communs aux utilisateurs de différents systèmes. Ces composants traitent de la capture

des propriétés de l'utilisateur, comme son squelette et ses mouvements. Ils permettent d'instancier son avatar virtuel et de gérer ses comportements mixtes, ou de lui associer un agent virtuel autonome. Concevoir et développer ce type de composant uniformise cette instanciation grâce à la réutilisation des composants développés au préalable. Cela permet ainsi de factoriser, simplifier et accélérer le développement de la partie virtuelle d'un utilisateur mixte.

Modéliser l'homme en tant que système interactif est nécessaire pour établir un modèle unifié des entités mixtes. Dans le cadre d'une approche systémique de l'humain, Gibson décrit en 1979 [Gib79] que la perception est préliminaire à l'action. De façon complémentaire, l'action est nécessaire à la perception et à l'exploration de notre environnement. Cette boucle sensori-motrice est donc nécessaire à notre capacité d'interaction avec notre environnement. Néanmoins, elle n'explique pas notre libre-arbitre et est ensuite étendue par MacKay [Mac90] à un système à trois corps, ces corps étant le sous-système perceptif, le sous-système cognitif et le sous-système moteur du corps humain. Depuis, la boucle perception-décision-action est couramment utilisée pour décrire les utilisateurs en tant qu'éléments de systèmes interactifs [FMG11]. Elle est donc tout à fait adaptée à la modélisation du comportement humain.

Un utilisateur interagissant avec un environnement virtuel nécessite des composants matériels et logiciels capables de réaliser ces fonctions de perception, de décision et d'action dans un environnement virtuel. Il ne peut pas interagir avec cet environnement sans eux. Il a dès lors besoin :

- d'interfaces de captation pour capter son activité. Par exemple, une caméra de profondeur de type Kinect permet de reconstruire un squelette simplifié de l'utilisateur et de capturer les mouvements de ce squelette ;
- d'un contrôleur exécutant une application pour instancier et contrôler sa partie virtuelle. Cette application utilise les informations fournies par les interfaces de captation. Par exemple, un capteur de type Kinect nécessite une application exécutée par un ordinateur auquel il est connecté afin de reconstruire le squelette de l'utilisateur ;
- d'interfaces de restitution pour restituer dans un environnement réel son avatar. Par exemple, l'utilisateur a besoin que l'application affiche en temps réel son squelette afin de savoir si l'application le détecte et supporte son avatar. Cette connaissance lui permet de connaître les limites du système de capture.

Ces interfaces utilisateur ont différentes portées. Elles peuvent :

- faire partie de leur environnement réel telles des caméras de profondeur (Microsoft Kinect, Leapmotion, ...) reliées à des unités de calcul reconstruisant grossièrement leur squelette afin d'analyser leurs mouvements. Par exemple, un utilisateur dont les gestes sont capturés par une Kinect peut manipuler un objet virtuel affiché sur un écran 3D [GARL13] ;
- être portées sur leur corps tels les casques Hololens ou Magic Leap. Les interfaces portées permettent à leur utilisateur de pouvoir voir en temps réel un environnement virtuel aligné au réel. Par exemple, une application Hololens peut afficher un modèle virtuel d'une cité détruite depuis plusieurs centaines d'années lorsque son utilisation se trouve dans ses ruines. Cette application permet alors de voir un modèle 3D de cette cité réalisé à partir des travaux d'historiens et d'archéologues ;
- être implantées dans leurs corps telles des électrodes implantées dans le cerveau. Ces électrodes permettent par exemple à une application d'interpréter les signaux émis par le cerveau. Ainsi, un utilisateur peut par exemple commander un bras robotisé par la pensée.

Ces interfaces et ces contrôleurs instancient la partie virtuelle de l'utilisateur mixte. Cette partie virtuelle des humains mixtes peut hybrider les notions d'avatar virtuel humanoïde [FALH20] et

d'agent autonome [CLLT10]. L'avatar virtuel humanoïde d'un utilisateur mixte représente l'utilisateur dans un environnement virtuel. Il clone son comportement, typiquement ses mouvements et sa localisation, éventuellement son apparence en fonction des capacités techniques et des choix de conception. Mais lorsque cette partie virtuelle de l'humain mixte n'est plus synchronisée en temps réel avec sa partie réelle, son corps, elle n'est plus un avatar. En effet, elle est alors capable de se comporter en tant qu'agent humanoïde autonome lorsqu'on lui attribue un comportement virtuel. Le comportement de cet agent peut être prédéfini par l'utilisateur. Par exemple, la partie virtuelle d'un humain mixte peut éteindre le téléviseur devant lequel se trouve l'humain mixte lorsqu'il dort. Elle peut également interagir vocalement et visuellement à une demande de communication interpersonnelle, à la manière d'un répondeur téléphonique étendu à la réalité mixte en somme.

La Figure 3.11 présente comment un Hololens peut instancier la partie virtuelle d'un humain mixte. Pour cela, il capte la localisation et l'orientation de la tête de l'utilisateur grâce à ses caméras à niveau de gris embarquées. Il localise également ses mains lorsqu'elles se trouvent dans le champ de vision de la caméra de profondeur de l'Hololens. Cette caméra est en effet en charge de les détecter et de les suivre. Ainsi, lorsque l'utilisateur approche sa main d'un objet virtuel, le visiocasque affiche cet objet virtuel, mais également le jumeau virtuel de la main de l'utilisateur. L'utilisateur perçoit alors que la présence de sa main dans son environnement est connue et a un impact sur l'objet virtuel lorsqu'il le touche. DOMIM permet de modéliser cette interaction, de décomposer l'utilisateur mixte en composants et de caractériser leurs fonctions et leurs comportements.

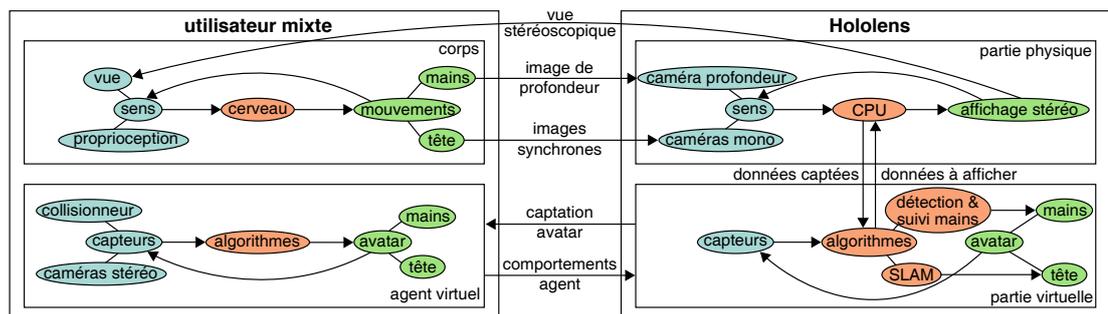


FIGURE 3.11 – Un exemple d'utilisateur mixte supporté par un Hololens. L'utilisateur nécessite une interface utilisateur naturelle pour pouvoir interagir avec le virtuel.

3.1.4 Synthèse

DOMIM permet de modéliser de façon unifiée les systèmes interactifs mixtes combinant réalité mixte, interfaces utilisateur naturelles et internet des objets. Ce modèle désambiguïse le réel du virtuel en les décorrélant. Il facilite une étude exhaustive de leurs interactions et de leurs fonctionnalités. Il permet de mettre en évidence et de traiter différents aspects des systèmes interactifs mixtes. Isoler chacun de ces aspects permet de mieux les appréhender durant les différentes phases de conception et de développement.

Modéliser un système interactif mixte à l'aide de DOMIM permet de distinguer les entités mixtes constitutives du système. Modéliser une entité mixte grâce à DOMIM permet de :

- décorréler la partie réelle de la partie virtuelle d'une entité mixte, afin de :
 - concevoir et étudier le comportement de chaque partie dans son environnement respectif ;

- contrôler la partie virtuelle de plusieurs entités mixtes par la partie réelle d'une seule entité mixte ;
 - contrôler la partie réelle de plusieurs entités mixtes par la partie virtuelle d'une seule entité mixte.
- décrire les interactions internes à une entité mixte, notamment afin de :
- décrire les capacités de perception, de contrôle et d'action de chacune de ces parties ;
 - étendre la connaissance d'un environnement à la connaissance de l'environnement dupliqué en les agrégeant. Dans le cas du ventilateur mixte, l'utilisateur perçoit en plus du ventilateur réel son bouton virtuel. Or, environnement réel et environnement virtuel ont des propriétés et contraintes très différentes. Ainsi, l'utilisateur peut s'affranchir des contraintes du monde réel pour interagir avec son environnement virtuel. Lors de sa synchronisation avec le réel, cette interaction est alors effectuée dans son environnement réel ;
 - augmenter ses capacités d'action en combinant actions réelles et virtuelles. Les interactions mixtes peuvent être complémentaires ou dupliquées. Les interactions dupliquées sont identiques pour les deux parties d'une entité mixte. Les interactions complémentaires sont distinctes mais coordonnées.

Les entités mixtes déterminent leur propre comportement dans chaque environnement réel ou virtuel qu'elles occupent. La boucle d'interaction contient une phase décisionnelle essentielle à la création de ce comportement. Cette phase s'appuie sur des processus de contrôle qui régissent le comportement des entités mixtes. Ce contrôle d'une entité mixte résulte non seulement de la boucle comportementale de chacune de ses parties, mais aussi de leur synchronisation. Modéliser et comprendre cet aspect d'une entité mixte sont nécessaires pour pouvoir concevoir et développer ses comportements. C'est pourquoi nous décrivons dans la section suivante comment ce contrôle s'opère en fonction de la nature de l'entité mixte.

3.2 Contrôle des entités mixtes

Avec DOMIM, le contrôle d'une entité mixte nécessite comme indiqué sur la Figure 3.6 une synchronisation des comportements des parties réelles et virtuelles. Il met en relation le processus décisionnel de la partie réelle avec celui de sa partie virtuelle. Suivant la nature de l'entité mixte, utilisateur mixte, interface pseudo-naturelle ou objet connecté, ce contrôle va se concrétiser différemment. Concevoir et mettre en œuvre ces comportements mixtes sont des aspects importants de ces systèmes interactifs. Cette tâche permet de décrire les comportements attendus de chaque entité d'un système. Elle est nécessaire à la mise en œuvre de son comportement dans le système. Le caractériser permet de l'adapter aux différentes entités mixtes. Pour cette raison, nous présentons dans les sous-sections suivantes ces différents types de contrôle. Néanmoins, nous ne décrivons pas d'implémentation de ces contrôles grâce à des technologies dites d'Intelligence Artificielle (IA), qui ne font pas partie de ce travail de thèse. Il est par ailleurs tout à fait possible d'intégrer ce domaine à nos travaux. Nous exposons par contre une technologie qui permet aux concepteurs et développeurs de décrire et de créer rapidement les comportements attendus sous la forme de scénarios.

3.2.1 Contrôle d'une partie virtuelle d'une entité mixte

La partie virtuelle d'une entité mixte n'a de comportement que si un développeur lui en attribue un. Le contrôle permettant d'instancier ce comportement doit être conçu et implémenté. Il est déterminé par l'application qui la supporte. Cette application collecte des informations, les analyse, détermine les actions à mener puis les exécute. Cette analyse permet de déclencher des

comportements pré-déterminés par le développeur. Diverses technologies comme l'apprentissage automatique de modèles de comportements interactifs [Mih15], ou encore les automates d'état fini probabilistes [HIM⁺16] permettent de déterminer ces comportements. DOMIM décrit comment ce processus décisionnel est dépendant des informations perçues par l'entité et limité par ses capacités d'action. Développer la richesse des comportements d'une entité nécessite donc de développer ses capacités de perception et d'action. Elles offrent alors à son processus décisionnel de multiples choix d'action.

De multiples techniques de mise en œuvre de ce processus décisionnel sont possibles. Une prise de décision peut ne nécessiter aucune perception ni analyse de son environnement. Elle peut n'être qu'une séquence d'actions pré-établies qui se répètent indéfiniment. Par exemple, une horloge virtuelle affiche l'heure selon une logique immuable.

Le comportement induit par ce processus décisionnel produit une suite d'actions séquençables en un graphe, tel un automate à états fini ou un réseau de Petri. Ce graphe décrit l'évolution de la partie virtuelle de l'entité dans son environnement. Il représente le scénario comportemental de cette partie. Les nœuds représentent les comportements de l'entité mixte. Par exemple, un scénario définit une simulation de remplacement d'une roue de voiture afin de former un mécanicien à cette procédure. Le scénario prévoit de rehausser la roue à l'aide d'un cric, de dévisser la roue, de finir de surélever la roue, de l'enlever, de mettre à la place la roue de secours, de rabaisser un peu la roue pour pouvoir la serrer, de la serrer, puis de la rabaisser complètement et d'enlever le cric. Ces étapes sont les nœuds du scénario. Les arcs sont conditionnés par les choix de comportement de la partie virtuelle. En fonction de son analyse du contexte perçu, le contrôle virtuel détermine le comportement à adopter et les actions à effectuer. Il fait alors appel à ses effecteurs virtuels pour réaliser ces actions et ainsi poursuivre sa progression dans le scénario. Par exemple, rehausser la roue nécessite de prendre le cric et la manivelle dans le coffre et de fixer le cric au châssis de la voiture à proximité de la roue à changer. Ces pré-conditions permettent ensuite d'actionner le cric à l'aide de la manivelle.

Un schéma de scénario basique est décrit par la Figure 3.12. Dans ce scénario, la partie virtuelle peut répondre de deux façons distinctes aux propriétés qu'elle perçoit de son environnement. Ce scénario simple s'arrête après ce choix binaire. Il montre comment la boucle d'interaction inclut à DOMIM permet de simuler le comportement d'une entité virtuelle. Cette boucle alterne perception, décision et action afin de déterminer le comportement de l'entité mixte. Cette boucle s'exécute lorsque le scénario est en phase de transition. Elle est en fait une heuristique, et a le choix entre deux solutions. En fonction des informations perçues, le processus décisionnel effectue un choix qui permet de passer à l'étape suivante, ou pas. Si l'information perçue n'est pas pertinente pour le processus décisionnel, la boucle continue de tourner en attente d'une information pour lui permettre de décider à quelle étape passer. v_p est la perception de la partie virtuelle de l'entité mixte, v_d sa décision, v_a son action. v_{pp,t_0} décrit les propriétés virtuelles perceptibles à l'instant t_0 . v_{pp,t_1A} décrit les propriétés virtuelles transmises à l'étape finale A à l'instant t_1 . v_{pp,t_1B} décrit les propriétés virtuelles transmises à l'étape finale B à l'instant t_1 .

Un TIES est constitué de multiples entités mixtes dont les comportements conjoints réalisent les fonctionnalités attendues du système. Il doit donc coordonner le comportement de ses entités mixtes pour être opérationnel. Or, la partie virtuelle de chaque entité mixte se comporte en fonction d'un scénario propre. Ces scénarios sont joués en parallèle dans l'environnement virtuel. Afin de réaliser les fonctionnalités attendues du système, ces scénarios doivent tenir compte du comportement des autres entités mixtes. Dès lors, les scénarios des parties impliquées deviennent interdépendants. En effet, ils doivent être coordonnés tandis qu'ils modifient tous des propriétés de l'environnement virtuel. Par exemple, deux robots virtuels jouent à cache-cache. Un robot déroule un scénario qui consiste à chercher l'autre robot tandis que l'autre robot exécute un scénario qui consiste à se cacher pour ne pas être trouvé. Il est donc possible de coordonner le comportement d'entités mixte contrôlées chacune par un scénario en mettant en œuvre des mécanismes de synchronisation entre instances de scénario.

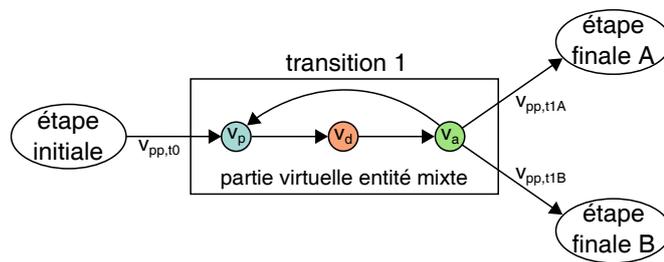


FIGURE 3.12 – Un scénario simple de choix d'action binaire. Ce choix s'offre à un contrôleur virtuel consécutivement à la perception par ses capteurs virtuels des propriétés de son environnement virtuel.

3.2.2 Contrôle des entités mixtes non-vivantes

Dans le cas des entité mixtes non-vivantes, typiquement un robot mixte comme le robot B15 de Microsoft ou un robot aspirateur, le contrôle est prépondéré par le comportement de sa partie virtuelle. En effet, sa partie réelle est essentiellement en charge du contrôle de la boucle sensori-motrice du robot. Elle mécanise, robotise son comportement virtuel. Sa partie virtuelle collecte les informations issues de l'environnement réel du robot. Elle les analyse, afin de déduire notamment leur localisation et leur orientation. Elle détermine les actions à simuler dans son environnement virtuel avant de les effectuer dans son environnement réel. Il n'y alors pas de concurrence entre les processus décisionnels réels et virtuels de l'entité mixte. Cela élimine les risques de blocage du contrôle mixte. Modéliser ce type d'entité avec DOMIM permet de déterminer et de caractériser les capteurs et effecteurs nécessaires à sa partie réelle afin de connaître ses capacités de perception et d'action à simuler. Développer ce type d'entité implique de développer l'application qui réalise ce comportement virtuel. Par exemple, à l'aide du mécanisme de scénario décrit dans la sous-section précédente. Cette mise en œuvre implique également d'interfacer cette application avec les composants réels de l'entité mixte.

3.2.3 Contrôle des entités mixtes vivantes

Le contrôle d'une entité mixte vivante est différent de celui d'une entité mixte non-vivante que nous avons décrit dans la section précédente. Le comportement d'une entité mixte vivante est déterminé par sa partie réelle complétée par sa partie virtuelle. Dans les systèmes interactifs mixtes, les entités mixtes vivantes sont principalement les utilisateurs. Ces systèmes ont pour fonction de permettre à leurs utilisateurs d'interagir avec leur environnement mixte. DOMIM distingue les entités constitutives d'un TIES et leur attribue des propriétés et des comportements réels et virtuels synchronisés. Modéliser grâce à DOMIM le contrôle de l'utilisateur permet donc de décrire la relation entre l'utilisateur et son environnement mixte. En particulier, DOMIM décrit la synchronisation de ses processus décisionnels réels et virtuels. Dans le cas d'une entité mixte vivante, le contrôle est supervisé par sa partie réelle, son corps. Sa partie virtuelle clone ses propriétés réelles mais ne les met pas à jour. DOMIM permet de modéliser le caractère unilatéral des synchronisations de propriétés entre parties réelles et virtuelles. Un utilisateur mixte peut néanmoins percevoir et conscientiser sa partie virtuelle lorsqu'elle est rendue. Elle constitue alors une marionnette immergée dans la partie virtuelle de son environnement mixte et est contrôlée par l'utilisateur grâce à des interfaces utilisateur naturelles. La partie virtuelle a alors la capacité de contrôler des objets mixtes et donc d'agir sur son environnement réel.

Si un système interactif mixte inclut a minima un utilisateur, il peut également inclure d'autres personnes, des animaux ou des plantes. Par exemple, un utilisateur peut connaître grâce à un casque de réalité mixte les besoins d'une plante en eau grâce à un capteur d'humidité dans le pot de la plante. Ou encore, un utilisateur étant lui-même une entité mixte, il doit coordonner les décisions de sa partie virtuelle et de son corps. Or, le processus décisionnel d'un utilisateur est assuré par sa cognition sauf situations exceptionnelles. Les comportements virtuels d'un utilisateur mixte ne sont généralement pas directement traduits en comportements réels afin de ne pas altérer l'intégrité de sa cognition. Néanmoins, ils peuvent être utiles pour protéger l'utilisateur, le guider ou améliorer sa performance dans le respect des limites éthiques. Dans le cas des exosquelettes, qui sont contrôlés par la cognition de leurs utilisateurs, les décisions de la partie virtuelle de l'exosquelette n'imposent pas à l'utilisateur de mouvements qu'il n'a pas sollicité sauf pour éviter des dangers, comme une chute ou un choc. Autre exemple, les fauteuils roulants électriques bénéficient d'une assistance à la conduite afin de limiter les risques liés à leur usage, notamment en cas de troubles cognitifs de leur utilisateur [VGD⁺20].

La partie virtuelle d'un utilisateur mixte est en charge d'interagir avec son environnement virtuel. Afin d'aider l'utilisateur à contrôler son interaction avec son environnement virtuel, il peut percevoir sa partie virtuelle afin de le guider pour toucher un objet virtuel. Par exemple, en affichant en stéréoscopie une partie de l'avatar virtuel de l'utilisateur comme sa main virtuelle afin de l'aider à saisir un objet virtuel. Les utilisateurs mixtes ont donc un contrôle prépondéré par leur cognition, leur processus décisionnel réel. Leur cognition peut être influencée par la perception de leur avatar virtuel. La partie virtuelle d'un utilisateur mixte peut par ailleurs interagir avec des objets mixtes. La Figure 3.13 décrit ce type de contrôle, généralement basé sur la détection et le suivi du squelette de l'utilisateur grâce à une caméra de profondeur, par exemple une Kinect.

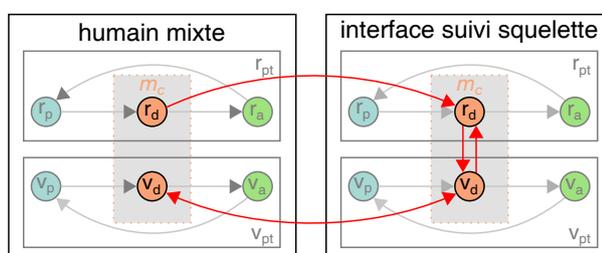


FIGURE 3.13 – Contrôle mixte d'une entité vivante. Dans le cas d'un humain mixte, le contrôle mixte ne met à jour que les propriétés virtuelles à partir des propriétés réelles de l'entité. La synchronisation des propriétés réelles et virtuelles est alors unilatérale. Néanmoins, les propriétés virtuelles peuvent être rendues afin de permettre à l'humain mixte de conscientiser leur état courant. Ou encore, d'effectuer des actions sur des objets connectés sous le contrôle de l'humain.

3.2.4 Synthèse

DOMIM permet de concevoir et développer le comportement virtuel d'une entité mixte en le scénarisant. Chaque partie virtuelle d'une entité mixte déroule un scénario propre qui décrit son comportement. Suivant le type d'entité mixte, leur contrôle est différent. Les entités mixtes vivantes comme les utilisateurs préservent leur cognition, mais peuvent néanmoins bénéficier d'un rendu de leur partie virtuelle dans le monde réel. Dans certains cas, leur partie virtuelle peut prendre la main sur leur cognition lorsqu'elles sont éthiquement acceptables et techniquement possibles, typiquement en situation de danger immédiat. Les entités mixtes non-organiques quant à elles ont

un contrôle virtuel prédominant, car leur environnement virtuel permet de déterminer leur comportement virtuel précurseur de leur comportement réel.

3.3 Interactions entre entités mixtes

Des interactions entre entités mixtes très distinctes ont lieu au sein des systèmes interactifs mixtes. Ces interactions sont externes à chaque entité et doivent être caractérisées, classifiées. Cela permet de mieux comprendre ces systèmes, de segmenter leur complexité et de factoriser les composants qui les réalisent. Ces interactions ont lieu soit entre entités réelles, soit entre entités virtuelles, soit entre entités réelles et entités virtuelles. Le graphe utilisateur-interface-environnement présenté dans la Figure 3.14 décrit les différents types d'interaction qui s'y déroulent.

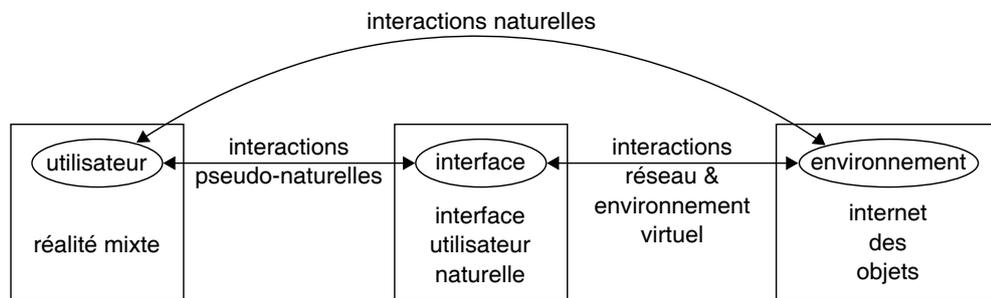


FIGURE 3.14 – Le graphe des types d'interaction des systèmes interactifs mixtes combinant réalité mixte, interfaces utilisateur naturelles et internet des objets.

Modéliser ces différents types d'interactions grâce à DOMIM permet de :

- décrire comment le fonctionnement interne de chaque entité mixte permet ses interactions externes avec les autres entités mixtes du système ;
- décrire comment les interactions entre entités peuvent permettre aux entités de se combiner et créer des synergies ;
- déterminer les spécificités de chaque type d'interaction.

Nous considérons quatre types d'interactions entre entités mixtes. Ce sont les interactions naturelles, les interactions pseudo-naturelles, les interactions en environnement virtuel et les interactions réseau. Nous décrivons ces interactions dans les paragraphes suivants.

Interactions naturelles

Une interaction naturelle se déroule entre des entités réelles sans médiation par des interfaces utilisateur ou réseau. Par exemple, une personne peut mettre en état de marche un ventilateur en appuyant sur son bouton de mise en route. Une interaction naturelle ne nécessite pas d'interface utilisateur ou de connexion, elle a lieu uniquement dans le monde réel. Si ce type d'interaction est modélisable par DOMIM lors de la conception du système, elle ne nécessite aucune implémentation d'une application pour se dérouler. Elle peut par contre être capturée et simulée dans un environnement virtuel par une application. Ce type d'interaction est décrit par la Figure 3.15 dans le cas d'un utilisateur qui prend à la main un téléphone intelligent, la préhension d'un objet étant une interaction naturelle. Décrire les interactions naturelles d'un TIES lors de sa conception peut être nécessaire à son bon fonctionnement. En effet, le système peut par exemple se mettre en action lorsqu'il détecte que l'utilisateur a saisi le téléphone intelligent.

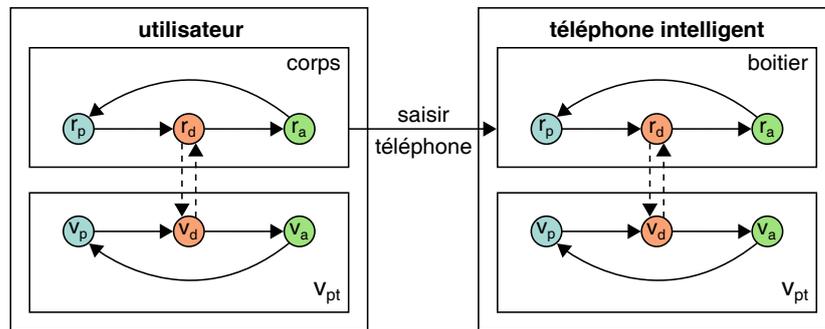


FIGURE 3.15 – Interactions naturelles entre deux entités mixtes. Les interactions naturelles ne requièrent pas d’interfaces utilisateur ou réseau, elles se déroulent entre les parties réelles des entités.

Interactions pseudo-naturelles

L’interface utilisateur naturelle est nécessaire pour instancier l’avatar virtuel de l’utilisateur. Les interactions pseudo-naturelles nécessitent une interface utilisateur naturelle pour intermédiaire les interactions entre l’utilisateur et la partie virtuelle de l’entité mixte. Cette interface utilisateur traduit les comportements de l’utilisateur dans son environnement virtuel.

Elle permet à l’utilisateur de percevoir cet environnement virtuel, y compris son propre avatar virtuel humanoïde. Cet avatar peut alors interagir avec son environnement virtuel sous le contrôle de l’utilisateur. Par exemple, un utilisateur peut toucher l’écran tactile de son téléphone intelligent afin de déplacer son avatar dans un jeu vidéo. Il peut également manipuler un objet virtuel avec sa main en effectuant les mêmes gestes que s’il manipulait cet objet réel.

Exemple d’interaction pseudo-naturelle, un utilisateur porte un Hololens qui instancie son avatar virtuel (cf. Figure 3.16). Le modèle DOMIM permet de concevoir et mettre en œuvre les interactions entre l’utilisateur et son environnement virtuel. Les développeurs peuvent réutiliser les composants logiciels implémentant ces interactions pseudo-naturelles.

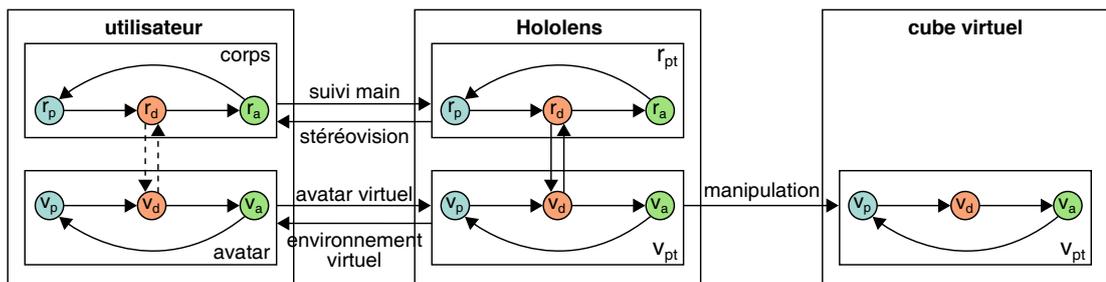


FIGURE 3.16 – Interactions pseudo-naturelles naturelles entre un utilisateur et une entité mixte. L’avatar de l’utilisateur interagit avec un cube virtuel que l’utilisateur perçoit en stéréovision grâce à l’Hololens.

Une interface utilisateur naturelle peut être intégrée à l’environnement ou portée par l’utilisateur. Dès lors, plusieurs utilisateurs ou objets peuvent s’en servir. Par exemple, une caméra Kinect est capable de détecter les mouvements de plusieurs utilisateurs simultanément. Ou encore un écran peut être perçu par les personnes qui lui font face. De la même façon, un Hololens peut être porté par plusieurs personnes consécutivement.

Interactions en environnement virtuel

Une interaction en environnement virtuel se déroule entre parties virtuelles de plusieurs entités mixtes. Par exemple, deux utilisateurs peuvent interagir à distance par l'intermédiaire de leurs avatars virtuels humanoïdes dans un environnement commun. Les interactions se déroulent entre les avatars virtuels des collaborateurs qui communiquent par la gestuelle et le vocal. La Figure 3.17 représente ce type d'interaction.

DOMIM permet de modéliser spécifiquement ces interactions, qui répondent à des contraintes très différentes du monde réel. Un environnement virtuel répond aux contraintes que le concepteur et le développeur créent. Il est hautement malléable et n'obéit pas à des lois physiques inaliénables. Généralement, ce type d'interaction vise à simuler des interactions qui ont lieu dans le monde réel. Par exemple, un avatar virtuel humanoïde peut toucher l'écran tactile d'un téléphone intelligent virtuel qu'il tient dans la main.

Lorsque certaines de ces entités virtuelles ne sont pas instanciées par la même partie réelle, elles nécessitent une interaction réseau entre les parties réelles qui les instancient afin de pouvoir interagir. DOMIM permet de caractériser ces interactions en décrivant les capteurs, processus décisionnels et effecteurs nécessaires à leur développement. DOMIM simplifie alors le développement d'environnements virtuels en permettant la réutilisation des entités précédemment développées ainsi que de leurs composants logiciels, comme des scripts ou des scénarios comportementaux.

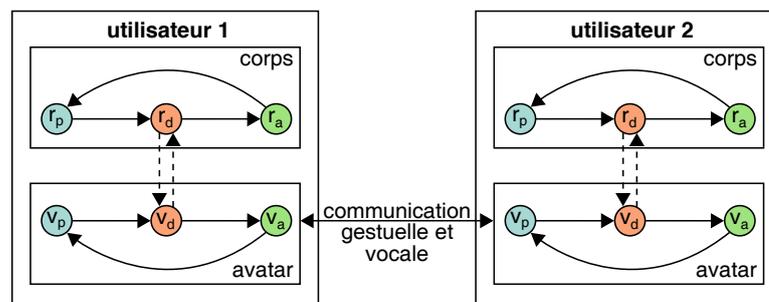


FIGURE 3.17 – Exemple de communication interpersonnelle en environnement virtuel.

