



HAL
open science

Métaphore d'interaction gestuelle en environnement virtuel : application à l'apprentissage de gestes

Florian Jeanne

► **To cite this version:**

Florian Jeanne. Métaphore d'interaction gestuelle en environnement virtuel : application à l'apprentissage de gestes. Autre [cs.OH]. Université de Technologie de Compiègne, 2017. Français. NNT : 2017COMP2398 . tel-01768757

HAL Id: tel-01768757

<https://theses.hal.science/tel-01768757>

Submitted on 17 Apr 2018

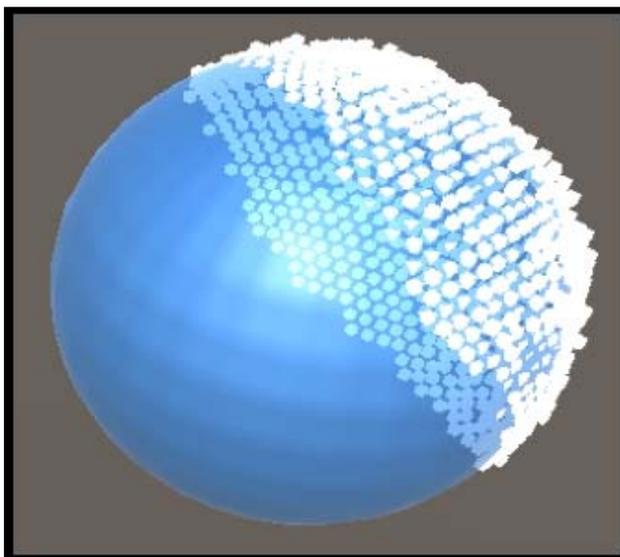
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

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.

Par Florian JEANNE

Métaphore d'interaction gestuelle en environnement virtuel : application à l'apprentissage de gestes

Thèse présentée
pour l'obtention du grade
de Docteur de l'UTC



Soutenue le 13 décembre 2017

Spécialité : Informatique : Unité de recherche Heudyasic
(UMR-7253)

D2398



Région
Hauts-de-France



UNIVERSITÉ DE TECHNOLOGIE DE COMPIÈGNE

LABORATOIRE HEUDIASYC - UMR CNRS 7253

MÉTAPHORE D'INTERACTION GESTUELLE EN ENVIRONNEMENT VIRTUEL : APPLICATION À L'APPRENTISSAGE DE GESTES

THÈSE DE DOCTORAT EN INFORMATIQUE

Présentée et soutenue publiquement le 13 décembre 2017 par

FLORIAN JEANNE

devant un jury composé de :

Rapporteurs	Benoît Bardy	Professeur, Université de Montpellier - EuroMov
	Frédéric Bevilacqua	Directeur de recherche CNRS, UPMC - IRCAM
Présidente	Catherine Pelachaud	Directrice de recherche CNRS, UPMC - ISIR
Examineurs	Dominique Lenne	Professeur, UTC - Heudiasyc
	Daniel Mestre	Professeur, Université de Aix-Marseille - ISM
Directrice	Indira Thouvenin	Enseignant-chercheur, UTC - Heudiasyc
Invité	Sébastien Destercke	Chercheur CNRS, UTC - Heudiasyc

Résumé

Au cours de nos travaux, nous nous sommes intéressés à l'apprentissage de gestes techniques en environnement virtuel. Ces environnements permettent en effet de se former à un geste réalisable dans un environnement réel, en ayant des aides qui ne seraient quant à elle pas disponibles, comme la trace en trois dimensions du geste par exemple, ou la possibilité de voir son geste d'un point de vue différent. Toutefois dans le cadre du guidage de geste, il s'avère que la plupart des métaphores visuelles usuelles ne tiennent pas ou peu compte du problème de dépendance qu'elles génèrent. L'hypothèse de guidage issue des théories sur l'apprentissage moteur, stipule en effet que l'utilisation continue d'aides pendant l'apprentissage génère une dépendance de l'apprenant vis-à-vis de ces aides. Ce dernier n'est alors plus capable de reproduire correctement le geste ou la tâche demandée sans l'aide de ces aides. Néanmoins, en réalité virtuelle l'utilisation d'aides d'apprentissage reste nécessaire pour la formation.

Pour répondre à cette problématique nous proposons une nouvelle métaphore dynamique d'interaction 3D en environnement virtuel, basée sur la modalité visuelle. Notre approche consiste à axer l'apprentissage sur le ressenti de l'utilisateur, en nous basant sur une pédagogie essai-erreur dans le cadre théorique du paradigme de l'énaction. Notre métaphore indique à l'apprenant ses erreurs de trajectoire en temps réel, l'incitant implicitement à se corriger. Notre hypothèse est que lorsque l'aide n'est plus affichée, l'apprenant est capable de reproduire le geste convenablement en se basant sur ses sensations passées. Nous avons conduit une première expérimentation afin de vérifier que notre métaphore permet bien de reproduire un geste présenté au préalable tout en corrigeant les erreurs de trajectoire. Puis, dans un second temps une deuxième expérimentation a démontré que la métaphore permettait d'obtenir de meilleures performances et qu'elle réduisait davantage la dépendance que des métaphores usuelles de guidage.

Mots clés : interaction gestuelle, guidage de geste, réalité virtuelle.

Abstract

In the light of our work, we were interested in precise gesture learning in virtual environment. In these environments, learners can train in safer and longer training sessions than in real environments, using augmented feedback that wouldn't be available otherwise—such as a 3D trace of canonical movements or the opportunity to see your gesture from a different point of view. However, in the context of gesture guidance, most of the usual visual metaphors do not tackle the dependency they generate. The guidance hypothesis, from motor learning theories, states that the continuous use of feedback during early phases of learning leads to a dependency on the feedback. Learners are no longer able to properly reproduce the requested gesture or task without the help of this feedback. However, in virtual environments the use of augmented feedback remains necessary for training.

To deal with this issue, we propose a new dynamic metaphor for 3D interaction in virtual environments, based on the visual modality. Our approach is focus users' training on their intrinsic feedback, by using a trial-and-error strategy. Our metaphor shows users' trajectory errors in real time, implicitly spurring them to correct these errors. We assume that when the metaphor is no longer visible, learners are able to properly reproduce the gesture using their past experiences. We conducted a first experiment to evaluate the capacity of our metaphor to properly guide learners during gesture training. Then, a second experiment showed that the metaphor allowed learners to improve their performance and that it further reduced the dependency than usual guiding metaphors.

Keywords : interaction, gesture guidance, virtual reality

Remerciements

Je tiens tout d'abord à remercier l'ensemble des membres de mon jury de thèse. Je souhaite exprimer mes sincères remerciements aux rapporteurs Benoît Bardy et Frédéric Bevilacqua pour l'intérêt qu'ils ont porté à ce travail. Je remercie Dominique Lenne, Daniel Mestre et Catherine Pelachaud qui ont également accepté de juger ce travail.

Mes remerciements les plus sincères à Indira Thouvenin pour avoir dirigé ces travaux. Je te remercie de m'avoir choisi pour cette thèse et de m'avoir fait confiance tout au long de ces trois années. Merci d'avoir été disponible et pour m'avoir transmis cette rigueur scientifique.

Je remercie également les différents membres du projet KIVA pour ces trois années de collaboration : Sébastien Destercke, Romain Lelong, Mehdi Sbaoui, Patrice Chaptal, Jérôme Martins. Je remercie chaleureusement Sébastien Destercke pour ses conseils et sa disponibilité, et pour toutes les parties réalisées lors des midi-jeux.

Merci à Kévin et Kévin, mes chargés de TD de RV01 pour m'avoir fait découvrir la réalité virtuelle et encourager lors de mon choix de faire une thèse. Merci également à toi Olivia pour les délires lors du projet RV01 (Jaffa, kree!).

Merci aux membres du CEA qui m'ont accueilli lors de mon stage de fin d'études à Cadarache, Delphine Keller, Fabien Ferlay, Céphise Louison, Arnaud Pilia, Kévin Pavy, Clément Coston et Tianchao Hu. Ce stage fut l'élément déclencheur pour mon envie de faire une thèse.

Je remercie bien évidemment les membres du labo (et ceux de Roberval) qui ont égayé ces journées et ces soirées, Rémy, Yann, Samba, David, Benoît, Stéphane, Idir, Freddy, Azzeddine, Ayyoub, Youcef et Subeer ; mes partenaires de bureau Minh Tien, Xuan et Marcio ; et les différents stagiaires du GI et du CI, Dorian, Sarah, Alban et Guillaume. Un grand merci à Rémy et à Yann pour m'avoir encouragé et m'avoir guidé. Votre aide m'a été vraiment précieuse et ces travaux n'auraient pas pu exister sans votre soutien.

Merci à tous mes amis pour votre soutien et l'éternelle question "Alors ta thèse ça avance?". Au moins, je ne l'entendrai plus celle-là.. Merci à mes amis parisiens Barbik, Dany, Emilie, Etienne, Zonkd, Nickun, Marina, Choop, Elodie, Mat', Mike, Alex, Gary, Pyro, Willy, Jojo la patate, Tristan, Guidis et Baptiste, aux néo-bordelais Flora, Benjamin et le p'tit dernier. Merci à mes amis de (ou via) l'UTC, Poupi, Marine, Pierre, Raggy, Gary, Guillaume, Camille, Adri, Lou, Julie, Lucie, Racha, Daphné, Lucas, Pierre, Louis, Gui,

Tom, Jules, Patty, Ali, Catala, Thibault, Louis, Julien, Kévin, Damien, Alexis et à tous ceux du Hand UTC. Merci à Cozian, Marine, Antho et Amandine. Merci pour votre amitié et pour tous ces moments..

Je termine en ayant une pensée toute particulière pour ma famille, mes parents et mon frère Sylvain qui m'ont toujours soutenu et encouragé. Vous m'avez toujours encouragé dans mes choix et vous m'avez permis d'entreprendre des études passionnantes.

Table des matières

Table des matières	9
Table des figures	13
Liste des tableaux	17
1 Introduction	21
1.1 Contexte des recherches	21
1.1.1 Contexte scientifique	21
1.1.2 Contexte applicatif	22
1.2 Apprentissage de gestes	23
1.2.1 Ergonomie et évaluation du geste	24
1.2.2 Différentes approches de l'apprentissage de gestes	25
1.3 Environnements d'apprentissage	27
1.3.1 Environnements virtuels informés	27
1.3.2 La réalité virtuelle pour la formation	28
1.3.3 Les feedbacks	28
1.3.4 Problème de dépendance aux feedbacks	29
1.4 Positionnement de la thèse et problématique de recherche	30
1.4.1 Étude du geste	30
1.4.2 Problématique de recherche	31
1.5 Organisation du manuscrit	32
2 Interaction gestuelle et guidage en environnement virtuel	35
2.1 Caractérisation du geste	35
2.1.1 Qu'est-ce que le geste?	36
2.1.2 Classifications du geste	37
2.1.3 Représentations du geste	41
2.1.4 Analyse biomécanique du mouvement humain	46
2.1.5 Acquisition de gestes	47
2.1.6 Reconnaissance de gestes	49
2.1.7 Conclusion sur la caractérisation du geste	52
2.2 Les interactions gestuelles	52
2.2.1 L'interaction et interaction gestuelle	53
2.2.2 L'interaction gestuelle et le paradigme de l'énaction	58

2.2.3	Conclusion sur les interactions gestuelles	60
2.3	Métaphores d'interaction pour le guidage	60
2.3.1	Généralités sur les feedbacks	60
2.3.2	Guidage non-dynamique	63
2.3.3	Guidage dynamique	67
2.3.4	Conclusion sur les métaphores d'interaction	73
3	EBAGG : une nouvelle métaphore de guidage	75
3.1	Travaux connexes	75
3.2	Modélisation de la métaphore	78
3.3	Calcul de l'erreur	80
3.4	Implémentations	82
4	Evaluation de la métaphore pour le guidage	87
4.1	Introduction	87
4.2	Hypothèses	87
4.3	Expérimentation	88
4.3.1	Tâche	88
4.3.2	Protocole expérimental	89
4.3.3	Participants	90
4.3.4	Dispositif expérimental	91
4.3.5	Evaluation	92
4.4	Résultats	92
4.4.1	Résultats pour le premier geste	94
4.4.2	Résultats pour le second geste	95
4.4.3	Résultats qualitatifs	96
4.5	Discussion	97
4.6	Conclusion	99
5	Evaluation de l'impact de EBAGG sur les performances des utilisateurs	101
5.1	Introduction	101
5.2	Métaphores de guidage choisies	102
5.3	Hypothèses	103
5.4	Expérimentation	105
5.4.1	Tâche	105
5.4.2	Participants	106
5.4.3	Protocole expérimental	107
5.4.4	Dispositif expérimental	109
5.4.5	Evaluation des performances	111
5.5	Résultats	112
5.5.1	Impact de l'entraînement avec et sans métaphore de guidage	112
5.5.2	Analyse de la phase de <i>post-training</i>	117
5.5.3	Analyse qualitative	120
5.6	Discussion	124

TABLE DES MATIÈRES

5.6.1	Groupe contrôle	124
5.6.2	Groupe EBAGG	124
5.6.3	Groupe trace	124
5.6.4	Groupe flèche 3D	125
5.7	Limitations	126
5.8	Conclusion	126
6	Conclusion et perspectives	129
6.1	Retours sur notre contribution	129
6.2	Limitations et Perspectives d'évolution	130
6.3	Applications futures	132
A	Taxonomies	133
A.1	Textes en environnement virtuel	133
A.2	Mouvement en environnement virtuel	134
A.3	Sélection et manipulation d'objets en environnement virtuel	135
B	Questionnaire et résultats de la première expérimentation	137
B.1	Questionnaire	137
B.2	Résultats	140
C	Résultats de la seconde expérimentation	141
C.1	Questionnaire	141
C.1.1	Immersion	141
C.1.2	Interaction	142
C.1.3	Feedback	142
C.1.4	Formation	146
C.1.5	Charge de travail	146
C.2	Résultats	148
	Bibliographie	157

