



HAL
open science

Vers une prise en compte du toucher social dans les interactions humain-agent en environnement virtuel immersif

Fabien Boucaud, Quentin Tafiani, Catherine Pelachaud, Indira Thouvenin

► To cite this version:

Fabien Boucaud, Quentin Tafiani, Catherine Pelachaud, Indira Thouvenin. Vers une prise en compte du toucher social dans les interactions humain-agent en environnement virtuel immersif. Journées de la Réalité Virtuelle (j•RV 2018), Oct 2018, Evry, France. hal-02097972

HAL Id: hal-02097972

<https://hal.science/hal-02097972>

Submitted on 12 Apr 2019

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

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

Vers une prise en compte du toucher social dans les interactions humain-agent en environnement virtuel immersif

Fabien Boucaud	Quentin Tafiani	Catherine Pelachaud	Indira Thouvenin
Sorbonne Universités	Sorbonne Universités	Sorbonne Universités	Sorbonne Universités
CNRS UMR 7222 ISIR	CNRS UMR 7222 ISIR	CNRS UMR 7222 ISIR	Université de Technologie de
UPMC Campus Jussieu	UPMC Campus Jussieu	UPMC Campus Jussieu	Compiègne
4 place Jussieu	4 place Jussieu	4 place Jussieu	CNRS UMR 7253 Heudiasyc
75252 Paris Cedex 05	75252 Paris Cedex 05	75252 Paris Cedex 05	57 Av. Landshut CS 60319
fabien.boucaud@etu.utc.fr	tafiani.quentin@gmail.com	catherine.pelachaud@upmc.fr	60203 Compiègne Cedex
			indira.thouvenin@utc.fr

RÉSUMÉ

Si le développement des travaux sur les agents sociaux artificiels leur permet de prendre en compte de plus en plus de modalités d'interactions (parole, gestuelle, expressions du visage, etc.), le toucher, sens pourtant particulièrement intéressant pour la communication sociale et émotionnelle, en reste aujourd'hui encore largement absent. Nous pensons que l'intégration du toucher aux modalités d'interaction entre humain et agent permettrait d'améliorer les canaux de communication empathique. Nous proposons donc une expérience pour mesurer l'apport et la crédibilité de l'utilisation du toucher social dans les modalités d'interaction humain-agent en environnement immersif de type CAVE

Les premiers résultats sont prometteurs quant à notre hypothèse et offrent de nouvelles pistes quant à l'amélioration de la prise en compte du toucher et son utilisation par l'agent.

Keywords: Réalité virtuelle, Interaction, CAVE, ECA, Agent Conversationnel Animé, Toucher Social, Emotion

1 INTRODUCTION

Avec la multiplication des recherches et des développements sur l'interaction humain-machine et plus particulièrement humain-agent, l'intérêt de la recherche se porte de plus en plus sur la variété des modalités d'interaction qu'il est possible de prendre en compte. Comme l'anthropologie a pu le montrer [1], le toucher a toujours été notre modalité d'interaction principale avec les objets techniques, du silex primitif à nos machines les plus perfectionnées aujourd'hui. A l'ère du numérique et de ses nouvelles technologies, cela est d'autant plus manifeste qu'on munit aussi bien nos smartphones que nos écrans d'ordinateur de fonctionnalités tactiles. [2]

Pourtant, de leur côté, là où les agents sociaux artificiels comme les agents conversationnels animés prennent en compte de plus en plus de modalités d'interaction, de l'interprétation du langage naturel à celle des expressions du visage ou de la gestuelle, le toucher est encore très peu intégré aux interactions sociales entre humain et agent. Si la recherche s'est tardivement intéressée à la question des significations sociales du toucher [2], pour diverses raisons aussi bien culturelles que techniques, il est toutefois clair aujourd'hui que le toucher est un sens social très important et notamment très utile à la communication empathique, c'est-à-dire la communication des émotions.

Nous proposons ici de nous intéresser à la façon dont on peut intégrer le toucher social aux modalités d'interaction humain-agent dans le contexte de la réalité virtuelle.

D'abord, il s'agira de définir précisément la notion de toucher social et de déterminer comment il est possible de mettre en place une interaction sociale humain-agent basée sur le toucher dans un environnement immersif de type CAVE. Plus précisément, on cherche ici à savoir dans quelle mesure l'utilisation du toucher social peut être considérée comme crédible dans un contexte d'interaction humain-agent en environnement virtuel immersif. Nous présenterons, dans cet objectif, l'expérience élaborée pour déterminer dans quelle mesure une telle interaction basée sur le toucher social entre humain et agent peut permettre d'améliorer leurs canaux de communication empathique.

2 DÉFINIR LE TOUCHER SOCIAL ET SES TECHNOLOGIES

2.1 Qu'est-ce que le toucher social ?

Par toucher social, on entend toutes les utilisations du toucher avec une intention sociale. Ainsi une poignée de main pour saluer quelqu'un, une tape d'encouragement dans le dos, n'importe quel type de toucher interpersonnel non-accidental (et même parfois accidentel) peut être considéré comme une occurrence de toucher social.

Sens essentiel au développement individuel et social de l'individu dès le plus jeune âge, le toucher remplit des fonctions communicatives [3] semblables à celles des autres types de communication non-verbale (gestes, expressions faciales, etc.) et est considéré comme un sens particulièrement utile à la communication empathique, c'est-à-dire la communication des émotions. Pourtant, le toucher a par le passé été très peu étudié, pour des raisons aussi bien culturelles que scientifiques. Ce n'est qu'à partir des années 1970-80 qu'un regain d'intérêt pour l'étude de ce sens s'est manifesté avec la remise au premier plan de l'importance du corps par les études féministes et la philosophie (Deleuze, Jean-Luc Nancy,...) [2]. Dans cette lignée, on s'intéresse aujourd'hui de plus en plus aux fonctions communicatives du toucher, et à la possibilité de les intégrer aux modalités d'interaction humain-agent, notamment dans le but d'améliorer les canaux de communication empathique.