Interactions réseau

Une interaction réseau unifiée entre entités mixtes est nécessaire pour simplifier la modélisation, créer des composants de développement réseau réutilisables et rendre les systèmes mixtes interactifs interopérables et extensibles. Une interaction réseau se produit entre les parties virtuelles d'entités mixtes grâce à une interface réseau. Cette interaction est également décomposable en deux interactions, une en environnement réel et une en environnement virtuel. En effet, l'interface réseau permet la capture et l'émission de propriétés de l'environnement réel. Cette interface échange avec d'autres interfaces réseau des signaux qui transmettent des messages sous le contrôle d'une application. Différentes technologies sont utilisées, certaines filaires, d'autres par ondes radio, ou encore par des signaux lumineux. Par exemple Z-Wave, Wi-Fi, Bluetooth, Ethernet ou encore Li-Fi sont des protocoles de communication qui s'appuient sur des mécanismes physiques de transmission d'une information numérisée. Ces interfaces sont équipées de capteurs qui permettent de percevoir un signal physique, par exemple une onde radio dans une bande de fréquence prédéterminée. Elles sont également équipées d'émetteurs qui permettent d'émettre des signaux qui transmettent des messages numérisés, généralement avec la même technologie réseau.

Une interaction réseau permet aux parties virtuelles de plusieurs entités mixtes d'interagir alors qu'elles ne sont pas supportées par le même contrôleur. Les interfaces réseau sont pilotées par des applications développées lors de la création du système interactif mixte. Modéliser ces interactions est nécessaire pour inclure à DOMIM le domaine de l'internet des objets, qui utilise systématiquement ces technologies réseau pour communiquer. Cela permet d'uniformiser les communications

qui s'opèrent entre utilisateurs, interfaces utilisateur naturelles et objets connectés. Modéliser ce type d'interaction à l'aide de DOMIM permet donc aux concepteurs de systèmes interactifs mixtes d'inclure l'internet des objets dans leurs systèmes et d'identifier les interactions réseau à implémenter.

Par exemple, un utilisateur mixte peut interagir avec un téléphone intelligent en lui adressant des requêtes vocales. Ces requêtes sont transmises par le réseau à un serveur qui analyse puis transmet sa réponse via le réseau. Ou encore deux collaborateurs peuvent travailler ensemble à distance en s'immergeant dans un environnement virtuel commun et en percevant l'avatar virtuel humanoïde de chacun. En modélisant ces interactions réseau avec DOMIM, les concepteurs créent des modèles-type d'interaction réseau. Ces modèles sont alors réutilisables par d'autres systèmes. Les développeurs peuvent également développer des composants réseau pérennes car génériques et réutilisables car unifiés. Exemple d'interaction réseau entre entités mixtes modélisé par DOMIM, un visiocasque et une prise mixte communiquent par le réseau Wi-Fi, chacun est équipé d'un émetteur et d'un récepteur Wi-Fi afin de pouvoir échanger des messages numériques. L'application Hololens instancie le jumeau virtuel de l'utilisateur en captant la position de sa tête et de ses mains. Elle instancie également la partie virtuelle de la prise connectée, avec laquelle elle communique par le Wi-Fi (cf. Figure 3.18). Il est également possible d'étendre ce schéma aux interactions naturelles entre l'utilisateur et la prise, car l'utilisateur peut également presser le bouton marche/arrêt du ventilateur.

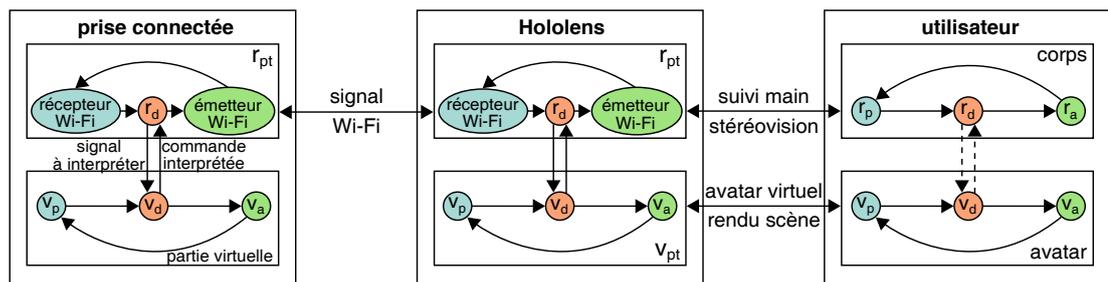


FIGURE 3.18 – Interactions réseau entre un utilisateur équipé d'un casque Hololens et une prise connectée via un geste dans l'air.

L'internet des objets permet d'utiliser les mécanismes de captation et d'actuation distribués dans un environnement réel. Une interface de captation ou d'actuation peut donc être commune à plusieurs objets mixtes. Les objets mixtes peuvent combiner leurs interfaces internes à celles d'autres entités mixtes ou encore communiquer avec des interfaces disséminées dans leur environnement.

Cette caractérisation des interactions des TIES combinée au modèle d'entité mixte DOMIM permet de décomposer ces systèmes en sous-systèmes jusqu'à isoler chacune de ses entités et d'étudier leurs mécanismes d'interaction tant internes qu'externes. DOMIM contribue ainsi à maîtriser la complexité de ces systèmes dans les phases de conception et de développement des systèmes interactifs mixtes. Il permet également de faciliter la réutilisation des composants et des interactions d'un système interactif mixte dans d'autres systèmes. Cette factorisation est simple lorsque plusieurs systèmes partagent des entités mixtes communes, des relations similaires ou encore lorsqu'ils font appel à des mécanismes communs d'échange d'information via un réseau. Il suffit alors de conserver les entités mixtes modélisées auparavant, de créer le modèle des nouvelles entités mixtes et de les relier soit avec des liaisons modélisées pour d'autres systèmes soit de modéliser les relations manquantes. Par exemple, le modèle DOMIM d'un utilisateur mixte de la Figure 3.11 et celui d'une prise mixte de la Figure 3.10 sont fusionnés dans la Figure 3.19. Ce

modèle indique comment un utilisateur équipé d'un casque de réalité mixte peut interagir avec une prise mixte de plusieurs manières, soit par une commande réseau soit par l'interaction entre leurs parties virtuelles. Il est également possible d'étendre ce schéma aux interactions naturelles entre l'utilisateur et la prise, car l'utilisateur peut également presser le bouton marche/arrêt du ventilateur.

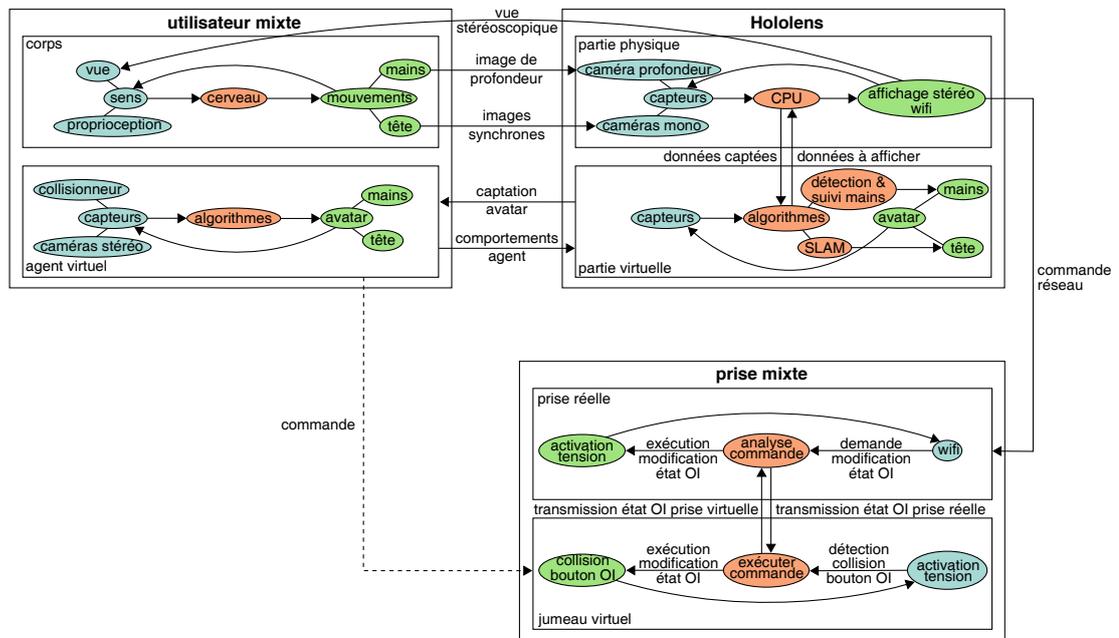


FIGURE 3.19 – Exemple d'une prise mixte commandée par un utilisateur mixte à l'aide d'un HoloLens. L'application exécutée par HoloLens instancie le jumeau virtuel de l'utilisateur afin de détecter ses intentions. Ces intentions sont ensuite transmises par le réseau à la prise mixte.

DOMIM permet de non seulement simplifier, unifier et accélérer la conception d'un système, mais également sa mise en œuvre. La flexibilité d'usage d'un modèle basé DOMIM provient de sa généralité. Elle permet de rendre interopératif et extensible un TIES. DOMIM offre également un cadre propice à l'unification des communications réseau entre applications basées DOMIM.

3.3.1 Synthèse

DOMIM décrit les interactions entre entités mixtes. Ces interactions dites externes sont la conséquence des interactions internes à chaque entité mixte. DOMIM permet de concevoir :

- les interactions simultanées entre entités mixtes et d'étudier leur complémentarité ;
- l'interaction directe de chaque partie avec son environnement. Par exemple, un utilisateur appuie sur le bouton marche/arrêt d'un ventilateur mixte, ou un avatar virtuel d'un utilisateur peut toucher le bouton virtuel de ce ventilateur mixte ;
- l'interaction indirecte d'une partie d'une entité via la synchronisation entre ses parties réciproques. Par exemple, lorsque l'avatar virtuel de l'utilisateur touche le bouton virtuel du ventilateur mixte, il déclenche le ventilateur virtuel, mais également le ventilateur réel via la synchronisation de l'état marche/arrêt du ventilateur mixte entre ses parties.

DOMIM classe les interactions externes des entités mixtes en quatre types d'interaction de la façon suivante :

- les interactions naturelles en environnement réel ;
- les interactions pseudo-naturelles reliant le réel au virtuel ;
- les interactions réseau reliant plusieurs environnements réels ;
- les interactions en environnement virtuel.

Cette classification permet d'optimiser et d'améliorer durant la conception chaque interaction ainsi que leur complémentarité en les caractérisant. Elle contribue à l'élaboration de patrons de conception. Elle structure les composants logiciels d'interaction nécessaires à la mise en œuvre d'un TIES, rendant possible leur réutilisation. DOMIM permet également de détecter d'éventuelles redondances dans les interfaces utilisateur naturelles présentes et donc de factoriser certaines fonctionnalités ou au contraire d'exploiter la fusion de leurs données.

DOMIM permet par ailleurs la conception d'interactions hybrides mono ou multimodales en combinant plusieurs interfaces. Nous décrivons ces interactions hybrides dans la section suivante.

3.4 Interactions utilisateur hybrides

Si la complexité des systèmes interactifs mixtes pose des défis quant à leur développement, ils n'en offrent pas moins de nouvelles possibilités d'interaction que nous allons aborder dans cette section. Les systèmes interactifs mixtes permettent aux utilisateurs d'interagir simultanément avec de multiples entités mixtes pouvant interagir entre elles simultanément. Parmi ces entités, plusieurs interfaces utilisateur naturelles sont potentiellement utilisables par l'utilisateur. La multiplicité des interfaces utilisateur naturelles au sein d'un système interactif mixte de grande dimension, par exemple un bâtiment ou une ville connectée, permet de les combiner. Cela a pour effet d'améliorer la qualité ou le nombre de canaux sensori-moteurs impliqués dans l'interaction utilisateur avec le virtuel, et donc l'expérience utilisateur. Combiner ces interfaces utilisateur est une piste prometteuse pour exploiter la complexité de ces systèmes. En effet, corréler différentes modalités d'interaction est typique des interactions naturelles entre l'homme et son environnement réel, comme décrit dans la section 2.1.3. La multimodalité permet alors d'améliorer l'efficacité des tâches réalisées par les utilisateurs de ces systèmes [PDG17]. La Figure 3.20 schématise les différents types d'interactions externes offertes par les TIES.

3.4.1 Hybridation des modalités

Une interaction hybride combine plusieurs interfaces utilisateur naturelles. Elle peut être mono ou multimodale. Plusieurs types d'associations entre interfaces utilisateur naturelles sont possibles, que nous allons décrire dans cette section.

Une interaction hybride peut être multimodale. Une modalité d'interaction associe un canal sensoriel à un canal moteur [Bou07]. Lorsque plusieurs interfaces utilisateur naturelles ont des modalités d'interaction distinctes, une interaction qui combine ces interfaces est dite multimodale. En effet, elle associe alors plusieurs canaux sensoriels à plusieurs canaux moteurs. Lors d'une interaction multimodale, l'information est transmise par la machine à l'homme suite à sa fission. La fission de l'information consiste à séparer en modalités une information, chaque modalité transmettant une partie de cette information. L'utilisateur transmet en retour à la machine une information issue de la fusion de ses modalités. Cette multimodalité peut prendre différents aspects [Mar95]. Elle est complémentaire lorsque les modalités sont utilisées conjointement, ou encore redondante lorsque des modalités équivalentes sont utilisées simultanément. Par exemple, une interaction gestuelle avec un objet virtuel ne permet pas de sentir par le toucher l'objet virtuel. Compléter l'interaction avec une interface tangible, comme un téléphone intelligent localisé au même endroit que l'objet virtuel, permet d'associer le toucher de l'écran du téléphone intelligent au contact de la main avec

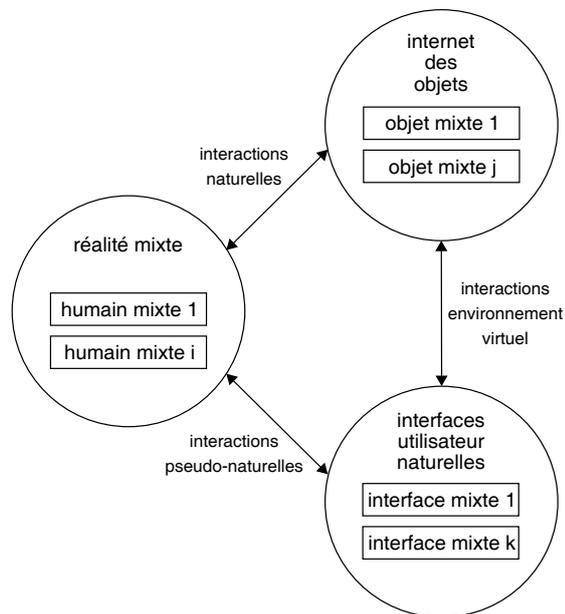


FIGURE 3.20 – Interactions multiples dans un système interactif mixte.

l'objet virtuel. Le téléphone peut également vibrer lors de ce contact, ce qui augmente le nombre de canaux sensoriels qui perçoivent ce contact. Par exemple, un TIES peut combiner le toucher et la gestuelle pour interagir avec un écran affichant une présentation lors de réunions, comme étudié par Bragdon et al. [BDHM11]. Sereno et al. combinent les interactions tactiles et tangibles pour la sélection de données spatialisées [SIAB18].

Une interaction hybride peut également être monomodale. Une interaction hybride monomodale bénéficie des qualités de chaque interface et contourne leurs limitations. Elle peut combiner une interface portée par l'utilisateur et une interface faisant partie de son environnement, ces interfaces utilisateur ayant toutes deux la même modalité d'interaction. Par exemple, combiner la détection et le suivi de la main de l'utilisateur par un HoloLens et par une Leapmotion est une interaction monomodale hybride. Lorsque la main de l'utilisateur est simultanément détectée par l'HoloLens et par la Leapmotion, elle permet à l'HoloLens d'afficher le jumeau virtuel de la main de l'utilisateur reconstruit par la Leapmotion. Cela a notamment pour utilité d'obtenir une reconstruction nettement plus précise de la main de l'utilisateur que celle fournie par l'HoloLens.

Une interface hybride peut en outre combiner multimodalité et monomodalité à interfaces multiples. Les systèmes interactifs mixtes offrent de multiples possibilités d'interaction hybride telles que :

- interagir avec plusieurs interfaces utilisateur intuitives simultanément, par exemple en combinant un casque de réalité mixte avec une interface haptique ;
- interagir avec plusieurs objets connectés simultanément, par exemple en commandant à l'aide d'un casque de réalité mixte l'ensemble des lumières d'une pièce ou d'un bâtiment ;
- interagir à plusieurs utilisateurs mixtes avec d'autres entités mixtes, par exemple plusieurs gestionnaires de parc automobile peuvent interagir simultanément, de façon collaborative, avec une flotte de véhicules connectés.

Si une interface utilisateur perçoit une information nécessaire à une autre interface utilisateur, il

est utile de privilégier le partage de l'information entre ces interfaces afin d'améliorer la performance du système interactif mixte. Dans cette optique, Golodetz et al [GCL⁺18] proposent d'accélérer la reconstruction dense d'un bâtiment grâce à une mutualisation des reconstructions denses effectuées par plusieurs utilisateurs équipés d'une caméra RGBD. La Figure 3.21 décrit ce système multi-utilisateurs à interactions monomodales multiples.

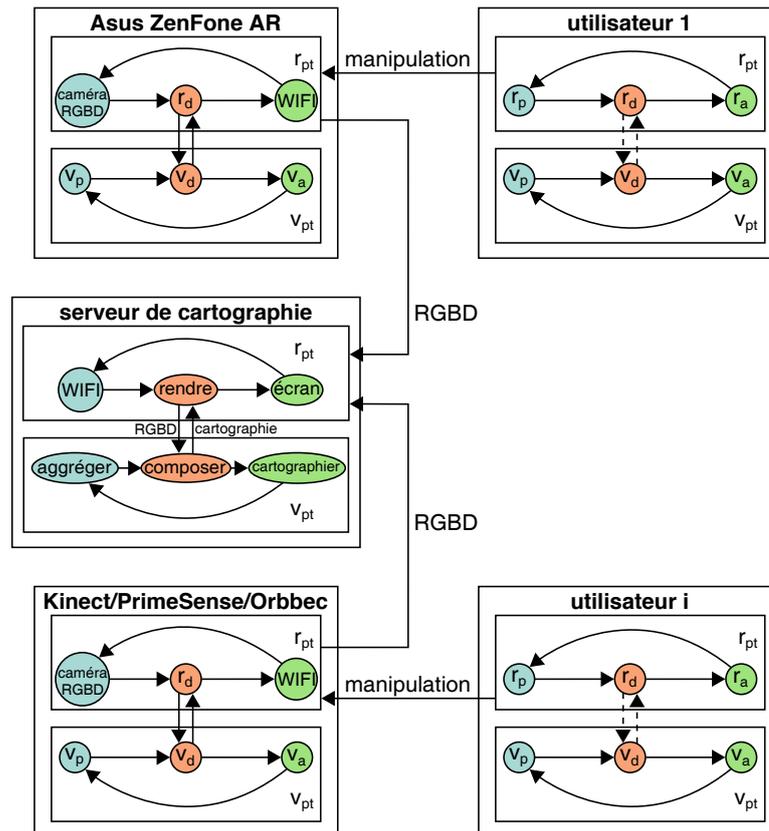


FIGURE 3.21 – Reconstruction dense mutualisée d'un bâtiment d'après [GCL⁺18]

Dans le cas d'un utilisateur mixte interagissant avec plusieurs interfaces utilisateur intuitives, la multiplicité des interfaces permet :

- de combiner les modalités d'interaction afin d'améliorer la performance ou l'expérience de l'utilisateur lors de son interaction avec un système ;
- de combiner plusieurs interfaces utilisateur afin de produire une interaction monomodale à interfaces multiples.

3.4.2 Interactions multi-sensorielles

Les interactions multi-sensorielles sont connues pour leur capacité à améliorer l'efficacité des interactions pseudo-naturelles [JPU19]. Il est donc utile lors de la conception d'un système interactif mixte de solliciter le plus largement possible les différents sens de l'utilisateur. Par exemple, un système interactif mixte est initialement pensé pour des interactions via un écran tactile entre

l'utilisateur et le jumeau virtuel d'un bâtiment intelligent. Nous améliorons cette interaction en combinant cette modalité d'interaction avec d'autres interfaces sensorielles, telles que le son binaural, la vibration de l'écran tactile ou encore la gestuelle combinée au tactile [HV16].

Déterminer les interactions multi-modales d'un système lors de la conception nécessite de comparer les multi-modalités possibles. Cette tâche peut être effectuée par un concepteur lors de la modélisation d'un système interactif mixte de dimension réduite, par exemple pour une pièce ou une maison connectée. Dans le cas de systèmes interactifs mixtes de grande dimension, comme une ville connectée, plusieurs solutions sont envisageables. Le domaine de l'internet des objets s'intéresse pour cette raison à la caractérisation automatique des services fournis à l'utilisateur. Le Pallec et al. [LPMN16] présentent une approche basée modèle visant à automatiser la composition de services offerts aux utilisateurs. Notre modèle est compatible avec cette approche. En effet, un TIES créé des interactions dynamiques entre entités mixtes capables de s'adapter en fonction du contexte. Si un développeur peut prévoir une adaptation d'une entité mixte à son contexte sans automatisation, il ne peut couvrir qu'un nombre restreint de configurations de TIES. Une alternative à cette approche est de segmenter un système de grande dimension en de multiples sous-systèmes et de permettre à des concepteurs, voir à des utilisateurs, de déterminer les meilleures interactions multi-modales possibles. DOMIM permet alors d'améliorer l'efficacité des interactions multi-modales que le concepteur souhaite réaliser. En effet, la décomposition des capacités d'interaction internes d'une entité mixte permet :

- de lister l'ensemble des capacités de perception et d'action des entités mixtes d'un système. Le concepteur peut alors associer certaines capacités du système en fonction de critères comme la proximité à l'utilisateur ou le gain apporté à une interaction utilisateur. Il peut également évaluer systématiquement l'opportunité de croiser ces capacités. Par exemple, lorsqu'un système intègre un téléphone intelligent et un casque de réalité mixte, l'interaction utilisateur n'utilise généralement qu'une petite partie des capacités d'interaction du téléphone. Or, ce type d'objet peut afficher, vibrer, émettre du son ou réagir au toucher. Il peut également estimer sa propre localisation, ou encore échanger avec d'autres objets connectés sa connaissance de l'utilisateur et de son environnement. Le concepteur peut alors choisir de coordonner une interaction gestuelle détectée par le casque de réalité mixte à une interaction tactile fournie par le téléphone ;
- d'augmenter le nombre de canaux sensori-moteurs lors d'une interaction entre un utilisateur et une ou plusieurs entités mixtes. L'analyse des capacités sensori-motrices du système permet alors de détecter des modalités non encore intégrées à cette interaction. Par exemple, lors de l'utilisation conjointe d'un téléphone intelligent et d'un casque de réalité mixte (HoloLens, Magic Leap,...), chacun de ces objets est contrôlé par une application. Ces applications peuvent communiquer mutuellement leurs interfaces utilisateur naturelles respectives et enrichir une interaction purement gestuelle avec une vibration, une interaction tactile ou encore en combinant leurs affichages respectifs, comme proposé par Normand et al. [NM18].

La Figure 3.22 montre cette complétion des modalités d'interaction visant à augmenter la crédibilité du contenu virtuel perçu par l'utilisateur.

Une complétion classique des modalités est par exemple d'associer la perception visuelle à la perception auditive de la partie virtuelle d'une entité mixte. Typiquement, deux utilisateurs qui interagissent en réalité virtuelle voient l'avatar humanoïde de l'autre, mais peuvent également se parler.

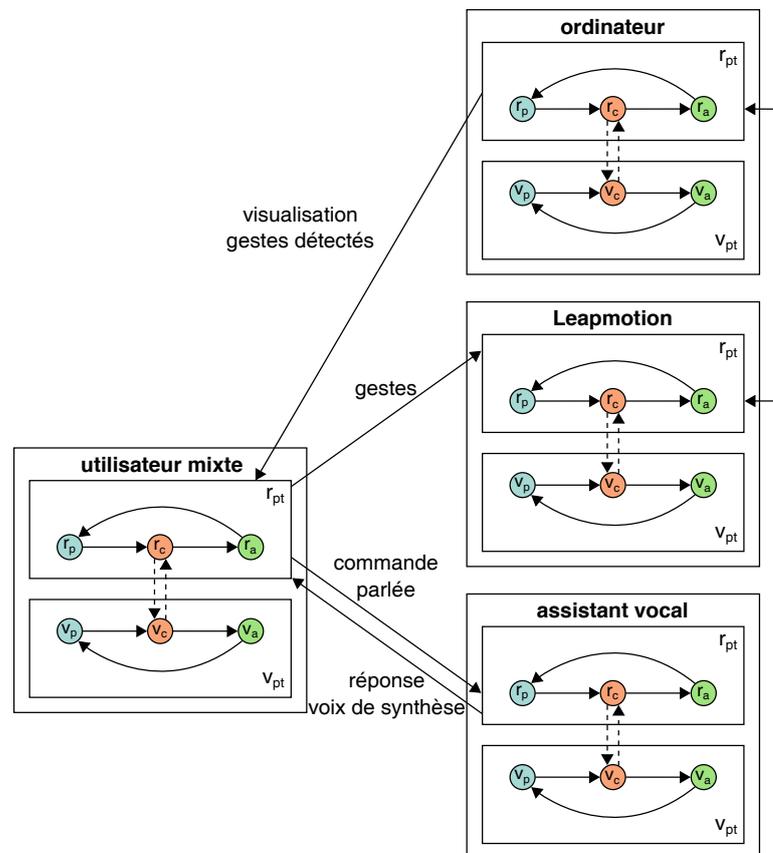


FIGURE 3.22 – Exemple d’interactions multi-sensorielles : joindre le geste à la parole permet de combiner plusieurs canaux sensori-moteurs. La performance et la crédibilité d’une interaction pseudo-naturelle sont alors améliorées. Par exemple, en pointant du doigt l’objet virtuel concerné par la commande vocale.

3.4.3 Interactions mono-modales à interfaces multiples

Combiner des modalités est utile pour améliorer une interaction pseudo-naturelle. En effet, chaque modalité permet d’atténuer les défauts des autres modalités en apportant ses qualités. Une alternative à cette approche permet d’améliorer une interaction pseudo-naturelle sans combiner plusieurs modalités. En effet, les interactions hybrides peuvent également combiner pour une même modalité d’interaction différentes interfaces utilisateur naturelles. Leur association permet d’augmenter l’information sensorielle échangée par un seul canal sensori-moteur. La même information est alors émise à partir de différentes localisations. DOMIM permet par exemple de concevoir et mettre en œuvre des bâtiments mixtes capables :

- de lister l’ensemble des interfaces utilisateur naturelles produisant la même modalité d’interaction. Chaque entité mixte du bâtiment déclare alors ses capacités au serveur qui supporte la partie virtuelle du bâtiment mixte. Le serveur détermine les interfaces à utiliser pour diffuser par exemple un message. Un bâtiment mixte peut notamment, en cas de danger, lister l’ensemble des écrans connectés qu’il contient en prévision de la diffusion d’un message d’évacuation par l’image ;
- de synchroniser ces interfaces utilisateur. Elles produisent alors simultanément la même interaction pseudo-naturelle. Cette synchronisation s’effectue par le réseau du bâtiment mixte

et est pilotée par la partie virtuelle du bâtiment. Par exemple, un téléphone intelligent porté par l'utilisateur et un écran posé dans son environnement complètent leurs affichages [LSH⁺ 15].

La Figure 3.23 présente un exemple d'interaction monomodale à interfaces multiples. Cet exemple associe deux caméras de profondeur Kinect reconstruisant le corps d'un utilisateur. La mutualisation de leur reconstruction améliore la robustesse du système. L'utilisateur perçoit le suivi de son corps à l'aide de deux écrans se trouvant contre des murs opposés. Plusieurs interfaces fournissent simultanément un stimulus identique pour l'utilisateur. Cela permet d'améliorer la robustesse du suivi de l'utilisateur lors des occultations. Radkowski et al. [RU18] combinent plusieurs Kinect 2 afin d'améliorer la précision de la reconstruction de leur environnement.

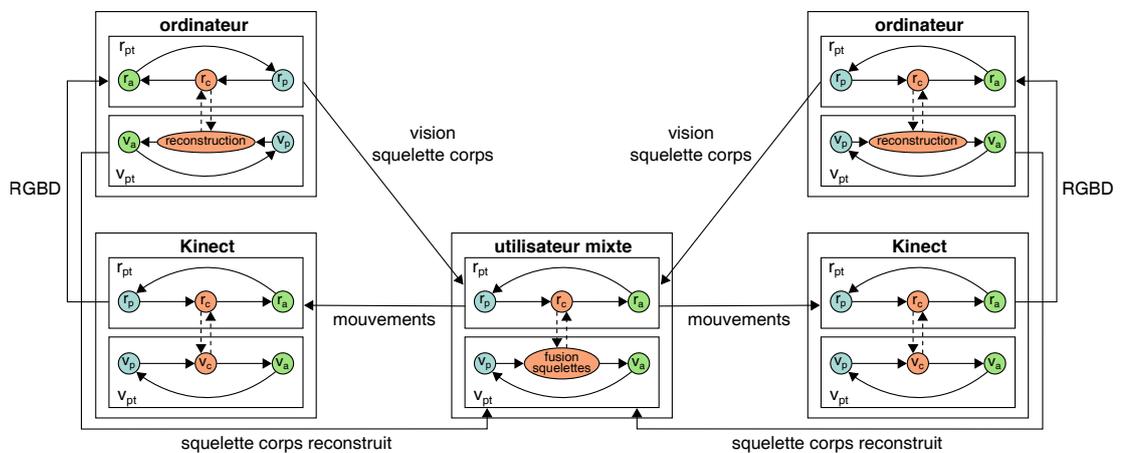


FIGURE 3.23 – Exemple d'interaction à modalité simple et interfaces multiples : deux caméras de profondeur reconstruisent le corps de l'utilisateur et suivent ses mouvements.

Les interactions utilisateur hybrides permettent l'interaction d'un utilisateur avec plusieurs interfaces utilisateur naturelles en combinant une ou plusieurs modalités sensorielles. Néanmoins, l'hétérogénéité des systèmes interactifs mixtes rend leur recherche difficile, en particulier lorsque ces systèmes sont de grande dimension. DOMIM facilite la découverte des modalités d'interaction disponibles dans ces systèmes en les caractérisant. Il augmente l'interopérabilité et l'extensibilité de ces systèmes en unifiant la modélisation et le développement de ces modalités. De cette façon, DOMIM favorise également la production d'interactions utilisateur hybrides.

3.5 Conclusion

Les systèmes interactifs mixtes combinant réalité mixte, interfaces utilisateur naturelles et internet des objets sont des systèmes hétérogènes, complexes lorsqu'ils sont de grande dimension, peu interopérables et peu extensibles. Ces aspects sont dus à un manque de modèle unifié permettant de les concevoir et de les développer de façon uniforme. Notre modèle DOMIM offre une approche unifiée de ces systèmes. Modéliser un système interactif mixte à l'aide de DOMIM permet de distinguer les entités mixtes constitutives du système et de décrire leurs interactions de façon unifiée. Il permet de désambiguïser le réel du virtuel au sein de chaque entité mixte. Il facilite l'étude exhaustive des interactions et des fonctionnalités de ces systèmes. Il simplifie et améliore la conception de ces systèmes grâce à son approche unifiée qui les rend plus interopérables et extensibles.

En bref, DOMIM soutient et facilite la conception et la mise en œuvre des systèmes interactifs mixtes combinant réalité mixte, interfaces utilisateur naturelles et internet des objets en permettant :

- d'identifier les entités mixtes constitutives d'un système interactif mixte ;
- de les modéliser avec une approche unifiée ;
- de décorrélérer la partie réelle de la partie virtuelle d'une entité mixte ;
- de décrire les interactions internes à une entité mixte ;
- de créer des interactions entre entités mixtes et d'étudier leur complémentarité ;
- de composer ces interactions afin de créer des interactions hybrides de plus haut niveau ;
- de segmenter ces systèmes en sous-systèmes afin de maîtriser leur complexité ;
- d'améliorer l'évolutivité, l'interopérabilité et l'extensibilité de ces systèmes grâce à son approche unifiée ;
- de créer des bibliothèques de composants et d'entités mixtes afin de factoriser, de simplifier et d'accélérer les pratiques de conception et de développement.

Dans le chapitre suivant, nous présentons un cadre de conception et de développement basé sur DOMIM. Ce cadre permet de créer des modèles d'application ainsi que des composants réutilisables. Il offre la possibilité aux créateurs de systèmes interactifs mixtes de simplifier leur production, de réutiliser les composants développés, de les rendre plus flexibles, d'améliorer leur capacité d'interconnexion et d'accélérer leur développement.

Chapitre 4

De la conception à l'application

Dans ce chapitre, nous présentons un *framework*¹ basé sur DOMIM. Ce framework offre une approche unifiée de la création des TIES permettant de simplifier et systématiser la gestion de leur complexité et de leur hétérogénéité. DOMIM donne un cadre d'architecture permettant de réaliser des systèmes complexes. Le framework facilite son implémentation et propose les outils pour le compléter et facilement développer ces systèmes. Nous souhaitons démontrer ainsi l'utilité de DOMIM de la conception à la mise en œuvre des TIES.