Table des figures

1.1	Environnement d'apprentissage du poste de moulage de Montupet	22
2.1	<i>Continuum</i> de Kendon en lien avec la parole	38
2.2	Classification des gestes libres utilisés en interaction gestuelle	39
2.3	Notation Laban	42
2.4	Benesh : exemple de partition	43
2.5	Conté : exemple de partition	44
2.6	Notation Eshkol-Wachman	44
2.7	Valérie Sutton et la correspondance <i>DanceWriting</i> de sa pose	45
2.8	Exemple de <i>SignWriting</i> écrit en colonnes	46
2.9	Interaction et rétroaction selon Cadoz	54
2.10	Vues exocentrique et égocentrique [Bowman et al., 1999]	56
2.11	Spectre de l'adaptation dans le domaine des interfaces homme-machine	58
2.12	Schéma général de la théorie sensorimotrice de la perception	59
2.13	Modèle étendu de la boucle fermée du contrôle moteur	61
2.14	Métaphore de l'autoroute	64
2.15	Poste de l'apprenant sur le système OSE avec différentes métaphores de guidage	64
2.16	Visualisation en réalité augmentée des zones cibles d'ablation, structures anatomiques critiques et voies d'insertion	65
2.17	<i>Guide spheres</i> pour l'apprentissage de gestes obstétriques	66
2.18	Feedback pour le guidage, succession d'ellipses représentant la trajectoire idéale	67
2.19	Mesure de l'ergonomie d'un poste de travail en environnement virtuel	68
2.20	<i>Gestu-Wan</i>	69
2.21	<i>YouMove</i>	70
2.22	Superposition d'avatars	71
2.23	Vue de l'apprenant. L'avatar de l'instructeur est situé en face de l'apprenant	71
2.24	Les trois types de flèches selon Tönnis	72
2.25	Aides visuelles 3D directionnelles	73
3.1	Guidage multimodal pour le geste	76

3.2	Manipulation d'un bras robotique et guidage visuel projeté sur une surface cachant le bras de l'utilisateur	77
3.3	<i>Hue Cue</i>	78
3.4	La métaphore EBAGG	83
3.5	Schéma de fonctionnement de la métaphore EBAGG	83
3.6	Comportement de la métaphore EBAGG. La trajectoire du geste de référence est représentée en noir, tandis que celle du geste utilisateur est en bleu. La distance les séparant est en rouge	84
4.1	Visualisation des deux gestes à réaliser. Le tracé rouge symbolise la forme du geste.	89
4.2	Différentes phases et nombre d'itérations correspondant pour la réalisation d'un geste	90
4.3	Caméra OptiTrack Prime 13W	91
4.4	Plateforme de l'expérimentation	92
4.5	Comportement de la métaphore EBAGG en fonction de l'écart de performance entre l'utilisateur et la référence. Le geste de référence est représenté par la trajectoire noire, tandis que celui de l'utilisateur est représenté en bleu. Les distances calculées entre deux points de chaque séquence sont symbolisées par les segments en pointillés rouges. Cette figure est présentée plus en détail au chapitre 3 et sert ici de rappel.	93
4.6	Diagramme en barres illustrant la performance des participants au cours du temps lors de la réalisation du premier geste. La phase de <i>training</i> est divisée en deux sous-parties.	94
4.7	Diagramme en barres illustrant la performance des participants au cours du temps lors de la réalisation du second geste. La phase de <i>training</i> est divisée en deux sous-parties.	95
4.8	Réponses des participants aux questions sur le comportement de la métaphore EBAGG.	97
4.9	Réponses des participants à la question : "Arrivez-vous à vous concentrer sur la tâche plutôt que sur l'aide visuelle?"	97
4.10	Réponses des participants à la question : "Avez-vous le sentiment de vous être amélioré dans la réalisation de ces gestes durant cette expérimentation?"	98
5.1	Tracé de la trajectoire du geste de référence à reproduire	103
5.2	Flèche rouge en trois dimensions montrant la direction dans laquelle se diriger pour suivre à nouveau la trajectoire idéale du geste de référence à reproduire	104
5.3	EBAGG et ses particules indiquant la direction et l'intensité de l'erreur	104
5.4	Moule : (1) chape injecteur, (2) appuis noyaux, (3) chape collecteur	106
5.5	Tracé de la trajectoire du geste de référence à reproduire	107
5.6	Feedbacks	108
5.7	Différentes phases de l'expérimentation	109

Table des figures

5.8	CAVE TRANSLIFE	110
5.9	Apprenant réalisant un geste dans le CAVE TRANSLIFE	110
5.10	Wand utilisé pour l'expérimentation associé au modèle 3D de la soufflette	111
5.11	Régression linéaire sur le coût cumulatif moyen des quinze premières itérations pour le groupe de contrôle	113
5.12	Régression linéaire sur le coût cumulatif moyen des quinze premières itérations pour le groupe utilisant la trace comme métaphore de guidage	114
5.13	Régression linéaire sur le coût cumulatif moyen des quinze premières itérations pour le groupe utilisant la flèche 3D comme métaphore de guidage	115
5.14	Régression linéaire sur le coût cumulatif moyen des quinze premières itérations pour le groupe utilisant EBAGG comme métaphore de guidage	116
5.15	Coûts cumulatifs moyens des vingt itérations répartis selon les trois phases d'entraînement et par groupes respectifs	118
5.16	Résultats du questionnaire	120
5.17	Charge de travail totale pour les différents groupes selon l'évaluation NASA-TLX	123
5.18	Scores moyen des différents groupes pour la question sur l'exigence mentale de l'étape 1 de l'évaluation NASA-TLX	123
A.1	Taxonomie de Bowman	133
A.2	Taxonomie de Bowman	134
A.3	Taxonomie de Bowman	135
A.4	Taxonomie de Poupyrev	136

Liste des tableaux

2.1	Intérêts et inconvénients des approches extrêmes d'uniformisation et d'adaptation [Bobillier-Chaumon et al., 2005]	58
5.1	Valeurs de la pente et de l'ordonnée à l'origine 5.1 pour chaque groupe	112
5.2	Coûts cumulés moyens estimés par le modèle et valeurs empiriques obtenues lors de l'expérimentation pour le groupe de la trace du geste	119
5.3	Coûts cumulés moyens estimés par le modèle et valeurs empiriques obtenues lors de l'expérimentation pour le groupe EBAGG	119
B.1	Moyennes et écarts-types du questionnaire de la première expérimentation (* voir tableau B.2)	140
B.2	Résultats de la question 9	140
C.1	Moyennes et écarts-types du questionnaire de la seconde expérimentation (* voir tableau C.2, ** voir tableau C.3, # question non proposée)	148
C.2	Résultats de la question 2	149
C.3	Résultats de la question 23	149
C.4	NASA-TLX : Réponses des questions 27 à 32 pour les participants des groupes Contrôle et Trace 3D	150
C.5	NASA-TLX : Réponses des questions 27 à 32 pour les participants des groupes Flèche 3D et EBAGG	151
C.6	NASA-TLX : nombre d'occurrences de chaque critère (questions 33 à 47) pour les participants des groupes Contrôle et Trace 3D	152
C.7	NASA-TLX : nombre d'occurrences de chaque critère (questions 33 à 47) pour les participants des groupes Flèche 3D et EBAGG	153
C.8	NASA-TLX : charge de travail moyenne du groupe de Contrôle	154
C.9	NASA-TLX : charge de travail moyenne du groupe Trace 3D	154
C.10	NASA-TLX : charge de travail moyenne du groupe Flèche 3D	155
C.11	NASA-TLX : charge de travail moyenne du groupe EBAGG	155

Liste des Algorithmes

1	DTW	81
2	k-DTW	82

Chapitre 1

Introduction

1.1 Contexte des recherches

1.1.1 Contexte scientifique

Nos travaux se positionnent dans le domaine de la réalité virtuelle. Nous nous intéressons plus particulièrement aux environnements d'apprentissage de gestes qui soulèvent un certain nombre de problématiques. En effet, la rencontre de ces deux domaines permet de créer des environnements virtuels centrés sur l'utilisateur en se focalisant sur le couplage "action - perception". L'utilisateur, ou apprenant, interagit dans un environnement qui réagit en retour, par le biais de retours sensoriels (ou *feedbacks*) de natures (visuels, sonores, haptiques, vibro-tactiles) et de temporalités (*a priori*, en temps réel, *a posteriori*) différentes.

L'équipe ICI (Information, Connaissance, Interaction) du laboratoire Heudiasyc de l'Université de Technologie de Compiègne, s'intéresse en partie aux environnements virtuels d'apprentissage, notamment les environnements virtuels informés qui sont définis par Thouvenin comme étant "des environnements virtuels dotés de modèles à base de connaissance dans lequel il est possible à la fois d'interagir et de permettre des comportements par interprétation de représentations dynamiques ou statiques." [Thouvenin, 2009]. Ces environnements permettent donc d'intégrer des modèles de représentation des connaissances permettant de créer des interactions et comportements à base de connaissance.

En effet, l'interaction entre l'utilisateur et l'environnement se base également sur des connaissances décrivant la sémantique des différents objets virtuels, on parle alors d'interaction informée. Cela permet notamment d'adapter les feedbacks proposés en se basant sur les informations disponibles, comme par exemple le contexte de l'interaction (quelle est l'action réalisée?) ou encore le contexte de l'utilisateur (niveau d'expertise, compétences, etc.). Cet apport de connaissances améliore ainsi l'interaction en renforçant le couplage entre l'utilisateur et le système.

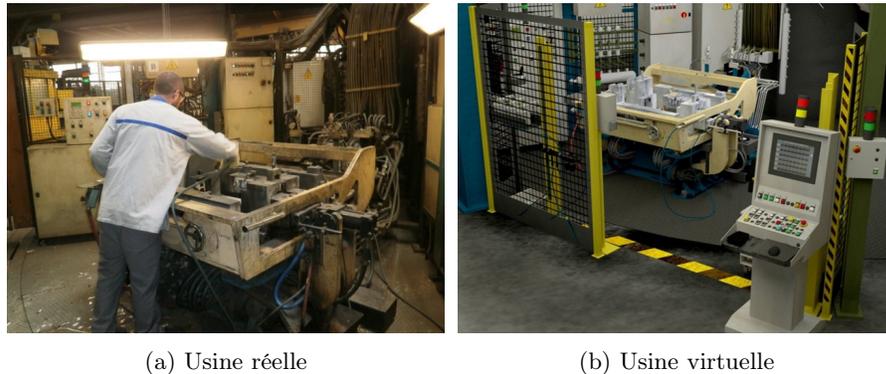


FIGURE 1.1 – Environnement d’apprentissage du poste de moulage de Montupet

1.1.2 Contexte applicatif

Notre recherche s’inscrit dans le cadre du projet KIVA (*Knowledge and Informed Virtual environment for gesture cApitalization*) qui est cofinancé par le Fonds européen de développement économique et régional (FEDER), ainsi que par la région Haut-de-France. L’objectif de ce projet est de créer un nouvel environnement d’apprentissage de gestes techniques (Figure 1.1), viable économiquement, permettant d’accélérer le développement des compétences professionnelles. Il vise également à capitaliser les connaissances gestuelles des opérateurs. Il associe en premier lieu le groupe Montupet, qui est un industriel intégré à la division *Light Metal Casting* du groupe canadien Linamar. C’est l’un des leaders mondiaux dans la conception et la fabrication de pièces de moteurs et de châssis automobile en aluminium. Montupet cherche via ce projet à améliorer son efficacité et à réduire ses taux de rebuts, qui sont des conséquences directes de la formation. En effet, la fabrication des culasses (partie supérieure d’un moteur à pistons), nécessite une haute technicité dans le geste. La société Reviattech quant à elle, est une jeune entreprise innovante dans le domaine de la réalité virtuelle et de la réalité augmentée. Dans ce projet, elle était en charge du développement de l’environnement virtuel et de la partie pédagogique de la formation. Enfin, le partenaire scientifique du consortium pour ce projet KIVA est le laboratoire Heudiasyc (UMR CNRS 7253), à travers cette thèse et d’un postdoctorat. A travers la thèse, nous nous sommes intéressés aux feedbacks de formation et aux problématiques qui y sont liées, tandis que le chercheur postdoctoral Yann Soullard était en charge de la reconnaissance de gestes.

De nos jours, la réalité virtuelle est beaucoup utilisée pour la conception et le développement d’environnements virtuels de formation, notamment pour l’apprentissage d’habiletés motrices comme par exemple dans le sport [Bideau et al., 2010]. Néanmoins, il s’avère que les environnements d’apprentissage de gestes précis (par exemple, les gestes techniques) ont été relativement peu étudiés principalement à cause de verrous technologiques mais également

scientifiques. Le projet KIVA est centré sur certains d'entre-eux. Le premier verrou scientifique du projet est déterminer comment reconnaître un geste technique et de déterminer comment l'intégrer dans l'environnement virtuel en tant que connaissance, ce qui implique de modéliser numériquement ce geste. Le second verrou scientifique est de proposer des retours sensoriels adaptatifs pour une formation basée sur la méthode essai-erreur.

Dans ce chapitre nous présentons les différentes notions qui nous ont permis d'aboutir à notre question de recherche. Tout d'abord nous nous intéressons à l'apprentissage de geste et aux études portant sur le geste (section 1.2). Ensuite dans une deuxième partie (section 1.3), nous présentons plus en détail les environnements virtuels d'apprentissage de gestes. Enfin, dans la section 1.4 nous nous positionnons par rapport aux différentes problématiques soulevées dans les parties précédentes et proposons notre question de recherche, puis nous présentons l'organisation du manuscrit dans la section 1.5.

1.2 Apprentissage de gestes

Nous nous situons dans le domaine de la formation aux gestes. Dans un premier temps, cela implique de tenir compte de la gestuelle de l'apprenant afin qu'il ne se blesse pas. Puis, dans un second temps cela implique de lui fournir les outils et retours appropriés pour que son interaction gestuelle soit la plus naturelle possible. Cadoz définit l'interaction comme étant une notion qui considère deux entités s'influçant mutuellement et où l'on ne peut pas les qualifier d'agissante ou de subissante [Cadoz, 2006]. Une interaction gestuelle, quant à elle, inclut la notion de geste qui est défini comme étant un mouvement physique d'une partie ou l'intégralité du corps visant à communiquer de l'information et du sens [Kurtenbach, 1990, Vatavu et al., 2005].