Dans cette optique, les travaux de définition, qualification et classification du toucher social « naturel », c'est-à-dire entre

humains, d'Hertenstein et al. [4] [5] ainsi que de Bianchi-Berthouze et Tajadura-Jiménez [6] sont une source d'informations théoriques essentielle pour élaborer des systèmes capables de produire un toucher technologique crédible. Leurs études mettent en évidence la façon dont on peut définir les caractéristiques des différents types de toucher et comment chaque type de toucher est plus particulièrement susceptible d'être utilisé dans l'expression de certaines émotions. En particulier, ces travaux mettent en évidence que le toucher est lui-même un sens particulièrement multi-modal : ses caractéristiques sont très variées, entre la pression, la force d'impact, la vitesse de déplacement le long de la peau (dans le cas des caresses par exemple), la durée totale du geste de toucher, etc. Pour autant, Hertenstein et al. montrent bien la façon dont ces seules caractéristiques ne suffisent pas à interpréter l'intention communicative d'un toucher. En effet, le toucher repose sur des principes d'équipotentialité et d'équifinalité. Un même type de toucher, comme un coup, peut aussi bien être utilisé dans une situation compétitive pour encourager un partenaire ou faire montre de camaraderie que dans une intention d'exprimer de la colère, c'est l'équipotentialité. De la même façon, deux types de toucher différents, comme donner un coup ou agripper quelqu'un avec force peuvent aussi bien exprimer la colère l'un que l'autre, c'est l'équifinalité. A ces éléments physiques de mesure du toucher s'ajoutent donc d'autres facteurs extérieurs entrant en compte dans l'interprétation sociale du toucher. Parmi ceux-ci : le contexte, la relation entre le toucheur et le touché, leurs cultures respectives, l'endroit du corps touché, etc.

Ce n'est qu'en prenant en compte toutes ces facettes du toucher qu'il devient possible de mettre au point des technologies capables de médiatiser et simuler le toucher de manière crédible.

2.2 Les technologies du toucher social

Sur le plan technique, les technologies du toucher (on peut parler de technologies haptiques ou tactiles) sont très variées [7] recouvrant aussi bien les vibrations que le retour de force, etc. Pour autant, il n'existe actuellement pas de technologie permettant de reproduire le toucher réel à l'identique et d'autant moins dans un contexte de réalité virtuelle où les éléments virtuels sont par définition dépourvus de matérialité. Afin de mettre au point un système qui permettrait de toucher et être touché par des agents conversationnels animés, c'est-à-dire munis d'un corps virtuel, il nous faut donc plutôt nous intéresser à des technologies capables de produire une sensation tactile sans corps réel à toucher. Nous envisageons ainsi des dispositifs tels qu'un gant ou une manchette vibrotactiles, capables de reproduire certains types de toucher à un endroit spécifique du corps, à l'exemple du TASST de Huisman et al. [8], ou des dispositifs à retour de force permettant de donner une impression de tenir des objets virtuels. Le pseudo-haptique tel que définit par Anatole Lécuyer et al. [9] est un autre concept intéressant en ce qui concerne le toucher en réalité virtuelle.

Dans ses travaux [10], Gijs Huisman revient en détails sur la variété des technologies du toucher social et sur leurs différences. Quand on parle de technologies du toucher social, on distingue la médiation (transmettre un toucher à un autre individu par un

intermédiaire technologique) de la simulation (génération autonome d'un comportement tactile) du toucher social. La distinction réside en ce que la simulation du toucher ré-utilise les technologies de la médiation mais y ajoute une capacité d'adaptation, c'est-à-dire une intelligence, un modèle de décision capable de déterminer en fonction de la situation et des différents paramètres de l'interaction en cours le comportement tactile à adopter.

Quant à savoir si la médiation et la simulation présentent les mêmes propriétés de transmission des émotions que le toucher social naturel, les travaux de Van Erp et Toet (2013) [11] mettent en évidence la crédibilité de trois principes :

- les émotions peuvent être communiquées en utilisant uniquement le toucher sans autre repère.
- la communication interpersonnelle peut également être accomplie avec un toucher médiatisé et n'est pas seulement limitée au contact physique direct.
- non seulement les humains mais aussi les systèmes peuvent utiliser le toucher médiatisé pour communiquer leurs émotions.

De la même manière, le travail de Bickmore et al. [12] met en évidence la façon dont la simulation de toucher social permet de renforcer la communication empathique entre humain et agent conversationnel animé. Le constat est le même, quoique nuancé, dans le cadre de l'utilisation de la réalité augmentée pour matérialiser les agents et dans une situation sociale de jeu coopératif. [13]

En ce qui concerne la crédibilité de la réponse tactile d'un agent à la simulation, il y a encore très peu de travaux qui s'y intéressent avec un agent d'apparence humaine. Nous pouvons toutefois citer les travaux de Yohanan [14] sur la « *Haptic Creature* » qui mettent en lumière que les humains s'attendent plutôt à ce que l'agent touché réagisse de façon mimétique avec l'émotion transmise. Si la créature (c'est-à-dire l'agent) est touchée joyeusement, le comportement attendu est plutôt un comportement joyeux qui indiquerait la bonne compréhension de l'émotion par la créature.

2.3 La détection du toucher dans le CAVE

Nous avons évoqué les différentes façons dont on peut utiliser la technologie pour produire des sensations tactiles à destination d'un humain. Il reste cependant à étudier la façon dont on peut faire en sorte que l'agent puisse lui aussi percevoir le toucher, et ce dans le contexte de la réalité virtuelle.