Premièrement, nous présentons un cas d'usage, le ventilateur mixte, utilisé dans ce chapitre pour illustrer les concepts présentés. Deuxièmement, nous donnons une vue d'ensemble de notre framework. Troisièmement, nous décrivons le processus de conception basé DOMIM. Quatrièmement, nous présentons le processus de développement basé DOMIM. Cinquièmement, nous analysons l'usage de ce framework. Enfin, nous décrivons son implémentation.

4.1 Cas d'usage du ventilateur mixte

A titre d'exemple, dans ce chapitre, nous concevons et développons un TIES de petite dimension à l'aide de notre framework visant à l'illustrer. Ce cas d'école nous permet de décrire ses propriétés, ses principaux concepts et les outils qu'il propose. Ce système permet à un utilisateur d'interagir avec un ventilateur mixte afin de contrôler sa mise en route. La partie réelle du ventilateur mixte est composée d'un ventilateur branché à une prise connectée. Sa partie virtuelle est un bouton virtuel marche/arrêt positionné sur sa surface. La Figure 4.1 donne une vue d'ensemble du système.

L'architecture matérielle de ce système, présentée par la Figure 4.2, est constituée de :

- un Hololens : il permet à l'utilisateur d'interagir en réalité mixte avec le ventilateur ;
- un routeur Wi-Fi : il assure la communication entre l'Hololens et l'ordinateur ;
- un ordinateur : il communique avec l'Hololens par Wi-Fi et avec le ventilateur par Z-Wave. Il fait office de serveur Z-Wave ;
- un dongle Z-Wave : branché sur l'ordinateur, il lui permet de communiquer en Z-Wave ;
- une prise connectée : compatible Z-Wave, elle communique avec le dongle Z-Wave contrôlé par l'ordinateur ;

1. Plusieurs traductions possibles mais néanmoins peu usuelles de ce terme sont : infrastructure logicielle, socle d'applications, canevas, infrastructure de développement, cadre de travail, boîte à outils, cadre d'applications, ou encore cadre de développement.



FIGURE 4.1 – Aperçu du système ventilateur mixte. A gauche, la vue de l'utilisateur augmentée par l'Hololens sur le ventilateur mixte. L'Hololens affiche en stéréoscopie le bouton virtuel du ventilateur mixte sur sa surface. A droite, vue d'ensemble de l'utilisateur équipé de l'Hololens et observant le ventilateur mixte branché à une prise connectée.

- un ventilateur : branché à une prise connectée interrupteur fermé, il se met en état de marche lorsque la prise lui délivre du courant.

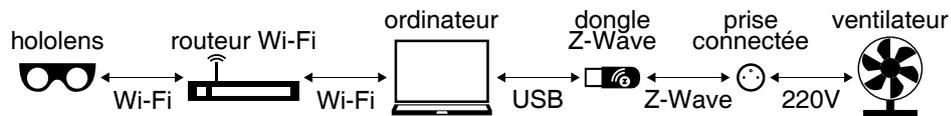


FIGURE 4.2 – Architecture matérielle du cas d'usage ventilateur mixte.

Le système permet à l'utilisateur d'observer et de contrôler un ventilateur mixte grâce à une interaction pseudo-naturelle. Pour être un TIES, il nécessite :

- d'être en présence d'un objet relié à un réseau, le ventilateur mixte ;
- de décorrélérer le réel du virtuel du ventilateur mixte afin de pouvoir décrire leurs comportements propres ;
- de percevoir la partie virtuelle du ventilateur et d'interagir avec elle.

Traiter chacun de ces aspects de manière méthodique permet de simplifier la conception et la mise en œuvre des TIES.

4.2 Vue d'ensemble du framework

Face à des TIES complexes, hétérogènes, peu interopérables ou extensibles, les concepteurs et les développeurs ont besoin d'unifier leurs pratiques et de disposer de méthodologies et d'outils adaptés. Afin de répondre à ces besoins, notre framework est composé de différents éléments qui sont :

- de s'appuyer sur le modèle unifié DOMIM pour créer un cadre de conception générique ;
- une méthodologie de conception unifiée par ce modèle, permettant de créer des modèles-type de conception ainsi que des architectures de référence afin de concevoir des systèmes simples, flexibles, réutilisables, interopérables et extensibles ;
- des outils de développement permettant de simplifier la production des TIES.

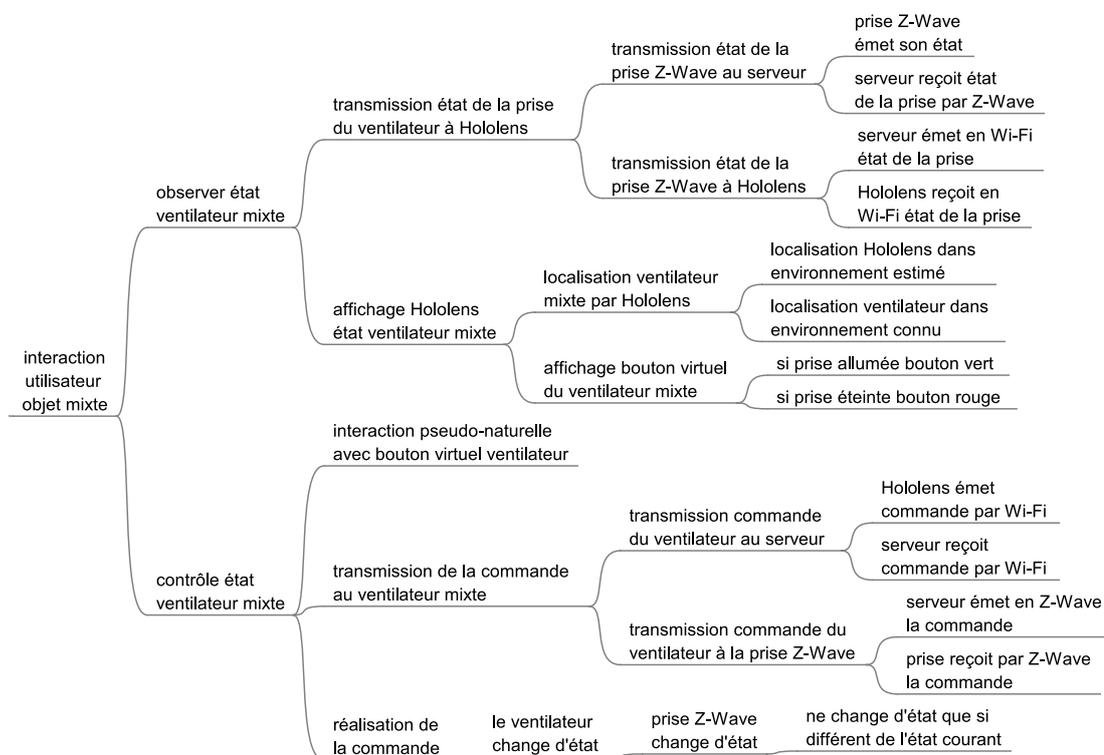


FIGURE 4.3 – Décomposition fonctionnelle du cas d'usage ventilateur mixte.

Notre framework guide les concepteurs et les développeurs dans la création de ces systèmes à travers une approche générique des domaines de la réalité mixte, des interfaces utilisateur naturelles et de l'internet des objets. Pour cela, il propose un cadre de conception et de développement qui s'appuie sur le modèle unifié DOMIM. La Figure 4.4 décrit l'ensemble du processus de conception et de développement dans le cas d'un système permettant à un utilisateur d'interagir à l'aide d'un Hololens avec un ventilateur mixte.

Dans un premier temps, la phase de conception modélise le TIES à l'aide du formalisme graphique décrit dans la section 3.1.2. La réalisation de ce modèle est effectuée à l'aide de modèles basés DOMIM réutilisables, qu'ils soient génériques ou issus de TIES antérieurs. Ce modèle permet de décrire l'architecture du système souhaité en déterminant les plateformes et applications nécessaires.

Dans un second temps, la phase de mise en œuvre réalise le système attendu. Les développeurs s'appuient sur le modèle et l'architecture établis afin d'implémenter le système. Ils peuvent itérer au besoin avec les concepteurs afin de faire évoluer modèle et architecture. Ensuite, les développeursinstancient les applications nécessaires au système. Un TIES nécessite généralement d'implémenter plusieurs applications interconnectées. Pour ce faire, notre framework d'implémentation offre des squelettes d'application pour différentes plateformes utilisées dans des systèmes de réalité mixte. Ces squelettes d'application fournissent une implémentation générique d'une entité mixte sous forme de réseau de Petri modifiable facilement grâce à un éditeur. Ils fournissent également des composants logiciels réutilisables afin d'accélérer, d'uniformiser et de simplifier le processus de développement. Par exemple, un composant logiciel fournissant un protocole réseau commun à un large nombre d'objets connectés comme le protocole Z-Wave, ou encore un algo-

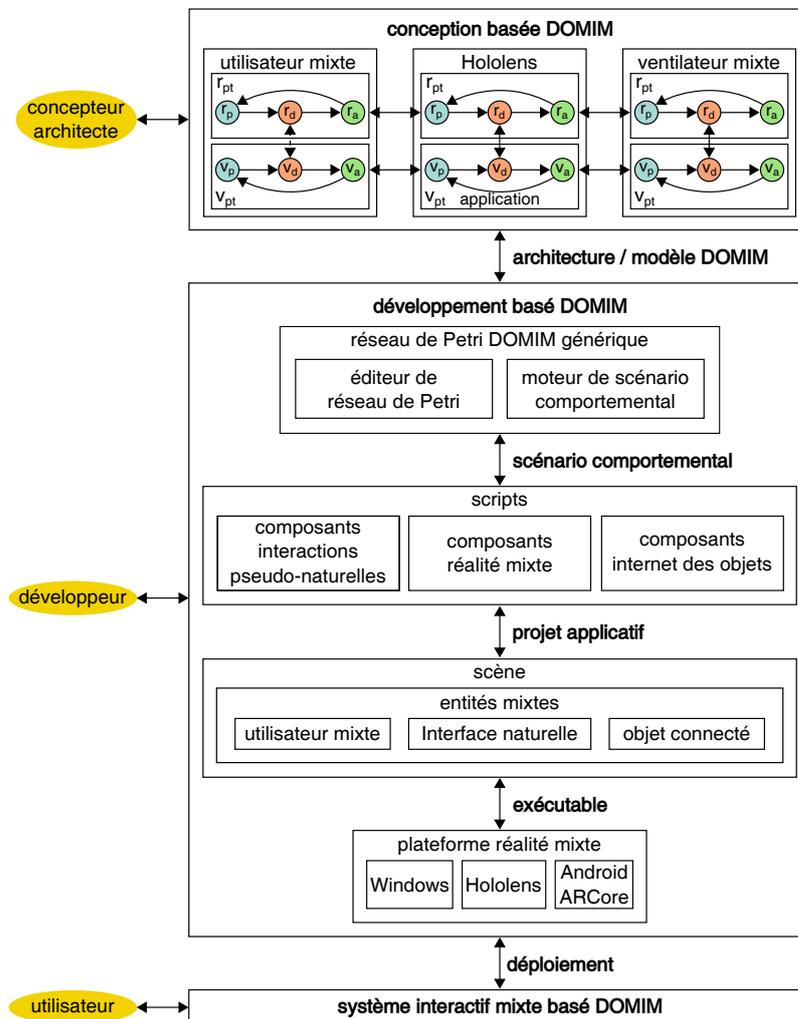


FIGURE 4.4 – Exemple d’usage du framework : un concepteur conçoit un ventilateur mixte. Un développeur implémente alors l’architecture résultante sous la forme de deux applications. Ces applications sont développées simplement et rapidement grâce à des squelettes d’application pour Hololens et Windows qui fournissent les composants d’interaction naturelle, de réalité mixte et de réseau nécessaires. Elles instancient les parties virtuelles de l’utilisateur et du ventilateur mixte.

rithme permettant à une entité mixte de se colocaliser avec les autres entités mixtes d’un TIES.

Après avoir déterminé et instancié les applications nécessaires, le développeur personnalise le comportement des entités mixtes instanciées par ces applications. Il utilise alors un éditeur permettant de personnaliser le scénario comportemental de chaque entité mixte constitutive du système. Ce scénario comportemental est relié à des scripts implémentant les comportements souhaités. Ces scripts sont des composants de l’application. Les composants pré-développés et fournis par notre framework sont dédiés aux interactions naturelles, au contrôle de la partie virtuelle des entités mixtes et à la communication réseau.

Notre framework accélère ainsi les transitions entre les différents types d’utilisateurs :

- grâce aux patrons de conception et aux composants réutilisables, un développeur retranscrit

facilement le travail d'un architecte ;

- grâce aux outils intégrés le développeur crée facilement une application qui va correspondre au système interactif d'un utilisateur donné.

Une fois l'ensemble des applications nécessaires implémentées, il ne reste plus qu'à déployer et utiliser le TIES.

4.3 Conception

Comme nous avons pu le décrire dans le chapitre 3, la phase de conception d'un système basé DOMIM repose principalement sur la réalisation d'un modèle et de l'architecture sous-jacente du système. Ce modèle inclut les entités mixtes constitutives du système et décrit leurs relations externes et internes. Nous n'avons pas développé de solution logicielle spécifique pour cette phase. Elle peut être réalisée à l'aide d'un modèleur UML comme Rational Rose, ArgoUML, Modelio ou encore Sparx Enterprise Architect. L'architecte logiciel devant concevoir le système avec DOMIM a la feuille de route suivante :

- énumérer les entités mixtes impliquées. Dans notre exemple du système ventilateur mixte on peut citer l'utilisateur, l'Hololens, le serveur et le ventilateur mixte, qui sont les quatre entités mixtes de ce TIES ;
- segmenter en sous-systèmes le système global afin de maîtriser sa complexité grâce à des sous-systèmes plus facilement appréhendables. La Figure 4.6 présente un sous-système qui fait abstraction du serveur en charge de la liaison entre Wi-Fi et Z-Wave. La Figure minimaliste 4.7 ne contient que deux entités mixtes, l'utilisateur et le ventilateur pour une compréhension encore plus simple du fonctionnel du système ;
- décrire les relations entre entités permettant la réalisation des fonctionnalités attendues, voir de repenser ces fonctionnalités. Ces relations sont décrites par la Figure 4.3 illustrant la décomposition fonctionnelle du système. Elles apparaissent plus simplement dans la Figure 4.5 issue d'une modélisation basée DOMIM. La Figure 4.7 présente d'une façon plus compréhensible les interactions entre entités mixtes et décompose plus clairement leur caractère bidirectionnel que l'arbre fonctionnel présenté par la Figure 4.3 ;

A cette étape, le fonctionnel du TIES est défini à haut niveau. Le concepteur poursuit sa feuille de route en détaillant ensuite les aspects techniques et technologiques de chacune des entités mixtes en :

- énumérant les capteurs et effecteurs nécessaires. Ils sont listés par exemple sur la Figure 4.5. Pour l'Hololens, les capteurs de la partie réelle sont ses caméras ainsi que sa centrale inertielle. Ses capteurs lui permettent de se localiser dans son environnement, de le reconstruire et de détecter les objets qui l'entourent ;
- complétant au besoin ces capteurs et effecteurs. Il est possible soit de les ajouter à l'entité mixte. Il est également possible d'utiliser les capteurs et effecteurs des autres entités mixtes du système dont les fonctionnalités sont accessibles par le réseau. Par exemple, si nous souhaitons interagir grâce à la reconnaissance vocale, nous ajoutons alors à l'entité mixte Hololens sur la Figure 4.5 le capteur *microphone* et sa prise en charge par un module de reconnaissance vocale intégré ou connecté à l'application ;
- décrivant la synchronisation, dont la complétion, entre partie réelle et partie virtuelle. Par exemple, la Figure 4.7 décrit comment la partie logicielle et la partie matérielle du ventilateur se combinent. En effet, elle fait apparaître comment l'utilisateur peut interagir par le geste avec le bouton virtuel afin d'émettre une commande par le réseau à la partie virtuelle du ventilateur mixte. Lors de la synchronisation entre la partie virtuelle et la partie réelle du ventilateur mixte, l'état marche/arrêt de sa partie réelle prend alors la même valeur que sa partie virtuelle.

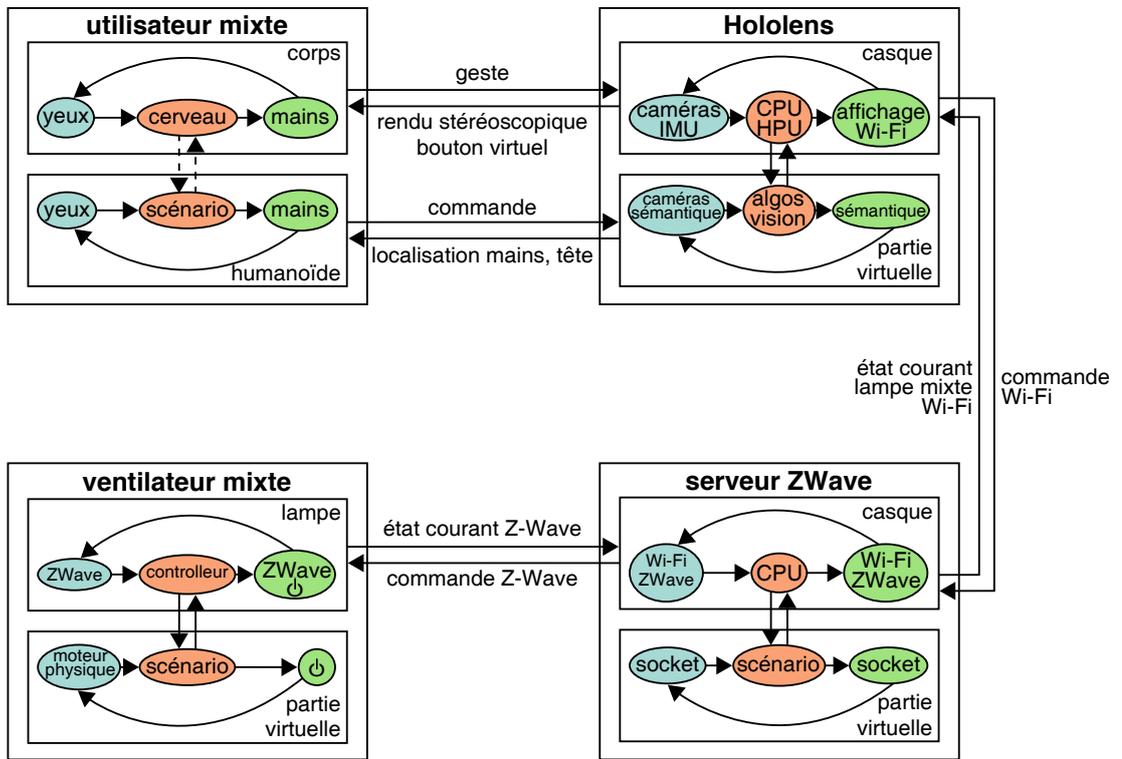


FIGURE 4.5 – Modèle basé DOMIM du système ventilateur mixte.

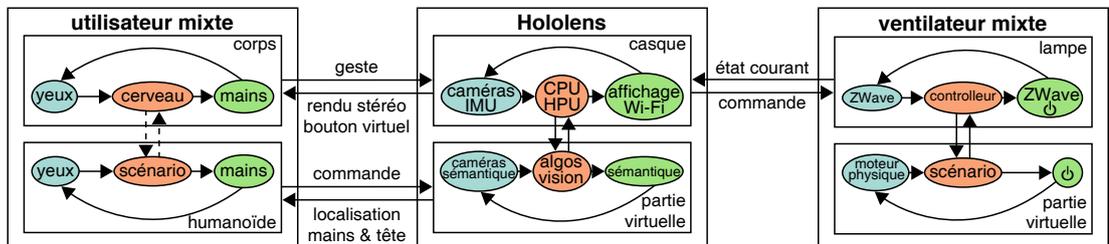


FIGURE 4.6 – Modèle basé DOMIM d'un sous-système ventilateur mixte simplifiant la compréhension du système. Les flèches en pointillé indiquent une interaction indirecte.

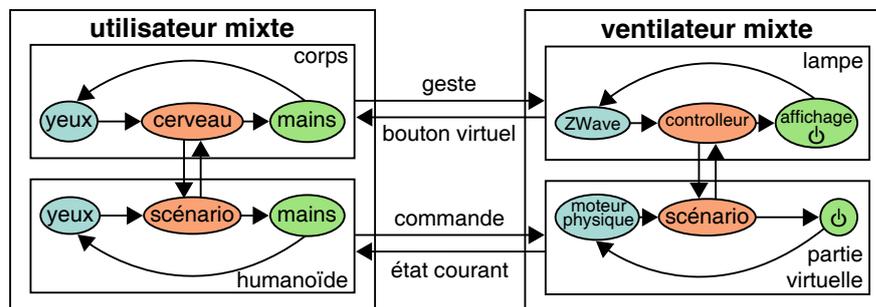


FIGURE 4.7 – Modèle basé DOMIM d'un sous-système ventilateur mixte réduit à deux entités mixtes, qui sont l'utilisateur et le ventilateur mixte.

A la sortie de cette étape de conception, le concepteur a créé des schémas d'architecture qui modélisent finement les entités mixtes constitutives du système interactif mixte, et leurs relations tant externes qu'internes. Le modèle le plus compact sera utilisé pour avoir une vue d'ensemble des fonctionnalités du TIES, par exemple lors d'une présentation du projet. Le modèle plus détaillé est celui qui sera utilisé par le développeur pour implémenter le système, car il présente l'architecture du système. Le concepteur peut également compléter cette description de l'architecture et des fonctionnalités du système par des documents écrits. Le développeur peut alors transposer ces schémas grâce aux outils de développement dédiés que nous présentons dans la section suivante.

4.4 Développement

Nos outils de développement sont conçus pour implémenter rapidement et facilement l'architecture produite lors de la conception du système. Ils permettent de transposer le comportement de chaque entité mixte dans l'environnement de développement des applications du système. Un concepteur peut effectuer cette transposition pour chaque entité mixte en laissant le développeur implémenter ses fonctionnalités, mais il n'aura pas la vue d'ensemble du système lors de cette phase. Une application pour Hololens porte alors les entités mixtes utilisateur et Hololens de notre exemple. Une autre application pour Windows porte le ventilateur mixte et communique avec la première application.

Le développeur associe un scénario comportemental générique à une entité mixte, par exemple un ventilateur mixte, comme indiqué dans la Figure 4.8. Il fournit également à chaque entité mixte un modèle 3D de sa partie virtuelle. L'édition et la mise en œuvre de ce scénario générique sont fournis par la partie tierce #SEVEN de Claude et al. [CGA15] intégrée au framework. Ce scénario modélise et implémente le modèle DOMIM sous la forme d'un réseau de Petri. Il est le lien entre le concept DOMIM et sa mise en œuvre. Il est éditable grâce à un éditeur interactif sous la forme de programmation graphique. Le développeur personnalise les nœuds transitions pré-établis du réseau de Petri qui implémente ce scénario. Un nœud transition possède des pré et post-conditions qui permettent de contrôler l'état courant de propriétés et de les modifier lorsque la transition s'est déroulée. Il peut également en ajouter afin de complexifier le comportement de l'entité. Cela lui permet alors d'associer plusieurs comportements afin de créer des comportements plus évolués. Cela permet également de s'adapter à des interactions utilisateur qui ne sont pas prévues dans le cas initial. Par exemple, un utilisateur peut exprimer son intention d'éteindre le ventilateur, mais le scénario peut également prévoir d'éteindre automatiquement le ventilateur si l'utilisateur ne fait plus face au ventilateur. Les nœuds place du réseau de Petri restent quant à eux inchangés.

Le développeur a à sa disposition des menus déroulants lui permettant de choisir le type de



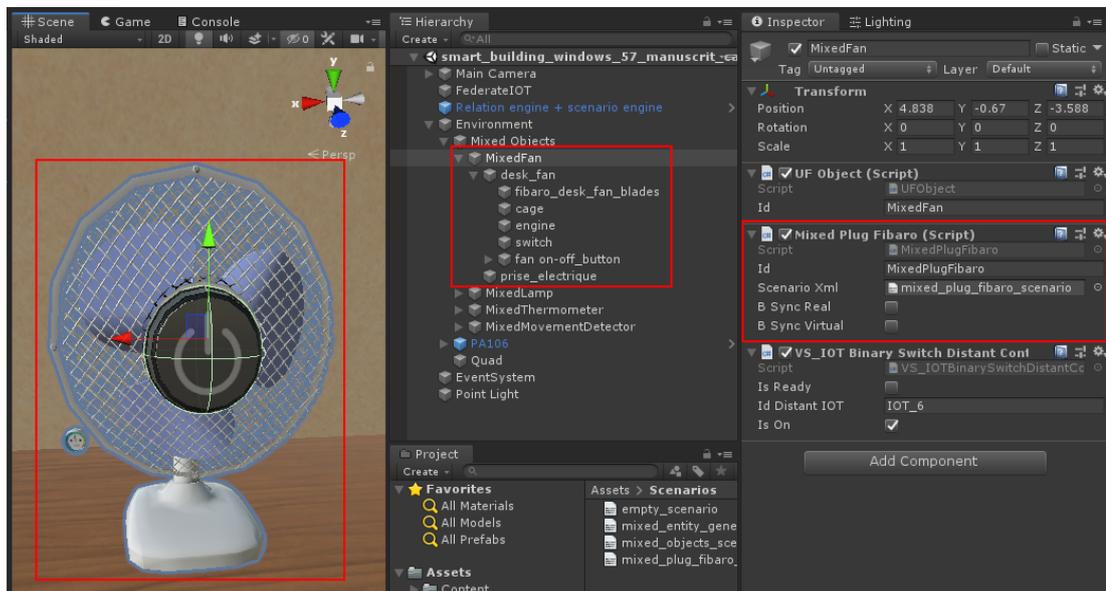


FIGURE 4.8 – Exemple de ventilateur mixte. Le ventilateur mixte fait partie du graphe de la scène, il est associé au moteur de scénarisation #SEVEN de Claude et al. [CGA15] via un script qui déclare son fichier de scénario.

capteur ou d'effecteur souhaité, quelle entité du graphe de scène est associée ainsi que la méthode d'un script associés à la pré ou post-condition de la transition. Le développeur peut soit utiliser des composants pré-développés dédiés à la RM, aux NUI et à l'IoT, fournis sous forme de scripts par le squelette d'application. Ces composants, détaillés par la suite dans ce chapitre, accélèrent et uniformisent les pratiques de mise en œuvre tout en évitant des implémentations de techniques complexes. Le développeur peut également développer ses propres scripts et ainsi enrichir le framework. Par exemple, le développeur utilise une application générique DOMIM pour Windows afin d'implémenter le ventilateur mixte. Il utilise un composant-type d'entité mixte générique qu'il nomme ventilateur mixte. Il ajoute alors la propriété marche/arrêt à ce composant. Il associe ensuite au scénario les composants logiciels nécessaire pour implémenter son comportement. Pour cela, il associe les composants mettant en œuvre les protocoles de communication réseau Wi-Fi et Z-Wave aux capteurs et effecteurs des nœuds transition du réseau de Petri grâce à des menus déroulants. Ainsi, ce scénario permet d'implémenter les échanges qui s'opèrent par le réseau de façon unifiée. Le scénario final de notre ventilateur mixte est représenté par la Figure 4.9. Il présente une symétrie entre partie réelle et partie virtuelle de l'entité mixte. En effet, il met en œuvre des boucles comportementales distinctes mais néanmoins modélisées de façon identique et synchronisées par une boucle. Les transitions suivantes décrivent le contrôle de l'entité mixte et donc son comportement :

- *set real power state* : met à jour les propriétés réelles du ventilateur mixte à partir de ses propriétés virtuelles. Par exemple, si sa partie virtuelle est en état de marche, une commande est alors transmise à sa partie réelle afin qu'il se mette en route ;
- *network communication* : phase de communication entre l'application mettant en œuvre le ventilateur mixte et d'autres applications ou des services du système, comme le service Z-Wave ;
- *check real power state* : capture l'état de fonctionnement du ventilateur afin de mettre à jour cette propriété réelle du ventilateur mixte, fournie par le serveur Z-Wave lors de la transition précédente ;

- *set virtual power state* : met à jour les propriétés virtuelles du ventilateur mixte à partir de ses propriétés réelles ;
- *semantics* : analyse les messages reçus par le réseau et prépare les messages à envoyer aux autres entités mixtes ;
- *check virtual power state* : observe l'état courant de la prise virtuelle qui peut avoir été activée/désactivée par son environnement virtuel, par exemple par un avatar.

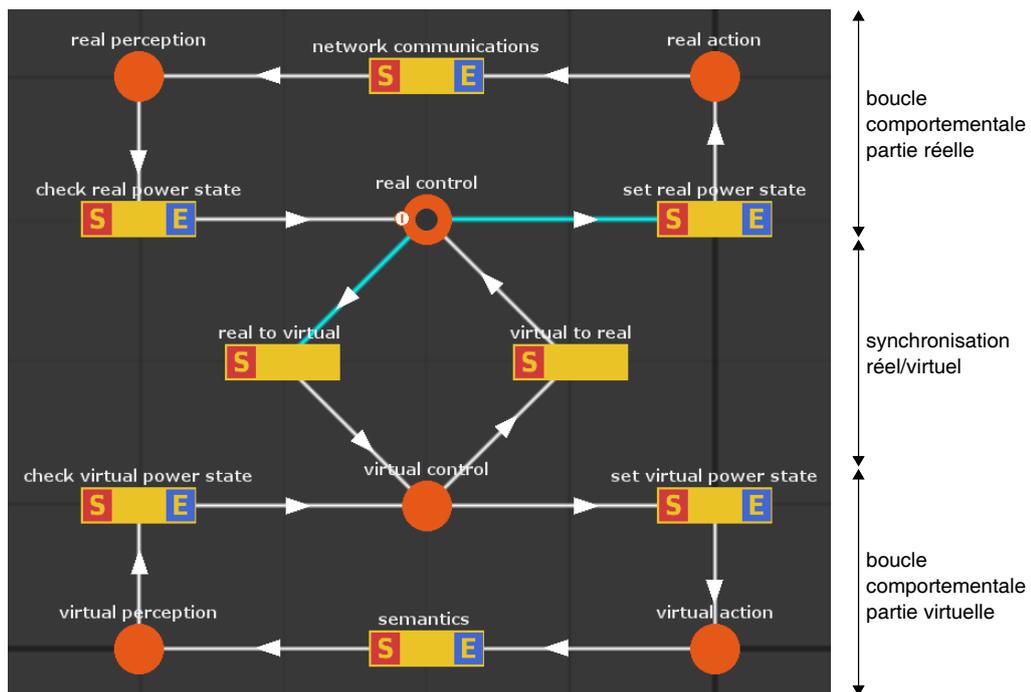


FIGURE 4.9 – Scénario comportemental du ventilateur mixte, affiché dans l'éditeur de réseaux de Petri intégré à Unity. Les cercles orange sont les nœuds de type place. Les rectangles jaunes sont les nœuds de type transition, le S indique un capteur et le E un effecteur. Le point noir est le jeton à sa place initiale. Les comportements des parties réelles et virtuelles DOMIM sont modélisés et implémentés de façon symétrique. Ils sont reliés par une boucle de synchronisation.

Si ce mécanisme d'implémentation est lourd lorsque le développeur implémente un TIES basique intégrant peu d'entités mixtes, il présente de multiples avantages dans le cas de TIES complexes :

- il permet de factoriser et de réutiliser les implémentations des comportements d'entités mixtes,
- la programmation graphique de ces comportements facilite leur réutilisation et leur compréhension rapide,
- en fournissant des composants pré-implémentés, nous unissons les TIES et les rendons plus interoperables et extensibles.

4.5 Implémentation

Notre framework nécessite l'environnement de développement (EDI) Unity 2018.4². Cet EDI est compatible avec un très grand nombre de plateformes, ce qui permet de développer des composants logiciels réutilisables pour les plateformes Android ARCore, Hololens, et Windows mais également d'étendre l'usage de ce framework à d'autres plateformes comme ARKit pour un coût de développement faible. Un projet générique est fourni pour chaque plateforme prise en charge. Ce projet contient :

- une implémentation générique d'un modèle interne DOMIM sous la forme d'un réseau de Petri [CGA15]. Cette implémentation donne un cadre de développement unifié d'une ou plusieurs entités mixtes instanciées par la plateforme cible du projet,
- un éditeur de réseau de Petri intégré à Unity [CGA15], qui fait le lien entre le scénario comportemental d'une entité mixte décrit dans la section 3.2.1 et les scripts implémentant son comportement et ses fonctionnalités,
- des composants pré-développés en C# sous forme de scripts dédiés à la réalité mixte, aux interactions pseudo-naturelles et à l'internet des objets. Ces composants accélèrent et uniformisent l'intégration de fonctionnalités comme la colocalisation des entités mixtes dans un référentiel spatial commun, le suivi de la main ou encore l'usage de protocoles réseau tels que le Wi-Fi ou Z-Wave.