Ces interactions gestuelles permettent d'utiliser des gestes pour interagir avec l'environnement ou de manière plus générale, le système. Selon Bowman, les interactions gestuelles sont plus naturelles pour l'utilisateur que les interactions dites traditionnelles. En effet, dans [Bowman et al., 2012], les auteurs définissent le naturalisme ou la fidélité d'interaction comme étant le degré objectif, pour lequel les actions utilisées dans une interface utilisateur pour un tâche précise, correspondent aux actions utilisées pour cette tâche dans le monde réel.

We define the naturalism, or the interaction fidelity, of a UI as the objective degree with which the actions (characterized by movements, forces, or body parts in use) used for a task in the UI correspond to the actions used for that task in the real world.

[Bowman et al., 2012, McMahan, 2011]

Pour un utilisateur, une interaction utilisant les mouvements du corps pour interagir avec des objets virtuels lui semblera plus proche de la réalité que les interactions traditionnelles, telles que les commandes prédéfinies via l'utilisation d'une clavier et d'un souris.

Dans notre contexte, il existe principalement deux approches pour aborder l'apprentissage de gestes requérant une haute précision. On peut tout d'abord s'intéresser aux conséquences des gestes. C'est le domaine de l'ergonomie du geste où l'on va chercher à évaluer le geste par rapport à la santé des personnes, notamment dans le milieu industriel. On peut également s'intéresser au contrôle et à l'apprentissage moteur. Dans ce cadre-là, on privilégiera d'avantage le geste en lui-même, la manière dont il est réalisé et les performances obtenues.

1.2.1 Ergonomie et évaluation du geste

L'ergonomie (ou Human Factors) est la discipline scientifique qui vise la compréhension fondamentale des interactions entre les humains et les autres composantes d'un système, et la profession qui applique principes théoriques, données et méthodes en vue d'optimiser le bien-être des personnes et la performance globale des systèmes.

[Falzon, 2004]

L'interaction gestuelle implique comme son nom l'indique une partie ou l'intégralité du corps de l'utilisateur. Certains de ces gestes demandent un effort physique trop important à l'utilisateur, d'autres une utilisation trop prolongée de leurs muscles ou trop fréquente, ce qui peut engendrer des troubles pour la santé. On parle alors de troubles musculo-squelettiques [Simoneau et al., 1996]. L'ergonomie du geste permet entre autre de mettre en avant les impacts négatifs des gestes et détermine ainsi la qualité du geste, notamment en détectant les facteurs de risque des troubles musculo-squelettiques du membre supérieur qui peuvent exister dans les activités professionnelles. Il y a tout d'abord les facteurs biomécaniques (la répétitivité des gestes, les efforts excessifs, le travail statique et les positions articulaires extrêmes) liés à des contraintes soutenues et/ou répétées sur des périodes plus ou moins longues allant de quelques semaines à plusieurs années. Ensuite, il faut prendre en compte les facteurs psychosociaux et le stress et enfin, il existe aussi les facteurs de risque individuels qui sont intrinsèquement liés aux individus (l'âge, la latéralité, le genre ou encore les antécédents médicaux).

Normes d'ergonomie

L'étude de ces troubles est notamment apparu dans le monde industriel. Les ergonomes ont alors analysé les gestes des opérateurs dans les ateliers et usines afin de les protéger dans leur activité professionnelles. Ces études ont abouti à la création de normes qui évaluent les mouvements physiques. David propose dans [David, 2005] une synthèse des méthodes utilisées pour évaluer les facteurs de risque des troubles musculo-squelettiques. Néanmoins, comme le précise Ahmed, "les spécifications d'évaluation des gestes sont soit similaires, soit inspirées des méthodes RULA, ISO et AFNOR" [Ahmed, 2016]. La méthode RULA (*Rapid Upper Limb Assessment*) est par exemple une

méthode d'évaluation des postures de travail. Elle permet d'exposer le risque des postures de travail, forces et charges musculaires pouvant provoquer des troubles musculo-squelettiques [McAtamney and Corlett, 1993].

Evaluation du bon geste en ergonomie

Évaluer un geste, revient à définir la qualité de ce geste et d'en exposer les défauts, notamment les aspects néfastes pour l'utilisateur c'est-à-dire réduire l'impact physique négatif. Différentes méthodes d'évaluation existent comme les questionnaires, les capteurs comme l'électromyogramme et différentes normes. Plus récemment, sont apparues avec la réalité virtuelle différentes applications visant à évaluer ergonomiquement le geste comme *ERGO-WIDE3* et *ErgDesign* (section 2.3.3). Bennes *et al.* ont également proposé la méthodologie *ASAP* qui est un *framework* utilisant la réalité virtuelle pour faire converger les besoins de différents domaines comme l'ergonomie ou la conception de produit, dans un environnement industriel soumis à de fortes contraintes économiques et temporelles [Bennes *et al.*, 2012]. Ce *framework* permet d'analyser une interaction en tenant compte des différents aspects et représentations des facteurs humains de manière collaborative.

En ergonomie, cette évaluation revient donc à caractériser la manière de faire un geste et à déterminer quels sont les risques physiques pour l'être humain résultants de la réalisation de ce même geste.

1.2.2 Différentes approches de l'apprentissage de gestes

L'apprentissage de gestes comprend différentes approches. Nous ne présentons ainsi pas l'ensemble de ces approches, mais nous revenons plutôt sur les deux principales : l'approche cognitive qui correspond à la théorie classique, et l'approche dynamique qui est plutôt contemporaine.

Approche cognitive

Dans le cadre d'une approche cognitive, l'hypothèse du traitement de l'information prédit la présence d'un processus cognitif entre la perception du stimulus et l'établissement de la réponse motrice. Selon Schmidt [Schmidt and Lee, 1988], ce processus est une suite d'opérations qui incluent l'identification de ce stimulus, le choix de la réponse à effectuer, la programmation de cette réponse, sa réalisation et son contrôle. Il définit alors l'apprentissage moteur, ou *motor learning* en Anglais, comme étant un changement durable des habilités motrices provoqué par la pratique et l'expérience [Schmidt *et al.*, 2005]. Il inclut le développement d'un programme moteur (succession d'actions et de poses permettant à un geste d'atteindre son but), ainsi que la réduction progressive des variations du geste via la boucle sensori-motrice et l'utilisation de feedbacks [Shmuelof *et al.*, 2012, Sigrist *et al.*, 2013]. L'apprenant commence donc son apprentissage par une phase où le geste est décomposé en sous-processus nécessaires à la réalisation du geste. Par exemple,

le geste du coup droit au tennis comprend le mouvement du bras, mais également le positionnement des pieds, le fléchissement des genoux, l'orientation des épaules ou encore la tenue de la raquette avec la main. L'ensemble de ces actions et poses forment le programme moteur du geste de coup droit. L'apprenant doit alors s'approprier ce programme moteur avant de se réguler grâce à des feedbacks afin de réduire la variabilité de son geste.

L'apprentissage peut être divisé en trois étapes selon [Fitts and Posner, 1967] [Schmidt and Wrisberg, 2008]. Premièrement, une phase cognitive où l'apprenant se fait une image mentale de la tâche à accomplir et comprend ce qu'il est supposé faire. Ensuite vient une phase associative. Ici l'apprenant comprend comment il doit faire cette tâche, notamment grâce à la détection et à la correction des erreurs qu'il a commises. Enfin, la troisième et dernière étape est la phase autonome où l'apprenant est capable de réaliser la tâche par lui-même, voire de manière automatique.

Approche dynamique

Dans le cadre d'une approche dynamique par contre, les habilités motrices sont considérées comme un ensemble de contraintes qui sont liées à l'organisme, l'environnement ou la tâche elle-même [Newell, 1986]. Ces contraintes jouent alors un rôle déterminant sur la limitation des degrés de liberté du système moteur. Elles ne causent pas les actions réalisées par l'humain, mais en excluent plutôt un certain nombre. En d'autres termes, elles régulent la dynamique du mouvement en restreignant l'ensemble des coordinations possibles [Kelso et al., 1980, Kugler et al., 1980].

L'apprentissage peut alors être défini comme étant le processus de changement de la dynamique de coordination produisant l'émergence du patron à apprendre à partir de l'information comportementale mémorisée correspondante [Schöner et al., 1992]. Ici, un patron désigne une solution à un problème (une contrainte) dans un contexte (l'environnement). L'apprentissage de gestes consiste ainsi à maîtriser les différents degrés de libertés liés à ces gestes. Selon Bernstein [Bernstein, 1967], dans une activité complexe le problème du débutant est de gérer la redondance de ses degrés de liberté. En effet, le système comprend alors plus de degrés de liberté que nécessaire pour réaliser la tâche. La maîtrise du geste correspond donc à une simplification du système en réduisant les degrés de liberté pour ne garder que ceux qui sont nécessaires à la réalisation du geste.

Bernstein divise alors l'apprentissage en trois étapes [Bernstein, 1967]. Premièrement, une phase de contrôle où l'apprenant réduit de lui-même la plupart des degrés de liberté et réalise alors la tâche avec des degrés de libertés résiduels. Puis une phase d'exploration, où cet apprenant relâche les degrés de libertés bloqués précédemment et compense à l'aide d'autres ensembles articulaires. Enfin, la phase d'exploitation correspond à une phase d'optimisation. L'apprenant est plus efficace et économique dans la réalisation de ses gestes. Il est alors capable de situer son geste dans un contexte environnemental.

Conclusions

L'apprentissage de gestes se distingue donc bien de l'étude ergonomique du geste. L'apprentissage s'intéresse davantage à la performance durant la formation, à la manière dont l'apprentissage doit se dérouler afin que le geste appris corresponde bien au geste de référence, tandis que l'ergonomie s'intéresse à la santé et à la sécurité de l'individu sur un poste de travail particulier. Les deux approches sont néanmoins complémentaires et nécessaires pour une formation sur le geste, notamment dans un cadre professionnel où les gestes seront répétés de nombreuses fois durant des années. Dans la prochaine partie nous nous intéresserons aux environnements virtuels d'apprentissage, notamment dans le cadre de l'apprentissage moteur.

1.3 Environnements d'apprentissage

Au cours de nos travaux, nous nous sommes intéressés à l'apprentissage de gestes techniques et aux environnements virtuels d'apprentissage de gestes. Ces environnements permettent en effet de se former à un geste réalisable dans un environnement réel, en ayant des aides qui ne seraient quant à elle pas disponibles, comme la trace en trois dimensions du geste par exemple, ou la possibilité de voir son geste d'un point de vue différent.

1.3.1 Environnements virtuels informés

Selon Bowman, un environnement virtuel est un monde artificiel généralement en trois dimensions, observé depuis une vue à la première personne et contrôlée par l'utilisateur en temps réel [Bowman et al., 2003]. Les environnements virtuels informés (ou EVIs) sont un type particulier d'environnements virtuels. Différentes définitions ont été proposées au cours du temps. Aylett et Luck [Aylett and Luck, 2000] définissent en 2000 les environnements virtuels intelligents comme étant la combinaison de l'intelligence artificielle, de la vie artificielle et de la réalité virtuelle, intégrée dans des créatures ou agents autonomes qui interagissent entre eux. Bowman quant à lui, propose en 2003 la notion d'*information-rich virtual environment* ou IRVE [Bowman et al., 2003]. Les IRVE sont des environnements virtuels réalistes enrichis par l'ajout d'informations abstraites relatives au contenu du monde. Ici, l'information est représentée sous forme de texte. Bowman propose notamment une taxonomie de la représentation des textes dans les environnements virtuels (Annexe A.1). Donikian introduit en 2004 la notion d'environnement virtuel informé. Un EVI est un environnement virtuel dont les modèles 3D contiennent non seulement la géométrie de la scène mais aussi toutes les informations pertinentes pour les entités comportementales à simuler comme les éléments symboliques ou sémantiques qui peuvent permettre à ces entités de percevoir, décider et agir [Donikian, 2004]. Enfin, Thouvenin [Thouvenin, 2009] considère un EVI comme étant un environnement virtuel doté de modèles à base de connaissances dans lequel il est possible à la fois d'interagir et de permettre des comportements

par interprétation de représentations dynamiques ou statiques. Cette définition élargit le concept d'EVI en intégrant l'utilisation de modèles de représentation des connaissances comme :

- les bases de données ;
- les ontologies ;
- les heuristiques ;
- les modèles de raisonnement.

Les EVIs permettent ainsi de représenter les connaissances dans les environnements virtuels. Ces connaissances peuvent être communiquées à l'utilisateur dans le cadre de la formation via les feedbacks en environnement virtuel par exemple. Aussi, certaines informations provenant de ce même utilisateur peuvent également être interprétées comme des connaissances (âge, latéralité, expertise, etc.) et être classifiées par le système. Cela est notamment bénéfique pour les environnements d'apprentissages où les connaissances permettent d'évaluer la qualité du geste de l'apprenant et d'effectuer un retour à l'apprenant afin qu'il se corrige.

1.3.2 La réalité virtuelle pour la formation

Les environnements virtuels sont utilisés depuis de nombreuses années pour la formation aux gestes dans de nombreux domaines comme la santé [Moreau et al., 2007], le sport [Bideau et al., 2010, Covaci et al., 2015], la maintenance industrielle [Louison et al., 2017] ou encore la navigation fluviale [Fricoteaux et al., 2014] (exemples détaillés dans le chapitre 2).

Les apprenants peuvent en effet s'exercer en toute sécurité dans des sessions d'entraînement plus longues qu'en environnement réel. Dans notre cas, c'est un avantage pour des scénarios de maintenance industrielle où les conditions de travail peuvent être difficiles. De plus, la formation peut avoir lieu sur des postes décentralisés, à l'écart des machines utilisées pour la production. Elle ne peuvent en effet plus être endommagées par des utilisateurs novices. De plus, grâce aux EVIs, il est possible de proposer des scénarios différents en fonction de l'utilisateur et des actions qu'il réalise [Carpentier, 2015], de rejouer une situation particulière, de changer de point de vue ou de proposer des feedbacks sous différentes formes [Bowman, 1999, Fricoteaux et al., 2014]. Il est également possible de s'exercer à des situations qui sont difficiles à mettre en place dans un environnement réel comme des situations dangereuses par exemple [Amokrane and Lourdeaux, 2009]. Néanmoins, dans un environnement de formation, il est nécessaire d'effectuer des retours à l'apprenant sur l'exercice qu'il vient de réaliser.