Pour faire en sorte de détecter le toucher de l'utilisateur sur l'agent, nous nous sommes inspirés des travaux de Nguyen et Wachsmuth [15] concernant la perception du toucher pour un agent virtuel au sein d'un CAVE. Rappelons ici qu'un CAVE (Cave Automatic Virtual Environment) est un système capable de produire un environnement immersif de réalité virtuelle, composé de trois murs et un sol sur lesquels sont projetés l'environnement et dans lequel l'utilisateur (sa localisation et ses mouvements) est suivi par des caméras de *motion capture* et garde la vision de son propre corps [16].

Nguyen et Wachsmuth proposent de couvrir le modèle de l'agent virtuel d'une « peau » virtuelle constituée de « cellules tactiles », c'est-à-dire des plaques de taille différentes qui

serviront de capteurs du toucher. Lorsqu'un élément du monde réel suivi par *tracking* dans le CAVE est détecté par le biais de ses coordonnées comme entrant en contact avec l'une de ces cellules, on peut dire qu'il y a collision (donc toucher) et enregistrer les informations liées à cet événement.

Ainsi il devient possible de déterminer précisément quelle cellule a été touchée et donc de quel endroit du corps il s'agit, ainsi que de suivre le déplacement potentiel du toucher, ce qui pourra permettre de déterminer le type de toucher dont il s'agit (caresse ? simple contact ?). Il s'agit cependant d'une implémentation simple qui ne permet que de distinguer s'il y a eu toucher ou pas : ce que Nguyen et Wachsmuth appellent un sens du toucher « passif ». Il n'y a pour le moment aucune autre information quant à l'intensité du toucher ou sa durée par exemple. Toutefois, cette implémentation basique permet déjà de « toucher » un agent virtuel et de le faire prendre conscience de ce toucher. En attribuant des valeurs de valence aux types de toucher en fonction de l'endroit du corps touché, Nguyen et Wachsmuth mettent en place un premier modèle émotionnel pour leur agent qui se met à jour en fonction de la manière avec laquelle il est touché et de l'endroit où il est touché.

En partant de ce principe, nous avons muni le modèle 3D de notre agent virtuel de *colliders* (formes géométriques capables de détecter lorsqu'il y a collision avec un autre élément de l'environnement muni de *colliders*) dans Unity, de sorte à reproduire le principe de ces cellules tactiles, comme illustré en figure 1. Lors d'une collision entre la représentation virtuelle de la main de l'utilisateur (dont la position est définie en fonction des données de suivi des caméras) et une cellule tactile, on enregistre les différents paramètres de l'événement tactile (vitesse au moment de la collision, durée, etc.) et on fait de même à chaque nouvel événement tactile pour constituer une séquence qui représentera le toucher sur toute sa durée (jusqu'à ce que tous les *colliders* cessent de détecter qu'il y a collision, c'est-à-dire que la main n'est plus en contact avec l'agent).

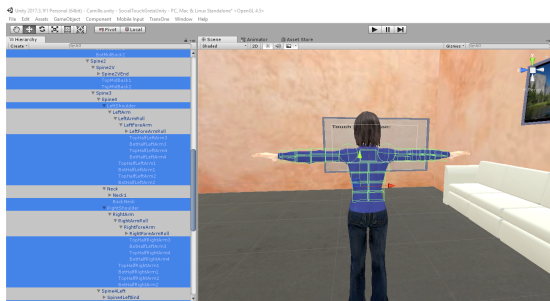


Figure 1. Les cellules tactiles de l'agent vu de dos

Cette approche a beaucoup de potentiel d'amélioration mais il faut bien tenir compte des limitations inhérentes à l'introduction d'un sens du toucher dans un environnement virtuel tel qu'un CAVE. S'il est effectivement possible, en suivant les travaux que nous venons de présenter, de « toucher » un agent virtuel, il n'en reste pas moins que l'agent n'a pas d'enveloppe physique et que

notre main ou notre outil de toucher va passer à travers la représentation visuelle de l'agent sans opposer la moindre résistance ou la moindre sensation tactile. Il est dès lors difficile d'envisager une mesure de pression ou la possibilité de tenir le bras de l'agent ou serrer sa main. Complémenter l'approche de Nguyen et Wachsmuth d'un dispositif capable d'offrir un retour haptique lors du toucher réalisé par l'humain nous a donc semblé essentiel pour rendre l'interaction un minimum crédible.

3 PRÉSENTATION DES PROTOTYPES D'INTERFACES HAPTQUES

Pour que toucher et être touché par l'agent soit crédible, il n'est pas envisageable de se contenter du retour visuel de la collision entre le corps de l'agent et le corps de l'utilisateur, d'autant moins dans la mesure où dans un CAVE la main de l'utilisateur passe à travers le corps de l'agent et inversement. Il faut mettre en place une suppléance sensorielle tactile pour éviter qu'il n'y ait dissonance entre ce que l'humain voit et ce qu'il s'attend à ressentir. En effet, notre mode d'interaction avec la réalité est basé sur nos cinq sens et du fait de notre habitude perceptive de ces interactions, nous avons des attentes sensorielles. Nous nous attendons à avoir une sensation de toucher quand nous voyons notre main entrer en contact avec un élément du monde extérieur. Si cette sensation s'avérait soudain absente, il y aurait dissonance et donc gêne et perte de crédibilité pour l'utilisateur. Pour fournir un retour sensoriel, une suppléance qui compensera du point de vue perceptif la nature immatérielle du virtuel, à l'utilisateur, nous nous sommes orienté vers l'utilisation des vibrations et la confection d'un manchon et d'un gant vibratoire. Ces deux dispositifs sont nécessaires pour à la fois simuler le toucher de l'agent (manchon) et offrir une suppléance perceptive au moment du toucher de l'utilisateur sur l'agent, pour lui indiquer qu'il a bien touché l'agent (gant).

3.1 Choix techniques

Parmi les nombreuses possibilités de technologies haptiques nous avons choisi les vibrations à la fois pour la légèreté d'appareillage sur le corps, la dense littérature existante sur la façon dont on peut les utiliser pour produire des sensations de toucher intéressantes [8] [17] et le faible coût de la technologie.