Nous décrivons maintenant l'implémentation générique DOMIM sous la forme d'un réseau de Petri, avant de détailler les différents composants logiciels fournis par le framework via ses squelettes d'application.

4.5.1 Réseau de Petri DOMIM

Les réseaux de Petri sont utilisés pour modéliser et implémenter des systèmes temps réel, comme des systèmes de réalité virtuelle [MKBK98] pouvant être connectés [ZMD00], des systèmes interactifs multimodaux en réalité augmentée [MSW⁺03], des processus cyber-physiques de construction [Cor18], des systèmes de réalité mixte permettant de contrôler les bâtiments connectés en réalité virtuelle [LLCVF19] ou encore l'édition de contenu interactif en réalité augmentée [LGR⁺19]. Ce sont des graphes biparti orientés, dont les nœuds sont soit des places soit des transitions. Un ou plusieurs jetons se déplacent dans ce graphe de place en place lorsque une ou plusieurs préconditions à une transition sont validées et donnent lieu à des post-conditions lors de ce passage.

Nous utilisons le moteur de scénario #SEVEN [CGA15] pour intégrer une implémentation des réseaux de Petri à Unity. Ces moteurs et leurs éditeurs associés, intégrés à Unity via un plugin, permettent de spécifier les scénarios comportementaux des entités mixtes, comme décrit dans la section 3.2.1. Un scénario générique implémente le modèle interne DOMIM décrit dans la section 3.1.2 notamment par la Figure 3.6. Le moteur #SEVEN exécute continuellement la synchronisation de la boucle en huit du modèle interne DOMIM. Pour cela, il exécute une boucle comportementale de sa partie réelle, suivi d'une traduction de ses propriétés réelles en propriétés virtuelles et d'une injection dans sa partie virtuelle, puis une boucle comportementale du virtuel, avant de traduire les propriétés virtuelles en propriétés réelles et de les injecter dans sa partie réelle. Ce processus est décrit par la Figure 4.10. Par exemple, le scénario du ventilateur mixte est décrit par la Figure 4.9. Les étapes de ce scénario sont décrites dans la section 4.4.

Plusieurs entités mixtes peuvent être instanciées par la même application. Chaque entité a alors son propre scénario comportemental, pouvant communiquer en temps réel avec d'autres scénarios.

2. <https://unity.com/fr>

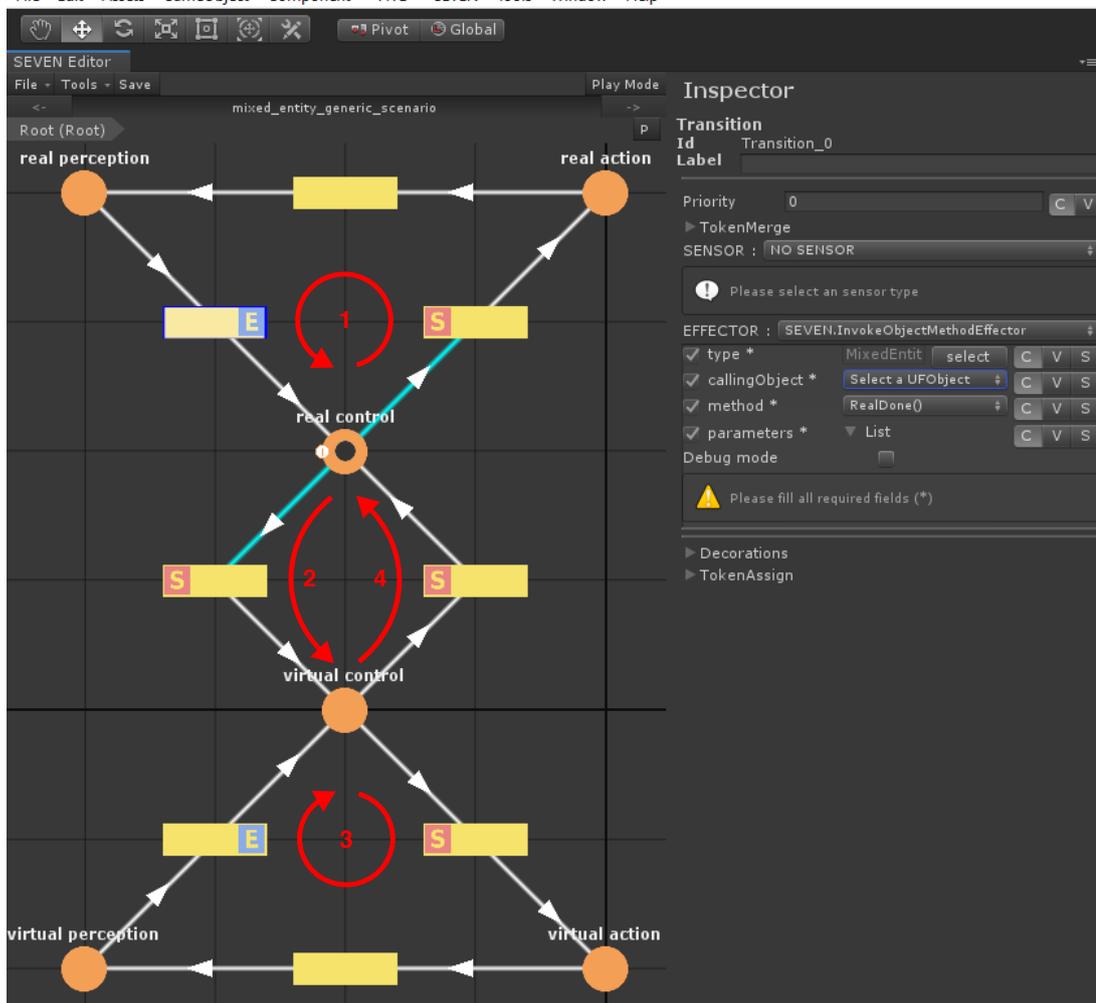


FIGURE 4.10 – Scénario générique implémentant une entité mixte sous la forme d’un réseau de Petri. La bouche comportementale du réel est la première séquence du graphe. Ensuite, l’état du réel est traduit en état du virtuel. Puis, la boucle comportementale du virtuel est exécutée. Enfin, l’état du virtuel est actualisé dans le réel. Chaque entité mixte du graphe de scène possède son propre scénario comportemental qui associe des scripts pré-développés ou non aux capteurs et effecteurs des transitions.

Par exemple, la même application peut instancier simultanément la partie virtuelle d’un ventilateur mixte et la partie virtuelle d’une lampe mixte. Cette communication permet d’assurer leur coordination, si nécessaire au bon fonctionnement du système et à la réalisation des interactions attendues. L’application peut, par exemple, éteindre simultanément la lampe et le ventilateur lorsque l’utilisateur sort de la pièce.

4.5.2 Composants

Développer des systèmes interactifs mixtes combinant réalité mixte, interfaces utilisateur naturelles et internet des objets fait appel à des compétences en développement très variées et pointues. Cela rend leur développement chronophage et difficile à maintenir aux performances de l'état de l'art dans des domaines aussi nombreux et dynamiques. De plus, les aspects liés à la communication réseau de ce type de système, comme les technologies de transmission réseau, les protocoles et les architectures réseau employés, rendent critique le besoin en uniformisation de ces communications afin de permettre à ces systèmes de s'interconnecter et de simplifier leur évolutivité et leur extensibilité. Proposer des composants logiciels pré-développés est donc un gain en temps de développement, en performance, en qualité et robustesse des applications développées ainsi qu'en extensibilité et en adaptabilité de ces systèmes. Nous donnons une vue d'ensemble des composants déjà fournis par notre framework dans la Figure 4.11.

Composants pour la réalité mixte

Deux composants de réalité mixte permettant de rendre cohérents plusieurs affichages d'un TIES sont disponibles. Ils sont utilisés pour colocaliser des interfaces utilisateur naturelles et leur permettre de combiner leurs modalités d'interaction. Le premier composant permet de colocaliser un Hololens et un téléphone mixte afin de compléter les interactions entre le téléphone et l'utilisateur à l'aide de l'Hololens. Ce composant est alors nécessaire pour que l'Hololens affiche de façon cohérente les informations du téléphone en fonction de sa pose. Le deuxième composant combine le rendu stéréoscopique d'un Hololens au rendu stéréoscopique d'un écran stéréoscopique, ce qui produit une réalité mixte multi-couches. Cette réalité mixte multi-couches permet d'améliorer les champs de vision et de regard du système interactif mixte, mais également d'améliorer la gestion des occultations du virtuel par le réel, par exemple d'une banane virtuelle par la main de l'utilisateur. Ces composants permettent à un développeur de ne pas avoir à re-développer la synchronisation du rendu visuel entre un Hololens et un téléphone mixte ou entre un Hololens et un écran stéréoscopique.

Co-localisation d'un Hololens et d'un mobile Android ARCore

Le premier composant permet de co-localiser deux interfaces utilisateur mixtes, une étant un Android ARCore et l'autre étant un Hololens, afin qu'ils puissent contribuer de façon collaborative à l'instanciation d'une entité mixte. L'Hololens et le mobile Android ARCore embarquent tous deux la capacité de reconstruire la géométrie de leur environnement réel et de s'y localiser. Ils sont également connectés par un réseau Wi-Fi local qu'ils peuvent instancier eux-mêmes. Cette connexion leur permet d'échanger en temps réel leurs poses respectives dans le même environnement. Néanmoins, le référentiel spatial de chacun ne leur permet pas d'interpréter correctement les coordonnées fournies par l'autre objet connecté. Afin d'estimer une matrice de passage $M_{m \rightarrow h}$ de la transformation du mobile t_m de son propre repère vers le repère de coordonnées de l'Hololens $t_{m \rightarrow h}$, l'Hololens estime la pose du mobile $t_{m \rightarrow h}$ en détectant et suivant la texture qu'il affiche. Simultanément, le mobile transmet à l'Hololens par le réseau sa matrice de transformation t_m dans son repère local. La matrice de passage est alors obtenue grâce au calcul suivant :

$$M_{m \rightarrow h} = t_h - t_{m \rightarrow h}$$

Lorsque l'Hololens ne détecte plus la texture affichée par le mobile, l'Hololens utilise alors cette matrice de passage pour calculer la transformation du mobile dans son propre repère. Ce type de composant est utile car l'estimation de la pose d'une texture par l'Hololens :

- est consommatrice de ressources calculatoires et peut alors nuire au bon fonctionnement de l'application Hololens durant l'interaction ;

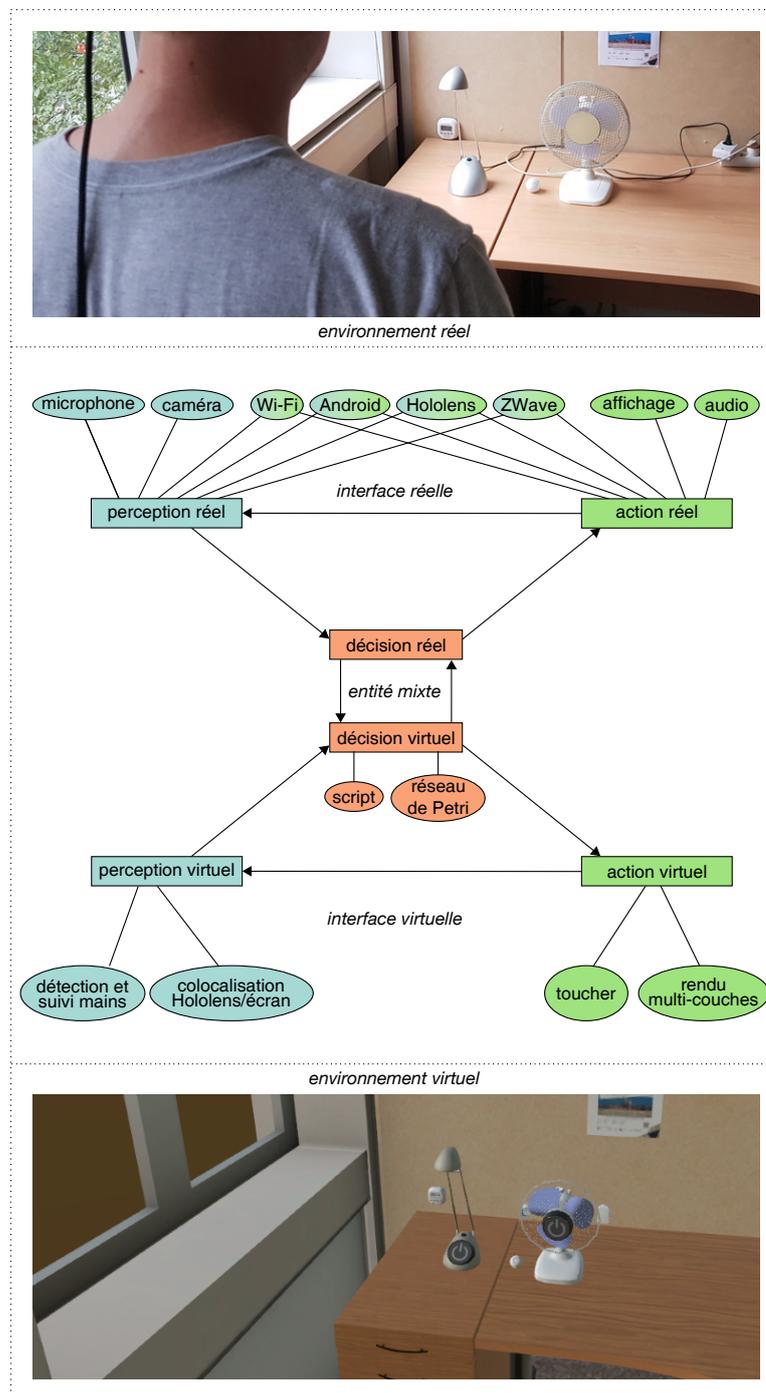


FIGURE 4.11 – Vue d’ensemble des composants de notre framework de développement basé DO-MIM. Tous les composants de cette figure sont développés et disponibles pour les utilisateurs du framework. D’autres pourraient venir enrichir cet ensemble : reconnaissance et synthèse vocale, haptique, BCI,... Ces composants peuvent être rapidement reliés à l’application via l’éditeur de réseau de Petri.

- ne peut être assuré que si l'utilisateur portant l'Hololens regarde l'écran du mobile à la distance adéquate pour sa bonne captation par la caméra de l'Hololens. Si l'écran du mobile n'est pas dans le champ de vision de l'Hololens sa pose ne peut pas être estimée. Ce calcul permet également à l'Hololens d'être capable par la suite de localiser un mobile se trouvant dans une autre pièce.

La mise en œuvre de ce composant requiert :

- la transmission par le mobile à l'Hololens de sa localisation estimée par ARCore ;
- l'intégration au scénario comportemental de l'application Hololens du script implémentant ce calcul afin qu'elle ait la connaissance de la localisation du mobile dans son environnement ;
- l'instanciation dans l'application Hololens du jumeau virtuel du mobile localisé.

Co-localisation d'un Hololens et d'un écran statique

Le second composant permet de colocaliser un Hololens et un écran statique stéréoscopique. Il reprend le principe du 1er composant de réalité mixte, mais en le simplifiant, car l'écran étant statique, sa pose ne nécessite d'être estimée qu'une seule fois. En effet, une fois la pose de l'écran estimée, l'Hololens peut alors instancier le jumeau virtuel de l'écran et le localiser dans son environnement sans nécessité de mise à jour de cette localisation durant le cycle de vie de l'application.

Composants pour l'interaction pseudo-naturelle

Rendu stéréoscopique multi-couches

Les affichages des systèmes interactifs mixtes sont des caractéristiques distinctes. Dans certains cas d'usage, leurs limitations peuvent être compensées en les associant à d'autres dispositifs de rendu. Par exemple, un Hololens a un champ de vision limité, un champ de regard élevé et simule avec plus ou moins de succès les occultations du virtuel par le réel. Le champ de regard est la zone totale pouvant être perçue par un capteur en mouvement³. A l'inverse, un écran statique offre un champ de vision large, un champ de regard réduit et est incapable de gérer les occultations du virtuel par le réel lorsque le réel occultant le virtuel se trouve entre les yeux de l'utilisateur et le virtuel. C'est par exemple le cas des salles immersives et des CAVes. La Figure 4.12 montre un utilisateur porter une paire de lunettes stéréoscopiques actives sous un casque Hololens afin de pouvoir voir simultanément en 3D le contenu de l'écran et le contenu de l'Hololens. Ce composant peut être réutilisé pour une meilleure gestion des occultations dans le cas de salles immersives par des développeurs travaillant sur ce type d'environnement. Il a pour fonction de fournir un rendu stéréoscopique multi-couches d'un ou plusieurs environnements virtuels. Ce rendu est réparti entre plusieurs affichages stéréoscopiques se trouvant simultanément dans le champ de vision de l'utilisateur. Ce composant permet de résoudre les problèmes d'occultation du virtuel par le réel, notamment dans le cas des écrans stéréoscopiques statiques lorsqu'ils sont associés à un casque de réalité mixte à optique semi-transparente. Il est également utile pour composer des environnements combinant contenus virtuels publics et privés.

L'ordinateur en charge d'effectuer le calcul du rendu affiché par l'écran calcule les matrices de projection des caméras virtuelles correspondant aux yeux de l'utilisateur. La localisation de ces caméras virtuelles est estimée par le casque de réalité mixte en fonction de sa propre localisation ainsi que de la localisation de l'écran estimée à l'aide du composant de colocalisation d'un Hololens et d'un écran statique décrit précédemment. Cette localisation est fournie à l'ordinateur contrôlant l'écran statique en temps réel par une boucle réseau. Chaque matrice de projection est estimée grâce à un script implémentant un algorithme classique d'estimation du frustum⁴. Les étapes de cet algorithme sont les suivantes :

3. https://en.wikipedia.org/wiki/Field_of_regard

4. https://fr.slideshare.net/N_Baron/view-frustum-in-the-context-of-head-tracking

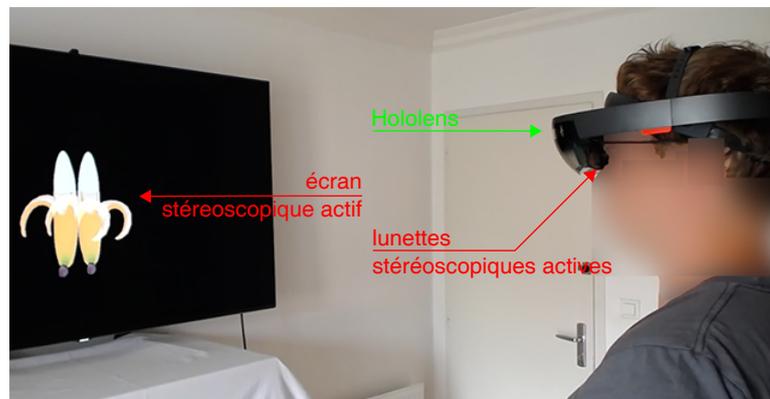


FIGURE 4.12 – Utilisateur portant simultanément des lunettes stéréoscopiques actives et un casque de réalité mixte à affichage semi-transparent, un HoloLens.

- définir la localisation des yeux de l'utilisateur relativement au centre de l'écran :
 - la localisation du casque est inversement transformée dans le repère de coordonnées de l'écran par l'application du casque ;
 - sa position relative est communiquée par réseau à l'application contrôlant l'écran statique ;
 - la localisation de chaque œil est estimée en fonction de la distance interpupillaire de l'utilisateur et de la localisation du casque dans le repère de l'écran ;
- calculer le frustum de chaque caméra virtuelle en fonction de la localisation de l'œil et de la taille de l'écran ;
- calculer la matrice de projection de chaque caméra virtuelle localisée au niveau de l'œil de l'utilisateur.

Détection et suivi de la main

Ce composant fournit une implémentation de la détection et du suivi de la main de l'utilisateur. Dès lors, cette implémentation scriptée est facilement intégrable au scénario comportemental d'une entité mixte via l'éditeur de réseau de Petri. Il permet de détecter la collision entre les jumeaux virtuels de plusieurs entités mixtes, ce qui permet par exemple de détecter le contact entre ces entités dans le monde réel, comme nous le décrivons dans le chapitre 5. L'extension de ce composant à d'autres plateformes que l'HoloLens, par exemple Magic Leap, Android ou encore iOS, pourrait venir enrichir le système.

Toucher mixte

Ce composant fournit une implémentation de la détection du toucher d'une surface physique. De la même manière que le composant précédent, cette implémentation scriptée est facilement intégrable au scénario comportemental d'une entité mixte via l'éditeur de réseau de Petri. Il permet de détecter par exemple la collision entre le doigt de l'utilisateur et l'écran tactile. Cela permet alors de compléter le nombre de canaux sensori-moteurs engagés par l'utilisateur afin d'interagir avec un système interactif mixte. Cette complétion améliore la perception du virtuel par l'utilisateur. Ce composant s'appuie sur les modules fournis par Unity par défaut, il est donc générique aux différentes plateformes compatibles avec Unity.

D'autres composants dédiés à l'interaction pseudo-naturelle sont également envisageables, comme un composant de reconnaissance ou de synthèse vocale, d'interaction haptique, d'inter-

action gestuelle ou encore d'interaction cerveau-machine s'appuyant la bibliothèque OpenVibe⁵.

Composants internet des objets

Deux composants dédiés à l'internet des objets sont disponibles. Ces deux composants permettent à des applications DOMIM de communiquer entre elles via le même protocole de communication.

Socket TCP

Le premier composant implémente la communication threadée d'un socket TCP entre un client et un serveur en s'appuyant sur .NET. Il permet à plusieurs applications d'un TIES de communiquer entre elles afin d'échanger leurs propriétés virtuelles. Il est utilisable aisément dans un réseau local combinant Wi-Fi et ethernet par exemple. Deux scripts sont fournis par le framework, l'un fournissant une implémentation d'un client et l'autre l'implémentation d'un serveur. Le code fourni ne nécessite que de fournir et de traiter les messages à envoyer ou recevoir ainsi que l'adresse IP du serveur. Ils sont implémentés sous la forme d'un thread dédié afin de ne pas bloquer les autres fonctionnalités de l'application durant les échanges continus en temps réel entre applications DOMIM. Nous n'avons pas développé de protocole spécifique, cette tâche sortant du cadre de cette recherche et requérant l'étude des différents protocoles standardisés et normalisés existant pour les trois domaines de la réalité mixte, des interfaces utilisateur naturelles et de l'internet des objets. Un exemple de démarche de standardisation de la communication entre interfaces utilisateur 3D est le consortium UMI3D⁶. UMI3D est un consortium qui développe un projet open source de protocole de communication basé interaction pour la collaboration en réalité mixte temps réel.

Composant ZWave

Notre composant Z-Wave permet à une application faisant office de serveur Z-Wave et équipée d'un module de communication Z-Wave de contrôler l'ensemble des objets connectés compatibles Z-Wave d'un système interactif mixte. Les objets connectés compatibles Z-Wave communiquent avec ce serveur en transmettant leur état courant et en exécutant les commandes transmises par le serveur. Ce composant s'appuie sur un service Z-Wave basé sur la bibliothèque logicielle Z-Wave .NET. Un serveur Z-Wave nécessite pour la plateforme Windows un périphérique Z-Wave connecté, par exemple un dongle branché sur port USB.

4.6 Conclusion

Les processus de conception et de développement que nous avons décrit dans ce chapitre sont identiques pour tous les systèmes interactifs mixtes. Notre framework orienté conception décrit le processus de conception de ces systèmes dans la section 4.3. Il permet de concevoir ces systèmes plus simplement et facilement en définissant les étapes successives de conception. Il fournit également un modèle générique d'entité mixte qui permet de décrire ces systèmes à différentes granularités qui sont, de la plus élevée à la plus basse :

- une description des interactions attendues entre plusieurs systèmes interactifs mixtes ;
- un inventaire des entités mixtes qui constituent le système et de leurs relations ;
- une description de sous-systèmes de ce système ;
- une description du comportement interne de chaque entité mixte.

5. <http://openvibe.inria.fr/>

6. <https://umi3d-consortium.org/>

Ces multiples niveaux de description permettent de modéliser finement le fonctionnement d'un système interactif mixte. Par exemple, nous avons décrit dans la section 4.3 un système ventilateur mixte à plusieurs niveaux de granularité.

Notre framework de développement permet aux développeurs de produire de façon accélérée le système. En effet, il n'a pas à créer un projet pour une plateforme logicielle à partie de zéro, tous les outils nécessaires et composants logiciels pré-développés sont disponibles dans le squelette d'application fourni par notre framework :

- le développeur n'a pas à développer de modèle d'entité mixte. C'est un facteur important de gain de temps et d'uniformisation des systèmes basés DOMIM. Le développeur déclare les propriétés de chaque entité mixte. Il intègre à chaque entité mixte le modèle 3D de sa partie virtuelle. Ce modèle 3D est généralement le jumeau virtuel de sa partie réelle. Il peut également être une métaphore complémentaire, ou une combinaison des deux. Dans le cas du ventilateur, il déclare sa propriété marche/arrêt. Il intègre également le modèle 3D du ventilateur, dont son bouton virtuel qui vient compléter le jumeau virtuel du ventilateur ;
- le développeur vient personnaliser un scénario comportemental générique basé DOMIM grâce à un éditeur de graphe inclus à l'EDI (Unity⁷ dans notre cas) en décorant ses nœuds de transition avec des composants logiciels pré-développés. Par exemple, l'entité mixte ventilateur utilise le composant prédéveloppé de communication Z-Wave pour s'interfacer avec la prise connectée qui contrôle son état de fonctionnement. Ces composants peuvent également fournir des techniques d'interaction pseudo-naturelles. La factorisation du code et l'usage d'une technique d'implémentation similaire à du block-coding sont des facteurs de gain de temps. Ce procédé fournit également une *visualisation* du comportement de l'entité mixte qui simplifie la compréhension du comportement de l'entité mixte. Par exemple, un graphe de Petri décrit par la Figure 4.9 présente le comportement du ventilateur mixte ;
- en cas de fonctionnalité non fournie par les composants logiciels pré-développés, le développeur la développe dans un cadre de développement très structurant qui assure sa réutilisabilité et sa portabilité sur différentes plateformes. Par exemple, la création du composant de communication Z-Wave permet de fournir rapidement à d'autres entités mixtes la connectivité Z-Wave sans nouvelle implémentation. Elle est attribuée simplement à une entité mixte via l'éditeur de réseau de Petri. Si ce composant n'est plus adapté, on peut le remplacer par un autre composant mettant en œuvre un protocole réseau équivalent. Pour cela, le développeur a juste à modifier un nœud du réseau grâce à l'éditeur fourni.

Une extension de ce framework à d'autres EDI comme le Unreal Engine est envisagée afin d'augmenter la généricité du framework, mais également de réduire sa dépendance à un EDI spécifique. Une autre alternative est la création d'un EDI entièrement dédié à notre problématique. L'avantage majeur d'un tel EDI serait de ne plus devoir utiliser un EDI par application, mais de regrouper l'ensemble des applications à implémenter au sein d'un même projet afin de simuler leurs interactions. Afin d'accélérer le développement du système interactif mixte, les concepteurs et les développeurs auraient alors la possibilité de le simuler dans un premier temps en réalité virtuelle. Ils pourraient ainsi tester le contrôle des parties virtuelles avant d'étendre ce système à la réalité mixte. En effet, la réalité mixte pose de multiples contraintes supplémentaires au développement et au déploiement du système. Elle requiert d'intégrer en plus la partie réelle de chaque entité mixte. Cela nécessite un espace pour tester le système. De plus, la synchronisation entre réel et virtuel est également consommatrice en temps. S'assurer que la partie virtuelle d'un système interactif mixte est fonctionnelle permet de gagner du temps avant de le déployer et de traiter d'autres problématiques de développement. Mais le coût de développement d'un tel EDI serait alors très élevé. De plus, il nécessiterait une collaboration étendue avec les différents acteurs de la réalité mixte, des interfaces utilisateur naturelles et de l'internet des objets.

7. <https://unity.com/>

Notre méthodologie de passage entre le formalisme graphique DOMIM et l'implémentation d'un TIES structure le passage de sa conception à sa mise en œuvre. Elle ouvre la porte à l'automatisation de ce processus, utile à l'interconnectivité et à l'extensibilité des TIES. Elle permet notamment d'automatiser des processus spécifiques à l'IoT, comme l'identification automatique des protocoles nécessaires, la découverte des entités mixtes disponibles ou manquantes, la caractérisation des services ubiquitaires possibles en fonction des besoins des utilisateurs, etc. Elle permet également d'automatiser la création d'interactions multimodales améliorant la naturalité d'une interaction en coordonnant les NUI disponibles.

Chapitre 5

Cas d'usage

Dans ce chapitre, nous présentons les cas d'usage étudiés dans le cadre de nos travaux de recherche sur les systèmes interactifs mixtes combinant réalité mixte, interfaces utilisateur naturelles et internet des objets. Ces prototypes nous permettent de mettre en avant les capacités du framework basé DOMIM, qui sont de :

- fournir un modèle générique applicable aux utilisateurs, aux interfaces utilisateur naturelles et aux objets connectés. Ce modèle assure la synchronisation entre réel et virtuel tout en les décorrélant. Il offre la flexibilité nécessaire à l'interopérabilité et à l'extensibilité des systèmes interactifs mixtes ;
- créer des squelettes d'application génériques par plateforme. Ces squelettes contiennent une implémentation-type d'une entité mixte dont le comportement est à décrire en décorant un réseau de Petri. Ces squelettes permettent non seulement de réduire le temps de développement, mais également d'uniformiser les pratiques de développement des systèmes interactifs mixtes ;
- fournir des bibliothèques de composants par domaine. Les domaines couverts sont la réalité mixte, les interfaces utilisateur naturelles et l'internet des objets. Les composants pré-développés fournis par notre framework de développement permettent d'accélérer le développement en factorisant l'effort de développement. Ils permettent également d'accéder à des techniques avancées de développement qui requièrent des compétences métier spécifiques.

En fournissant un modèle générique, DOMIM permet à notre framework d'associer RM, NUI et IoT de façon homogène de la conception à la mise en œuvre. Il n'existe pas actuellement de modèle générique de ce type dans la littérature, ce qui rend la conception et la mise en œuvre des TIES hétérogène, complexe, peu interopérable et extensible. La fourniture de squelettes d'application permet d'obtenir un environnement de développement intégrant des outils unifiés et dédiés ainsi que des implémentations-type d'une entité mixte et de composants logiciels régulièrement utilisés par les TIES. Cela permet d'homogénéiser les pratiques de développement grâce à des outils communs, et de les accélérer en réduisant la charge de travail pour les développeurs grâce à la factorisation et à la réutilisation du code produit. L'implémentation-type d'une entité mixte basée DOMIM fournit un formalisme graphique sous la forme d'un réseau de Petri générique. Ce réseau implémente les différents aspects de DOMIM et les rend facilement et rapidement programmables grâce à une programmation graphique. Sans ce type d'outils qui permet de gagner du temps de développement tout en simplifiant la mise en œuvre des TIES, les systèmes produits restent hétérogènes, peu ou pas interopérables, décorrélant réel et virtuel de manière variable, ce qui est une source de confusions et bloque l'extensibilité de ces systèmes. Notre framework, en procurant des composants logiciels spécifiques qui requièrent du temps et des connaissances avancées en RM, NUI et IoT, permettent à des développeurs non-spécialistes de mettre en œuvre des TIES simplement et rapidement.

Nous présentons dans ce chapitre six cas d'usage réalisés grâce à notre framework basé DOMIM. Le premier permet de contrôler par le geste un environnement domotique. Le second met en œuvre un contrôle tangible de cet environnement. Le troisième autorise une distribution du rendu de l'interface utilisateur entre un visiocasque de réalité mixte et un téléphone intelligent tout en fournissant une interaction tangible avec cette interface. Le quatrième cas étend ce TIES en fournissant des interactions hybrides qui associent interactions multimodales combinant gestuelle, tangible et tactile à un rendu visuel distribué. Le cinquième cas déploie ces interactions hybrides au contrôle d'un bâtiment connecté par l'intermédiaire de sa maquette numérique. Le sixième cas s'intéresse au retour multisensoriel de la localisation des mains d'un utilisateur de visiocasque de réalité mixte.

5.1 Contrôle gestuel des environnements domotiques

Nous décrivons dans cette section comment notre framework peut être utilisé pour produire des systèmes interactifs mixtes permettant à l'utilisateur d'interagir par le geste avec l'internet des objets. Ce cas d'usage montre comment DOMIM peut être utilisé pour modéliser de façon générique le TIES mis en œuvre. Il permet également de montrer l'utilisation des squelettes d'application incluant les outils et les composants pré-intégrés socket TCP, composant Z-Wave et interaction gestuelle.

Ce cas d'usage est une extension du cas d'usage ventilateur mixte décrit dans la section 4.1, auquel il ajoute un deuxième objet mixte, une lampe connectée. Ce cas d'usage permet à un utilisateur de contrôler un ventilateur connecté en interagissant de façon pseudo-naturelle avec sa partie virtuelle. Notre objectif est de montrer l'uniformité des systèmes basés DOMIM, ainsi que le gain fourni par la réutilisation des travaux issus de la conception et du développement du cas d'usage ventilateur mixte. Les objets mixtes utilisés dans cette section sont un ventilateur et une lampe. Ce cas d'usage permet à un utilisateur d'interagir avec un objet connecté grâce aux gestes reconnus par le système de réalité mixte utilisé. Nous utilisons le mode d'interaction gestuel reconnu par Hololens, que nous allons tout d'abord caractériser.