1.3.3 Les feedbacks

Pour l'apprentissage de gestes, un instructeur est nécessaire pour corriger les gestes ou la posture des apprenants. Néanmoins, les environnements virtuels d'apprentissage permettent de réduire sa présence en utilisant des aides

pour l'apprentissage, telles que des techniques de feedforward, qui sont des stimuli proposés *a priori* du geste à l'apprenant, et des feedbacks. Il faut cependant distinguer les feedbacks intrinsèques des feedbacks extrinsèques. Les feedbacks intrinsèques sont inhérents à l'exécution ou au résultat d'un mouvement [Schmidt et al., 2005], on parle de *inherent feedback* en Anglais. Par exemple, vous ressentez le mouvement de vos membres ou lors d'un tir au basket-ball, vous savez que vous avez échoué si le ballon ne rentre pas dans le panier. Les feedbacks extrinsèques (ou *augmented feedback* en Anglais) sont des informations sur le geste ou la tâche réalisée qui complètent ou qui augmentent les feedbacks intrinsèques [Schmidt et al., 2005]. Par exemple, la visualisation numérique de la trajectoire du geste que l'on vient de réaliser. Cette information n'est pas nécessaire et en temps normal non-disponible durant la réalisation d'un geste, mais elle vient enrichir la compréhension du geste à réaliser.

1.3.4 Problème de dépendance aux feedbacks

En général, les feedbacks utilisés en temps réel pendant la réalisation du geste permettent d'améliorer les performances de l'apprenant pendant les sessions d'apprentissage. En outre, il s'avère qu'il est très difficile pour les apprenants de reproduire les gestes appris en conditions réelles, sans aucune aide d'apprentissage. Ceci est expliqué par l'*hypothèse de guidage* qui stipule que l'utilisation de ces feedbacks en continu pendant la phase d'apprentissage génère une dépendance de l'apprenant vis-à-vis de ces feedbacks [Salmoni et al., 1984, Schmidt, 1991, Sigrist et al., 2013]. Ils ne sont plus capable de reproduire correctement ou avec la même précision la tâche motrice demandée sans l'aide de ces feedbacks.

Lors de l'apprentissage d'un geste, les apprenants ont besoin de s'approprier ce geste via le ressenti des feedbacks intrinsèques qu'ils reçoivent, c'est-à-dire la proprioception. Elle est définie comme étant la sensation des positions relatives des membres du corps et le ressenti de l'effort nécessaire à réaliser pour réaliser un mouvement. Or, cette dépendance force les apprenants à ignorer ces sensations [Schmidt and Wulf, 1997] et à se focaliser plutôt sur le feedback. Par conséquent, en conditions réelles lorsqu'il n'y a plus de feedback, ils ont beaucoup de difficulté à reproduire le geste correctement puisqu'ils s'étaient habitués à le faire en présence du ou des feedbacks envers lequel/lesquels ils sont dépendants.

Différentes solutions ont été proposées à ce jour. Bayart *et al.* par exemple ont étudié la variété des feedbacks comme source de réduction de la dépendance [Bayart et al., 2005]. Les résultats de l'expérience ont montré que cette diversification permet de réduire la dépendance contrairement au groupe utilisant le même feedback tout au long de l'apprentissage. Ronsse *et al.* quant à eux se concentrent sur l'utilisation de modalités (visuelle et sonore) différentes pour réduire cette dépendance [Ronsse et al., 2010]. Leur étude neurologique montrerait que l'utilisation de feedbacks sonores plutôt que visuels favoriserait la perception des stimuli proprioceptifs et qu'elle permettrait de réduire la dépendance. Néanmoins, peu de travaux se sont intéressés à la diminution de

cette dépendance en utilisant des feedbacks visuels. Or dans un contexte de formation industrielle soumise à de fortes contraintes environnementales (bruit, chaleur, etc.), il semble que la modalité visuelle soit la plus adaptée.

1.4 Positionnement de la thèse et problématique de recherche

1.4.1 Étude du geste

Nous avons vu qu'il existait différentes manières d'étudier l'apprentissage de gestes, que ce soit du point de vue de l'ergonomie et de la santé au travail ou de l'apprentissage moteur. Il existe néanmoins d'autres méthodes que nous n'avons pas encore abordées dans ce manuscrit comme une étude biomécanique du geste, qui analyse les propriétés mécaniques du mouvement au niveau du tissu osseux, des muscles, des joints et des articulations, ou encore l'étude des gestes pour la parole et le langage des signes. Le projet KIVA ayant pour ambition de créer un environnement virtuel d'apprentissage de gestes techniques, nous avons fait le choix pour cette thèse de nous positionner dans le cadre de l'apprentissage moteur qui vise à évaluer les performances réalisées par l'apprenant. Nous nous intéressons donc que très peu aux autres types d'études cités, sans pour autant les délaisser totalement dans le cadre du projet. L'objectif de la thèse est donc d'étudier l'apprentissage de gestes d'un point de vue moteur en nous basant principalement sur les travaux de Schmidt [Salmoni et al., 1984, Schmidt, 1991, Schmidt et al., 2005].

Le cadre du projet KIVA est un environnement industriel. Comme mentionné précédemment, certaines modalités s'avèrent difficilement exploitables dans un milieu soumis à de fortes contraintes comme la chaleur et le bruit. La formation pourrait certes se dérouler dans un environnement plus clément mais cela ne refléteraient pas nécessairement la réalité et pourrait alors perturber les nouveaux opérateurs lors du passage en conditions réelles qui auraient une nouvelle information à traiter, à savoir ces désagréments.

De plus, nous nous intéressons au problème soulevé par l'hypothèse de guidage de l'apprentissage moteur. Différentes solutions ont été étudiées par le passé dans le domaine des tuteurs intelligents, comme la variété des feedbacks proposés par exemple [Bayart et al., 2005, Frenoy, 2016]. En effet, l'essentiel des réponses trouvées en réalité virtuelle s'intéressent à d'autres modalités des feedbacks ou visent à réduire la fréquence d'utilisation d'un feedback en ne le proposant pas qu'à certaines itérations du geste. Dans le cadre de nos travaux de recherche, nous proposons donc une nouvelle piste de réflexion afin de répondre à cette problématique de la dépendance vis-à-vis de la modalité visuelle en environnement virtuel.

Par ailleurs, il faut également tenir compte de la grande variabilité du geste. En effet, reproduire parfaitement un mouvement n'est pas réalisable à cause des différentes variations physiologiques des systèmes squelettique et neuromusculaire [Hatze, 1986]. Si l'on étudie les gestes de plusieurs personnes devant

reproduire un geste bien précis, il s'avère qu'elles le réaliseront toutes d'une manière différente. Il existe donc une variabilité interpersonnelle du geste. Cependant, si l'on se concentre sur une seule de ces personnes répétant plusieurs fois le geste, on remarque que chaque itération de ses gestes est différente des autres, il y a toujours au moins un infime changement dans sa réalisation. Ainsi il existe également une variabilité intrapersonnelle. Il nous faut alors en tenir compte lors de l'apprentissage et proposer un modèle de formation au geste qui s'adapte à chaque utilisateur.

1.4.2 Problématique de recherche

Notre problématique concerne donc l'interaction et plus précisément le guidage de l'utilisateur pendant l'apprentissage de gestes techniques. Nous souhaitons proposer un guidage visuel qui vise à diminuer cette dépendance aux feedbacks durant l'apprentissage tout en tenant compte de la diversité des utilisateurs.

Comment guider visuellement la gestuelle de l'utilisateur en réduisant la dépendance générée par le feedback, tout en tenant compte de la variabilité du geste ?

Il en résulte quatre grandes questions qui s'axent sur l'interaction entre l'humain et le système informatique.

Caractérisation du geste

Afin d'entraîner un utilisateur à réaliser un geste il faut que le système soit capable de le reconnaître et de l'évaluer.

— Qu'est-ce qu'un bon geste ?

Il nous faut tout d'abord déterminer si l'on doit évaluer la manière dont le geste est réalisé ou si l'on ne doit regarder que le résultat, c'est-à-dire si l'objectif est atteint ou non.

Variabilité du geste

Comme expliqué précédemment, il existe une grande variabilité intrapersonnelle du geste, mais également interpersonnelle. Le système informatique doit être capable de reconnaître le geste réalisé malgré les petites variations possibles.

— Comment doit/doivent se comporter le(s) feedback(s) retourné(s) par le système par rapport aux variations du geste ?

Guidage de la gestuelle

Il existe un certain nombre de feedbacks disponibles pour l'apprentissage moteur en environnement virtuel et issus de différents paradigmes (avatar, trace du geste, aide directionnelle, etc.). Il est ainsi important de déterminer quelle est la méthode la plus appropriée.

- Quelle méthode doit-on utiliser pour guider l'apprenant lors de la réalisation de ses gestes ?

Cette problématique dépend en grande partie des réponses aux questions précédentes. Le guidage sera d'avantage renforcé si l'on s'intéresse au geste dans sa globalité et pas uniquement au résultat final.

Dépendance au feedback

Enfin, l'objectif principal est de notre recherche est de réduire le problème de la dépendance lors de l'apprentissage de gestes.

- Comment limiter au maximum la dépendance générée par les feedbacks lors de l'apprentissage ?

Pour répondre à l'ensemble de ces questions, nous avons proposé une nouvelle métaphore d'interaction visant à guider visuellement en environnement virtuel les gestes des utilisateurs. Au cours de nos travaux, nous avons mené des expérimentations visant à confirmer nos hypothèses sur l'efficacité de notre feedback à améliorer les performances lors d'exercices de reproduction de gestes.

1.5 Organisation du manuscrit

Suite à cette introduction présentant le contexte de nos recherches, nous nous intéressons à la littérature sur l'étude des gestes et l'interaction gestuelle, ainsi que sur les techniques de feedforward et les feedbacks pour l'apprentissage de gestes en environnement virtuel. Ce chapitre 2 est focalisé sur la caractérisation du geste en abordant l'étude du geste et différentes classifications de celui-ci, mais également les différentes représentations et les méthodes de reconnaissance des gestes. Nous précisons également ce qu'est l'interaction gestuelle en abordant notamment le principe de l'énaction dont s'inspire nos travaux. Enfin nous reviendrons plus longuement sur les différents types de feedbacks ainsi que sur les différentes métaphores utilisées pour le guidage de gestes en environnement virtuel.

Le chapitre 3 est consacré à notre proposition de feedback comme métaphore de guidage pour l'apprentissage de geste. Notre approche se base sur l'apprentissage par l'erreur, où l'apprenant va apprendre le geste en le répétant plusieurs fois et en corrigeant ses erreurs au fur et à mesure jusqu'à atteindre la performance du geste souhaitée.

Ensuite, les chapitres 4 et 5 décrivent les deux expérimentations que nous avons menées au cours de cette thèse. La première expérimentation a consisté à évaluer la capacité de notre métaphore de guidage à améliorer les performances des utilisateurs lors de l'apprentissage de gestes. Les participants devaient s'entraîner à réaliser deux gestes (un geste simple et un geste plus complexe) dans une usine virtuelle. Puis, dans une seconde expérimentation nous avons comparé notre feedback à d'autres métaphores de guidage issues de la littérature. Nous avons également comparé ses performances avec celles d'un groupe de

contrôle ne disposant d'aucune aide pendant l'apprentissage afin de déterminer si la répétition des gestes pouvait suffire à améliorer significativement les performances. L'objectif global de cette expérimentation était donc de comparer les performances d'un groupe d'utilisateurs ayant notre métaphore comme aide pour le guidage, avec trois autres groupes ayant des conditions expérimentales différentes. Ces deux contributions expérimentales permettent de valider les hypothèses précédemment énoncées.

Enfin, le chapitre 6 conclut ce manuscrit de thèse. Il propose une synthèse et une discussion des résultats obtenus, identifie les limites de notre travail et propose différentes perspectives d'évolution et d'amélioration.

Chapitre 2

Interaction gestuelle et guidage en environnement virtuel

Afin de répondre aux problématiques soulevées dans l'introduction de ce mémoire de thèse, nous nous sommes intéressés aux approches existantes qui sont liées à l'interaction gestuelle. Nous nous sommes donc tout d'abord focalisés sur le geste, notamment au sens du geste expert dans un contexte bien précis, et plus particulièrement sur sa caractérisation afin de pouvoir répondre à la question "Qu'est-ce qu'un bon geste?". Nous présentons les différentes classifications existantes ainsi que les approches possibles pour la reconnaissance d'un geste. Nous présentons également la distinction entre les différents types de gestes. Dans un second temps, nous nous intéressons à l'interaction et plus particulièrement à l'interaction gestuelle en environnement virtuel. Nous y abordons le problème de la variabilité du geste, puis, nous explorons le domaine des feedbacks. Nous explicitons les travaux de recherche qui ont été effectués sur les différents types de feedback existants, ainsi que sur leurs effets sur l'interaction ou plus particulièrement sur l'apprentissage (dans le cadre de l'apprentissage moteur). Enfin nous étudierons plus en détail différentes métaphores qui peuvent être utilisées pour le guidage, puis plus précisément celles utilisées pour le guidage de gestes.

2.1 Caractérisation du geste

Pour analyser et caractériser le geste, il faut tout d'abord disposer des outils nécessaires pour l'enregistrer, le reconnaître puis comprendre sa signification. Cette section présente ainsi deux types de caractérisation du geste, tout d'abord d'un point de vue sémantique où l'on étudie le sens et l'objectif d'un geste (sous-sections 2.1.1, 2.1.2 et 2.1.3), puis d'un point de vue qualitatif et quantitatif où l'on va chercher à décrire le geste afin qu'il soit perçu correctement et interpréter par le système informatique (sous-sections 2.1.5 et 2.1.6).