De même, on a ici choisi de réaliser un manchon à porter sur l'avant-bras à la fois pour le fait que cela facilite la mise en place et le port du dispositif et pour le fait que les études montrent que le bras est un endroit où le toucher social est plus facilement toléré, même entre personnes n'ayant pas de relation proche [18]. C'est donc particulièrement intéressant pour nous dans la mesure où la relation de départ entre l'utilisateur et l'agent est probablement de l'ordre de celle qu'auraient deux inconnus.

3.2 Description technique des dispositifs

Les prototypes réalisés, visibles dans les figures 3 et 4, utilisent la technologie Arduino pour sa facilité d'utilisation, son faible coût et sa polyvalence. Notre manchon est ainsi conçu comme une matrice de deux colonnes et quatre lignes de moteurs vibrants similaires à ceux que l'on trouve dans un téléphone portable, tandis que le gant est équipé de quatre de ces vibreurs, chacun

pour un « coin » de la paume de la main. En utilisant les principes du *Tactile Brush algorithm* [17] et en manipulant les durées et niveau d'intensité des vibrations, nous avons pu mettre en place sur le manchon, malgré les limitations inhérentes aux vibrations pour reproduire le toucher humain, quatre types de touchers inspirés de la catégorisation d'Hertenstein et al. [4] [5], choisis pour leur potentiel de transmettre des émotions différentes. Il s'agit du coup, du tapotement, de la caresse et du contact simple, dont nous avons définis les caractéristiques, en nous basant à nouveau sur les travaux d'Hertenstein et al., telles que suit : pour le coup un toucher court (400 ms), très localisé, sans déplacement et avec une vitesse élevée ; pour la caresse, un toucher plus doux en termes de vitesse, avec déplacement, long ($4 \times 200 \text{ms} = 800 \text{ms}$) et avec une grande surface de contact ; pour le tapotement, deux touchers courts ($2 \times 200 \text{ms}$), très localisés, mais avec une vitesse plutôt élevée ; pour le contact simple, un toucher très localisé avec vitesse peu élevée mais avec une longue durée (1000 ms). Il convient de préciser que le toucher naturel prend en compte des caractéristiques de toucher bien plus nombreuses que celles sélectionnées ici, mais les vibrations ne nous permettent pas, par exemple, de simuler de la pression. Ces quatre types de toucher présentent l'avantage de ne pas nécessiter de prise en compte de la pression ou de propriétés dynamiques comme une secousse, d'être très localisés, et donc de pouvoir être réalisés par un manchon vibrotactile. Nous estimons que les paramètres pris ici en compte sont suffisants pour fournir un retour sensoriel compréhensible par un humain.

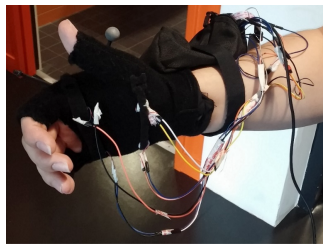


Figure 2. Le prototype du gant



Figure 3. Le prototype du manchon

Pour que l'interaction tactile soit réellement crédible cependant, il faut également faire en sorte que le geste réalisé visuellement par l'agent soit correctement synchronisé avec le signal vibratoire, sans quoi on risque de provoquer une autre dissonance perceptive qui ôterait tout réalisme à l'interaction. Le timing est donc réglé au sein de l'application Unity au moment de l'envoi des signaux d'animation de l'agent et d'activation des vibrations.

4 OBJECTIFS

Notre questionnement initial étant de déterminer dans quelle mesure il est possible d'améliorer les canaux de communication empathique humain-agent par l'intégration du toucher social aux

modalités d'interaction, nous proposons ici une expérience et un système qui permettraient de vérifier l'hypothèse qu'un échange de toucher social entre humain et agent, accompagné d'expressions faciales et gestuelles appropriées, permet une communication empathique à même d'augmenter la crédibilité de l'interaction. Nous nous concentrons sur l'étude de trois aspects fondamentaux de ce sujet : le type de toucher, et ses caractéristiques, effectué par l'humain sur l'agent pour transmettre une émotion donnée ; à l'inverse, l'évaluation par l'humain des caractéristiques, du type et de l'émotion « transmise » par le toucher de l'agent ; et enfin la crédibilité générale de l'agent au cours des deux parties précédentes.

L'ambition de nos travaux étant de proposer une intégration du toucher social pour améliorer la boucle interactive sociale humain-agent, nous nous positionnons dans le cadre de la simulation du toucher social. Il nous a dès lors fallu déterminer comment mettre en place un système qui permette de toucher et d'être touché par un agent dans un contexte de réalité virtuelle. Cela nécessite d'une part une interface technologique permettant à l'utilisateur de toucher et se sentir être touché, et d'autre part un modèle de décision qui permette à l'agent de choisir les touchers à réaliser. Nous avons montré précédemment les différentes possibilités envisageables pour simuler le toucher et ce que nous avons mis en place pour la détection par l'agent du toucher de l'humain. Nous allons maintenant revenir sur le protocole expérimental mis en place et les interfaces technologiques choisies pour produire les sensations tactiles.

5 PROTOCOLE EXPÉRIMENTAL

5.1 Hypothèse opérationnelle

On s'attend à ce que l'utilisation du toucher pour transmettre une émotion à l'agent et la réaction adaptée (en termes de gestuelle et d'expressions faciales) de ce dernier soient crédibles pour le participant.

De la même manière, on s'attend à ce que l'utilisation du toucher par l'agent, associée à une gestuelle et des expressions faciales cohérentes, permette une transmission d'émotion reconnaissable et crédible pour le participant.

5.2 Procédure générale

L'expérience se déroule en deux phases distinctes au cours desquelles le participant est amené à toucher et être touché par un agent virtuel.

Phase 1. D'abord, le sujet doit toucher l'agent virtuel avec le gant vibrotactile enfilé sur sa main dominante, afin d'exprimer une émotion, et l'agent virtuel répondra par un comportement adapté à l'émotion, tout ceci à l'intérieur d'un environnement virtuel de type CAVE.