La technique d'interaction gestuelle fournie par défaut par Hololens est indirecte. Elle combine un air-tap, décrit par la Figure 5.1 d'après [TAL18], et un pointeur permettant de sélectionner l'objet de l'interaction. Cette technique permet de contourner les limites d'un champ de vision de l'Hololens ne couvrant qu'une petite partie du champ de vision humain, ainsi que les limites d'un cône de détection de la main également restreint. L'air-tap tient compte de la position des doigts de la main tandis que le pointeur est dirigé par le regard. Les interactions gestuelles que nous décrivons dans ce chapitre s'appuient toutes sur la détection et le suivi de la main lors d'un air-tap.

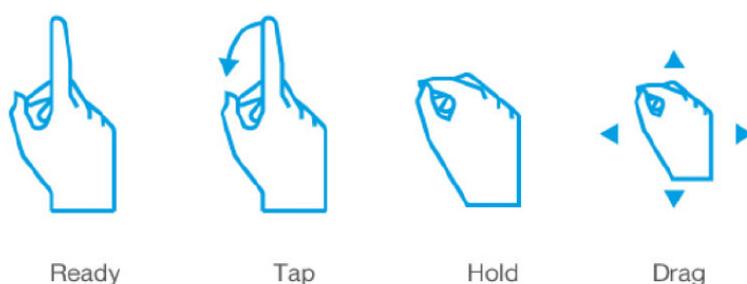


FIGURE 5.1 – Le geste air-tap utilisé par l'Hololens d'après [TAL18]. Le geste *ready* est préliminaire à la détection et au suivi de la main. Un autre geste ou une main qui sort du cône de capture gestuelle de l'Hololens met fin à la détection d'une séquence gestuelle de type air-tap.

Ce cas d'usage permet de commander par le geste un objet connecté en réalité mixte. Il fait appel à l'interaction gestuelle standard de l'Hololens, qui consiste à regarder l'objet à commander puis activer la commande via un air-tap. Les deux objets commandés sont un ventilateur et une lampe. Chacun de ces objets est branché à une prise connectée compatible Z-Wave. La nécessité d'un serveur Z-Wave équipé d'un dongle Z-Wave implique d'inclure à ce système un ordinateur ayant cette fonction, en l'occurrence un PC Windows, mais qui peut également être un Raspberry, un hub Z-Wave... Ce système nécessite en outre un Hololens, deux prises connectées, un ventilateur et une lampe.

La phase de conception de ce système est basée sur notre framework de conception. Elle reprend les étapes décrites dans la section 4.3. Le système étant extensible à plusieurs objets mixtes, elle étend l'architecture du système décrite par la Figure 4.5 en ajoutant l'entité mixte ventilateur et en décrivant ses interactions avec le reste du système. L'action utilisateur dans ce schéma est la gestuelle conjuguée au regard. La factorisation d'une conception basée DOMIM est donc réelle. La Figure 5.2 présente le modèle de conception basé DOMIM du système ventilateur et lampe mixte. Les parties virtuelles de l'utilisateur, de la lampe et du ventilateur sont instanciés par l'Hololens. L'Hololens effectue leur rendu visuel et les interactifs grâce aux NUI embarquées par l'Hololens. Le serveur Z-Wave permet à l'Hololens de communiquer avec le ventilateur et la lampe mixtes à l'aide de la technologie radio Z-Wave alors que Hololens communique via Wi-Fi avec lui. Le serveur Z-Wave sert de plateforme de communication entre les technologies réseau Wi-Fi et Z-Wave. L'Hololens peut alors transmettre des commandes aux objets mixtes et s'informer de leur état courant.

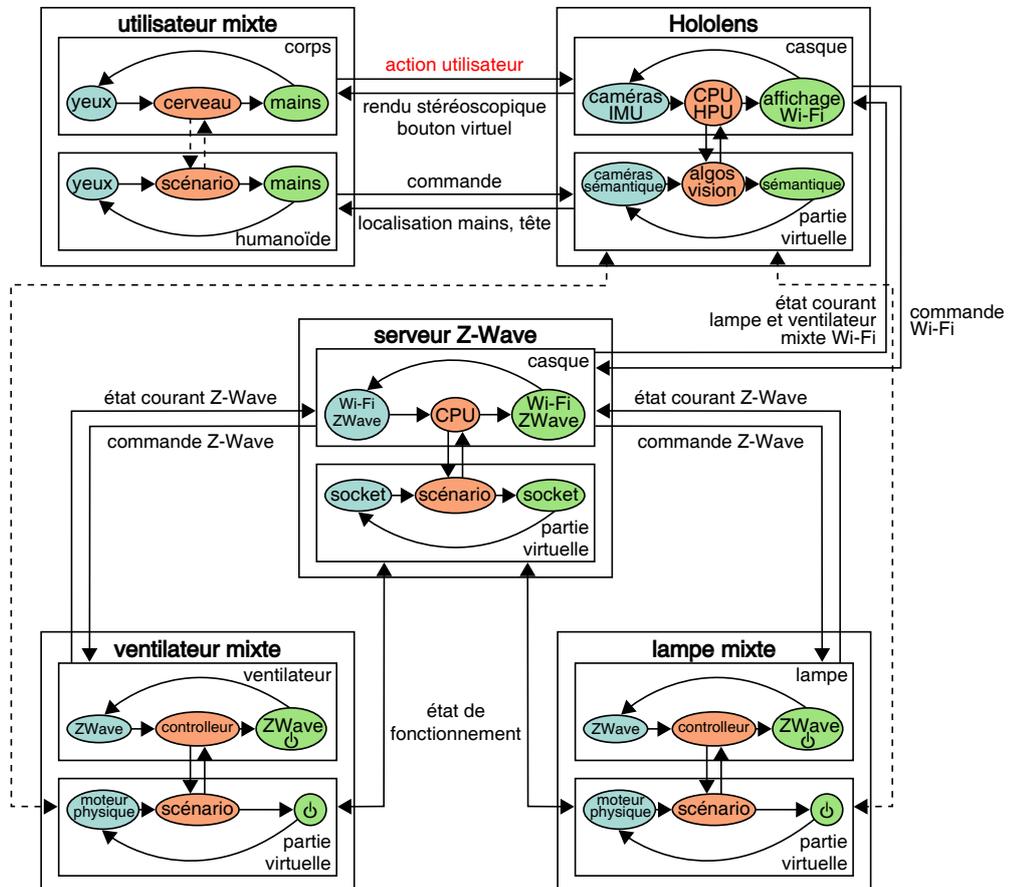


FIGURE 5.2 – Modèle basé DOMIM d'un système lampe et ventilateur mixte interactif par l'action utilisateur gestuelle combinée au regard.

Nous décrivons comment l'usage de notre framework permet de factoriser et d'accélérer le développement de notre cas d'usage. Nous reprenons également le développement issu de la section 4.4. L'implémentation du système nécessite le développement de deux applications, une pour HoloLens et l'autre pour le serveur Z-Wave. Or, ces objets mixtes ont déjà été développés pour le cas d'usage du chapitre 4.

Chaque application est développée sous Unity avec notre framework de développement à l'aide de nos squelettes d'application HoloLens et Windows. Le serveur Z-Wave communique avec l'HoloLens par le réseau Wi-Fi tandis qu'il communique avec les prises connectées des objets mixtes via le réseau Z-Wave. Les prises connectées fournissent au serveur leur état courant, qui transmet cette information à l'application HoloLens. L'application HoloLens affiche la partie virtuelle de chaque objet connecté, un bouton marche/arrêt, sur la surface de cet objet. Les outils de notre framework de développement utilisés par ces applications développées auparavant sont :

- l'implémentation générique d'une entité mixte DOMIM sous forme de réseau de Petri générique. Elle contient les éléments internes réels et virtuels d'une entité mixte ainsi que leurs interactions ;

- l'éditeur de réseau de Petri pour décorer à l'aide de scripts les nœuds transition du graphe générique ;
- les composants :
 - socket TCP ;
 - plateforme Windows : serveur Z-Wave.

Le système est montré par la Figure 5.3. Les parties virtuelles du ventilateur et de la lampe mixte sont visibles sur les images de droite. Ces parties sont rendues en stéréoscopie. Elles viennent compléter leur partie réelle. L'utilisateur peut alors mettre en route ou arrêter ces objets mixtes à distance. Par rapport au cas d'usage initial, l'extension du développement de ce système a consisté à :

- ajouter à l'application Hololens la partie virtuelle de la lampe mixte dans le jumeau virtuel de la pièce. La reconstruction d'un jumeau virtuel n'étant pas couverte par le scope de cette thèse, la partie virtuelle de la lampe, un bouton marche/arrêt, ainsi que le jumeau virtuel de la pièce sont issus d'une modélisation par un infographiste. Néanmoins, le jumeau virtuel de la pièce pourrait également être obtenu par la reconstruction 3D de la pièce fournie par un Hololens, une caméra RGBD associée à un algorithme de reconstruction 3D, un LIDAR, un scanner laser ou autre ;
- ajouter au réseau de Petri la perception de la lampe mixte et son contrôle ;
- étendre la sémantique des messages échangés entre Hololens et le serveur Z-Wave à la prise en compte de la lampe mixte. Ces messages permettent de connaître l'état courant de la lampe et de la commander par le réseau ;
- ajouter à l'application du serveur Z-Wave la présence de la lampe mixte. Cela nécessite de déclarer au serveur Z-Wave l'identité réseau de la prise connectée associée à la lampe mixte. Elle est alors capable de communiquer avec la prise connectée de la lampe mixte, afin de connaître son état courant et de la commander. L'application intègre également à ses messages avec Hololens ces informations.

Le coût en temps de l'extension du système à la lampe mixte est alors faible pour le développeur. Les communications réseau, de même que les mécanismes d'interaction gestuelle sont identiques. Il est possible d'étendre facilement et rapidement à l'ensemble des prises connectées d'un bâtiment ce type d'interaction. Étendre un TIES à plus d'objets connectés ne pose plus de problème de complexité dans la mesure où il est possible de segmenter un TIES en sous-systèmes facilement compréhensibles grâce à notre formalisme graphique et notre modèle unifié DOMIM. Décrire le comportement d'une entité mixte issu de la composition des processus décisionnels réels et virtuels est notamment couvert par ce formalisme. Notre framework basé DOMIM structure la mise en œuvre du TIES, ce qui le rend interopérable et extensible. Les squelettes d'application accélèrent et uniformisent l'implémentation des applications du système. Les bibliothèques de composants fournies par notre framework évitent des développements fastidieux et pointus.

Sans DOMIM, un développeur doit définir une méthodologie unifiant les domaines couverts par les TIES s'il souhaite les rendre interopérables et extensibles. Il doit également développer des composants logiciels qui sont dans le cas de notre framework basé DOMIM fournis et conformes à notre méthodologie. Généralement le développeur va au plus pressé sans s'occuper d'homogénéité des TIES, ni d'interopérabilité ou d'extensibilité, il produit alors un prototype non réutilisable. Les TIES produits sont basiques et ne tirent pas partie de l'ensemble des opportunités d'interactions innovantes qu'offrent les TIES. Typiquement, leur architecture logicielle génère des confusions entre réel et virtuel qui se concrétisent par des problèmes de cohérence entre propriétés réelles et virtuelles, sources de bogues du TIES. Par exemple, si au lancement du système ventilateur mixte aucune synchronisation n'est effectuée alors que le ventilateur fonctionne, sa partie virtuelle indiquera qu'il est à l'arrêt. De même, sa partie virtuelle ne peut afficher qu'il est à l'arrêt que lorsque la commande est transmise à sa partie réelle, puis appliquée, et non dès réception de la commande.

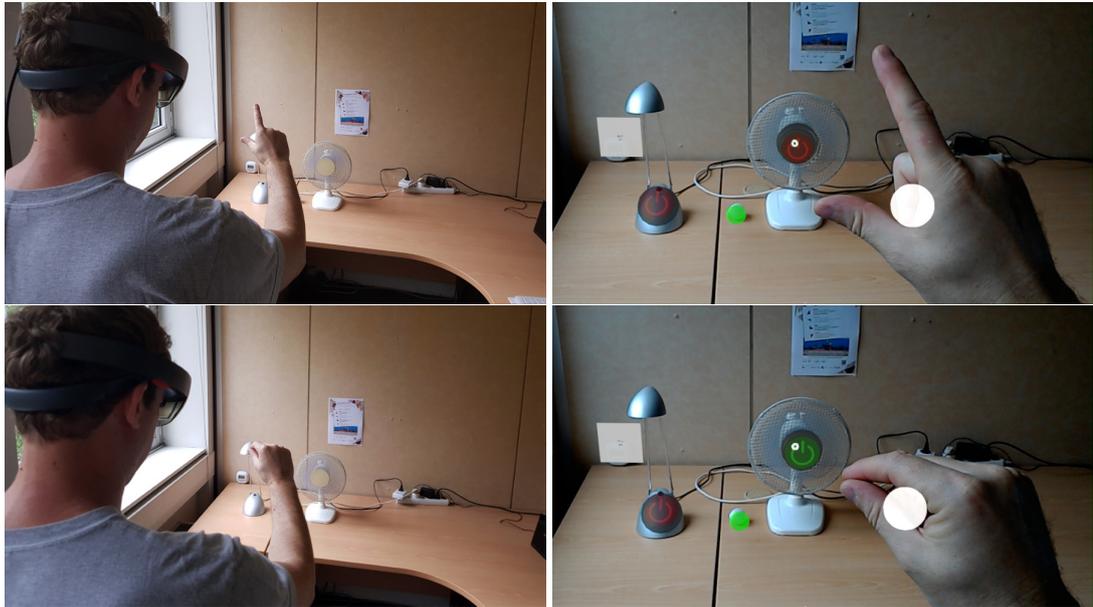


FIGURE 5.3 – Interaction gestuelle avec un ventilateur mixte. En haut, le regard de l'utilisateur fixe le bouton virtuel marche/arrêt du ventilateur mixte, l'utilisateur est prêt à activer le ventilateur à l'aide d'un geste air-tap. En bas, l'utilisateur a activé le ventilateur. L'interaction gestuelle avec la lampe mixte se déroule de la même manière.

L'extensibilité du système permet par exemple d'ajouter d'autres objets compatibles Z-Wave à ce système à moindre coût, comme un ventilateur mixte. La communication réseau utilisant le protocole standardisé Z-Wave, l'inclusion de cet objet ne nécessite aucune action particulière, ce qui le rend interopérable. Il reste alors à :

- donner la connaissance de ce nouvel objet mixte aux autres entités du système. Nous intégrons à l'environnement virtuel connu des autres entités ce nouvel objet. L'usage des technologies IoT de découverte automatique des entités d'un système peut par exemple être utilisé afin d'automatiser cette tâche. Des bibliothèques d'objets 3D pourraient également fournir la géométrie et l'apparence d'un objet à détecter et identifier dans l'environnement réel ;
- attribuer à tous les objets du système un comportement incluant ce nouvel objet. Les technologies IoT de caractérisation de service pourraient être utiles à l'automatisation contextuelle de la mise en œuvre des services ;
- étendre les modalités d'interaction disponibles en fonction de la nature et des fonctions de l'objet.

5.2 Contrôle tangible des environnements domotiques

Nous montrons dans cette section comment la technique d'interaction pseudo-naturelle décrite dans la section précédente peut être facilement et rapidement remplacée par une autre technique grâce à notre framework. Ce cas d'usage fait appel à notre composant toucher mixte. Les interactions gestuelles ont pour bénéfice de ne pas toucher de surface, ce qui est pertinent dans un espace public pour des raisons sanitaires et hygiéniques. Elles contournent les problèmes liés à l'interaction directe. Néanmoins, l'absence de toucher rend ces interactions peu naturelles. De plus, elles requièrent un apprentissage des gestuelles reconnues par le système. Or, les interactions avec la main peuvent prendre plusieurs formes. Elles peuvent être tangibles en touchant et manipulant

directement l'objet de l'interaction. L'utilisateur cherche dès lors à interagir avec un objet virtuel au plus près de l'interaction réciproque dans un environnement réel, par exemple en touchant avec la main un cube virtuel.

Il est possible d'interagir avec un objet mixte en utilisant sa partie réelle comme une interface tangible, ce que permet DOMIM grâce à la synchronisation entre la partie virtuelle et la partie réelle d'une entité mixte. Si l'intérêt n'est pas évident dans le cas d'un ventilateur ou d'une lampe comparé à l'usage de leur interrupteur, il permet pour des objets de plus grande dimension de rendre plus accessibles les commandes de l'objet. Par exemple, une machine industrielle a ses boutons d'arrêt d'urgence situés en des points précis. Actuellement, une commande mobile peut être utilisée afin de déplacer ces boutons, mais nécessite de porter la commande. Un bouton virtuel peut être affiché n'importe où à proximité ou à la surface de la machine et donc être plus rapidement accessible tout en permettant à l'utilisateur de la machine d'avoir les mains libres. Pour cela, l'application Hololens peut déplacer ces commandes virtuelles au plus près de l'utilisateur, et donc de gagner en efficacité et en rapidité d'interaction. Le dispositif de ce cas d'usage est similaire au cas d'usage 5.1 précédent. Il n'en diffère que par la technique d'interaction utilisée. Dans le cas présent, la technique consiste à détecter le toucher du bouton virtuel par l'utilisateur. Le système est modélisé au format DOMIM dans la Figure 5.4. On peut constater que la modification du modèle du système ne nécessite que de changer le terme "action utilisateur" de *geste combiné au regard* à *toucher*, ce qui montre la souplesse et la généralité des modèles basés DOMIM.

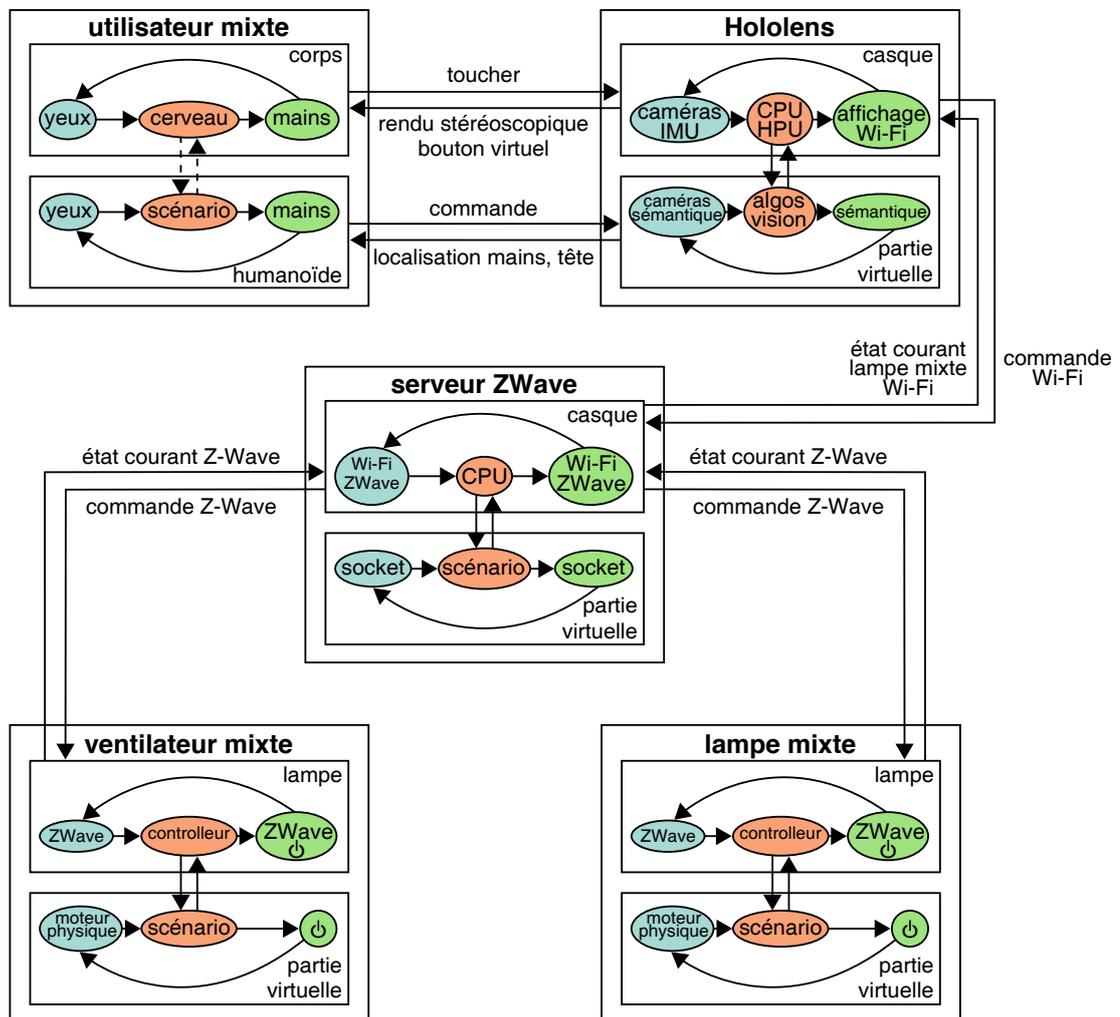


FIGURE 5.4 – Modèle basé DOMIM d'un système lampe et ventilateur mixte interactif utilisant l'objet de l'interaction comme interface tangible.

Notre framework de conception, de même que notre modèle DOMIM, sont suffisamment génériques et unifiés pour que la conception d'un système soit modifiable facilement et rapidement. Le développement de ce système bénéficie également de la souplesse de notre framework de développement. Pour être intégrée, elle ne nécessite que de modifier l'application Holens du système précédent. Cette technique est fournie par un composant pré-développé de notre framework de développement. Elle est intégrée au réseau de Petri, éditable sans recompilation du système, en remplaçant le capteur et l'effecteur du composant d'interaction gestuelle par ceux qui fournissent le composant d'interaction tangible. Le bouton virtuel étant rendu sur la surface de l'objet mixte, l'utilisateur touche alors cette surface, ce qui rend cette interaction tangible. Les outils utilisés sont identiques aux outils du cas d'usage précédent. Le système opérationnel est présenté par la Figure 5.5.

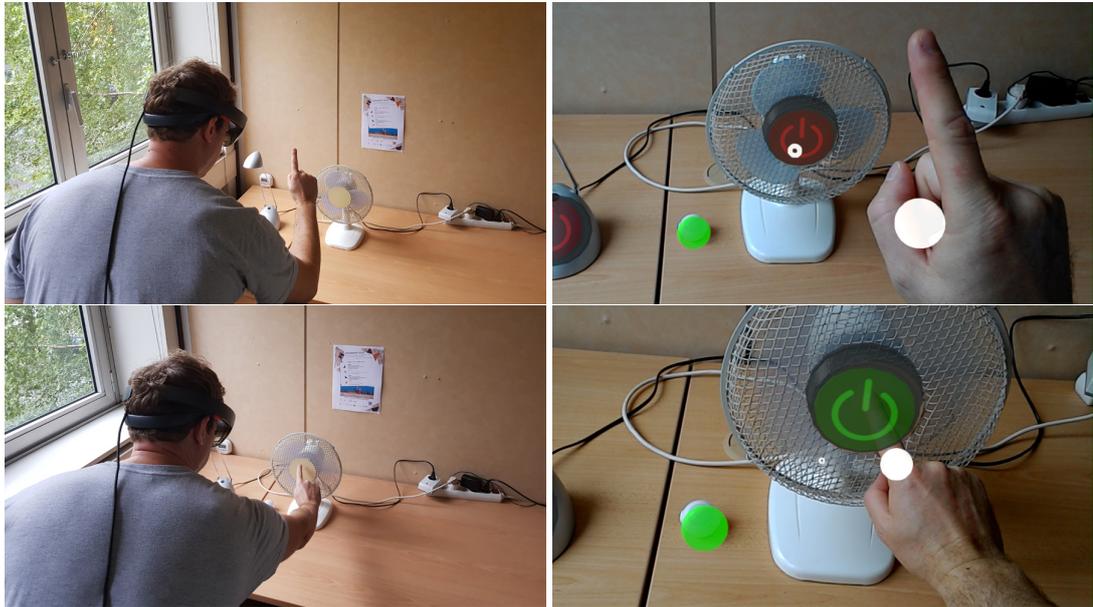


FIGURE 5.5 – Interaction tangible avec un ventilateur mixte. En haut, l'utilisateur s'apprête à toucher le bouton virtuel marche/arrêt du ventilateur mixte pour l'activer. En bas, l'utilisateur a activé le ventilateur.

Ce cas d'usage montre la capacité de notre modèle DOMIM à uniformiser et réutiliser simplement la modélisation des TIES. Il montre également comment les briques logicielles fournies par notre framework accélèrent et facilitent le remplacement de techniques d'interaction par d'autres techniques grâce à l'intégration d'un éditeur de réseau de Petri à l'environnement de développement.

5.3 Téléphone mixte, interactions tangibles et rendu distribué

Dans chaque cas d'usage précédent, nous n'avons employé qu'un seul périphérique de rendu visuel. Ce cas d'usage met en perspective la capacité de notre framework à créer des rendus distribués qui font appel à l'interconnexion des affichages d'un TIES ainsi qu'à leur colocalisation en faisant appel à des composants pré-intégrés dédiés. Ce cas d'usage met en œuvre une interaction tangible avec rendu distribué entre un objet mixte actif et l'utilisateur. Cet objet est un téléphone intelligent. En effet, il a la capacité de se localiser dans l'espace et de reconstruire son environnement. Il exécute des applications permettant d'interagir avec sa partie virtuelle. Il sait communiquer avec plusieurs technologies réseau. L'utilisateur peut également interagir pseudo-naturellement avec lui grâce au tactile, au tangible, à la vue, au son, à la voix ou encore aux vibrations. Ce cas d'usage montre que notre framework est capable de décrire précisément et de développer rapidement ce type de système innovant grâce aux composants qu'il fournit. Contrairement aux preuves de concept traitant de problématiques d'interaction équivalentes dans la littérature, il permet de produire ce type d'interaction grâce à des applications interopérables et extensibles.

Une interface utilisateur tangible permet de manipuler des propriétés virtuelles d'un objet mixte par la préhension et le mouvement. Nous décrivons dans cette section une technique d'interaction tangible permettant d'interagir intuitivement avec la partie virtuelle d'un téléphone mixte. Ce cas d'usage s'inscrit dans une logique similaire aux travaux de Normand et al. [NM18], ainsi que de Mohr et al [MTL⁺19]. Le premier problème est le champ de vision (FoV) restreint de l'écran

d'un téléphone intelligent. Le second problème est l'absence de rendu visuel stéréoscopique. Ces problèmes contraignent et limitent fortement la visualisation de la partie virtuelle du téléphone intelligent. Afin de compenser ces limitations, nous associons un casque de réalité mixte semi-transparent Hololens à un téléphone intelligent. La Figure 5.6 montre un utilisateur portant un Hololens interagir avec un téléphone intelligent. Dans cet exemple, l'Hololens permet d'afficher les contacts de l'utilisateur stockés dans le téléphone afin de bénéficier d'un champ de vision étendu ainsi que d'un rendu stéréoscopique de ce nuage de données.

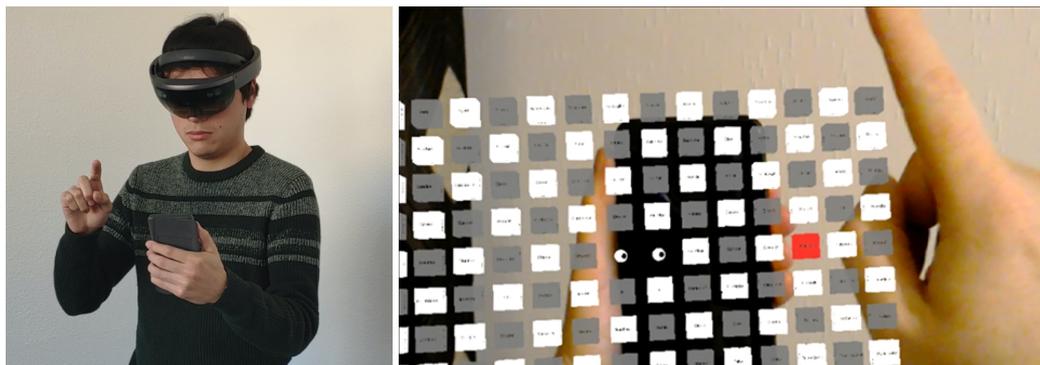


FIGURE 5.6 – Un utilisateur portant un Hololens interagit avec un téléphone intelligent. L'hololens vient enrichir les modalités d'interaction de l'utilisateur avec son téléphone intelligent en affichant les contacts de l'utilisateur stockés dans le téléphone sous la forme d'un nuage de données.

Ces deux objets connectés, qui sont en l'occurrence des interfaces utilisateur intuitives, ont la capacité de reconstruire leur environnement et de s'y localiser. Pour ce faire, l'Hololens utilise une implémentation d'un algorithme de type SLAM sur un DSP dédié appelé HPU (Holographic Processing Unit). Le téléphone intelligent est quant à lui une plateforme Android compatible ARCore, qui fournit également une implémentation d'un SLAM. Dans ce cas d'usage, ces deux objets connectés s'auto-localisent dans le même environnement, mais dans des référentiels distincts. Il est donc nécessaire de calibrer ce système interactif mixte afin qu'ils puissent interpréter leurs localisations réciproques. Connaître cette information leur permet de produire un rendu mutualisé et cohérent. De cette manière, Hololens peut afficher autour du téléphone intelligent certaines de ses propriétés virtuelles tout en utilisant les capacités d'interaction du téléphone, comme sa dalle tactile. Chaque interface communique à l'autre ses coordonnées en temps réel via le réseau. La Figure 5.7 présente le modèle de conception basé DOMIM de ce système. Elle décrit comment les rendus du téléphone et de l'Hololens se complètent. Elle indique également quels sont les composants nécessaires à la création de ce système. En outre, elle montre les interactions entre entités mixtes à réaliser pour produire le système attendu. Ce procédé a été breveté par Bataille et al. [BAG19a].

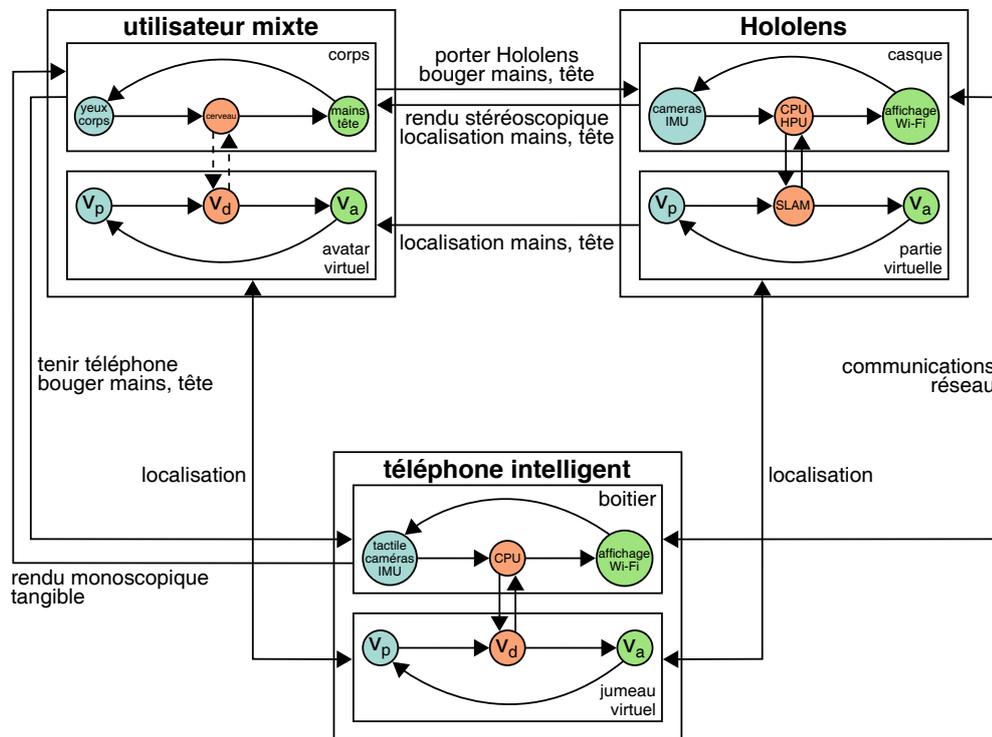


FIGURE 5.7 – Modèle basé DOMIM du système téléphone intelligent mixte.