2.1.1 Qu'est-ce que le geste ?

Définitions

Dans notre vie de tous les jours, les gestes font partie intégrante de notre communication, qu'ils soient intentionnels ou non. Les gestes peuvent notamment, remplacer la parole, enrichir une information en l'illustrant, ou tout simplement renseigner sur l'état émotionnel du locuteur.

Mouvement du corps, principalement de la main, des bras, de la tête, porteur ou non de signification.

Manière de mouvoir le corps, les membres et, en particulier, manière de mouvoir les mains dans un but de préhension, de manipulation.

Ces définitions du dictionnaire Larousse¹ montrent que le geste n'est pas nécessairement réalisé par la ou les main(s), mais plutôt par une partie quelconque du corps, ou dans certain cas l'ensemble du corps. Les gestes sont au cœur de nos actions. Différentes approches caractérisent et classent les gestes selon leur signification, leur description, leur lien à la parole, leur relation avec l'environnement ou tout simplement selon leur finalité. Ainsi, nous étudierons dans les parties suivantes, différentes classifications afin de mieux comprendre le geste.

Si l'on se place dans le domaine de l'informatique, le geste a été introduit comme modalité d'interaction pour compléter les interfaces traditionnelles, comme le clavier et la souris par exemple. Kurtenbach définit alors le geste en 1990 comme étant un mouvement du corps contenant une information [Kurtenbach, 1990]. Une nouvelle fois la notion du corps est présente et importante. En effet, pour Kurtenbach, dire au revoir en faisant un mouvement latéral avec sa main est un geste, mais l'appui sur une touche du clavier n'est pas un geste puisque dans la finalité de cette action, le mouvement du doigt n'est ni observé, ni important pour le système : "*All that matters is which key was pressed*" [Kurtenbach, 1990]. En réalité virtuelle, le geste permet effectivement de transmettre de l'information à l'environnement virtuel. Néanmoins, comme nous le verrons par la suite, il peut également permettre de capter de l'information en provenance de l'environnement, lors d'une exploration haptique par exemple.

Geste technique

Dans [Bril and Roux, 2002], Bril définit le geste technique comme étant une habileté acquise par apprentissage permettant la réalisation d'une tâche orientée vers un but spécifique. Cet apprentissage peut créer des automatismes permettant le contrôle du geste "à un niveau infra conscient" [Tomás, 2008].

Selon Le Bellu la notion de geste professionnel ou de métier, renvoie à l'idée d'une expertise acquise au cours du temps. Le geste se décompose alors

1. <http://www.larousse.fr>

selon trois composantes qui sont les opérations physiques (manuelles et kinesthésiques), attentionnelles et cognitives [Le Bellu et al., 2010]. Ainsi, le geste technique correspond à la maîtrise d'une action, d'une ressource matérielle ou immatérielle. Le geste professionnel quant à lui, est plus conséquent et plus riche puisqu'il comprend l'histoire d'un métier. Il s'est construit au cours du temps.

En outre, Tomás différencie le geste de métier du simple geste réalisé dans un cadre professionnel [Tomás, 2008]. Il doit comprendre une manière de faire qui est considérée comme acceptable et acceptée par les professionnels du milieu correspondant, pour être appelé geste de métier. Ces manières de faire peuvent alors être apparentées à "une mémoire professionnelle collective" qui dirige et régule l'activité.

2.1.2 Classifications du geste

Le geste peut être classé selon différents critères : sa sémantique où l'on va étudier l'information transmise, sa fonction où l'interaction avec l'environnement prime ou encore sa description. Nous présentons dans cette sous-section ces différentes classifications.

Classification selon la sémantique du geste

Pour comprendre l'information transmise par un geste, Kendon s'est intéressé aux gestes associés à la parole [Kendon, 1988]. En ne considérant que les mouvements des bras, des mains et de la tête, il propose d'analyser ces gestes en fonction de leur symbolique. Il distingue alors cinq types de gestes :

- gesticulation (*gesticulation*) : un mouvement qui accompagne le discours, sans convention de forme particulière ;
- geste encadré par le discours (*speech-framed gestures*) : il fait partie intégrante de la phrase et en est indissociable s'il on veut en comprendre le sens. Par exemple : "Rémy est allé par là [en pointant du doigt]" ;
- emblème (*emblems*) : signe conventionnel admis par une large communauté. Par exemple : le pouce vers le haut, le signe "OK" en formant un cercle avec le pouce et l'index en ayant les trois autres doigts tendus ;
- pantomime (*pantomime*) : geste ou séquence de geste sans paroles ayant une ligne narrative ;
- langage des signes (*signs*) : les gestes permettant une communication linguistique complète sans paroles.

McNeill baptise ces classifications comme étant le "Continuum de Kendon" (*Kendon's Continuum*) dans [McNeill, 1992], et propose une autre classification qui n'est cette fois-ci plus uniquement en lien qu'avec le discours. Il les classe selon la manière dont ils se réfèrent à l'information véhiculée. McNeill propose une classification similaire à celle de de Rimé et Schiaratura [Rimé and Schiaratura, 1991], mais différencie les gestes iconiques des gestes de battements [McNeill and Levy, 1980, Levy and McNeill, 1992] :

2.1. CARACTÉRISATION DU GESTE

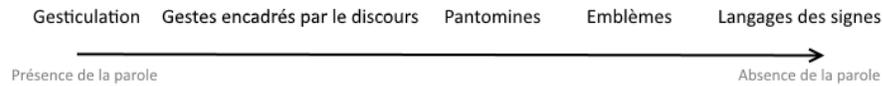


FIGURE 2.1 – *Continuum* de Kendon en lien avec la parole

- les gestes iconiques (*iconic*) : décrivent les caractéristiques d’une entité, un objet ou une action ;
- les gestes métaphoriques (*metaphoric*) : imiter une action pendant le discours ;
- les gestes déictiques (*deictic*) : désignation d’un objet, d’un lieu ou même d’une personne, à l’aide d’une partie de son corps (doigt, main ou encore tête) ou d’un objet tenu ;
- les pulsations ou battements (*beats*) : gestes qui rythment le discours, sans sens propre particulier ;
- les gestes symboliques (*emblems*) : signes conventionnels admis par une large communauté.

Ici, les gestes iconiques et métaphoriques sont liés au discours ou à un objet dans le discours, tandis que les gestes déictiques et symboliques sont en relation avec l’objet directement. Toutefois, seuls les gestes symboliques peuvent être interprétés indépendamment du contexte puisqu’ils véhiculent une grande quantité d’information que les autres types de geste [Billinghurst and Buxton, 2011].

Puis, en 2004, Kendon s’appuie sur la démarche de McNeill et complète son *continuum* en y incluant le rapport à la parole (fig. 2.1). On y voit décroître la présence de la parole en le parcourant de la gesticulation vers la langue des signes. En effet, elle est absente dès la pantomime, mais n’est pas nécessaire pour la compréhension.

En ce qui concerne l’informatique et plus particulièrement l’interaction homme-machine, les classifications sont sensiblement les mêmes mais vont plus loin en se limitant pas qu’à la parole ou au discours. Karam et Schraefel proposent alors la taxonomie suivante [Karam and m. c. schraefel, 2005] :

- les gestes déictiques (*deictic*) : désignation à l’aide du doigt afin d’établir l’identité ou la localisation d’un objet en fonction du domaine d’application (applications d’ordinateur, réalité virtuelle ou smartphones par exemple) ;
- les gestes de manipulation (*manipulative*) : le but est de contrôler une entité en créant une relation ou un lien entre le mouvement (de la main ou du bras) et l’entité manipulée [Quek et al., 2002] ;
- les gestes sémaphoriques (*semaphoric*) : les sémaphores sont des systèmes utilisant des drapeaux, lumières ou les bras d’une personne pour alerter. Ainsi, par extension les gestes sémaphoriques sont des signes codifiés qui peuvent être statiques ou dynamiques [Quek et al., 2002] ;

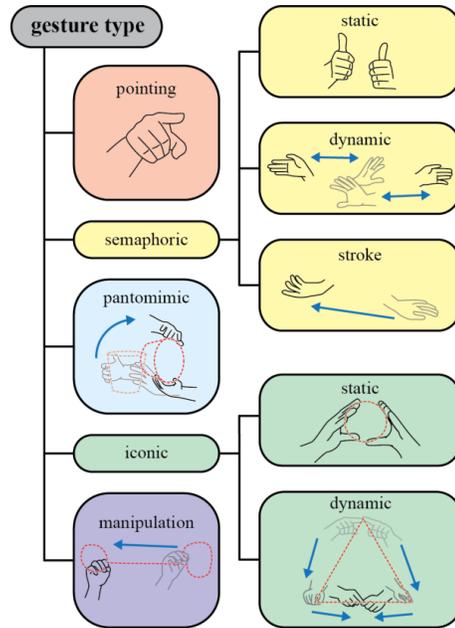


FIGURE 2.2 – Classification des gestes libres utilisés en interaction gestuelle [Aigner et al., 2012]

- gesticulation (*gesticulation*) : mouvement qui accompagne le discours, sans convention de forme particulière [Bolt and Herranz, 1992] ;
- langage des signes (*language gestures*) : les gestes permettant une communication linguistique complète sans paroles ;
- combinaison de différents types de gestes (*multiple gesture styles*) : il est possible de combiner plusieurs des autres catégories. Gestes déictiques et de manipulation [Fisher et al., 1987, Rekimoto, 1997], gestes sémaphoriques et de manipulation [Grossman et al., 2004, LaViola Jr et al., 2001, Ou et al., 2003] ou encore des gestes déictiques, sémaphoriques et de manipulation [Cao and Balakrishnan, 2003, Rekimoto, 2002, Wu and Balakrishnan, 2003].

Cette taxonomie est reprise par Aigner *et al.* dans [Aigner et al., 2012] qui proposent alors une classification des gestes libres (*mid-air*) utilisés dans la plupart des interactions gestuelles (figure 2.2). Ces gestes constituent alors le vocabulaire de l'interaction.

Classification selon la fonction du geste

Cadoz quant à lui, classe les gestes par rapport à leur fonction et à leur rapport avec l'environnement [Cadoz, 1994]. Il distingue notamment différentes fonctions :

- la fonction ergotique qui concerne la manipulation, la transformation ou la modification de l’environnement à travers le changement des paramètres des objets (position, orientation, taille, forme, etc.) ;
- la fonction épistémique qui permet de prendre conscience de l’environnement et de recevoir de l’information de celui-ci (e.g., le cas d’une exploration haptique) ;
- la fonction sémiotique où les gestes transmettent des informations vers l’environnement.

L’utilisateur agit sur l’environnement et celui-ci lui renvoie un retour sensoriel. La majorité des interactions seraient donc de nature sémiotique. Néanmoins, les retours sensoriels semblent jouer un rôle primordial dans le comportement de l’utilisateur, ce qui relèverait plutôt d’une fonction ergotique [Quek et al., 2002]. Le geste ne serait alors plus uniquement un simple moyen de communication unidirectionnel avec le système. Cette classification introduit la notion de relation entre l’utilisateur et l’environnement dans lequel il évolue (environnement virtuel dans notre cas).

Classification selon la description du geste

Enfin, dans certains domaines, comme l’écriture ou le sport par exemple, la description de la forme du geste peut également être un sujet d’étude intéressant. On y distingue la topocinèse et la morphocinèse. [Cole et al., 2002] définissent une théorie nommée *embodied of integrative theory of gesture*, qui définit l’intégration des théories motrices et communicatives du geste. Elle distingue et identifie notamment les rôles joués par les processus moteurs et les processus communicatifs du geste :

- la topocinèse est le contrôle moteur, dépendant principalement de la proprioception ;
- la morphocinèse correspond au contrôle communicatif, lié à la sémantique du geste.

Cependant, Serre [Serre, 1984] a quant à lui défini différemment ces processus. La sémantique du geste n’est plus comprise dans la morphocinèse, mais est rattachée à une troisième catégorie, la sémiocinèse.

- la topocinèse est une activité motrice dont le but est d’atteindre un point de l’espace, et où le moyen est un mouvement pour atteindre un point, comme dans les sports collectifs par exemple ;
- la morphocinèse est une activité motrice dont le but est de produire une forme conforme à un code, et où l’espace, voire le temps, sont des moyens pour produire cette forme. Nous prendrons pour exemple la gymnastique et l’écriture ;
- la sémiocinèse est une forme corporelle pour produire du sens et des émotions chez les spectateurs, comme dans le cadre de la danse.

Comme les précédentes classifications, cette description de la forme du geste reste qualitative et difficilement interprétable par un système informatique. Il s'avère alors nécessaire d'utiliser des systèmes de notation afin de retranscrire le geste numériquement et permettre ainsi au système de le reconnaître et de le comprendre.

2.1.3 Représentations du geste

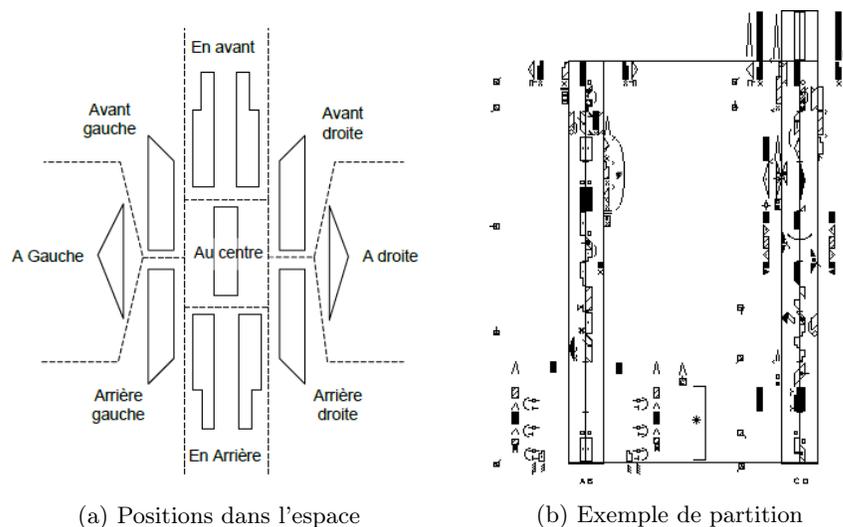
Le geste est considéré dans de nombreux domaines, notamment dans l'industrie, comme une connaissance qu'il faut capitaliser afin que le savoir implicite ne soit pas perdu. De nombreuses méthodes existent pour retranscrire le geste. Différents systèmes de notations sont notamment apparus, dans le cadre de la chorégraphie pour capitaliser la gestuelle et la transmettre. De même pour le langage des signes. L'objectif de ces notations est d'arriver à retranscrire les composantes essentielles du geste : l'espace, le temps et le rythme.