Lors de cette phase, des scénarios émotionnels seront lus au participant en tant que mise en contexte. Ces scénarios ont pour objectif d'indiquer au participant l'émotion qu'il doit transmettre et son intensité. Les trois émotions choisies pour être transmises par le participant et l'agent sont la colère, la sympathie et la tristesse. Elles présentent l'avantage d'être des émotions à la fois

très différentes en termes de polarités et *a priori* facilement représentables pour les participants.

Phase 2. Ensuite, l'agent virtuel prodiguera à son tour un toucher au niveau du second dispositif de l'utilisateur, c'est-à-dire le manchon vibrotactile, placé sur l'avant-bras de sa main non dominante, avec une expression faciale et gestuelle adaptée à l'émotion qu'il souhaite transmettre.

Pour que l'agent soit capable d'exprimer ces émotions par le toucher, nous avons sélectionné les quatre types de toucher mentionnés précédemment : le coup, le tapotement, le contact simple et la caresse. Par ailleurs, selon les travaux de classification des types de toucher d'Hertenstein mentionnés plus tôt [4] [5], ces quatre types de toucher et ces trois émotions montrent bien l'équifinalité et l'équipotentialité du toucher : on peut aussi bien utiliser le coup que le tapotement pour exprimer la colère, mais on peut aussi utiliser le tapotement et la caresse pour exprimer la sympathie, et la caresse et le contact simple pour la tristesse. Comme nous l'avons vu précédemment cependant, les propriétés physiques d'un toucher perçu ne sont pas suffisantes pour interpréter socialement un toucher, de nombreux autres facteurs sont nécessaires pour rendre le toucher interprétable sans ambiguïté. Pour cette raison, nous avons décidé d'ajouter aux touchers de l'agent des expressions faciales correspondant à l'émotion qu'il cherche à transmettre.

Les scénarios proposés aux utilisateurs comme mise en contexte proviendront des travaux de Bänziger et al. (2006) [19] (2010) [20], et de Scherer et al. (1991) [21]. L'orientation de certaines questions sera inspirée quant à elle des travaux de Demeure et al. (2011) [22].

A titre d'exemple, les scénarios émotionnels suivants correspondent à la sympathie :

Scénario à intensité émotionnelle faible : « Vous croisez une amie à vous, Camille, que vous n'aviez pas vu depuis un moment, et lui exprimez ce que vous ressentez »

Scénario à intensité émotionnelle forte : « Votre meilleure amie, Camille, que vous n'avez pas vu depuis un certain temps et avec laquelle vous êtes complices depuis toujours, vous appelle pour vous inviter à une soirée qu'elle organise. Une fois sur place, elle vous ouvre la porte et vous lui exprimez ce que vous ressentez ».

Les touchers associés à cette émotion pour l'agent virtuel sont la caresse et le tapotement.

Les participants sont ainsi divisés en trois groupes selon l'émotion qu'ils doivent transmettre et l'émotion que l'agent doit leur transmettre.

5.3 Matériel

En termes de dispositifs, nous utilisons donc un CAVE avec le système de détection du toucher présenté précédemment en 2.3 et les prototypes présentés en 3. A cela s'ajoute la plateforme logicielle GRETA [23] [24] qui permet de contrôler le comportement social d'agents conversationnels animés (parole, gestuelle, expressions du visage), ici virtuels et dans une veine photoréaliste. Enfin, nous avons mis au point deux prototypes de dispositifs de retour haptique basés sur les vibrations.

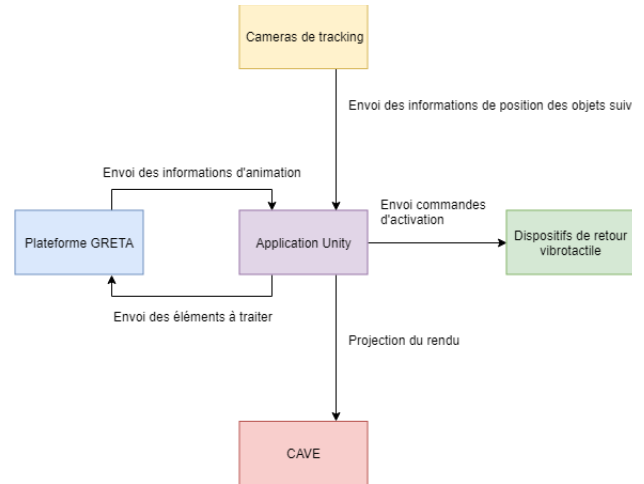


Figure 4. Architecture du système

La figure 4 illustre les différents composants du système réalisé et les relations qu'ils entretiennent les uns avec les autres. Au centre, l'élément principal du système est l'application Unity qui est tout simplement l'application développée sous le moteur de jeu Unity 3D qui contrôle le rendu de l'environnement et de l'agent, la gestion de la physique, etc. Le moteur Unity 3D a été choisi pour développer cette application pour la facilité d'intégration avec le dispositif CAVE du laboratoire Heudiasyc, où un module TRANSONE a été développé spécifiquement pour prendre en charge le lancement d'applications Unity. Cette application communique avec tous les autres éléments du système pour obtenir les informations nécessaires à la mise à jour du rendu et pour fournir les informations nécessaires telles que les données d'entrée détectées (détection du toucher sur l'agent notamment).

Avec le système ainsi mis en place, il est donc possible d'interagir avec un agent par le toucher dans le CAVE en s'équipant des deux prototypes réalisés (le manchon et le gant) et des lunettes de 3D stéréoscopiques. A l'aide des caméras de suivi on peut détecter quand le gant touche les cellules tactiles de l'agent, ce qui permet d'enregistrer les données du toucher réalisé et déclencher une animation de réaction (expression faciale et/ou gestuelle) chez l'agent ainsi qu'un retour sensoriel au niveau des vibrations du gant. Réciproquement, l'agent peut réaliser un geste de toucher synchronisé avec les vibrations du manchon pour toucher l'utilisateur et accompagner son geste de paroles et/ou d'expressions faciales.