L'implémentation du système nécessite le développement de deux applications, une pour Hololens et l'autre pour Android ARCore. Chaque application est développée sous Unity avec notre framework de développement à l'aide de nos squelettes d'application Hololens et Android ARCore. Utiliser ces squelettes d'application permet de gagner du temps lors du développement. Tout d'abord, car il évite la recherche de bibliothèques permettant d'interagir avec chacune des plateformes. Ensuite, car il fournit une implémentation générique du modèle DOMIM. Cette implémentation est fournie sous la forme d'un réseau de Petri que le développeur n'a plus qu'à décorer à l'aide de composants pré-développés. Les composants pré-développés utilisés par ces applications sont :

- l'implémentation générique d'une entité mixte DOMIM sous forme de réseau de Petri générique. Elle contient les éléments internes réels et virtuels d'une entité mixte ainsi que leurs interactions ;
- l'éditeur de réseau de Petri pour décorer à l'aide de scripts les nœuds transition du graphe générique ;
- les composants :
 - plateforme Hololens : co-localisation d'un Hololens et d'un mobile Android ARCore ;
 - socket TCP.

Le résultat final est présenté par la Figure 5.8. Cette technique d'interaction permet à l'Hololens de compléter le rendu de la partie virtuelle du téléphone intelligent et de bénéficier d'un champ de vision accru et permettant la stéréovision. Elle permet également de conserver l'aspect tangible de la manipulation du téléphone à la main pour pouvoir déplacer sa partie virtuelle, ce qui augmente le nombre de canaux sensori-moteurs impliqués comparativement à une interaction sans préhension avec uniquement la partie virtuelle du téléphone. En outre, les yeux virtuels du téléphone mixte fournissent à l'utilisateur un retour visuel de la détection et du suivi de la main de l'utilisateur, ce

qui améliore également la performance de l'interaction. Nous avons décrit dans le chapitre 3 comment cette augmentation du nombre de canaux sensori-moteurs impliqués dans une interaction pseudo-naturelle permet d'améliorer la performance et l'expérience de l'utilisateur.

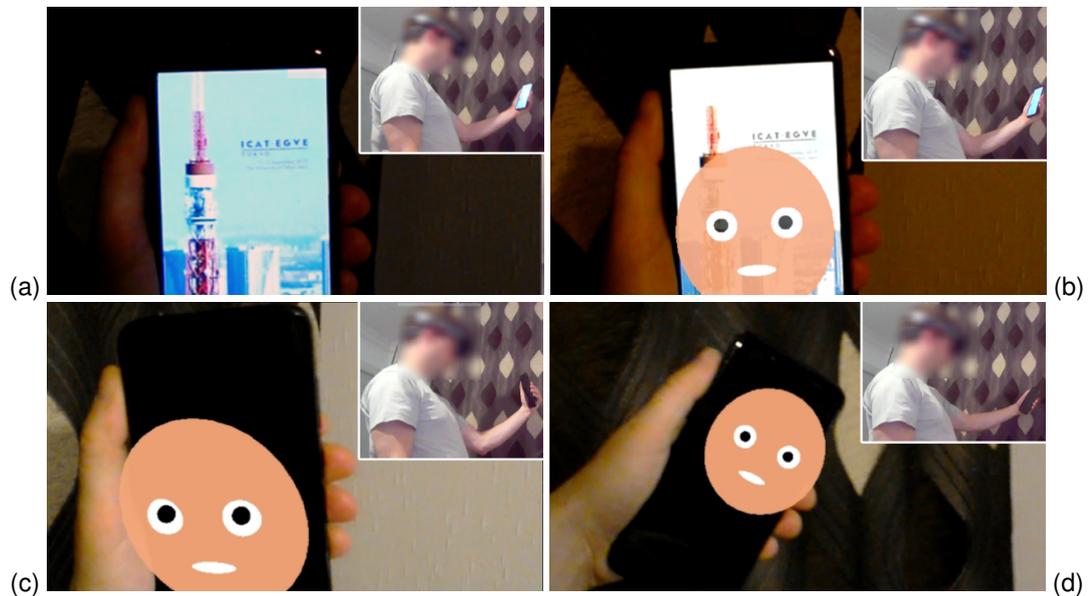


FIGURE 5.8 – Exemple de téléphone intelligent mixte : (a) l'utilisateur tient son téléphone à la main et l'observe avec un Hololens sur sa tête, (b) Vuforia sur Hololens détecte la texture affichée par le téléphone et estime la matrice de transfert des coordonnées du téléphone et de l'Hololens, (c) and (d) le suivi visuel du téléphone par Hololens est perdu, les communications réseau entre téléphone et Hololens permettent à chacun de connaître la position de l'autre dans un référentiel commun grâce à la matrice de transfert.

Par rapport aux cas d'usage précédent, ce cas d'usage met en œuvre deux applications basées DOMIM sur deux interfaces utilisateur qui ont la capacité de s'auto-localiser, de déterminer un référentiel spatial commun, et de distribuer le rendu de la partie virtuelle entre ces interfaces. L'intégration de cette distribution à un TIES est simplifiée et unifiée par notre framework.

5.4 Téléphone mixte et interactions hybrides mêlant interactions gestuelles, tangibles et tactiles

Dans la section précédente, nous avons décrit un système interactif mixte qui permet à l'utilisateur d'interagir par la main et le regard avec un téléphone intelligent. Un intérêt de DOMIM est de pouvoir modéliser et mettre en œuvre simplement plusieurs modalités d'interaction associées à un rendu distribué. Nous présentons dans cette section une extension de cette technique d'interaction en hybridant l'interaction tangible du système initial avec la gestuelle et le tactile, en associant plusieurs composants d'interaction pré-intégrés à notre framework de développement. Notre framework basé DOMIM permet de modéliser clairement cette interaction hybride à partir du modèle antérieur. Il permet également de factoriser la tâche de développement en réutilisant le système antérieur et en lui ajoutant grâce à des composants pré-développés des modalités d'interaction.

Les interfaces utilisateur naturelles offrent généralement des modalités d'interaction mono-sensorielles, par exemple interagir par le geste. Cela a pour conséquence de limiter la perception

que les utilisateurs ont de l'environnement virtuel avec lequel ils interagissent. Interagir avec la partie virtuelle d'entités bénéficie de l'usage d'un maximum de canaux sensori-moteurs, comme nous l'avons décrit dans la section 3.4, qui présente la notion d'interaction hybride. Les interactions hybrides combinent plusieurs interfaces ayant des modalités complémentaires ou non. Elles améliorent la performance de l'utilisateur dans son interaction avec le système. Comme dans l'exemple précédent, nous combinons un Hololens à ce téléphone afin de rendre plus intuitives ses interactions avec l'utilisateur. Le champ de vision stéréoscopique de l'Hololens vient compléter le rendu monoscopique de l'écran. L'Hololens affiche en stéréoscopie une représentation des contacts de l'utilisateur stockés sur son téléphone intelligent. Cette représentation spatiale est relative à la position du téléphone. Les applications Hololens et Android ARCore échangent leurs localisations respectives après l'estimation de la matrice de passage entre leurs repères de coordonnées respectifs de la façon décrite dans le paragraphe 4.5.2. L'interaction gestuelle permet de cibler un contact avec lequel interagir. Lorsque la main se rapproche de ce contact, les contacts se déplacent afin de placer le contact ciblé au centre de l'écran du téléphone. Ainsi, lorsque l'utilisateur touche l'écran tactile du téléphone, il touche ce contact. Le contact est alors sélectionné et les applications peuvent passer à un autre mode d'interaction spécifique à un contact, comme de déclencher une visioconférence avec un rendu stéréoscopique. Le fait de pouvoir manipuler les contacts en manipulant le téléphone permet d'interagir de façon tangible avec le nuage de contacts. Le procédé décrit a été breveté par Bataille et al. [BAG19b]. La Figure 5.9 présente le modèle DOMIM de ce cas d'usage. La Figure 5.10 montre un exemple d'interaction obtenu grâce à ce système.

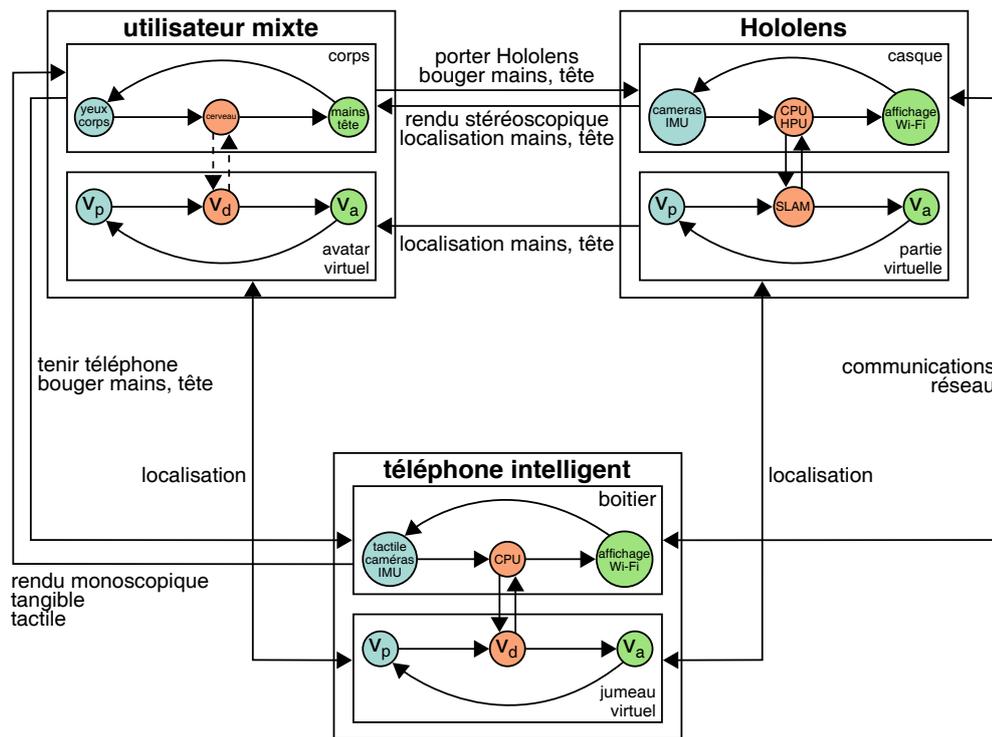


FIGURE 5.9 – Exemple de modèle basé DOMIM d'un système téléphone mixte multimodal. Rendu monoscopique et stéréoscopiques combinés, interactions gestuelles par Hololens, et tactiles et tangibles par le téléphone intelligent.

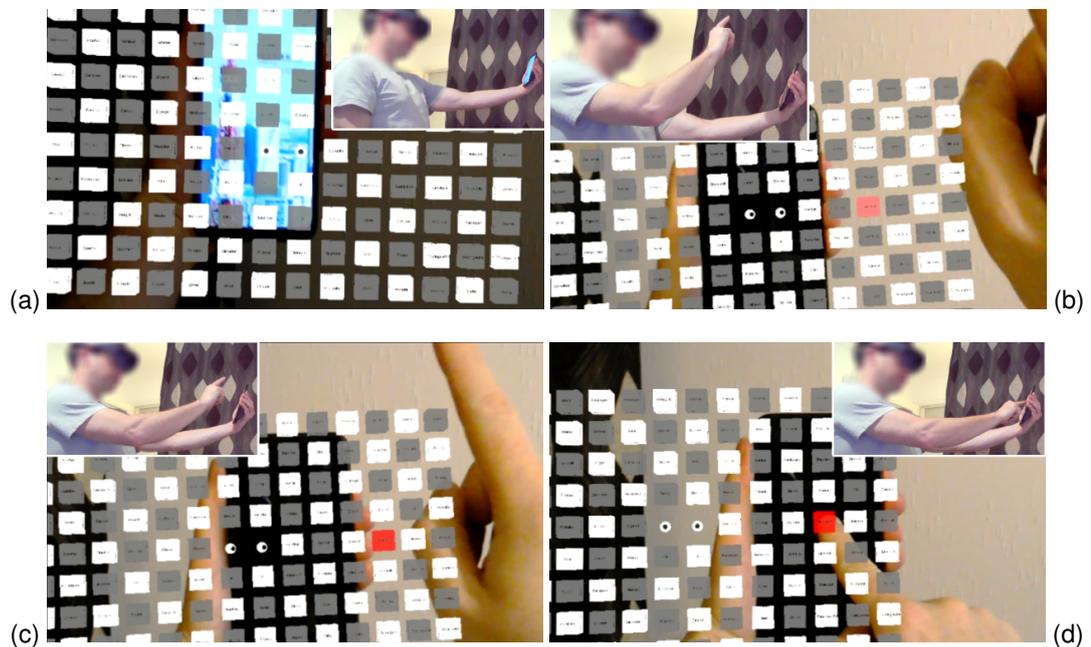


FIGURE 5.10 – Exemple d’interaction avec des contacts d’un utilisateur : chaque cube correspond à un contact de l’utilisateur et affiche son nom. Ces cubes peuvent être remplacés par la partie virtuelle du contact. (a) l’utilisateur observe le nuage de ses contacts rendu par l’Hololens et manipulé de façon tangible en déplaçant le téléphone, (b) l’utilisateur sélectionne un contact en approchant sa main du contact ciblé, (c) le contact sélectionné converge vers le centre de l’écran tactile du téléphone, (d) l’interaction entre l’utilisateur et contact sélectionné devient tactile.

Ce cas d’usage a permis d’explorer les possibilités d’interactions multimodales à rendu distribué offertes par la combinaison d’un téléphone mixte et d’un Hololens grâce à notre framework. Il montre la capacité à associer par le réseau différentes modalités d’interaction, à les synchroniser dans l’espace et à distribuer leur rendu grâce aux composants logiciels fournis par notre framework et à leur simplicité d’intégration. Cette exploration a bénéficié d’une vision claire du système. Cette vision nous a permis de détecter des modalités d’interaction disponibles qui n’étaient pas exploitées et d’étudier comment et pourquoi les composer. Elle nous a donc permis de concevoir une technique hybride d’interaction innovante tout en offrant un gain de temps à la réalisation de ce système. Par ailleurs, l’association de plusieurs composants du framework a permis d’améliorer l’extensibilité de chacun de ces composants. Il a également permis de détecter des limitations actuelles du suivi de la main par un Hololens à proximité d’une surface. En effet, lorsque la main est à quelques centimètres de l’écran, l’Hololens ne détecte plus la main de l’utilisateur et l’écran tactile prend alors le relais de l’interaction.

5.5 Interactions hybrides avec un bâtiment connecté

Ce cas d’usage montre la capacité de notre framework DOMIM à gérer des TIES plus complexes que les cas d’usage précédents ainsi qu’à les rendre interopérables, voir à fusionner plusieurs TIES faisant appel à la RM, aux NUI et à l’IoT. C’est pourquoi nous avons choisi un bâtiment connecté comme environnement. En effet, un bâtiment connecté contient de multiples objets connectés avec lesquels interagir. Interagir avec une liste d’objets connectés, ce qui est le cas de la plupart des systèmes de domotique utilisés, est peu naturel et intuitif, lent et n’indique souvent pas la localisation de ces objets. Afin de permettre au gestionnaire d’un bâtiment d’interagir intuitivement

tivement avec le bâtiment connecté, nous associons les cas d'usage 5.2 et 5.4 et remplaçons le nuage de contact de l'utilisateur par le jumeau virtuel d'une pièce d'un bâtiment mixte. Dans cette pièce, se trouvent notamment un ventilateur et une lampe mixtes. Nous utilisons l'Hololens pour afficher en stéréoscopie, sur la pose du téléphone, le jumeau virtuel de la pièce. Lorsque l'utilisateur approche sa main du téléphone mobile, il sélectionne l'objet mixte le plus proche de sa main. Lorsqu'il touche l'écran du téléphone, il active/désactive l'objet mixte commandé. Notre framework de conception permet de réutiliser l'ensemble des travaux de conception présentés précédemment dans ce chapitre en les regroupant au sein d'une même architecture. Cela démontre sa capacité à factoriser, unifier, interconnecter et étendre ces systèmes, grâce à l'usage du modèle DOMIM permettant de décrire les relations réel/virtuel tant internes qu'externes qui animent les entités mixtes. Ce cas d'usage a permis de s'assurer de la cohésion de notre framework, tant d'un point de vue conception que de son développement. En effet, il a permis de fusionner plusieurs prototypes en un seul et donc de s'assurer de l'interopérabilité des composants logiciels fournis par notre framework. La Figure 5.11 présente le modèle DOMIM associé à ce bâtiment. La Figure 5.12 montre l'utilisateur interagissant avec une pièce du bâtiment mixte.

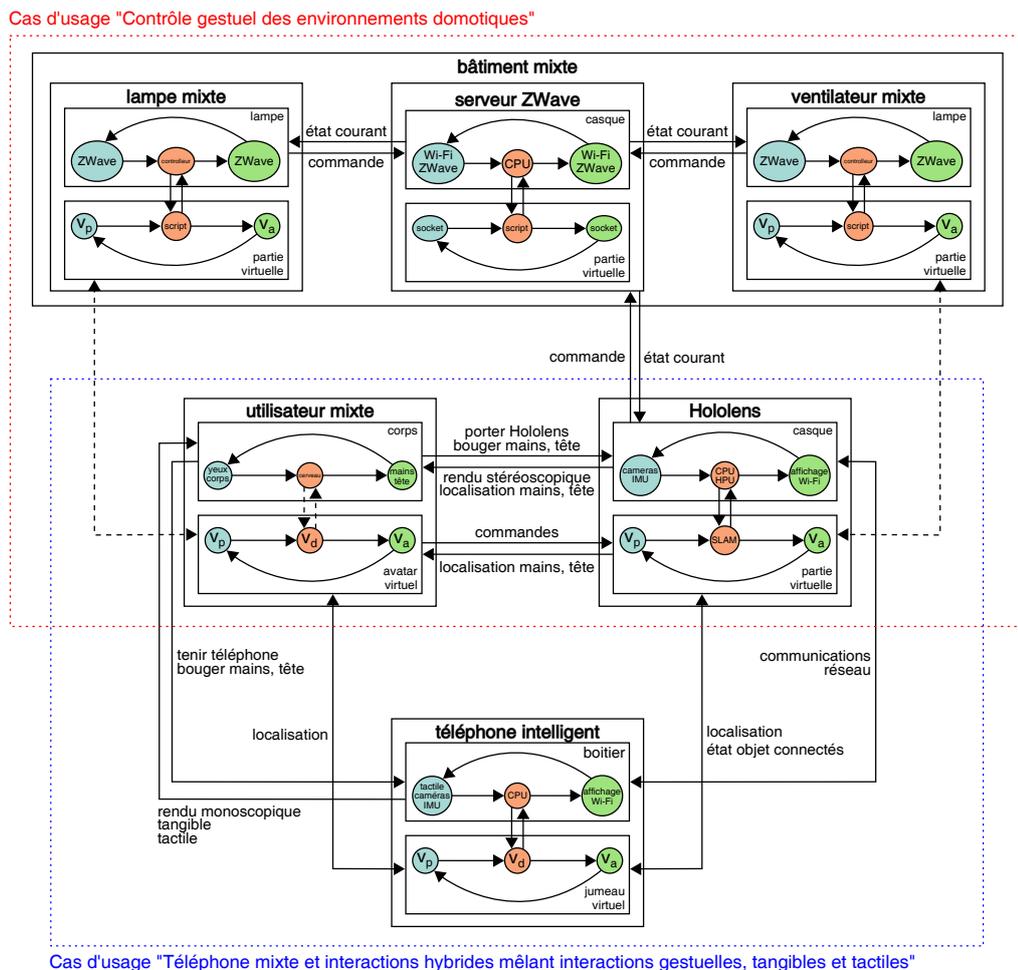


FIGURE 5.11 – Exemple de modèle basé DOMIM d'un bâtiment mixte contrôlé par une interaction hybride. Il combine les cas d'usage des sections 5.1 et 5.4.

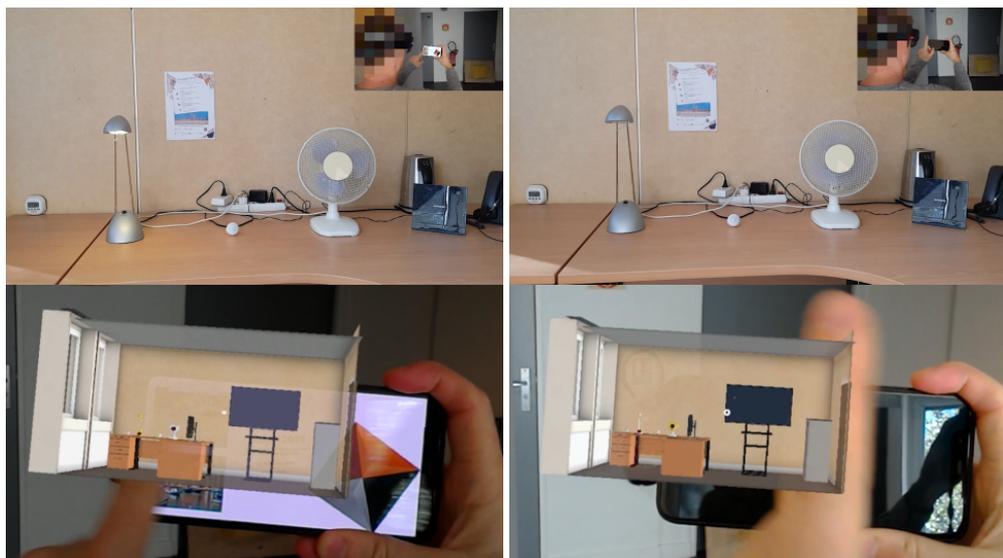


FIGURE 5.12 – Exemple d'interaction avec un bâtiment connecté. En haut, l'utilisateur observe le jumeau virtuel d'une pièce mixte affiché par l'Hololens sur la pose de l'écran du téléphone. Cette pièce contient notamment un ventilateur mixte et une lampe mixte. En bas à gauche, l'utilisateur touche avec le doigt l'interrupteur virtuel de la lampe mixte, ce qui l'allume. En bas à droite, l'utilisateur touche avec le doigt l'interrupteur virtuel du ventilateur mixte, ce qui le met en état de marche.

5.6 Retour multisensoriel de la localisation des mains d'un utilisateur

Ce cas d'usage montre la capacité de notre framework DOMIM à créer des retours d'information multimodaux sur la connaissance qu'un TIES a de son environnement. Il combine un rendu multimodal visuel et sonore de la position de la main de l'utilisateur à sa détection et à son suivi. Ce rendu multimodal permet de contourner les défauts de chaque modalité de rendu tout en associant

leurs affordances perçues.

Les systèmes interactifs mixtes que nous étudions font appel à des interfaces utilisateur naturelles. Elles permettent à un utilisateur d'interagir avec un environnement virtuel, et par extension avec un environnement mixte. Cette interaction nécessite pour être efficace que l'utilisateur connaisse les limites de cette interface utilisateur. Si l'utilisateur ne sait pas quand l'interface utilisateur le perçoit, il ne sait pas quand sont prises en compte ses tentatives d'interaction, ni si elles se déroulent correctement. Il en résulte des interactions non-intentionnelles qui perturbent le fonctionnement des systèmes interactifs mixtes. Dans cette section, nous décrivons un cas d'interaction entre un utilisateur et l'Hololens et les problèmes qu'il soulève. Nous proposons une solution basée sur une interaction multi-sensorielle. Cette solution est incluse à notre framework de développement DOMIM sous la forme d'un composant logiciel. Ce composant logiciel peut être intégré au scénario comportemental d'une entité mixte grâce à notre éditeur de scénario intégré à l'EDI Unity. Elle permet d'améliorer le retour sensoriel des interfaces utilisateur naturelles qui détectent la position des mains des utilisateurs en ajoutant le canal auditif au canal visuel.

Dans le cas d'un Hololens, l'interaction gestuelle n'est possible que lorsque la main de l'utilisateur se trouve dans le cône de vision de la caméra de profondeur de l'Hololens située entre ses optiques. Lors d'une interaction gestuelle directe, le champ de vision réduit de l'Hololens ne permet pas d'afficher les limites du cône de détection de la main de l'Hololens. L'utilisateur ne peut pas connaître précisément avec les techniques d'interactions usuelles de l'Hololens les limites perceptuelles de l'Hololens. Ce problème génère des erreurs lors d'une interaction gestuelle entre un utilisateur et l'Hololens. En effet, il arrive alors fréquemment que la main de l'utilisateur sorte du champ de vision de l'Hololens alors que la détection et le suivi de la main de l'utilisateur capturent encore les gestes de sa main. L'utilisateur pense que sa main n'est plus détectée alors qu'elle l'est encore, ce qui peut déclencher des interactions non souhaitées par l'utilisateur. Par exemple, le geste de fermeture de l'application peut être perçu par l'Hololens alors que l'utilisateur pense que sa main n'est plus suivie. Afin de résoudre ce problème, nous décrivons maintenant une modalité d'interaction multi-sensorielle entre l'utilisateur et l'Hololens qui permet de contourner ces limitations. Cette modalité d'interaction combine gestuelle et audio spatialisé.

Nous utilisons le son spatialisé afin de savoir si la main de l'utilisateur est détectée par l'Hololens et où elle est localisée, en complément de l'affichage de sa position par l'Hololens. Le son spatialisé émit par le jumeau virtuel de la main fournit quant à lui une information beaucoup plus précise et intuitive sur cette limite dans le suivi de la main tout en réduisant la quantité d'informations visuelles affichées par l'Hololens. L'ouïe est un canal perceptuel très performant tant en terme de précision de la localisation qu'en terme de rapidité de la perception. Ce mécanisme est représenté dans la Figure 5.13. Un brevet sur ce procédé de retour sensoriel multiple a été déposé en 2019 [Bat19].

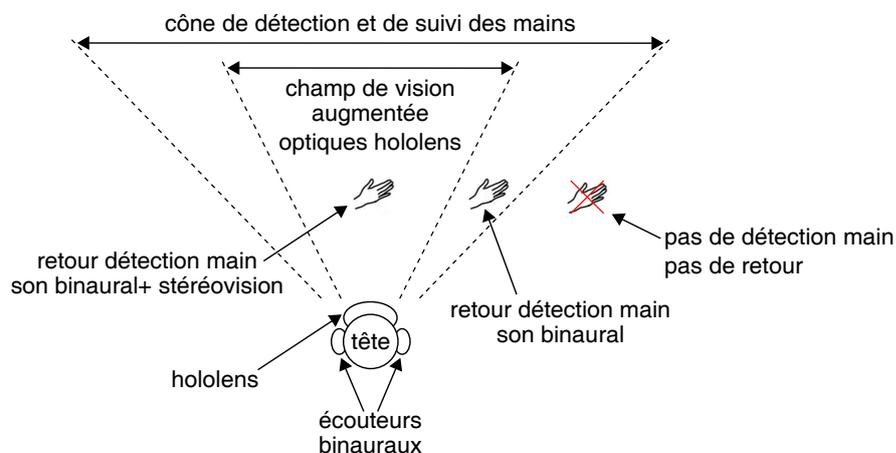


FIGURE 5.13 – Exemple d'interaction multimodale : retour visuel et sonore ait à l'utilisateur par un Hololens de la détection et du suivi de sa main.

DOMIM permet de modéliser et de développer simplement des systèmes interactifs multimodaux. En effet, durant la conception cela revient à augmenter le nombre de canaux sensori-moteurs utilisés par une interaction pseudo-naturelle. Cette augmentation du nombre de canaux sensori-moteurs s'effectue en ajoutant des interfaces utilisateur naturelles à l'architecture du système. Les applications permettant leur synchronisation sont facilement identifiables, car ce sont les applications des entités mixtes qui sont reliées par ces interfaces utilisateur naturelles. Elles peuvent être regroupées en une seule application qui instancie l'ensemble des parties virtuelles des entités mixtes reliées, dont celle de l'utilisateur. Le lien avec l'internet des objets est évident dans la mesure où les interfaces utilisateur naturelles impliquées sont potentiellement connectées à ce réseau. Dans la mesure où l'interaction avec l'internet des objets requiert l'utilisation d'interfaces utilisateur, cet exemple montre comment cette interaction peut être rendue plus naturelle et intuitive.

La Figure 5.15 modélise ce système interactif mixte en se basant sur DOMIM.

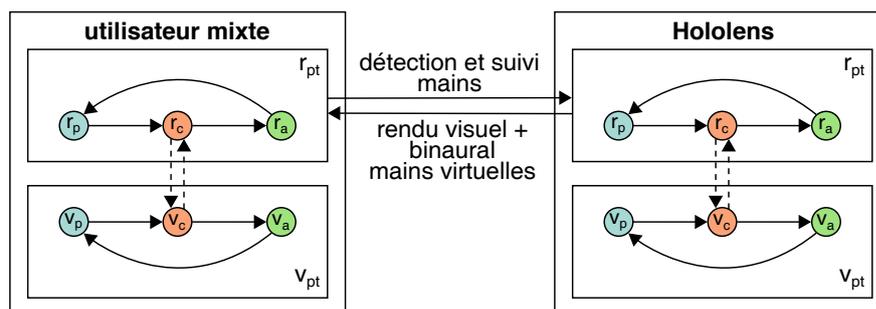


FIGURE 5.14 – Interactions multiples dans un système interactif mixte.

Une application Hololens basée DOMIM baptisée HoloHands a été développée pour implémenter ce système. Elle a permis de développer le composant de notre framework pour la plateforme Hololens de détection et de suivi de la main permettant des interactions directes entre la main de l'utilisateur et son environnement virtuel. Ce composant a été réutilisé dans tous les prototypes proposés dans ce chapitre. La Figure 5.15 montre ce système interactif mixte en action. Lorsqu'une main se trouve dans le champ de vision de l'HoloLens, la position de la main est indiquée par une sphère, tandis qu'un son spatialisé est rendu par le système audio de l'HoloLens à l'emplacement de la sphère. Lorsque cette même main sort du champ de vision de l'HoloLens mais est toujours

détectée et localisée par l'Hololens, la sphère n'est plus visible, mais le son spatialisé est toujours joué. Lorsque la main n'est plus détectée par l'Hololens, le son n'est plus joué. L'utilisateur peut ainsi évaluer intuitivement et naturellement l'espace de détection de la main par l'Hololens. Cet espace est beaucoup plus important que le champ de vision de l'Hololens, comme nous pouvons le voir sur la Figure 5.15.

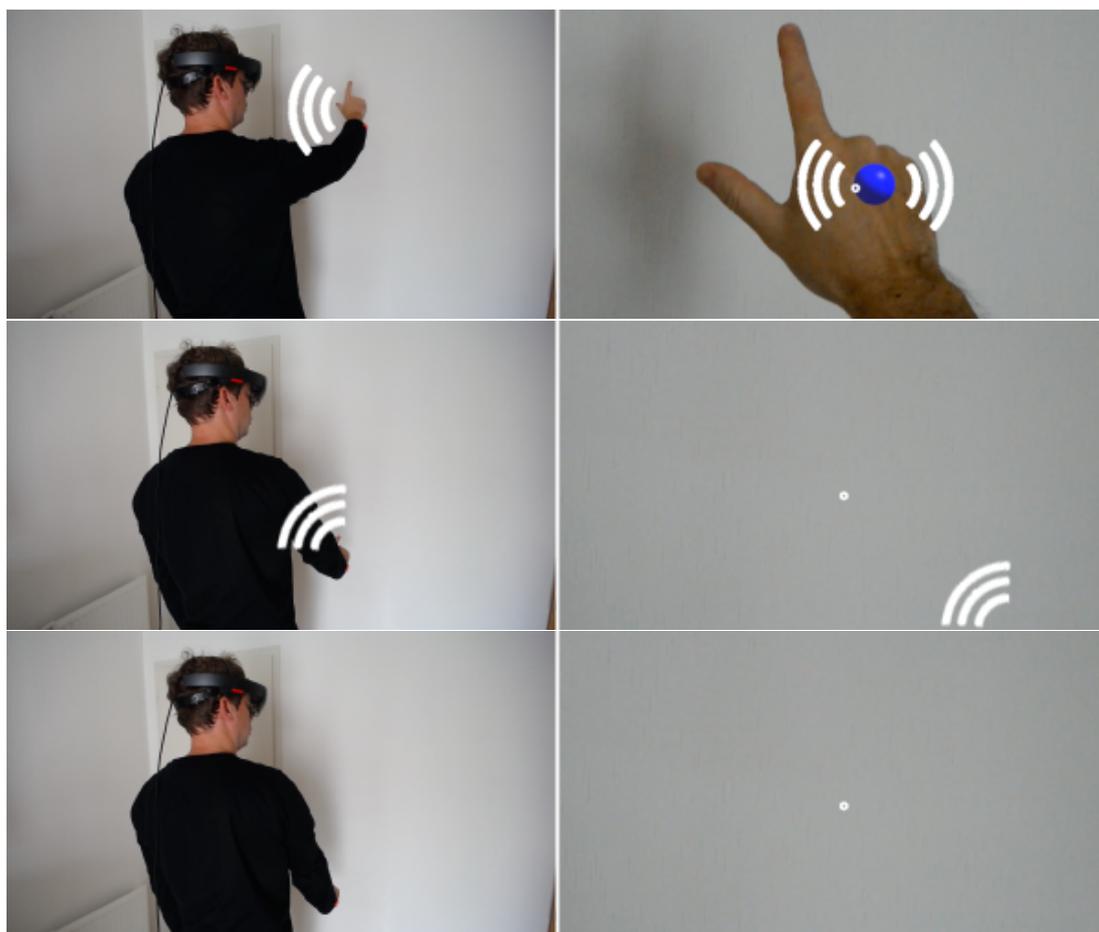


FIGURE 5.15 – L'application HoloHands fournit un retour multimodal stéréoscopique et binaural de la localisation d'une main détectée par l'Hololens. En haut, la main est détectée et sa position estimée est visible et audible. Au milieu, la main sort du champ de vision de l'Hololens mais est toujours perçue par lui, le son spatialisé est toujours joué, l'utilisateur est informé de cette détection. En bas, la main n'est plus détectée par l'Hololens, le son spatialisé n'est plus joué par l'Hololens.