La notation Laban

Également appelée *labanotation*, la notation Laban [Laban and Ullmann, 1971, Loke et al., 2005], permet de transcrire les gestes du corps humain dans le cadre de la danse, à la manière d'une partition de musique. Ce système, créé en 1928, est l'œuvre de Rudolf Laban, et décrit les gestes des jambes et des bras d'un danseur (figure 2.3).

D'autre part, la *Laban movement analysis* est une méthode pour décrire, analyser et interpréter le mouvement humain dans l'espace. Cette méthode d'analyse utilise la notation Laban pour retranscrire et analyser le geste. Elle comprend quatre paramètres : *Motif*, *Effort*, *Shape* et *Structural* [Loke et al., 2005].

- le paramètre *motif* est une simple description de la fonction ou de la raison d'un mouvement, qui décrit les aspects essentiels d'un mouvement dans un contexte précis. Par exemple, cela peut être la description des pas effectués lors d'un ballet, sans représenter les autres aspects du mouvement ;
- le paramètre *effort* (ou l'énergie) d'un mouvement est décrit selon les dimensions suivantes : le poids, l'espace, le temps et la fluidité. Chaque dimension est représentée par deux polarités :
 - poids : léger / fort ;
 - espace : direct / indirect ;
 - temps : soudain / maintenu ;
 - fluidité : lié / libre.
- le paramètre *shape* décrit la mise en forme du corps dans l'espace lors d'un mouvement ;
- le paramètre *structural* fournit la description la plus complète et la plus spécifique d'un mouvement, dans des termes définis et mesurables : les différentes parties du corps, l'espace (direction, niveaux, distances, degrés

FIGURE 2.3 – Notation Laban ²

de liberté), le temps et la dynamique (qualité ou texture, e.g. fort, lourd, élastique, accentué, souligné).

Le système de notation Laban est intéressant dans le sens où il permet de transcrire les mouvements du corps symboliquement. Le système de portées verticales permet notamment de séparer l'ensemble des éléments du corps (jambes, haut du corps, bras, doigts, tête, etc.) dans la largeur. Au delà de la danse, ce système s'est généralisé pour être aujourd'hui utilisé dans différents domaines s'intéressant au mouvement humain.

La notation Benesh

Conçue par Rudolf Benesh en 1955, elle est principalement utilisée pour la danse (figure 2.4). Elle permet, tout comme la notation Laban, de transcrire les mouvements humains sur des partitions [Benesh et al., 1956, McGuinness-Scott, 1983].

La notation Benesh est une partition (voir figure 2.4) contenant des pictogrammes, qui correspondent à des parties précises du corps. Leur succession sur une portée de cinq permet de retranscrire un mouvement de la chorégraphie, et de le lire de la gauche vers la droite. Ces cinq lignes sont directement inspirées du schéma des proportions du corps humain dessiné par Léonard de Vinci, et permettent une représentation formelle du mouvement. En effet, il est possible d'y représenter les déplacements et les positions des différentes parties du corps. De plus, les relations entre les différents danseurs, entre eux ou avec l'environnement, y est facilement retranscrit et perceptible par le lecteur de la

² M. Bastien & E. Mirzabekiantz - <http://notation.free.fr/>

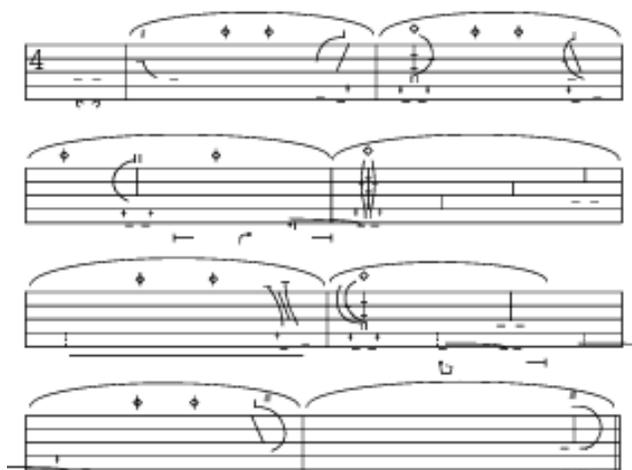


FIGURE 2.4 – Benesh : exemple de partition ²

partition. Elle permet une analyse du mouvement qui n'est cependant pas assez fine. Tout comme la notation Laban, elle s'attarde sur les bras et les jambes de l'individu, et plus particulièrement sur leurs extrémités.

La notation Conté

Cette notation, est l'œuvre de Pierre Conté [Conté, 1954]. Il s'intéressa aux travaux de Laban sur la notation, mais, tout en reconnaissant les progrès accomplis par ce système, n'y adhéra pas. Il proposa donc son propre système de notation, proche du système musical.

Son système est basé sur le solfège traditionnel. Il utilise des portées pour chaque partie du corps : les pieds, le tronc, les bras et la tête. Il ajouta cependant certains pictogrammes qui n'étaient pas présent dans le système musical pour l'adapter au danseur, comme les directions des pas par exemple (voir figure 2.5). Contrairement aux deux notations précédentes, cette notation ne symbolise pas le geste par des pictogrammes ressemblant quelque peu à la réalité, mais bien par des notes musicales.

La notation Eshkol-Wachman

Ce système de notation a été créé en 1958, par la chorégraphe Noa Eshkol et l'architecte Abraham Wachman [Eshkol and Wachmann, 1958]. Contrairement aux notations précédentes, il n'est pas exclusif à la danse et à la notation chorégraphique, mais vise plutôt à annoter les mouvements du corps humain. Il décompose les membres du corps pour les retranscrire sur une partition totalement différente de celles vues précédemment (voir figure 2.6). .

Il est utilisé dans de nombreux domaines comme la danse, la physiothérapie, l'étude du comportement animalier ou encore pour le diagnostic précoce de l'au-



FIGURE 2.7 – Valérie Sutton et la correspondance *DanceWriting* de sa pose³

un premier système de notation chorégraphique sous le nom de *DanceWriting* [Sutton, 1981], [Sutton, 1982]. Ce système intéressa d'autres chercheurs, notamment pour le langage des signes, et fut donc adapté. De nouvelles notations ont donc vu le jour :

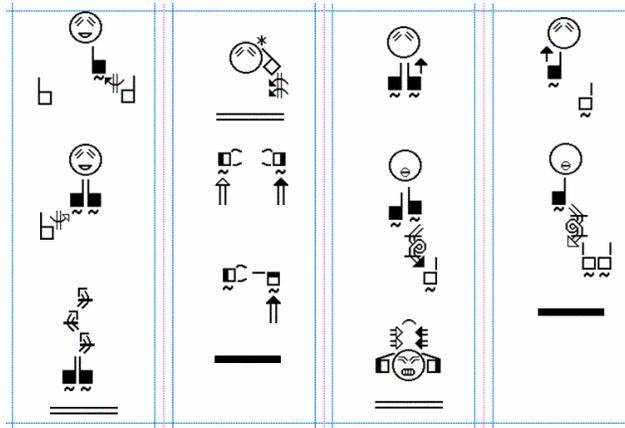
- *SignWriting* qui fut la première notation du langage des signes ;
- *MimeWriting* pour le mime ;
- *SportsWriting* pour le mouvement sportif ;
- *ScienceWriting* pour l'étude du mouvement, pas nécessairement humain, dans le cadre de la science, de l'ergonomie et du médical.

DanceWriting (voir figure 2.7) possède certaines similitudes avec le système de Benesh. Les pieds et les mains du danseur sont représentés sur une portée de cinq lignes, ainsi que différentes parties du corps comme la tête, les épaules, les hanches et les genoux, mais également le sol.

SignWriting (voir figure 2.8) est le premier système d'écriture du langage des signes. Il consiste en une représentation symbolique du visage et des mains, et permet de retranscrire les expressions du visage grâce aux variations de la bouche et des sourcils, et de la position des mains. Des pictogrammes sont aussi utilisés pour représenter les mouvements des mains.

3. <http://www.dancewriting.org/>

4. <http://www.signwriting.org/>

FIGURE 2.8 – Exemple de *SignWriting* écrit en colonnes ⁴

Conclusion

Ces systèmes de notation correspondent à un vocabulaire et une grammaire du geste. Elles sont spécifiques à chaque domaine, mais ne peuvent s'appliquer de façon générique à tous les gestes. Toutefois, le mouvement peut également être décrit selon une analyse biomécanique, expliquant alors d'avantage le fonctionnement du corps humain.

2.1.4 Analyse biomécanique du mouvement humain

Le mouvement humain peut ainsi être décrit selon les propriétés mécaniques du corps ou selon les articulations et segments déplacés. Cette analyse biomécanique de l'organisme est réalisée à l'aide de méthodes cinématiques (mesure des déplacements, vitesses et accélérations des segments du corps) et dynamique (mesure des forces et moments générés par la posture et le mouvement du corps) [Fuchs et al., 2006]. La cinématique directe correspond à la description et au positionnement de chaque articulation afin d'obtenir la pose souhaitée, tandis que pour la cinématique inverse on a la pose désirée et il faut alors déterminer la configuration d'articulations correspondante. De même, la dynamique directe est le calcul des accélérations produites par les efforts et la dynamique inverse est le calcul des efforts lorsque l'on connaît le mouvement.

On y définit également la chaîne cinématique comme étant une modélisation du corps, qui est alors lui-même décrit comme étant un ensemble de corps (ou de segments dans le cas du corps humain) reliés entre eux par des articulations. Les degrés de liberté d'une chaîne correspondent alors aux paramètres nécessaires pour décrire une configuration du mouvement : " *Tous les degrés de liberté définissent une direction de mouvement selon deux sens opposés (i.e. flexion et extension) ; ils sont tous caractérisés par un axe de rotation [...]*" [Boutet, 2008].

Les degrés de liberté articulaires et les paramètres biomécaniques

des segments corporels conditionnent les modalités d'exécution des mouvements corporels, et donc la réalisation des tâches motrices. Déterminer et modéliser l'enchaînement des mouvements des différentes parties du corps dans son environnement naturel favorise leur simulation dans un environnement virtuel

[Fuchs et al., 2006]

Ainsi, dans le domaine de la réalité virtuelle, l'analyse du mouvement humain est principalement utilisée pour définir des modèles du mouvement en temps réel, afin de les appliquer à des personnages virtuels (section 2.3.3) pour simuler un comportement réel de ces avatars. Jauregui *et al.* par exemple étudie dans [Jauregui et al., 2014] comment l'animation visuelle de son propre avatar en temps réel peut être modifiée artificiellement afin de générer différentes sensations haptiques. Leur technique peut notamment être utilisée en réalité virtuelle pour un entraînement à des gestes sportifs ou industriels.

Conclusion

En réalité virtuelle, la représentation biomécanique du mouvement humain permet l'animation d'un avatar. Elle est également à la base des outils d'ergonomie pour protéger l'humain des mauvaises postures et des troubles musculo-squelettiques. Mais elle ne répond pas à la question posée dans nos travaux, c'est-à-dire le guidage de l'apprenant dans le cadre de l'apprentissage d'un geste précis. Nous sommes par contre bien conscients de l'importance de cette analyse qui doit être effectuée avant toute formation.

En outre, avant de comprendre le sens et l'objectif de chaque geste, le système informatique doit être capable de le détecter et de le capturer, puis de l'analyser et de le reconnaître. Nous nous intéressons donc dans la sous-section 2.1.5 suivante à l'acquisition des gestes et à leur reconnaissance dans la sous-section 2.1.6.

2.1.5 Acquisition de gestes

Pour qu'un geste soit reconnu et interprété par le système, il faut tout d'abord en obtenir les données brutes comme la position à chaque instant des objets ou parties du corps dont les mouvements sont suivis, leur orientation, leur accélération, etc. C'est la phase d'acquisition où l'on effectue un premier tri des données en enlevant le bruit. Différentes techniques sont utilisables pour capturer le geste (positions et orientations dans l'espace) :

- capture inertielle : ces capteurs (accéléromètres, gyroscopes, magnétomètres, etc.) permettent d'acquérir le déplacement et les orientations des parties du corps [Vignais et al., 2013] ;
- capture magnétique : ils sont généralement composés d'une base et de capteurs constitués de bobines perturbant le champ magnétique. Le calcul de la différence de potentiel entre la base et les capteurs, permet d'obtenir la position et l'orientation de ces derniers ;

- capture mécanique : l'utilisation d'exosquelette permet d'obtenir avec une grande précision la position et l'orientation du corps via des codeurs angulaires situés sur chaque articulation. Néanmoins, étant une capture assez intrusive pour l'utilisateur, cette méthode est peu adaptée à la capture des mouvements humains ;
- capture optique [Han et al., 2017] :
 - caméras infrarouges et marqueurs passifs réfléchissants : système largement utilisé en réalité virtuelle et dans l'industrie du cinéma, il est composé de caméras projetant des rayonnements infrarouges, qui sont réfléchis par les marqueurs et ainsi de nouveau captés par ces caméras. Les marqueurs sont associés à un objet ou à une partie du corps. Il est alors possible de suivre leurs mouvements et donc d'obtenir leurs positions et orientations en temps réel [Covaci et al., 2015, Louison et al., 2017] ;
 - cellules photosensibles et marqueurs actifs : différents marqueurs constitués de diodes électroluminescentes sont placés sur les objets ou les parties du corps dont on désire suivre les mouvements. Ces marqueurs sont alors détectés par au moins trois cellules photosensibles et permettent alors d'obtenir la position et l'orientation de ces objets ;
 - caméras à temps de vol : la *Kinect* de Microsoft ² par exemple, première version en 2010 puis une seconde en 2014, permet de détecter une ou plusieurs personnes, et de reconstruire leurs chaînes cinématiques. La première version de la *Kinect* utilise la technologie *Light Coding*, une projection de motifs semi-aléatoires via une lumière proche de l'infrarouge, pour déterminer la position des objets ou corps détectés les uns par rapport aux autres [Biswas and Basu, 2011, Smisek et al., 2013, Suma et al., 2011]. La seconde version quant à elle, utilise une technique de mesure par temps de vol. Les différentes caméras illuminent la scène et les objets à l'aide d'une lumière, et calculent le temps qu'elle prend pour effectuer le trajet entre l'objet et la caméra. La distance est alors directement proportionnelle au temps mis pour effectuer ce trajet. Pour chacune des versions, on obtient ainsi un squelette en trois dimensions de chaque personne, nous indiquant les positions et orientations de certaines articulations du corps.