Figure 5. Une interaction tactile humain-agent

5.4 Procédure détaillée

Tout d'abord, le participant est familiarisé avec l'environnement virtuel, un agent virtuel, différent de l'agent présent dans la suite de l'expérience, et le fonctionnement du matériel afin de supprimer l'effet de nouveauté et de s'appropriier le matériel. Lors de cette phase, l'utilisateur touche l'agent autant de fois et de la manière dont il le souhaite et l'agent touche l'utilisateur de manière neutre, avec une vibration forte.

Dans le détail, la scénarisation est la suivante. Le participant entre dans la pièce, voit Camille (l'agent virtuel) de dos, se positionne au niveau de la marque blanche et l'interpelle par un toucher avec la main gantée.

Camille se retourne, voit l'utilisateur, a une réaction et se présente.

C'est la première session de la première phase qui commence : suite à la lecture du scénario d'intensité faible correspondant, ils doivent soit transmettre leur sympathie à l'agent (groupe C1), soit transmettre leur colère à l'agent (C2), soit transmettre leur tristesse à l'agent (C3), uniquement par le toucher, quatre fois. Suite à cela, ils sont invités à répondre à quelques questions avant de réaliser la deuxième session, identique si ce n'est que cette fois c'est le scénario émotionnel d'intensité forte qui est lu. Suite à cela, on invite à nouveau le participant à répondre à quelques questions, puis c'est la seconde phase qui commence.

Lors de la seconde phase, c'est à l'agent de soit « transmettre sa colère » à l'utilisateur (C1), soit « transmettre sa tristesse » à l'utilisateur (C2), soit « transmettre sa sympathie » à l'utilisateur (C3), avec l'expression faciale correspondante. Il y a à nouveau deux sessions de quatre touchers réalisés par l'agent, entrecoupées de réponse aux questionnaires, mais cette fois ce n'est pas l'intensité de l'émotion qui change entre les deux sessions mais le type de toucher utilisé pour la transmettre. Ainsi en C2, l'agent transmet d'abord sa tristesse en touchant le bras de l'humain quatre fois pour la première session, puis en caressant son bras quatre fois pour la seconde.

5.5 Participants

L'expérience, d'une durée d'une heure, a été menée avec douze participants, dont huit femmes et quatre hommes. Neuf n'ont aucune expérience de la réalité virtuelle. Dix des participants se sont considérés réceptifs au toucher et deux n'ont pas su dire. Tous les participants ont entre 18 et 39 ans et sont de culture occidentale.

6 RÉSULTATS

Les données subjectives ont été recueillies via des échelles de Likert à 5 artefacts/items. Leurs valeurs commencent à 1 pour la valeur « pas du tout », puis 2 pour « pas vraiment », 3 pour « moyennement », 4 pour « plutôt » et 5 pour « tout à fait ». À la vue du nombre restreint de participants à l'expérience (12 au total, dont 3 groupes de 4), et aux nombreux chevauchements des écart-types constatés dans les résultats, la significativité générale des résultats n'est pas assurée. De nouvelles expériences avec plus de participants et des protocoles améliorés permettront d'infirmer ou de confirmer les éléments obtenus ici. Nous considérons toutefois que les réponses des participants aux questionnaires nous confortent dans l'idée que le toucher social améliore la communication empathique entre humain et agent. De même, ces résultats apportent d'ores et déjà de précieuses pistes d'amélioration de notre système pour l'intégration du toucher aux interactions sociales humain-agent.

6.1 Première phase: toucher du participant sur l'agent

Tous les participants ont bien identifié l'émotion qu'ils devaient transmettre à partir des scénarios qui leur ont été lus. En revanche, tous les concernés ont exprimé une grande difficulté à décider comment transmettre la tristesse par le toucher.

Etonnamment, et bien qu'ils aient été clairement informés que seul le gant était pris en compte pour la détection du toucher par l'agent, tous les participants ont utilisé au moins une fois un type de toucher jugé inadapté à la réalité virtuelle pour transmettre leur émotion. Dans le cas de la sympathie et de la tristesse et il s'agissait principalement d'enlacement de l'agent virtuel, et dans le cas de la colère les participants ont essayé de pousser ou de secouer l'agent.

Dans l'ensemble, la plupart des participants ont trouvé la réaction de l'agent à leurs touchers peu crédible, notamment parce que ses expressions faciales (sourire pour la sympathie, affaissement des coins des lèvres pour la tristesse, froncement des sourcils pour la colère, à titre indicatif) étaient peu vues ou jugées trop subtiles.

6.2 Deuxième phase: toucher de l'agent sur le participant

Pour ce qui est du toucher de l'agent sur les participants, on peut s'apercevoir que sept des huit participants concernés ont su reconnaître explicitement les vibrations correspondant à une caresse, et plus de la moitié des concernés ont su identifier celles correspondant respectivement au coup et au tapotement. En revanche, le contact simple n'a été mentionné et donc reconnu par aucun des participants concernés.

On remarquera également que si la colère et la sympathie ont été dans l'ensemble plutôt bien reconnues (plus de la moitié des participants concernés), la tristesse n'a quant à elle pas bien été reconnue et souvent confondue avec du réconfort ou de la compassion.

En revanche, lorsque l'agent les a touchés, les participants l'ont dans l'ensemble trouvé crédible ou du moins beaucoup plus crédible que dans les réactions qu'il avait lorsqu'ils l'ont touché.