En fournissant un composant multimodal de retour visuel et sonore de la localisation d'une entité mixte dans l'espace, notre framework améliore la naturalité de l'interaction entre l'utilisateur et le virtuel. Cette amélioration provient de l'association des affordances perçues par plusieurs canaux sensoriels.

5.7 Rendu stéréoscopique multicouche d'un contenu virtuel

DOMIM permet de concevoir des TIES bénéficiant d'affichages visuels multiples grâce à leur colocalisation rendue possible par leur communication réseau. Notre framework fournit des com-

posants logiciels qui permettent de combiner plusieurs rendus stéréoscopiques dans le champ de vision de l'utilisateur de façon cohérente.

Nous décrivons dans cette section comment un rendu multi-couches d'un environnement virtuel permet de contourner les inconvénients des périphériques en charge du rendu des différentes couches en associant leur qualités et en atténuant leurs défauts. Par exemple, un écran statique stéréoscopique ne peut pas afficher un objet virtuel occulté par le réel, par exemple par une main ou une boîte. La Figure 5.16 représente cette situation. Or, cet objet virtuel peut être localisé entre les yeux de l'utilisateur dirigés vers l'écran et l'objet qui occulte l'écran. Cela perturbe alors la stéréopsis de l'utilisateur et génère de la cynétose. Nous présentons dans le paragraphe suivant une solution à ce problème en combinant cet écran à un casque de réalité mixte semi-transparent stéréoscopique.

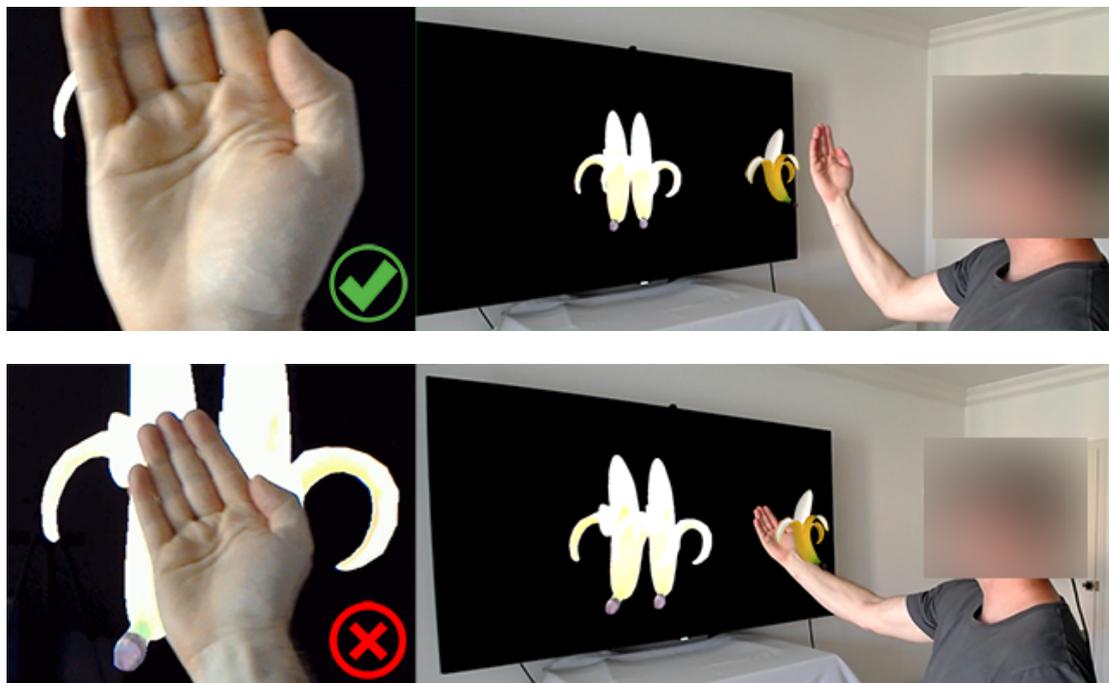
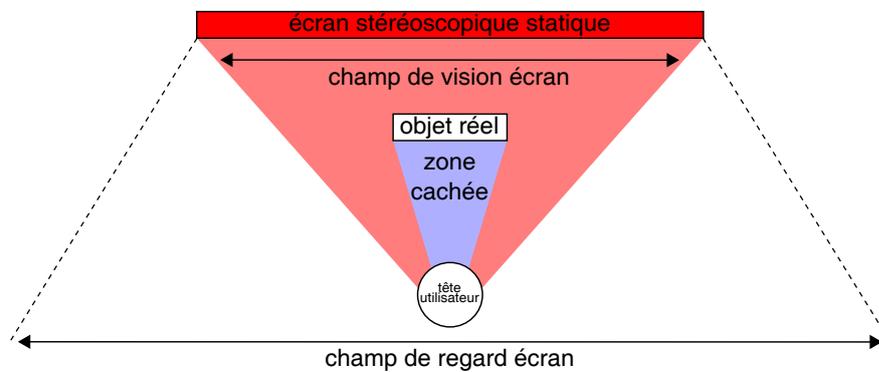


FIGURE 5.16 – Ecran statique stéréoscopique et occultations. Ce type d'écran ne permet pas de gérer les occultations de l'écran par son environnement.

Afin de résoudre le problème de l'occultation des écrans stéréoscopiques statiques, présents dans de nombreux foyers, mais également utilisés par les salles immersives de type CAVE [CNSD⁺92], nous lui associons un casque de réalité mixte semi-transparent à rendu stéréoscopique, comme l'Hololens. En effet, l'Hololens permet d'afficher systématiquement un environnement virtuel malgré les occultations du réel. Mais lorsque la zone qui devrait être masquée par le réel est affichée, cela produit une incohérence de la perception visuelle de l'utilisateur. Sa stéréopsis est alors perturbée.

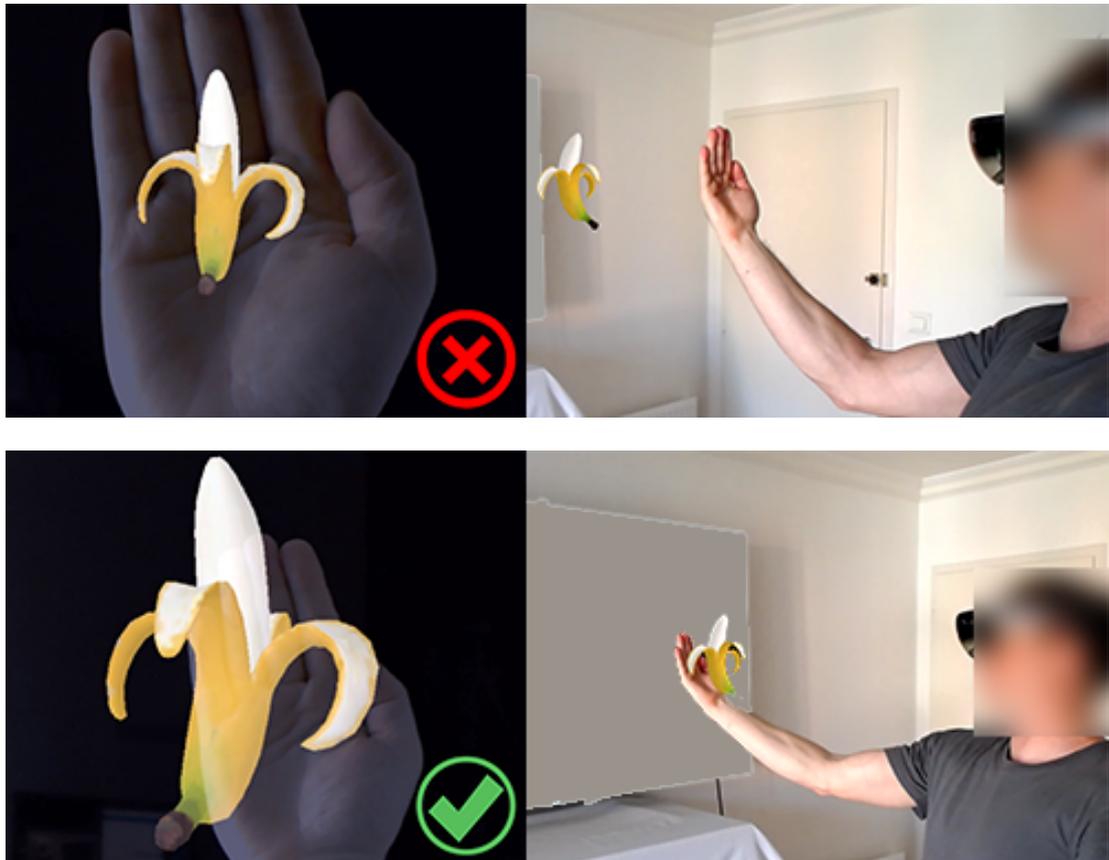
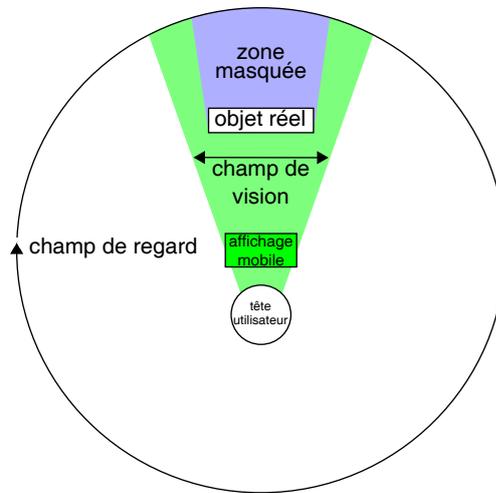


FIGURE 5.17 – Ecran stéréoscopique mobile. La zone occultée est par défaut systématiquement affichée sauf simulation des occultations réelles dans l’environnement virtuel.

Nishimoto et al. [NJ19] proposent de combiner la stéréoscopie d’un casque Hololens à une salle immersive afin d’augmenter le champ de regard de la salle immersive. A partir d’un dispositif aux propriétés similaires, nous avons réalisé un TIES permettant la gestion des occultations. Ce TIES a la capacité de mêler une stéréovision au plus près des yeux de l’utilisateur à une stéréovision fournie par un écran placé au fond d’une pièce, et de les combiner pour imbriquer occultations

virtuelles et réelles de façon cohérente. Ce cas d'usage est pertinent dans le contexte des salles immersives afin de limiter les ruptures dans la stéréopsis des utilisateurs. L'utilisateur peut alors percevoir ces deux environnements virtuels simultanément. Il est nécessaire de les colocaliser afin que l'utilisateur puisse voir en même temps 2 rendus stéréoscopiques cohérents d'environnements virtuels alignés. Lorsque l'utilisateur est situé devant l'écran statique, il perçoit :

- un environnement virtuel qui n'est jamais masqué par un objet réel, car son affichage est porté par l'utilisateur ;
- un environnement virtuel systématiquement masqué par un objet réel localisé entre l'utilisateur et l'écran statique, car son affichage est contre le mur.

Le modèle basé DOMIM de ce système est présenté par la Figure 5.18. Il montre que le système final utilise l'Holens pour capturer la position et les interactions de l'utilisateur, tandis que l'écran stéréoscopique statique réplique l'environnement rendu par Holens en tenant compte du point de vue utilisateur.

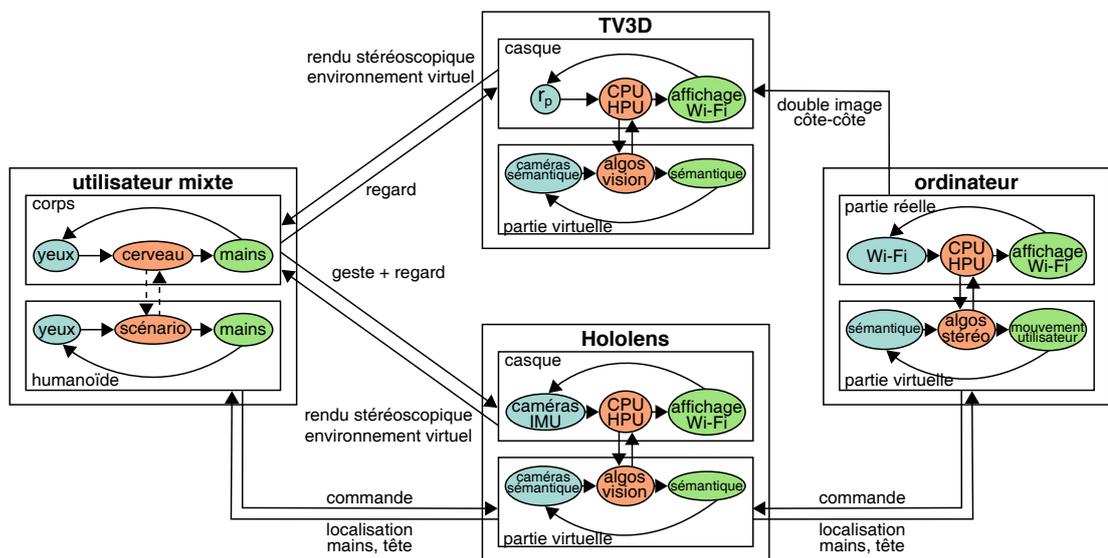


FIGURE 5.18 – Rendu stéréoscopique multi-couches : un système permettant de multiples rendus stéréoscopiques permet de percevoir un environnement virtuel cohérent malgré les occultations de l'écran statique, ainsi que de combiner les bénéfices de l'écran stéréoscopique statique et de l'écran stéréoscopique mobile.

Ce système permet de bénéficier du champ de vision du téléviseur, plus large que celui du casque de réalité mixte, tout en conservant la capacité d'afficher un objet virtuel même lorsque le téléviseur est masqué par un objet ou une partie du corps de l'utilisateur. La Figure 5.19 décrit le principe de ce système stéréoscopique multi-couches.

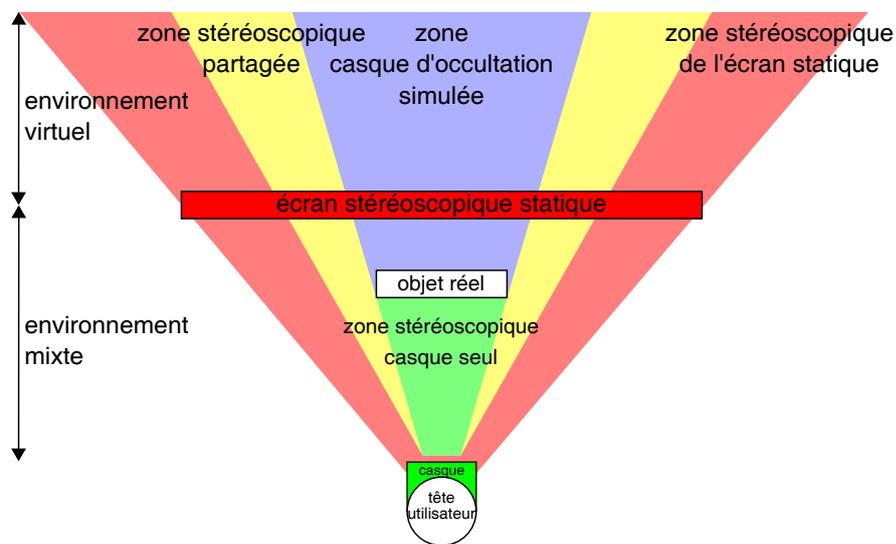


FIGURE 5.19 – Rendu stéréoscopique multi-couches : un système permettant de multiples rendus stéréoscopiques permet de percevoir un environnement virtuel cohérent malgré les occultations de l'écran statique, ainsi que de combiner les bénéfices de l'écran stéréoscopique statique et de l'écran stéréoscopique mobile.

Par exemple, dans le cas d'une salle immersive ou d'un CAVE, si un objet réel se trouve entre l'utilisateur et les murs de l'espace immersif, il masque systématiquement l'environnement virtuel affiché par le mur, y compris lorsqu'une partie de cet environnement virtuel se trouve entre l'utilisateur et l'objet. Par contre, si un casque de réalité mixte permet d'afficher cette partie de l'environnement virtuel situé entre l'objet réel et l'utilisateur alors il n'y a plus cette rupture de perception de l'environnement virtuel. Ainsi, l'utilisateur bénéficie de la perception d'un même environnement virtuel de plusieurs points de vue distincts.

Le développement du système consiste à reprendre le développement du système décrit dans la section 5.6. Une seconde application Windows est nécessaire afin d'effectuer le rendu stéréoscopique de l'écran statique. Une phase initiale au bon fonctionnement du système consiste à positionner l'écran dans l'espace de l'Hololens. Pour cela, nous avons utilisé le composant de colocalisation d'un Hololens et d'un écran statique décrit dans la section 4.5.2. Ensuite, l'Hololens transmet à l'ordinateur la position de la tête de l'utilisateur relative au centre de l'écran. L'Hololens a également la tâche de simuler l'occultation de l'objet virtuel par la main de l'utilisateur. L'application Windows développée grâce à notre framework de développement peut alors simuler le point de vue l'utilisateur sur son environnement virtuel et le transmettre à l'écran au format côte-côte nécessaire à la stéréoscopie. On ajoute alors au système un ordinateur Windows pour lequel on développe une application à l'aide de notre squelette d'application pour Windows. L'objet mixte implémenté est l'écran stéréoscopique. La partie virtuelle de l'écran mixte est en charge de calculer le rendu en fonction de la position des yeux de l'utilisateur. Notre composant Wi-Fi effectue cette transmission à partir de Hololens. La Figure 5.20 montre ce système interactif mixte. Les scripts logiciels produits pour l'implémentation de ce système sont inclus à notre framework.

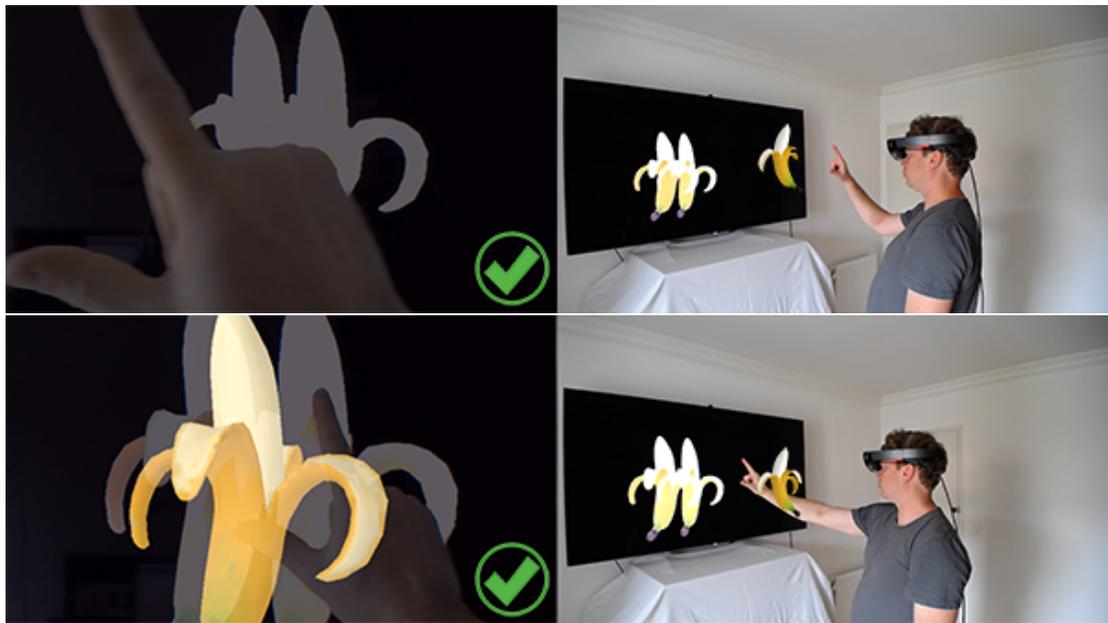


FIGURE 5.20 – Le système HoloTV en action : quelle que soit la position de la main de l'utilisateur elle n'interfère pas avec la perception de la banane virtuelle.

5.8 Conclusion

Nous avons dans ce chapitre présenté les cas d'usage qui nous ont permis durant la thèse de confronter le modèle DOMIM à des cas d'usage afin d'évaluer l'utilité de notre framework et de l'améliorer continuellement. Ces cas d'usage nous ont également permis de montrer dans ce chapitre que ces outils fonctionnent et ont des avantages. Notre framework conceptuel est suffisamment souple et générique pour s'adapter aux différents systèmes interactifs mixtes. Il offre un cadre conceptuel qui permet de réutiliser des travaux existants et de les adapter à de nouveaux systèmes. La conception basée DOMIM permet de réutiliser rapidement et simplement des modèles de systèmes interactifs mixtes combinant réalité mixte, interfaces utilisateur naturelles et internet des objets. Elle permet d'énumérer les entités mixtes constitutives de ce type de système, de caractériser leur comportement et leurs capacités d'interaction internes et externes. Le modèle unifié DOMIM rend les systèmes interopérables et extensibles. Il offre un cadre favorable à l'uniformisation de ces systèmes nécessaire à leur extensibilité. Par exemple, plusieurs systèmes permettant de contrôler un bâtiment peuvent être associés, permettant un passage à l'échelle à des systèmes interactifs mixtes de plus grande dimension. L'intégration de l'internet des objets à notre modèle rend possible un passage à l'échelle des systèmes produits grâce à notre framework à des systèmes de plus grande dimension. Ces dimensions sont liées aux capacités des structures de l'internet des objets, qui reste un paradigme très évolutif et en devenir. Egalement, le développement basé DOMIM s'appuie sur des composants, squelettes d'application et implémentation générique d'une entité mixte DOMIM créés pour réaliser ces cas d'usage. Ces composants permettent de réduire le temps de développement des systèmes basés DOMIM, car ils fournissent des implémentations qui leur sont nécessaires. Une implémentation générique du modèle DOMIM permet de fournir un cadre de développement unificateur aux développeurs. Ce caractère unifié des systèmes basés DOMIM les rend interopérables et par conséquent extensibles en un réseau de systèmes. Ces possibilités d'extensibilité des systèmes basés DOMIM sont directement liés à l'évolution des technologies et architectures réseau fournies par le domaine de l'internet des objets.

Chapitre 6

Conclusion et travaux futurs

Les travaux présentés portent sur la conception et le développement des systèmes interactifs mixtes combinant réalité mixte, interfaces utilisateur naturelles et internet des objets. Ces systèmes intègrent chaque jour un peu plus à notre environnement le numérique. L'objectif de notre recherche est de fournir :

- un modèle unifié de la réalité mixte, des interfaces utilisateur naturelles et de l'internet des objets ;
- des outils de conception et de développement basés sur ce modèle afin de les rendre plus homogènes, interopérables et extensibles ;
- d'explorer de nouvelles techniques d'interaction favorisant la naturalité et l'intuitivité de l'interaction entre les utilisateurs et ces systèmes.

Dans ce chapitre, premièrement, nous listons nos contributions à la conception, au développement et à l'exploration des systèmes interactifs mixtes combinant réalité virtuelle, interfaces utilisateur naturelles et internet des objets.

6.1 Contributions

Cette thèse a donné lieu à des contributions à la conception et à la mise en œuvre des TIES que nous décrivons dans cette section. Dans un premier temps, nous décrivons le modèle DOMIM unifiant réalité mixte, interfaces utilisateur naturelles et internet des objets. Ce modèle permet une approche homogène des TIES en les décomposant en entités combinant réel et virtuel. Dans un second temps, nous résumons notre framework de conception et de développement qui s'appuie sur notre modèle DOMIM afin de produire des TIES interopérables et extensibles. Dans un troisième temps, nous résumons les différentes techniques d'interactions que nous avons exploré. Ces techniques mettent en œuvre différentes modalités d'interaction et combinent des interfaces utilisateur naturelles afin de rendre plus naturelles les interactions entre utilisateurs et environnements connectés.

Le modèle DOMIM

Le modèle DOMIM permet de décrire de façon unifiée les utilisateurs, interfaces utilisateur naturelles et objets connectés. Pour cela, il s'appuie sur la notion d'entité mixte, constituée d'une partie réelle et d'une partie virtuelle synchronisées. Les interactions internes à chacune de ces parties sont modélisées grâce à la boucle perception-décision-action. Le comportement de chaque partie est synchronisé avec l'autre, ce qui produit un contrôle mixte de ces entités à la fois réelles et virtuelles. Les interactions externes des entités mixtes avec les autres entités mixtes constitutives

des systèmes interactifs mixtes résultent de leurs comportements internes. Les TIES sont composés d'entités mixtes interagissant avec leurs environnements réels et virtuels. Combiner comportements réels et virtuels permet de faire appel aux qualités de ces environnements tout en atténuant les aspects négatifs de chacun.

Framework basé DOMIM

Notre framework s'appuie sur le modèle DOMIM qui a été conçu pour unifier les TIES.

Notre framework de conception permet de produire des architectures et modèles de systèmes homogènes, interopérables et extensibles. Il permet également de maîtriser la complexité de TIES pouvant réunir de nombreuses entités lorsqu'ils sont de grande dimension, par exemple des villes ou des usines connectées. La capacité de notre framework à segmenter ces systèmes en sous-systèmes homogènes à de multiples granularités est un aspect essentiel de cette maîtrise.

Notre framework de développement fournit des outils basés DOMIM afin de soutenir la mise en œuvre des TIES. Nos outils sont conçus pour réduire le coût de développement de ces systèmes en simplifiant, accélérant et homogénéisant leur mise en œuvre. Ils contribuent à la mise en œuvre des TIES en :

- fournissant des squelettes d'application par plateforme qui implémentent le modèle d'entité mixte de DOMIM ;
- décrivant le comportement des entités mixtes à l'aide d'un scénario sous la forme d'un réseau de Petri, sans recompilation systématique des applications d'un TIES ;
- fournissant des composants pré-développés dédiés à la réalité mixte, aux interfaces utilisateur naturelles et à l'internet des objets. Ces composants fournissent des solutions efficaces à des problèmes requérant des compétences poussées dans ces domaines.

Techniques d'interaction

Afin de rendre plus intuitives et naturelles les interactions entre l'utilisateur et les systèmes interactifs mixtes, nous avons exploré des interactions multi-modales combinant gestuelle, tangible et tactile. Les prototypes développés permettent d'interagir avec des objets mixtes comme un ventilateur ou une lampe par le geste, le toucher ou encore la manipulation d'une interface tangible. Ils combinent également les modalités expérimentées afin d'augmenter le nombre de canaux sensorimoteurs impliqués lors de l'interaction. Ils permettent d'interagir localement ou à distance avec ces objets mixtes.

Nous avons également exploré à l'aide de prototypes des interactions mono-modales à interfaces multiples produisant des rendus visuels multi-couches, dont la gestion des occultations entre virtuel et réel grâce à une double stéréoscopie. Ce type d'interaction est utile pour la gestion des occultations mutuelles entre réel et virtuel dans le cas notamment des murs d'affichage d'environnements virtuels. Elle contribue ainsi à réduire un peu plus le fossé séparant des environnements aux propriétés très différentes mais néanmoins complémentaires.

6.2 Limitations et recherches à court terme

Nous n'avons pas créé d'outil dédié de conception basée DOMIM, comme un éditeur de modèle de TIES permettant de simuler et de générer de façon centralisée les différentes applications constitutives du système. Il aurait pour utilité d'avoir une approche holistique de la conception d'un TIES basé DOMIM. Un éditeur couvrant l'ensemble du TIES permettrait de modéliser l'ensemble du système, de générer dans la foulée les applications par plateforme et de les déployer en pouvant contrôler et simuler leurs interactions.

Certains aspects de l'internet des objets n'ont pas été pour le moment intégrés au framework de développement. En particulier, des mécanismes de découverte automatique des objets connectés

au sein d'un réseau favoriseraient l'automatisation de leur interopérabilité. Ils seraient de plus utiles à la colocalisation des entités du systèmes, qui pourraient échanger leurs localisations respectives. Egalement, des mécanismes de caractérisation des objets connectés pourraient permettre de créer automatiquement des services combinant ubiquité et réalité mixte. De très nombreuses technologies réseau et protocoles liés sont également intégrables à notre framework. De même que de nombreuses techniques d'interaction pseudo-naturelles sont intégrables à notre framework, comme l'haptique, la reconnaissance et la synthèse vocale.

L'enrichissement de nos outils d'édition et de nos composants réutilisables à l'ensemble du spectre technologique couvert par les TIES est un axe possible à court terme, qui représente néanmoins un travail de production très lourd. Nous envisageons préférentiellement l'exploration de nouvelles techniques d'interaction. Pour cela, nous souhaitons approfondir les possibilités d'interactions multimodales offertes par nos travaux. Nous souhaitons également expérimenter des cas d'usage relatifs aux besoins de notre entreprise.

6.3 Recherches à moyen terme

A moyen terme, l'établissement d'un groupe de travail avec des concepteurs est envisageable dans la mesure où notre éco-système en entreprise est centré conception et ergonomie. En particulier, nous allons accompagner ces métiers dans l'intégration la réalité mixte dans leurs activités professionnelles. Nous avons la possibilité de faire évoluer notre framework basé DOMIM en le mettant à la disposition de concepteurs et de développeurs. Les différentes propriétés de DOMIM seront confrontées alors aux réalités du terrain ainsi qu'aux méthodologies existantes.

DOMIM offre également de multiples possibilités d'associations transverses avec des domaines connexes. Combiner DOMIM avec d'autres domaines de recherche comme la robotique, les interfaces cerveau-machine ou l'intelligence artificielle sont envisagés. En effet, le contrôle des entités mixtes se prête à l'introduction de l'intelligence artificielle dans nos travaux et ouvre la porte à une autonomie plus importante des entités mixtes, y compris les NUI. De même, la robotique permet d'imbriquer réel et virtuel grâce aux actuateurs embarqués par les robots, ce qui permet non pas de percevoir un virtuel rendu, mais de traduire directement dans un environnement réel un comportement acté dans son environnement virtuel associé. Ce type d'association est plus naturel, car il permet des interactions physiques avec des objets matérialisés. Egalement, les interfaces cerveau-machine sont un domaine intéressant à explorer dans le cadre de la synchronisation des processus décisionnels réels et virtuels et du comportement qui en découle. Ce domaine est utile à moyen terme dans le cadre du handicap, mais offre également des perspectives d'estimation de la qualité de l'expérience utilisateur ou de détection de la disponibilité mentale de l'utilisateur.

Un autre axe prometteur est la communication interpersonnelle en réalité mixte et ses multiples ramifications dont le contrôle distant d'objets connectés. Actuellement, les communications vocales sont la principale modalité de communication interpersonnelle distante, secondées par la visioconférence que la crise sanitaire COVID-19 a fortement stimulé. L'extension de ces modalités d'interaction à des environnements mixtes distants paraît logiquement être la prochaine étape dans l'évolution des technologies de communication interpersonnelle. De plus, si l'internet des objets connaît une expansion certaine, le concept d'internet des humains¹ vise à replacer l'humain au cœur de la technologie en tenant compte de ses besoins essentiels comme la sécurité ou encore la confiance. Ce concept est de plus compatible avec les travaux menés lors de cette thèse, en particulier lorsque l'on considère l'humain en tant qu'entité mixte.

1. <https://ec.europa.eu/digital-single-market/en/blogposts/internet-humans>

6.4 Recherches à long terme

Les perspectives à long terme sur les systèmes interactifs mixtes sont prometteuses et multiples. En effet, le développement de ces systèmes et leur intégration grandissante au sein de la société semblent s'accélérer. Des travaux de recherche transverses combinant multi-modalités, intelligence artificielle, systèmes cyber-physiques [GBWG19] et robotique semblent être la meilleure orientation pour l'exploitation optimale du modèle DOMIM. En effet, DOMIM est adapté à la conception et à la mise en œuvre d'interactions multimodales, inclue le concept de processus décisionnel et comportemental d'une entité virtuelle, à l'interaction simultanée avec les parties réelles et virtuelles d'un environnement mixte, et requièrent des capteurs et effecteurs réels pour la réalisation de tâches dans un environnement réel. Typiquement, rendre perceptible et interactive la partie virtuelle d'un utilisateur mixte. Ou encore, créer des objets mixtes comme un robot mixte capable d'assurer des communications interpersonnelles en réalité mixte. Une convergence toujours plus forte entre domaines laisse entrevoir une fusion entre différents domaines de recherche initialement assez distants, tout comme l'apparition de nouveaux domaines de recherche traitant des problématiques que posent les TIES.