Conclusion

Ces différentes méthodes ont chacune leurs avantages et leurs inconvénients. Les systèmes optiques permettent basés sur l'utilisation de marqueurs passifs ou actifs permettent d'avoir un *tracking* (suivi des mouvements) avec une très grande précision et une latence très faible mais pour un coût élevé, tandis

2. <https://www.microsoft.com>

que les systèmes comme la *Kinect* ont une précision bien inférieure pour un coût assez faible. Les capteurs inertiels ne permettent pas d’obtenir la position des objets, mais uniquement leur déplacement et leur orientation. La capture magnétique corrige le problème d’occlusion des systèmes optiques mais est facilement perturbable par des modifications de champs magnétique.

2.1.6 Reconnaissance de gestes

Une fois le geste capturé, il doit alors être reconnu par le système afin d’être interprété. On cherche alors à identifier le type de geste réalisé ou la séquence de gestes. Il existe différentes méthodes qui ont été étudiées par le passé pour la reconnaissance de gestes [Ke et al., 2013, Lara and Labrador, 2013]. Certaines se basent sur un apprentissage d’un modèle à partir d’exemples de gestes de différents types, on parle de classe de gestes. Par la suite, pour un geste réalisé par un utilisateur, le système lui attribue une classe, c’est-à-dire un geste que le système connaît. D’autres sont un peu plus naïves, et sont utilisées pour déterminer la correspondance entre le geste réalisé et un geste de référence enregistré au préalable.

Nous distinguons dans les deux prochaines sous-parties les approches basées sur les algorithmes de correspondance et celles basées sur un apprentissage statistique ou apprentissage automatique (*machine learning*). Nous y présentons les méthodes les plus courantes en nous basant sur divers états de l’art [Ke et al., 2013, Murthy and Jadon, 2009, Rautaray and Agrawal, 2015].

Algorithmes de correspondance

La manière la plus simple pour reconnaître un geste est de le comparer directement à un ou des exemples de gestes que l’on veut reconnaître. Les algorithmes de correspondance comparent les représentations du geste réalisé avec un ou des exemples de gestes labellisés, c’est-à-dire dont on connaît le type et dont on a une représentation. Ils retournent alors une mesure de similarité comme une mesure de distance par exemple.

L’algorithme de déformation temporelle dynamique (*dynamic time warping* noté DTW) est un algorithme de programmation dynamique mesurant la similarité entre deux séquences temporelles [Alon et al., 2009, Myers and Rabiner, 1981, Ten Holt et al., 2007]. Il permet de trouver l’alignement optimal entre deux séquences temporelles qui peuvent varier en durée ou vitesse [Myers and Rabiner, 1981], et de calculer la distance entre ces deux séries selon cet alignement optimal. Par exemple, les deux séquences temporelles peuvent être les positions au cours du temps du geste de référence et du geste de l’utilisateur. Dans un cas plus général, prenons deux séries temporelles de longueurs N et M $X = (x_1, \dots, x_N)$ et $Q = (q_1, \dots, q_M)$. Le but du DTW est de calculer le coût cumulé $C(n, m)$ entre ces deux séquences. Il est défini tel que :

$$C(n, 1) = \sum_{i=1}^n d(x_i, q_1) \quad (2.1)$$

$$C(1, m) = \sum_{i=1}^m d(x_1, q_i) \quad (2.2)$$

où $d(x_n, q_m)$ est la distance euclidienne entre les deux points x_n et q_m . De plus pour $1 < n \leq N$ et $1 < m \leq M$, le coût $C(n, m)$ est calculé selon la formule suivante :

$$C(n, m) = \min\{C(n-1, m), C(n-1, m-1), C(n, m-1)\} + d(x_n, q_m) \quad (2.3)$$

Les conditions initiales et finales sont telles qu'elles assurent que le processus débute et se termine avec l'alignement des premiers et des derniers éléments de X et de Q . L'équation 2.3 est alors appliquée récursivement sur l'ensemble des observations du geste réalisé. Cet algorithme s'emploie généralement pour de la reconnaissance de gestes en *offline*, c'est-à-dire une fois que le geste est entièrement réalisé. Il existe néanmoins des versions dites de *early recognition* (ou de reconnaissance de gestes incomplets) où l'on va vouloir déterminer quel geste est réalisé avant qu'il ne soit entièrement terminé [Mori et al., 2006, Muscillo et al., 2011, Sabinas et al., 2013].

Il existe aussi la méthode des automates finis ou automates à nombre fini d'états (*finite state machine* noté FSM). Ici, le geste est segmenté en une séquence d'états [Davis and Shah, 1994, Ramey et al., 2011]. Chaque état est défini par un ensemble de caractéristiques (positions, vitesses, etc.) et correspond à une partie du mouvement. Un geste est ainsi reconnu lorsque l'ensemble des états sont séquentiellement atteints [Hong et al., 2000]. Les FSM sont peu coûteux en temps de calcul mais sont malheureusement peu robustes. En effet, cette méthode atteint ses limites avec la variabilité du geste. Si un ou des états ne sont pas atteints, le geste ne sera pas reconnu. Il faut noter cependant qu'il existe également des versions d'apprentissage automatique qui estiment leur représentation des états à partir d'exemples de gestes [Hong et al., 2000, Li and Jarvis, 2009].

Apprentissage automatique et apprentissage statistique

La méthode basée sur les modèles de Markov cachés (*hidden Markov models* noté HMM), qui est une extension des chaînes de Markov, est la plus répandue dans la reconnaissance de gestes [Rabiner, 1989]. Dans cette approche, le geste est représenté comme étant une séquence d'états liés entre eux par des transitions. Contrairement aux chaînes de Markov classiques, les observations et les états sont dupliqués, et les états sont ici inconnus (on parle d'états cachés). Ainsi, un HMM est entraîné à modéliser des probabilités de transitions entre états et des probabilités d'émissions d'une observation dans un état, estimées à partir d'exemples d'apprentissage. La probabilité d'une séquence d'observations X s'exprime comme la somme sur l'ensemble des séquences d'états H possibles des probabilités jointes de X et H :

$$P(X|\Theta) = \sum_H P(X, H|\Theta) = \sum_H P(X|H, \Theta)P(H|\Theta) \quad (2.4)$$

où $X = (x_1, \dots, x_T)$ est une séquence d'observations, $H = (h_1, \dots, h_T)$ une séquence d'états et Θ l'ensemble des paramètres du modèle. L'algorithme de Baum-Welch, s'appuyant sur la procédure *Forward-Backward*, est utilisé pour l'estimation des paramètres du modèle. [Baum and Eagon, 1967]. En outre, les HMM ont besoin d'un certain nombre de données pour l'apprentissage afin de modéliser les gestes [Carmona and Climent, 2012]. Bevilacqua *et al.* ont proposé une version simplifiée du HMM qui permet de faire de la reconnaissance de gestes en temps réel [Bevilacqua et al., 2010]. Ce système a été utilisé dans [Bevilacqua et al., 2007] et permet de faire de la *early recognition* à partir de peu de données d'apprentissage.

Les réseaux de neurones sont souvent utilisés pour des problèmes de classification et d'étiquetage. Ils sont constitués de neurones qui sont regroupés en couches. Chaque entrée d'un neurone est reliée aux sorties de tous les neurones de la couche précédente par des liaisons appelées synapses, auxquelles est associé un poids. Actuellement, deux types de réseaux de neurones se distinguent de l'ensemble des différents modèles : les réseaux convolutifs (ou convolutionnels) sont principalement utilisés dans la reconnaissance d'image [Demuth et al., 2014, LeCun et al., 1990, Hijazi et al., 2015] ou les systèmes de recommandation [Van den Oord et al., 2013] ; les réseaux récurrents sont quant à eux plutôt utilisés pour analyser des séries temporelles comme la parole, l'écriture ou le geste [Murakami and Taguchi, 1991] par exemple. Les réseaux de neurones récurrents peuvent notamment s'utiliser en temps réel pour la reconnaissance de gestes [Bailador et al., 2007] ou de séquences de gestes, néanmoins ils nécessitent un très grand nombre de données pour l'apprentissage.

Enfin, les machines à vecteurs de support (*support vector machine* noté SVM) est un ensemble de techniques assez populaire en *machine learning*. Les SVM sont utilisés pour résoudre des problèmes de discrimination (de l'Anglais *classification*) et de régression. Cette méthode a été employée pour reconnaître des mouvements de la main et de la tête en utilisant une Kinect [Biswas and Basu, 2011], mais également pour les gestes de la main réalisés dans le cadre du jeu *Pierre-feuille-ciseaux* en analysant les images provenant de trois caméras [Chen and Tseng, 2007]. Il existe également une variante pour faire de la *early recognition* qui utilise des *Structured SVM* [Hoai and De la Torre, 2014] où le système s'entraîne sur des événements partiels.

Conclusion sur les méthodes de reconnaissance de gestes

Le DTW et les HMM semblent être les méthodes les plus appropriées pour faire de la reconnaissance de gestes en temps réel. Le DTW se base sur une approche déterministe tandis que les HMM sont basés sur une approche probabiliste. Néanmoins, en ce qui concerne la reconnaissance de gestes, ils

semblent avoir des performances équivalentes [Carmona and Climent, 2012]. En revanche, le DTW n'a besoin que d'un seul exemple de geste pour reconnaître correctement un geste tandis que les HMM ont besoin d'une quantité plus importante de données pour estimer l'ensemble des paramètres [Carmona and Climent, 2012]. Il existe cependant des méthodes qui permettent d'avoir une faible quantité d'exemple pour le HMM (*one-shot learning* [Bevilacqua et al., 2010]), néanmoins il est alors nécessaire de réduire le nombre de paramètres en contraignant le modèle ce qui réduit sa capacité de description. Toutefois, pour le temps de calcul les HMM semblent plus avantageux que le DTW.

2.1.7 Conclusion sur la caractérisation du geste

Dans cette première partie de l'état de l'art, nous avons vu qu'il existait différentes classifications des gestes permettant de décrire leur sémantique ou encore leurs fonctions, ainsi que différentes représentations de ces mêmes gestes. Elles nous permettent de comprendre quelle est la signification du geste et peuvent décrire la manière dont le geste peut être réalisé.

Le geste peut alors être capturé par différentes techniques, la capture optique restant la plus utilisée en réalité virtuelle puisqu'elle permet une grande précision dans les mesures, malgré un coût assez élevé. Une fois le geste numérisé, il existe plusieurs méthodes de reconnaissance, certaines étant basées sur la similitude par rapport à une référence et d'autres étant basées sur un apprentissage à partir d'exemples. Les versions *early recognition* semblent être les plus performantes pour de la reconnaissance en temps réel. L'algorithme DTW ainsi que les HMM correspondent donc le plus à nos besoins au vu de leurs performances respectives en temps réel.

La caractérisation du geste peut être vue comme une classification, une notation, une mesure de paramètres physique, une capture de point dans l'espace ou encore la détection d'une similitude avec une référence connue par le système. Il est alors possible de passer du signal au symbole et de rendre le geste compréhensible par le système de façon à tenir compte de sa variabilité. L'interaction avec l'humain peut alors être possible. Le chapitre suivant traite donc de l'interaction entre l'humain et le système, et plus particulièrement de l'interaction gestuelle.

2.2 Les interactions gestuelles

En introduction de ce manuscrit, nous avons présenté l'interaction comme étant une influence entre deux entités et l'interaction gestuelle comme étant un mouvement du corps, et porteuse de sens (section 1.2) [Cadoz, 2006, Kurtenbach, 1990]. Dans le cadre de notre recherche, cette interaction n'est pas définie entre deux individus, mais bien entre un utilisateur humain et l'environnement virtuel. Nous nous intéressons donc à la manière dont cet humain peut agir sur l'environnement, quelles sont les conséquences de ses actions sur

l'environnement et comment ce dernier réagit. Pour cela nous nous basons notamment sur le paradigme de l'énaction.

La section 2.1 de cet état de l'art présentait la caractérisation du geste et des méthodes pour reconnaître le geste. Ici, nous abordons le geste en tant qu'interaction avec un environnement qui réagit et retourne des feedbacks. Néanmoins, la variabilité du geste étant importante, le système doit donc s'adapter à l'utilisateur afin de lui proposer des retours qui lui correspondent.

Dans cette partie, nous nous intéressons alors à l'humain au sein de l'interaction, notamment au travers des gestes qu'il réalise dans l'environnement virtuel. Tout d'abord, nous distinguons les notions d'interaction et de rétroaction, puis nous revenons sur la problématique de la variabilité du geste dans le cadre de l'interaction gestuelle. Enfin, nous présentons le principe de l'énaction et plus précisément l'interaction basée sur la perception dans un monde éenactif.

2.2.1 L'interaction et interaction gestuelle

Dans le cadre général des interfaces et de l'interaction homme-machine, deux grandes approches se dégagent pour agir avec un environnement : la manipulation directe [Shneiderman, 1993] et les agents conversationnels [Franklin and Graesser, 1996]. Dans la première approche, l'utilisateur agit directement dans l'environnement et contrôle donc ses actions, tandis que dans la seconde, l'utilisateur dialogue avec un agent qui contrôle l'environnement. Dans un cas l'interaction est directe, et dans le second cas elle est indirecte. Avec la manipulation directe, l'utilisateur a le sentiment de contrôler ce qu'il fait, ce qui favorise le sentiment d'immersion, notamment en réalité virtuelle. La première approche est donc largement utilisée dans ce domaine, où l'interaction et l'immersion priment avant tout autre chose. Dans nos travaux, nous abordons donc l'interaction selon l'approche de la manipulation directe.