6.3 Discussion

Il apparaît d'abord que la tristesse n'a pas été réellement considérée comme étant transmissible par le toucher : les participants ont ainsi mis en évidence le fait que lorsqu'ils sont tristes, ils s'attendent plutôt à ce qu'autrui viennent les toucher pour les réconforter ou exprimer leur compassion, mais qu'ils n'iraient pas eux-mêmes toucher quelqu'un. Il semble donc qu'il serait plus pertinent dans le cadre du toucher social de s'intéresser à une émotion de l'ordre du *sorry-for* (être triste pour quelqu'un).

Il est également intéressant de remarquer que si les résultats sont enthousiasmants quant à la reconnaissance des types de toucher simulés à l'aide du manchon vibrotactile, tous les participants ont malgré tout jugé que les vibrations n'étaient pas du tout crédibles pour imiter le toucher naturel.

Une autre chose intéressante et que nous n'avions pas forcément anticipée est que les participants n'hésitent pas à utiliser des types de toucher que nous aurions jugé non adaptés à cette expérience de réalité virtuelle, comme le câlin, les secousses, et autres touchers de cet ordre qui appellent a priori une certaine forme de résistance physique de l'objet touché. C'est-à-dire que les participants se rendent tout à fait compte que seule leur main dominante est suivie et reçoit un retour sensoriel, mais ils essaient malgré tout d'utiliser les types de toucher qui leur sembleraient le plus naturel, dans un contexte réel, à l'expression de l'émotion concernée.

Au vu des éléments évoqués et selon les remarques faites par les utilisateurs, il semble que les vibrations ont dans l'ensemble été jugées très peu crédibles car trop éloignées de la sensation du toucher réel. Lorsqu'on leur a demandé ce qu'ils auraient préféré avoir comme suppléance perceptive, leurs descriptions étaient plus proches de ce qu'on obtient avec des dispositifs à retour d'effort. Un exosquelette restreignant les mouvements de la main serait sans doute plus à même de donner la sensation de heurter ou tenir quelque chose de résistant, de physique.

Toutefois, ces genres de dispositifs impliquent des considérations très différentes quant à la physique de l'environnement 3D qui ne nous semblent pas encore indispensables. D'autres pistes pourraient s'avérer plus fructueuses quant à notre objectif de recherche.

Parmi celles-ci, l'un des participants a évoqué le fait que la sensation vibrotactile lui aurait paru peut-être moins surprenante et plus crédible s'il y avait eu médiation du toucher et du retour sensoriel à travers un objet physique tel qu'une Wiimote (manette de la console de jeu Wii de Nintendo) ou autre dispositif de commande de ce genre, plutôt qu'un gnt. Il aurait alors été moins dissonant de ressentir des vibrations puisqu'elles auraient été

médiatisées par un outil qui n'imité pas le toucher naturel directement.

Cette remarque est particulièrement intéressante si l'on considère que le toucher social est effectivement une modalité d'interaction sociale interpersonnelle déjà relativement rare dans nos interactions quotidiennes, en dehors des relations les plus intimes, et donc potentiellement étrange à utiliser également dans le contexte de l'interaction avec des agents virtuels (surtout lorsqu'il n'y a pas encore de relation entre l'humain et l'agent). Le toucher est en revanche le sens principal par lequel nous interagissons avec les objets, et en particulier les objets techniques et les outils. André Leroi-Gourhan [1] montre la façon dont le redressement de l'homme sur ses membres antérieurs et la main ainsi libérée, qui est notre organe principal pour toucher, ont pu être les éléments déterminants du développement technique, physiologique et cognitif des premiers hommes. Passer par une médiation de notre toucher en environnement virtuel pourrait donc s'avérer une manière pertinente de produire une sensation de toucher social cohérente même avec un retour sensoriel éloigné de celui du toucher naturel, encore extrêmement difficile à reproduire avec les moyens technologiques actuels.

Il reste cependant à déterminer quel genre d'"outil" de médiation pourrait être pertinent pour une interaction en réalité virtuelle et à étudier précisément dans quelle mesure l'utilisation d'un tel intermédiaire pourrait préserver les caractéristiques du toucher social, et dans quelle mesure cela pourra mettre l'humain à distance de l'agent ou non.

7 CONCLUSION

En résumé, notre objectif était ici de déterminer dans quelle mesure il est possible de mettre en place une interaction sociale crédible entre humain et agent conversationnel animé virtuel basée sur le toucher dans un environnement immersif de type CAVE. A travers l'expérience proposée, basée sur l'étude d'un échange de touchers sociaux émotionnels entre un participant et un agent conversationnel animé réagissant de manière appropriée à l'émotion transmise par l'humain puis utilisant à son tour des expressions gestuelles, faciales et un type de toucher approprié pour transmettre une autre émotion. Les questionnaires étaient concentrés sur l'étude de la crédibilité des touchers d'un point de vue sensoriel et sur l'étude de la crédibilité générale du comportement de l'agent par rapport au toucher.

Dans l'ensemble, l'expérience a permis de mettre en évidence qu'il semble effectivement possible de procéder à une communication empathique crédible et intéressante entre humain et agent en passant par une modalité d'interaction tactile. En particulier, il a été mis en lumière que l'utilisation de dispositifs vibratoires permettent à l'utilisateur de reconnaître les types de touchers réels simulés et que dans l'ensemble il est possible de reconnaître les émotions transmises par le toucher par un agent.

Toutefois notre agent ne répond pas encore aux critères évoqués dans la littérature [25] qui le qualifieraient en tant qu'agent social autonome. S'il possède déjà les moyens de percevoir le toucher et de le réaliser, il lui manque la capacité d'interpréter les touchers

reçus et de prendre en compte ces interprétations pour faire évoluer son comportement. Ce sera le rôle de la mise en place d'un modèle de décision adapté que de lui permettre de faire preuve d'autonomie pour intégrer le toucher à ses moyens d'interaction avec l'humain.

REMERCIEMENTS

Ces travaux de recherche sont possibles grâce au financement de l'ANR dans le cadre du projet Social Touch (ANR-17-CE33-0006).