Les limitations technologiques dont souffrent les TIES ne se résoudront pas sur le court et le moyen terme. En terme de NUI, la multimodalité ainsi que le rendu multicouches nécessitent l'amélioration des techniques existantes ainsi que leur miniaturisation. Ce qui permettrait d'intégrer un maximum de modalités complémentaires dans des interfaces utilisateur et des objets toujours plus petits et fondus à leur environnement. Nous sommes également attentifs à l'émergence de nouvelle NUI, comme l'holographie.

La prise en compte des problématiques environnementales actuelles semble en outre être un axe déterminant pour le succès à long terme des TIES et en particulier de l'IoT. La mise en place de mécanismes et méthodologies vérifiant la pertinence et l'utilité des TIES déployés, mais également la réduction de leur empreinte écologique sera probablement nécessaire pour que ces nouvelles technologies trouvent leur place dans notre société pour un coût écologique raisonné.

Un autre questionnement concerne la capacité des NUI et autres technologies augmentant l'humain comme les interfaces cerveau-machine à créer de nouvelles formes de communication entre humains plus performantes qu'une interaction purement naturelle. Ce qui implique d'attendre la maturation de ces technologies encore très jeunes, ainsi que de traiter des enjeux moraux et éthiques majeurs liés à leur usage.

Annexe A

Bibliographie de l’auteur

A.1 Publications

Articles longs, conférences internationales avec comité de lecture :

Guillaume Bataille, Valérie Gouranton, Jérémy Lacoche, Danielle Pelé, and Bruno Arnaldi. A Unified Design & Development Framework for Mixed Interactive Systems. In *VISIGRAPP, GRAPP, the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, pages 49–60, Valetta, Malta, 2020.

Guillaume Bataille, Valérie Gouranton, Jérémy Lacoche, and Bruno Arnaldi. Managing Mutual Occlusions between Real and Virtual Entities in Virtual Reality. To be published in *VISIGRAPP, GRAPP, the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 2021.

A.2 Brevets

- Guillaume Bataille. Enhanced feedback of a user hands detection in mixed reality with Holo-lens. Demande de brevet français FR1902242 déposée le 5 mars 2019. Demande de brevet international PCT/FR2020/050438 déposée le 5 mars 2020.
- Guillaume Bataille, Bruno Arnaldi, and Valérie Gouranton. Design-Oriented Mixed-Reality Internal Model (DOMIM). Demande de brevet français FR1902910 déposée le 21 mars 2019. Demande de brevet international PCT/EP2020/056209 déposée le 9 mars 2020.
- Guillaume Bataille, Bruno Arnaldi, and Valérie Gouranton. Virtual and tangible hybrid interactions. Demande de brevet français FR2001408 déposée le 13 février 2020. Demande de brevet international PCT/EP2020/057357 déposée le 17 mars 2020.
- Guillaume Bataille, Bruno Arnaldi, Valérie Gouranton, and Jérémy Lacoche. Multiple stereoscopic rendering for hybrid mixed reality. Demande de brevet français FR1914557 déposée le 17 décembre 2019.

Bibliographie

- [ADL10] Laurent Aguerreche, Thierry Duval, and Anatole Lécuyer. Reconfigurable Tangible Devices for 3D Virtual Object Manipulation by Single or Multiple Users. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology, VRST '10*, pages 227–230, New York, NY, USA, 2010. ACM.
- [AIM10] Luigi Atzori, Antonio Iera, and Giacomo Morabito. The Internet of Things : A Survey. *Computer Networks*, 54(15) :2787–2805, October 2010.
- [AIM17] Luigi Atzori, Antonio Iera, and Giacomo Morabito. Understanding the Internet of Things : definition, potentials, and societal role of a fast evolving paradigm. *Ad Hoc Networks*, 56 :122–140, March 2017.
- [ARE⁺17] G. Alce, M. Roszko, H. Edlund, S. Olsson, J. Svedberg, and M. Wallergård. [POSTER] AR as a User Interface for The Internet of Things #x2014;Comparing Three Interaction Models. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pages 81–86, October 2017.
- [Ash11] Kevin Ashton. That ‘internet of things’ thing. *RFID Journal*, 22(7), 2011.
- [Azu97] Ronald T. Azuma. A Survey of Augmented Reality. *Presence : Teleoperators and Virtual Environments*, 6(4) :355–385, 1997.
- [BAG19a] Guillaume Bataille, Bruno Arnaldi, and Valérie Gouranton. Design-Oriented Mixed-Reality Internal Model (DOMIM), 2019.
- [BAG19b] Guillaume Bataille, Bruno Arnaldi, and Valérie Gouranton. Virtual and tangible hybrid interactions, 2019.
- [Bat19] Guillaume Bataille. Enhanced feedback of a user hands detection in mixed reality with Hololens, 2019.
- [BCMP18] Elodie Bouzekri, Alexandre Canny, Célia Martinie, and Philippe Palanque. A Generic Software and Hardware Architecture for Hybrid Interactive Systems. In *EICS 2018 – Workshop on Heterogeneous Models and Modeling Approaches for Engineering of Interactive Systems*, Paris, France, June 2018.
- [BDBL⁺07] Stefano Baraldi, Alberto Del Bimbo, Lea Landucci, Nicola Torpei, Omar Cafini, Elisabetta Farella, Augusto Pieracci, and Luca Benini. Introducing Tangerine : A Tangible Interactive Natural Environment. In *Proceedings of the 15th ACM International Conference on Multimedia*, MM '07, pages 831–834, New York, NY, USA, 2007. ACM.
- [BDHM11] Andrew Bragdon, Rob DeLine, Ken Hinckley, and Meredith Ringel Morris. Code Space : Touch + Air Gesture Hybrid Interactions for Supporting Developer Meetings. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, ITS '11*, pages 212–221, New York, NY, USA, 2011. Association for Computing Machinery. event-place : Kobe, Japan.
- [BFA⁺17] Caroline Baillard, Matthieu Fradet, Vincent Alleaume, Pierrick Jouet, and Anthony Laurent. Multi-device mixed reality TV : a collaborative experience with joint use of

- a tablet and a headset. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology - VRST '17*, pages 1–2, Gothenburg, Sweden, 2017. ACM Press.
- [BH07] Alethea Blackler and Jörn Hurtienne. Towards a unified view of intuitive interaction : Definitions, models and tools across the world. *MMI-Interaktiv*, 13, 2007.
- [BH09] Galen V Bodenhausen and Kurt Hugenberg. Attention, perception, and social cognition. In Fritz Strack and Jens Förster, editors, *Social Cognition : The Basis of Human Interaction*, pages 1–22. Psychology Press, 2009.
- [Bho13] Achintya K. Bhowmik. Natural and Intuitive User Interfaces with Perceptual Computing Technologies. *Information Display*, 29(4) :6–10, July 2013.
- [Bla12] J Blake. The natural user interface revolution. In *Natural User Interfaces in. Net*, pages 1–43. Manning publications edition, 2012.
- [BMR12] Doug A. Bowman, Ryan P. McMahan, and Eric D. Ragan. Questioning Naturalism in 3D User Interfaces. *Communications of the ACM*, 55(9) :78–88, September 2012.
- [Bou07] Guillaume Bouyer. Rendu multimodal en Réalité Virtuelle : Supervision des interactions au service de la tâche. page 202, 2007.
- [BOZW15] Hrvoje Benko, Eyal Ofek, Feng Zheng, and Andrew D. Wilson. FoveAR : Combining an Optically See-Through Near-Eye Display with Projector-Based Spatial Augmented Reality. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*, pages 129–135, Daegu, Kyungpook, Republic of Korea, 2015. ACM Press.
- [CGA15] Guillaume Claude, Valérie Gouranton, and Bruno Arnaldi. Versatile Scenario Guidance for Collaborative Virtual Environments. In *Proceedings of 10th International Conference on Computer Graphics Theory and Applications (GRAPP'15)*, berlin, Germany, March 2015.
- [CLLT10] Linqin Cai, Huijuan Li, Zhiyong Luo, and Tang Xianlun. Modeling autonomous virtual agent in virtual environment for risk behavior. In *2010 2nd International Conference on Computer Engineering and Technology*, pages V4–451–V4–455, Chengdu, China, 2010. IEEE.
- [CN06] Céline Coutrix and Laurence Nigay. Mixed reality : a model of mixed interaction. In *Proceedings of the working conference on Advanced visual interfaces - AVI '06*, page 43, Venezia, Italy, 2006. ACM Press.
- [CNR05] Céline Coutrix, Laurence Nigay, and Philippe Renevier. Modèle d'interaction mixte : La Réalité Mixte à la lumière des modalités D'interaction. In *Proceedings of the 2nd French-speaking Conference on Mobility and Ubiquity Computing*, UbiMob '05, pages 153–160, New York, NY, USA, 2005. ACM.
- [CNSD⁺92] Carolina Cruz-Neira, Daniel J. Sandin, Thomas A. DeFanti, Robert V. Kenyon, and John C. Hart. The CAVE : audio visual experience automatic virtual environment. *Communications of the ACM*, 35(6) :64–72, June 1992.
- [Cor18] Fabiano R. Correa. Cyber-physical systems for construction industry. In *2018 IEEE Industrial Cyber-Physical Systems (ICPS)*, pages 392–397, St. Petersburg, May 2018. IEEE.
- [Cou87] Joëlle Coutaz. PAC : AN OBJECT ORIENTED MODEL FOR IMPLEMENTING USER INTERFACES. *ACM SIGCHI Bulletin*, 19(2) :37–41, October 1987.
- [DBAG14] Emmanuel Dubois, Christophe Bortolaso, Damien Appert, and Guillaume Gauffre. An MDE-based framework to support the development of Mixed Interactive Systems. *Science of Computer Programming*, 89 :199–221, September 2014.

- [DBP16] Shital Desai, Alethea Blackler, and Vesna Popovic. Intuitive Interaction in a Mixed Reality System. In *Proceedings of DRS 2016*, page 17, Brighton, UK, June 2016.
- [DGN10] Emmanuel Dubois, Philip Gray, and Laurence Nigay. *The engineering of mixed reality systems*. Human–Computer Interaction Series. Springer Science & Business Media, 2010.
- [FALH20] Rebecca Fribourg, Ferran Argelaguet, Anatole Lecuyer, and Ludovic Hoyet. Avatar and Sense of Embodiment : Studying the Relative Preference Between Appearance, Control and Point of View. *IEEE Transactions on Visualization and Computer Graphics*, 26(5) :2062–2072, May 2020.
- [FMB06] Philippe Fuchs, Guillaume Moreau, and Jean-Marie Burkhardt. *Le traité de la réalité virtuelle Volume 2 : L'interfaçage : l'immersion et l'interaction en environnement virtuel*. Presse des Mines, 2006.
- [FMG11] Philippe Fuchs, Guillaume Moreau, and Pascal Guitton. *Virtual reality : concepts and technologies*. CRC Press, 2011.
- [Fuc17] Philippe Fuchs. *Virtual reality headsets-a theoretical and pragmatic approach*. CRC Press, 2017.
- [GARL13] P. Gaucher, F. Argelaguet, J. Royan, and A. Lécuyer. A novel 3D carousel based on pseudo-haptic feedback and gestural interaction for virtual showcasing. In *2013 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 55–58, March 2013.
- [GBWG19] Christopher Greer, Martin Burns, David Wollman, and Edward Griffor. Cyber-physical systems and internet of things. Technical Report NIST SP 1900-202, National Institute of Standards and Technology, Gaithersburg, MD, March 2019.
- [GCB⁺10] Guillaume Gauffre, Syrine Charfi, Christophe Bortolaso, Cédric Bach, and Emmanuel Dubois. Developing Mixed Interactive Systems : A Model-Based Process for Generating and Managing Design Solutions. In Emmanuel Dubois, Philip Gray, and Laurence Nigay, editors, *The Engineering of Mixed Reality Systems*, pages 183–208. Springer London, London, 2010.
- [GCL⁺18] Stuart Golodetz, Tommaso Cavallari, Nicholas A. Lord, Victor A. Prisacariu, David W. Murray, and Philip H. S. Torr. Collaborative Large-Scale Dense 3D Reconstruction with Online Inter-Agent Pose Optimisation. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–1, 2018.
- [Gib77] James J. Gibson. The theory of affordances. In John Bransford Robert E Shaw, editor, *Perceiving, acting, and knowing : toward an ecological psychology*, pages pp.67–82. Hillsdale, N.J. : Lawrence Erlbaum Associates, 1977.
- [Gib79] James J Gibson. *The ecological approach to visual perception : classic edition*. Psychology Press, 1979.
- [Gri19] Michael W Grieves. Virtually Intelligent Product Systems : Digital and Physical Twins. *Complex Systems Engineering : Theory and Practice*, pages 175–200, 2019.
- [GS12] Edward Glaessgen and David Stargel. The digital twin paradigm for future NASA and US Air Force vehicles. In *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, page 1818, 2012.
- [HC97] A. Hussey and D. Carrington. Comparing the MVC and PAC architectures : a formal perspective. *IEE Proceedings - Software Engineering*, 144(4) :224–236, August 1997.
- [HD15] Lone Koefoed Hansen and Peter Dalsgaard. Note to Self : Stop Calling Interfaces "Natural". In *Proceedings of The Fifth Decennial Aarhus Conference on Critical Alternatives*, AA '15, pages 65–68, Aarhus, Denmark, 2015. Aarhus University Press.

- [Hei60] M. L. Heilig. US Patent 2955156, Stereoscopic-television apparatus for individual use, October 1960.
- [Hei67] Martin Heidegger. *What is a Thing?* Gateway / henry regnery edition, 1967.
- [HIM⁺16] Ruth Hoffmann, Murray Ireland, Alice Miller, Gethin Norman, and Sandor Veres. Autonomous Agent Behaviour Modelled in PRISM – A Case Study. In Dragan Bošnački and Anton Wijs, editors, *Model Checking Software*, pages 104–110, Cham, 2016. Springer International Publishing.
- [HPD17] Simon Hilt, Charles Pontonnier, and Georges Dumont. Model Based Compensation for Low Mass Objects Haptic Manipulation in Virtual Environments. In Jernej Barbic, Mirabelle D’Cruz, Marc Erich Latoschik, Mel Slater, and Patrick Bourdot, editors, *Virtual Reality and Augmented Reality*, volume 10700, pages 87–101. Springer International Publishing, Cham, 2017.
- [HV16] Wolfgang Hürst and Kevin Vriens. Multimodal Feedback for Finger-based Interaction in Mobile Augmented Reality. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, ICMI 2016, pages 302–306, New York, NY, USA, 2016. ACM.
- [Ish08] Hiroshi Ishii. Tangible Bits : Beyond Pixels. In *Proceedings of the 2Nd International Conference on Tangible and Embedded Interaction*, TEI ’08, pages xv–xxv, New York, NY, USA, 2008. ACM.
- [IU97] Hiroshi Ishii and Brygg Ullmer. Tangible Bits : Towards Seamless Interfaces Between People, Bits and Atoms. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, CHI ’97, pages 234–241, New York, NY, USA, 1997. ACM.
- [JBOW15] Brett R. Jones, Hrvoje Benko, Eyal Ofek, and Andrew D. Wilson. IllumiRoom : immersive experiences beyond the TV screen. *Communications of the ACM*, 58(6) :93–100, May 2015.
- [JGH⁺08] Robert J.K. Jacob, Audrey Girouard, Leanne M. Hirshfield, Michael S. Horn, Orit Shaer, Erin Treacy Solovey, and Jamie Zigelbaum. Reality-based Interaction : A Framework for post-WIMP Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’08, pages 201–210, New York, NY, USA, 2008. ACM.
- [JLW11] Jhilmil Jain, Arnold Lund, and Dennis Wixon. The Future of Natural User Interfaces. In *CHI ’11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA ’11, pages 211–214, New York, NY, USA, 2011. ACM.
- [JPU19] Céline Jost, Brigitte Le Pévédic, and Gérard Uzan. MulseBox : new multisensory interaction device. In *Proceedings of the 31st Conference on l’Interaction Homme-Machine - IHM ’19*, pages 1–13, Grenoble, France, 2019. ACM Press.
- [KBH⁺18] Kangsoo Kim, Luke Bölling, Steffen Haesler, Jeremy Bailenson, Gerd Bruder, and Greg F. Welch. Does a Digital Assistant Need a Body? The Influence of Visual Embodiment and Social Behavior on the Perception of Intelligent Virtual Agents in AR. pages 105–114, Munich, 2018.
- [KD99] Sing Bing Kang and Huong Quynh Dinh. Multi-layered image-based rendering. In *Graphics Interface*, volume 1, pages 2–13. Citeseer, 1999. Issue : 7.
- [KFMR17] Mareike Kritzler, Markus Funk, Florian Michahelles, and Wolfgang Rohde. The Virtual Twin : Controlling Smart Factories Using a Spatially-correct Augmented Reality Representation. In *Proceedings of the Seventh International Conference on the Internet of Things*, IoT ’17, pages 38 :1–38 :2, New York, NY, USA, 2017. ACM.
- [KKSF10] Gerd Kortuem, Fahim Kawsar, Vasughi Sundramoorthy, and Daniel Fitton. Smart objects as building blocks for the Internet of things. *IEEE Internet Computing*, 14(1) :44–51, January 2010.

- [KN14] Hiroyuki Kawakita and Toshio Nakagawa. Augmented TV : An augmented reality system for TV programs beyond the TV screen. In *2014 International Conference on Multimedia Computing and Systems (ICMCS)*, pages 955–960, Marrakech, Morocco, April 2014. IEEE.
- [KSF⁺12] Dimitris Kelaidonis, Andrey Somov, Vassilis Foteinos, George Poullos, Vera Stavroulaki, Panagiotis Vlacheas, Panagiotis Demestichas, Alexander Baranov, Abdur Rahim Biswas, and Raffaele Giaffreda. Virtualization and Cognitive Management of Real World Objects in the Internet of Things. In *2012 IEEE International Conference on Green Computing and Communications*, pages 187–194, Besancon, France, November 2012. IEEE.
- [LCL⁺17] S. Lin, H. F. Cheng, W. Li, Z. Huang, P. Hui, and C. Peylo. Ubii : Physical World Interaction Through Augmented Reality. *IEEE Transactions on Mobile Computing*, 16(3) :872–885, March 2017.
- [LGR⁺19] Flavien Lécuyer, Valérie Gouranton, Adrien Reuzeau, Ronan Gaugne, and Bruno Arnaldi. Authoring AR Interaction by AR. In *ICAT-EGVE 2019 - International Conference on Artificial Reality and Telexistence - Eurographics Symposium on Virtual Environments*, pages 1–8, Tokyo, Japan, September 2019.
- [LLCVF19] Jérémy Lacoche, Morgan Le Chenechal, E Villain, and A Foulonneau. Model and Tools for Integrating IoT into Mixed Reality Environments : Towards a Virtual-Real Seamless Continuum. In *ICAT-EGVE 2019 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, Tokyo, Japan, September 2019.
- [LPMN16] Michel Le Pallec, Mohamed Omar Mazouz, and Ludovic Noirie. Physical-Interface-Based IoT Service Characterization. In *Proceedings of the 6th International Conference on the Internet of Things, IoT'16*, pages 63–71, New York, NY, USA, 2016. ACM.
- [LSH⁺15] Sang-won Leigh, Philipp Schoessler, Felix Heibeck, Pattie Maes, and Hiroshi Ishii. THAW : Tangible Interaction with See-Through Augmentation for Smartphones on Computer Screens. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction, TEI '15*, pages 89–96, New York, NY, USA, 2015. ACM.
- [Mac90] D. G. MacKay. Perception, Action, and Awareness : A Three-Body Problem. In Odmir Neumann and Wolfgang Prinz, editors, *Relationships Between Perception and Action*, pages 269–303. Springer Berlin Heidelberg, Berlin, Heidelberg, 1990.
- [Mar95] Jean-Claude Martin. TYCOON : Theoretical Framework and Software Tools for Multimodal Interfaces. page 25, 1995.
- [Mas70] MORI Masahiro. Bukimi no tani genshō. *Le phénomène de la vallée de l'étrange*, *Energy*, 7(4) :33–35, 1970.
- [MBA⁺04] Jun Murayama, Laroussi Bougrila, Katsuhito Akahane, Shoichi Hasegawa, B at Hirsbrunner, and Makoto Sato. SPIDAR G&G : A Two-Handed Haptic Interface for Bimanual VR Interaction. June 2004.
- [Mih15] Alaeddine Mihoub. Apprentissage statistique de mod eles de comportement multimodal pour les agents conversationnels interactifs. page 153, 2015.
- [MK94] Paul Milgram and Fumio Kishino. A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information Systems*, E77-D, 1994.
- [MKBK98] R. Mascarenhas, D. Karumuri, U. Buy, and R. Kenyon. Modeling and analysis of a virtual reality system with time Petri nets. In *Proceedings of the 20th International Conference on Software Engineering*, pages 33–42, Kyoto, Japan, 1998. IEEE Comput. Soc.

- [MM09] Pranav Mistry and Pattie Maes. SixthSense : A Wearable Gestural Interface. In *ACM SIGGRAPH ASIA 2009 Art Gallery & Emerging Technologies : Adaptation*, SIGGRAPH ASIA '09, pages 85–85, New York, NY, USA, 2009. ACM.
- [MMK12] Masahiro Mori, Karl MacDorman, and Norri Kageki. The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*, 19(2) :98–100, June 2012.
- [MRK11] Andreas Möller, Luis Roalter, and Matthias Kranz. Cognitive Objects for Human-computer Interaction and Human-robot Interaction. In *Proceedings of the 6th International Conference on Human-robot Interaction*, HRI '11, pages 207–208, New York, NY, USA, 2011. ACM.
- [MSW⁺03] Asa MacWilliams, Christian Sandor, Martin Wagner, Martin Bauer, Gudrun Klinker, and Bernd Bruegge. Herding Sheep : Live System Development for Distributed Augmented Reality. In *Proceedings of the 2Nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '03, pages 123–, Washington, DC, USA, 2003. IEEE Computer Society.
- [MTL⁺19] Peter Mohr, Markus Tatzgern, Tobias Langlotz, Andreas Lang, Dieter Schmalstieg, and Denis Kalkofen. TrackCap : Enabling Smartphones for 3D Interaction on Mobile Head-Mounted Displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*, pages 1–11, Glasgow, Scotland Uk, 2019. ACM Press.
- [MUS16] Eric Marchand, Hideaki Uchiyama, and Fabien Spindler. Pose Estimation for Augmented Reality : A Hands-On Survey. *IEEE Transactions on Visualization and Computer Graphics*, 22(12) :2633–2651, December 2016.
- [NBB⁺19] Nahal Norouzi, Gerd Bruder, Brandon Belna, Stefanie Mutter, Damla Turgut, and Greg Welch. A Systematic Review of the Convergence of Augmented Reality, Intelligent Virtual Agents, and the Internet of Things. In Fadi Al-Turjman, editor, *Artificial Intelligence in IoT*, pages 1–24. Springer International Publishing, Cham, 2019.
- [NCLN11] Steven Neale, Winyu Chinthammit, Christopher Lueg, and Paddy Nixon. Natural Interactions Between Augmented Virtual Objects. In *Proceedings of the 23rd Australian Computer-Human Interaction Conference*, OzCHI '11, pages 229–232, New York, NY, USA, 2011. ACM.
- [NJ19] Arthur Nishimoto and Andrew E Johnson. Extending Virtual Reality Display Wall Environments Using Augmented Reality. In *Symposium on Spatial User Interaction*, pages 1–5, New Orleans LA USA, October 2019. ACM.
- [NM18] E. Normand and M. J. McGuffin. Enlarging a Smartphone with AR to Create a Hand-held VESAD (Virtually Extended Screen-Aligned Display). In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 123–133, Munich, Germany, October 2018.
- [Nor99] Donald A. Norman. Affordance, Conventions, and Design. *interactions*, 6(3) :38–43, May 1999.
- [Nor10] Donald A. Norman. Natural User Interfaces Are Not Natural. *interactions*, 17(3) :6–10, May 2010.
- [NPCA16] Michele Nitti, Virginia Pilloni, Giuseppe Colistra, and Luigi Atzori. The Virtual Object as a Major Element of the Internet of Things : A Survey. *IEEE Communications Surveys & Tutorials*, 18(2) :1228–1240, 2016.
- [OHM⁺13] Kenton O'hara, Richard Harper, Helena Mentis, Abigail Sellen, and Alex Taylor. On the Naturalness of Touchless : Putting the “Interaction” Back into NUI. *ACM Transactions on Computer-Human Interaction*, 20(1) :5 :1–5 :25, April 2013.
- [PDG17] Sara Price, Sam Duffy, and Monica Gori. Developing a pedagogical framework for designing a multisensory serious gaming environment. In *Proceedings of the 1st ACM*

- SIGCHI International Workshop on Multimodal Interaction for Education - MIE 2017*, pages 1–9, Glasgow, UK, 2017. ACM Press.
- [Pin07] Claudio Pinhanez. Ubiquitous Services. In *Proceedings - 2007 IEEE International Conference on Services Computing, SCC 2007*, pages 146–153, 2007.
- [Pos11] Stefan Poslad. *Ubiquitous computing : smart devices, environments and interactions*. John Wiley & Sons, 2011.
- [PPL18] Thies Pfeiffer and Nadine Pfeiffer-Leßmann. Virtual Prototyping of Mixed Reality Interfaces with Internet of Things (IoT) Connectivity. *i-com*, 17(2) :179–186, August 2018.
- [PS09] Nils Petersen and Didier Stricker. Continuous Natural User Interface : Reducing the Gap Between Real and Digital World. In *Proceedings of the 2009 8th IEEE International Symposium on Mixed and Augmented Reality, ISMAR '09*, pages 23–26, Washington, DC, USA, 2009. IEEE Computer Society.
- [PZCG14] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos. Context Aware Computing for The Internet of Things : A Survey. *IEEE Communications Surveys Tutorials*, 16(1) :414–454, 2014.
- [RBMN06] Stéphane Renouard, Jullien Bouchet, Mounir Mokhtari, and Laurence Nigay. Conception et développement de l'interaction multimodale Comparaison des plateformes Scylla et ICARE. page 9, 2006.
- [RH17] Joan Sol Roo and Martin Hachet. One Reality : Augmenting How the Physical World is Experienced by combining Multiple Mixed Reality Modalities. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology - UIST '17*, pages 787–795, Quebec City, QC, Canada, 2017. ACM Press.
- [RU18] Rafael Radkowski and Iowa State University. HoloLens Integration into a Multi-Kinect Tracking Environment. page 5, 2018.
- [Sch11] Daniel L Schacter. *Psychology*. Worth Publishers, second edition edition, 2011.
- [SHN19] Maximilian Speicher, Brian D. Hall, and Michael Nebeling. What is Mixed Reality? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*, pages 1–15, Glasgow, Scotland Uk, 2019. ACM Press.
- [SIAB18] Mickael Sereno, Tobias Isenberg, Mehdi Ammi, and Lonni Besançon. Combining tactile and tangible input for 3D selection. In *Proceedings of the 29th Conference on Interaction Homme-Machine - IHM '17*, pages 207–212, Poitiers, France, 2018. ACM Press.
- [Sim72] Herbert A Simon. Theories of bounded rationality. *Decision and organization*, 1(1) :161–176, 1972. Publisher : North-Holland.
- [SLRHM12] Tomás Sánchez López, Damith C. Ranasinghe, Mark Harrison, and Duncan McFarlane. Adding Sense to the Internet of Things. *Personal Ubiquitous Comput.*, 16(3) :291–308, March 2012.
- [SMS17] H. N. Saha, A. Mandal, and A. Sinha. Recent trends in the Internet of Things. In *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 1–4, January 2017.
- [SRC⁺18] Joao Santos, Joel J. P. C. Rodrigues, Joao Casal, Kashif Saleem, and Victor Denisov. Intelligent Personal Assistants Based on Internet of Things Approaches. *IEEE Systems Journal*, 12(2) :1793–1802, June 2018.
- [SS12] C. Sulisz and P. Seeling. An Off-the-shelf Wearable HUD System for Support in Indoor Environments. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia, MUM '12*, pages 60 :1–60 :4, New York, NY, USA, 2012. ACM.

- [Sut68] Ivan E. Sutherland. A Head-mounted Three Dimensional Display. In *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*, AFIPS '68 (Fall, part I), pages 757–764, New York, NY, USA, 1968. ACM.
- [SWM⁺02] Christian Sandor, Martin Wagner, Asa MacWilliams, Martin Bauer, and Gudrun Klinker. SHEEP : The Shared Environment Entertainment Pasture. page 1, January 2002.
- [TAL18] Ym Tang, Km Au, and Yohana Leung. Comprehending products with mixed reality : Geometric relationships and creativity. *International Journal of Engineering Business Management*, 10 :184797901880959, January 2018.
- [Tur05] P. Turner. Affordance as context. *Interacting with Computers*, 17(6) :787–800, December 2005.
- [Uim18] Mikael Uimonen. Accessing BIM-Related Information through AR. page 2, 2018.
- [VGD⁺20] Guillaume Vailland, Yoren Gaffary, Louise Devigne, Valérie Gouranton, Bruno Arnaldi, and Marie Babel. Vestibular Feedback on a Virtual Reality Wheelchair Driving Simulator : A Pilot Study. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, pages 171–179, Cambridge United Kingdom, March 2020. ACM.
- [Wal08] Stephen Walther. *ASP.NET 3.5 unleashed*. Sams, Indianapolis, Ind, 2008. OCLC : ocn180989631.
- [Wei99] Mark Weiser. The Computer for the 21st Century. *Mobile Computing and Communications Review*, 3(3) :3–11, July 1999.
- [WG93] Ian H. Witten and Saul Greenberg. User Interfaces. *Encyclopaedia of Computer Science*, pages 1411–1414, 1993.
- [ZMD00] Y. Zhou, T. Murata, and T.A. DeFanti. Modeling and performance analysis using extended fuzzy-timing Petri nets for networked virtual environments. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, 30(5) :737–756, October 2000.

Titre : De la conception à l'exploration des systèmes interactifs combinant Réalité Mixte, Interfaces Utilisateur Naturelles et Internet des Objets

Mots clés : réalité mixte, interfaces utilisateur naturelles, internet des objets

Résumé : La réalité mixte, les interfaces utilisateur naturelles et l'internet des objets permettent d'interagir simultanément avec le réel et le virtuel de façon intuitive et ubiquitaire. Leur convergence donne lieu à l'émergence de nouvelles formes de systèmes interactifs. Complexes et hétérogènes, ils requièrent des modèles et méthodologies de conception et d'implémentation unifiés. Nous proposons une approche permettant de maîtriser leur complexité, de simplifier et d'accélérer leur création, d'améliorer leur inter connectivité et leur extensibilité, et d'explorer les capacités d'interaction qu'ils ont à offrir.

Cette approche est basée sur un modèle unifié qui modélise les utilisateurs, les interfaces utilisateur naturelles et les objets connectés sous la forme d'entités mixtes interconnectées. Ce modèle détaille également leur capacité à développer des comportements mixtes issus de la relation entre réel et virtuel. Nous proposons un framework basé sur ce modèle permettant de simplifier, factoriser et uniformiser la conception et l'implémentation de ces systèmes. Nous explorons également des techniques d'interaction mono et multi-sensorielle hybrides.

Title : From design to exploration of interactive systems combining Mixed Reality, Natural User Interfaces, and the Internet of Things

Keywords : mixed reality, natural user interfaces, internet of things

Abstract : Mixed reality, natural user interfaces, and the internet of things enable simultaneous interaction with the real and the virtual in an intuitive and ubiquitous manner. New forms of interactive systems emerge from their convergence. Complex and heterogeneous, they require models and methodologies dedicated to their conception and their implementation. We propose an approach allowing to manage their complexity, to simplify and accelerate their creation, to enhance their interconnectivity and their scalability, and to explore innovative interaction techniques they enable.

This approach is based on our unified model. We model users, natural user interfaces, and connected objects as interconnected mixed entities. Our model details their capability to develop mixed behaviors. We propose a framework based on this model simplifying, factoring, and unifying the design and implementation of such systems. We also explore new mono and multi-sensorial hybrid interaction techniques they offer.

AVIS DU JURY SUR LA REPRODUCTION DE LA THESE SOUTENUE

Titre de la thèse:

De la conception à l'exploration des systèmes interactifs combinant Réalité Mixte, Interfaces Utilisateur Naturelles et Internet des Objets

Nom Prénom de l'auteur : BATAILLE GUILLAUME

Membres du jury :

- Monsieur AMMI Mehdi
- Monsieur CASIEZ Géry
- Monsieur LACOCHÉ Jérémy
- Monsieur ARNALDI BRUNO
- Madame GOURANTON Valérie
- Monsieur GUITTON Pascal

Président du jury : Monsieur CASIEZ Géry

Date de la soutenance : 13 Novembre 2020

Reproduction de la these soutenue

- Thèse pouvant être reproduite en l'état
 Thèse pouvant être reproduite après corrections suggérées

Fait à Rennes, le 13 Novembre 2020

Signature du président de jury



Le Directeur,

M'hamed DRISSI