RÉFÉRENCES

- [1] Leroi-Gourhan, André. Le geste et la parole. Paris, A. Michel. Tome I: Technique et langage. 1964, 325 p.
- [2] Anne Cranny-Francis (2011) Semefulness: a social semiotics of touch, *Social Semiotics*, 21:4, 463-481, DOI: 10.1080/10350330.2011.591993
- [3] M. J. Hertenstein, J. M. Verkamp, A. M. Kerestes, and R. M. Holmes. 2006. "The communicative functions of touch in humans, nonhuman primates and rats: A review and synthesis of the empirical research". In *Genetic, Social, and General Psychology Monographs*.
- [4] M. J. Hertenstein, D. Keltner, B. App, B. a. Bulleit, and A. R. Jaskolka. 2006. "Touch communicates distinct emotions". In *Emotion*.
- [5] M. J. Hertenstein, R. Holmes, M. McCullough, and D. Keltner. 2009. "The Communication of Emotion via Touch". In *Emotion*.
- [6] N. Bianchi-Berthouze, and A. Tajadura-Jiménez. 2014. "It's not just what we touch but also how we touch it". In Paper presented at the TouchME workshop, CHI Conference on Human Factors in Computing Systems.
- [7] M. Teyssier, G. Bailly, É. Lecolinet, C. Pelachaud. 2017. « Revue et Perspectives du Toucher Social en IHM ». In *AFIHM. 29ème conférence francophone sur l'Interaction Homme-Machine*.
- [8] Huisman, Gijs and Darriba Frederiks, Aduen and van Dijk, Betsy and Heylen, Dirk and Krose, Ben, The TaSST: Tactile Sleeve for Social Touch (April 14, 2013). IEEE World Haptics Conference 2013, 14-18 April, Daejeon, Korea.
- [9] D.A. Gómez Jáuregui, F. Argelaguet Sanz, A.H. Olivier, M. Marchal, F. Multon, A. Lécuyer, "Toward "Pseudo-Haptic Avatars": Modifying the Visual Animation of Self-Avatar Can Simulate the Perception of Weight Lifting", *IEEE Transactions on Visualization and Computer Graphics (Proceedings of IEEE VR 2014)*, pp.654-661, 2014
- [10] G. Huisman. 2017. "Social Touch Technology". In *IEEE Transactions on Haptic*
- [11] Jan B. F. van Erp, and Alexander Toet. 2013. "How to Touch Humans Guidelines for Social Agents and Robots that can Touch". In *Human Association Conference on Affective Computing and Intelligent Interaction*.
- [12] T. W. Bickmore, R. Fernando, L. Ring, and D. Schulman. 2010. "Empathic Touch by Relational Agents". In *IEEE Transactions on Affective Computing*.
- [13] G. Huisman, J. Kolkmeier, D. Heylen. 2014. "With Us or Against Us: Simulated Social Touch by Virtual Agents in a Cooperative or Competitive Setting". In *International Conference on Intelligent Virtual Agents*.
- [14] Yohanan, S., "The Haptic Creature: Social Human-Robot Interaction through Affective Touch," Ph.D. Thesis, University of British Columbia, 2012.
- [15] N. Nguyen, I. Wachsmuth, and S. Kopp. 2007. "Touch perception and emotional appraisal for a virtual agent". In *Proceedings Workshop Emotion and Computing*.
- [16] Carolina Cruz-Neira, Daniel J. Sandin and Thomas A. DeFanti. "Surround-Screen Projection-based Virtual Reality: The Design and Implementation of the CAVE", *SIGGRAPH'93: Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 135-142, DOI:10.1145/166117.166134
- [17] Ali Israr and Ivan Poupyrev. 2011. "Tactile Brush: Drawing on Skin with a Tactile Display". Disney Research Pittsburgh.
- [18] J. T. Suvilehto, E. Glereana, R. I. M. Dunbar, R. Haria, and L. Nummenmaa. 2015. "Topography of social touching depends on emotional bonds between humans". In *National Academy of Sciences*.
- [19] T. Bänziger, H. Pirker, and K. Scherer. 2006. "GEMEP-Geneva Multimodal Emotion Portrayals: A corpus for the study of multimodal emotional expressions". In *Proceedings of LREC*.
- [20] T. Bänziger and K. R. Scherer. 2010. "Introducing the Geneva Multimodal Emotion Portrayal (GEMEP) Corpus". In K. R. Scherer, T. Bänziger, & E. B. Roesch (Eds.), *Blueprint for affective computing: A sourcebook* (pp. 271-294).
- [21] K. R. Scherer, R. Banse, H. G. Wallbott, T. Goldbeck. 1991. "Vocal cues in emotion encoding and decoding". In *Motivation and Emotion*.
- [22] V. Demeure, R. Niewiadomski and C. Pelachaud. 2011. "How is Believability of Virtual Agent Related to Warmth, Competence, Personification and Embodiment?". In *Presence* (pp. 431-448).
- [23] I. Poggi, C. Pelachaud, F. de Rosis, V. Carofiglio, B. De Carolis. 2005. « Greta. A Believable Embodied Conversational Agent ». In *Multimodal Intelligent Information Presentation*.
- [24] E. de Sevin, R. Niewiadomski, E. Bevacqua, AM. Pez, M. Mancini, C. Pelachaud. 2010. « Greta, une plateforme d'agent conversationnel expressif et interactif ». In *Technique et Science Informatiques*
- [25] Gijs Huisman, Merijn Bruijnes, Jan Kolkmeier, Merel Jung, Adu'en Darriba Frederiks, et al.. "Touching Virtual Agents: Embodiment and Mind." Yves Rybarczyk; Tiago Cardoso; João Rosas; Luis M. Camarinha-Matos. 9th International Summer Workshop on Multimodal Interfaces (eINTERFACE), Jul 2013, Lisbon, Portugal. Springer, *IFIP Advances in Information and Communication Technology, AICT-425*, pp.114-138, 2014, *Innovative and Creative Developments in Multimodal Interaction Systems*.