Interaction et rétroaction

Les rétroactions ou *feedback* en Anglais, sont définies par Kluger & DeNisi [Kluger and DeNisi, 1996], comme étant des actions réalisées par un ou des agents extérieurs, afin de fournir des informations concernant un ou certains aspects de l'exécution d'une tâche.

Feedback intervention is defined as actions taken by (an) external agent(s) to provide information regarding some aspect(s) of one's task performance

[Kluger and DeNisi, 1996]

Cadoz quant à lui, établit une distinction entre rétroaction (figure 2.9b) et interaction (figure 2.9a), qui sont relativement proches.

La notion de rétroaction sous-entend une dissymétrie, un système orienté, avec une "entrée" et une "sortie", puis une "réinjection",

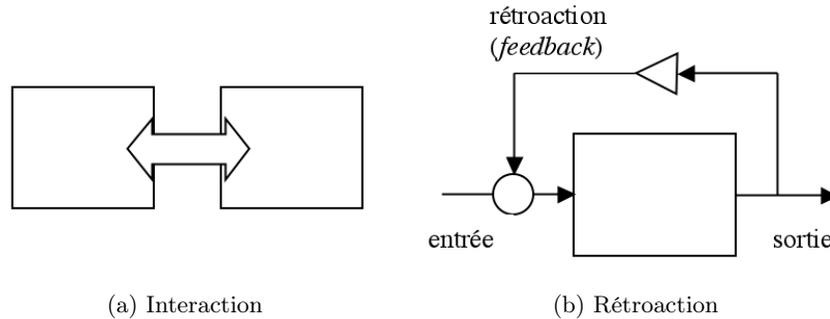


FIGURE 2.9 – Interaction et rétroaction selon [Cadoz, 2006]

additionnée à l'entrée avec un certain coefficient, de ce qui est produit en sortie, tandis que l'interaction considère deux dispositifs s'influençant réciproquement l'un l'autre, sans que l'on puisse considérer que l'un est agissant l'autre subissant

[Cadoz, 2006]

L'interaction peut ainsi être comparée à l'approche systémique du monde où elle est perçue comme une relation entre plusieurs systèmes imbriqués qui échangent de l'information [Cadoz, 2006]. La rétroaction informe alors l'utilisateur sur les résultats de son action. L'information peut comprendre ces résultats à l'état brut ou dans une toute autre forme. Par exemple, lors d'un entraînement chronométré, l'athlète peut avoir le retour sur sa performance en ayant le temps exact ou tout simplement un commentaire de son coach lui indiquant si sa performance était satisfaisante ou non.

Toutefois en réalité virtuelle, une interaction se réalise généralement dans un espace en trois dimensions. Ayant pour optique de guider l'apprenant lors de son apprentissage de gestes, nous nous intéresserons désormais uniquement aux interactions 3D.

Interaction 3D

Pour définir les interactions en environnements virtuels, différentes taxonomies ont été proposées par Bowman, et par Poupyrev et Ichikawa à la fin des années 1990. Bien qu'étant relativement anciennes et incomplètes au vu des nombreuses nouvelles techniques d'interaction qui ont vu le jour depuis, elles demeurent néanmoins des bases solides et pertinentes pour catégoriser l'interaction 3D en environnement virtuel.

En plus de la taxonomie sur les textes en environnement virtuel, Bowman présente dans [Bowman, 1999] deux taxonomies, l'une sur les techniques de déplacement en environnement virtuel, l'autre sur la sélection et la manipulation d'objet 3D. Poupyrev et Ichikawa décident quant à eux de représenter l'interaction non pas comme Bowman en s'axant sur l'action en elle-même, mais plutôt

en proposant une taxonomie basée sur la place de l'humain vis-à-vis de cette interaction [Poupyrev and Ichikawa, 1999].

Taxonomie sur le déplacement La première taxonomie de Douglas A. Bowman est centrée sur le mouvement et plus précisément sur le déplacement de l'utilisateur dans l'environnement virtuel [Bowman, 1999]. Au regard des interfaces sensori-motrices utilisées pour l'interaction, leur utilisation ne permet pas nécessairement le déplacement physique de l'utilisateur. Par exemple, un CAVE est limité par les écrans entourant l'individu, une manette de console de jeux est relié via un câble à la console ou n'est plus détectée à partir d'une certaine distance, un système de capture optique des mouvements est quant à lui restreint par la portée des caméras. Il nous faut donc utiliser les boutons ou autre joystick situés sur le(s) *wand(s)*, ou bien des métaphores d'interaction permettant de se déplacer dans l'environnement virtuel, sans pour autant se déplacer physiquement (ou très peu) dans l'environnement réel.

La figure A.2 de l'annexe A.2 montre une classification du déplacement en fonction du choix de la destination, de la vitesse ou de l'accélération, et des conditions requises pour que le mouvement se fasse. La sélection de la destination peut se faire à l'aide du regard, de gestes ou même d'une interface incluse dans l'environnement virtuel. La vitesse de déplacement peut-être constante, progressive ou même infinie dans le cadre d'une téléportation. Enfin, les conditions d'entrée concernent le moyen de déclencher le déplacement. Est-il continu ? Doit-on appuyer sur un bouton pour commencer à se déplacer et s'arrêter ?

Si l'on prend l'exemple de la téléportation qui est la technique de déplacement mise en avant dans les visiocasques récents (*head-mounted display* noté HMD) comme le HTC Vive³, la sélection de la destination peut s'effectuer via le *head tracking* ou en pointant l'un des *wands* dans la direction souhaité. La sélection de la vitesse est alors automatique puisque le déplacement est instantané, et seul l'appui sur un des boutons du *wand* permet de se téléporter.

Taxonomies sur la sélection et la manipulation d'objets La seconde taxonomie de Bowman sur l'interaction concerne la sélection et la manipulation d'objet [Bowman, 1999]. La figure A.3 de l'annexe A.3 décrit les différentes actions et conditions nécessaires à la prise en main d'un objet virtuel qu'il soit bidimensionnel ou tridimensionnel. Ce type d'interaction est décomposé en trois étapes. Il y a tout d'abord la sélection de l'objet qui peut s'effectuer via des feedbacks de modalités diverses, en indiquant directement l'objet via un système de pointage, le toucher ou un système de menu, ou tout simplement à l'aide des interfaces motrices comme les boutons du *wand*. Ensuite vient la manipulation de l'objet. Cela concerne la position et l'orientation de l'objet, mais également des critères sur la manière dont il est relié virtuellement à l'utilisateur et quels types de feedback cette manipulation retourne-t-elle pour que l'utilisateur la ressente. Enfin, la dernière étape est la libération de l'objet,

3. <https://www.vive.com/fr/>



(a) Vue exocentrique du monde (carte)

(b) Vue égocentrique du monde

FIGURE 2.10 – Vues exocentrique et égocentrique [Bowman et al., 1999]

notamment sur la manière d'indiquer à l'environnement virtuel que l'on désire lâcher l'objet et où il doit se trouver une fois qu'il est séparé de l'utilisateur. Cette taxonomie est donc focalisée sur la manière dont l'action de manipulation se déroule, sur toutes les étapes nécessaires au bon fonctionnement technique de l'interaction.

En outre, Poupyrev et Ichikawa s'intéressent à la place de l'utilisateur humain dans cette interaction et non pas à l'action qui en découle [Poupyrev and Ichikawa, 1999]. Ils distinguent alors les métaphores exocentriques des métaphores égocentriques. Les premières sont portées sur l'extérieur comme leur nom l'indique, c'est-à-dire sur la manipulation de l'environnement depuis l'extérieur sans être directement dans cet environnement. Une vue exocentrique sera par exemple la représentation du monde sous la forme du carte (figure 2.10a). Les métaphores égocentriques sont quant à elle centrée sur l'utilisateur. Une vue à la première personne de l'environnement virtuel par exemple est une vue égocentrique (figure 2.10b).

La figure A.4 de l'annexe A.3 présente donc différentes techniques d'interaction selon le point de vue exocentrique ou égocentrique. Cette taxonomie ne montre que certaines techniques proposées à la fin des années 1990 et est en grande partie incomplète. Elle permet de voir l'interaction d'une manière différente et replace l'humain au centre de l'interaction avec le système. Toutefois, elle ne tient pas compte des différences qu'il peut y avoir entre les utilisateurs, notamment de la variabilité des gestes.

Variabilité du geste

Comme discuté dans l'introduction (section 1.4), le geste comporte une très grande variabilité et présente de nombreux défis dans l'optique de guider le geste de l'apprenant. En effet, dans une approche standard, c'est le système qui impose aux utilisateurs les interactions possibles, par exemple l'utilisation de la téléportation dans une application utilisant un visiocasque au lieu de l'utilisation d'un *joystick* ou du déplacement physique de l'utilisateur. L'inter-

action doit néanmoins pouvoir être intuitive afin qu'elle puisse toucher le plus grand monde. Selon [Jego, 2013], une interaction standard fait appel aux stéréotypes, qui sont définis par [Rich, 1989] comme étant "*le modèle d'un groupe d'individus aux comportements similaires*". Il faut donc arriver à déterminer quel type d'interaction convient le mieux pour notre système. Toujours selon [Jego, 2013], les interactions standardisées sont d'avantage accessibles au plus grand nombre [Bobillier-Chaumon et al., 2005] et permettent un transfert d'apprentissage entre différentes interfaces. Néanmoins, ce type d'interaction, bien que destiné à un usage générique, est assez restrictif et ne tient donc pas compte de la variabilité du geste. Il faudrait donc que le système puisse comprendre les différentes manières de réaliser un geste pour pouvoir s'adapter aux différentes situations et aux différents utilisateurs.

Systems that allow the user to change certain system parameters and adapt their behaviour accordingly are called adaptable. Systems that adapt to the users automatically based on the system's assumptions about user needs are called adaptive

[Oppermann, 1994]

Oppermann qualifie un système d'adaptable lorsqu'il peut être paramétré par l'utilisateur afin de se comporter selon les préférences de ce même utilisateur. Par exemple, certaines personnes préfèrent avoir des feedbacks sur leurs performances pendant qu'elles réalisent l'action, tandis que d'autres voudront avoir des retours uniquement à la fin. Un système adaptable permettrait de choisir entre les deux types de feedback avant de commencer l'exercice. En revanche, il sera adaptatif s'il s'adapte au comportement des différents utilisateurs sans que ces derniers n'aient besoin de paramétrer le système (voir figure 2.11). Ce dernier apporte un confort à l'utilisateur dont les préférences seront déterminées par le système. En outre, cela permet une plus grande flexibilité et une plus grande souplesse en tenant compte du contexte utilisateur, ce qui dans le cadre d'un apprentissage de gestes par exemple, permet de différencier les retours proposés en fonction du niveau d'expertise. Un novice n'aura pas les mêmes feedbacks qu'un expert et les recevra à des fréquences différentes. Frenoy par exemple, propose un système dans le domaine des tuteurs intelligents qui sélectionne les feedbacks les plus pertinents pour l'apprenant, afin de maximiser le gain en matière d'apprentissage en s'adaptant à chaque utilisateur [Frenoy, 2016].

Dans le tableau 2.1, Bobillier-Chaumon *et al.* résument les différents avantages et inconvénients de l'interaction standard ou présentant de l'adaptation. Ainsi, un système adaptatif semble d'avantage correspondre à l'interaction gestuelle, notamment dans le cadre de l'apprentissage de gestes en environnement virtuel. Néanmoins, ce genre de système semble difficile à mettre en place puisqu'il doit pouvoir comprendre les différentes variations d'un même geste et répondre en conséquence. Dans la partie suivante, nous introduisons le principe de l'énaction qui présente une conception cognitive sur la manière dont interagit l'utilisateur dans et avec un environnement.

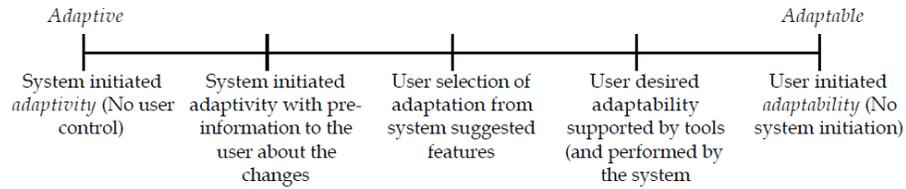


FIGURE 2.11 – Spectre de l’adaptation dans le domaine des interfaces homme-machine [Oppermann and Rashev, 1997]

TABLE 2.1 – Intérêts et inconvénients des approches extrêmes d’uniformisation et d’adaptation [Bobillier-Chaumon et al., 2005]

	Standard	Adaptation
Intérêts	<ul style="list-style-type: none"> — référence communautaire voire culture de masse — transfert d’apprentissage entre applications — réutilisation et interopérabilité — simplicité de mise en œuvre et de maintenance — sûreté et contrôle du processus 	<ul style="list-style-type: none"> — ergonomie, confort, efficacité — interaction personnalisée — flexibilité et souplesse — prise en compte du contexte et situation d’usage
Inconvénients	<ul style="list-style-type: none"> — non prise en compte de la variabilité — décontextualisation — rigidité 	<ul style="list-style-type: none"> — risque d’isolement — complexité de mise en œuvre — effets négatifs lors d’une mauvaise adaptation

2.2.2 L’interaction gestuelle et le paradigme de l’énaction

Le neurobiologiste Varela propose dans ses travaux [Varela et al., 1992] la notion d’énaction qui analyse l’adaptation du comportement d’un individu, en réaction à l’évolution de l’environnement dans lequel il se situe. Ce paradigme est basé sur le principe biologique de l’autopoïèse qui vise à définir l’être vivant. Varela définit ce principe ainsi :

Un système autopoïétique est organisé comme un réseau de processus de production de composants qui régèrent continuellement par leurs transformations et leurs interactions le réseau qui les a produits, et qui constituent le système en tant qu’unité concrète dans l’espace où il existe, en spécifiant le domaine topologique où il se réalise comme réseau. Il s’ensuit qu’une machine autopoïétique engendre et spécifie continuellement sa propre organisation. Elle accomplit ce processus incessant de remplacement de ses compo-

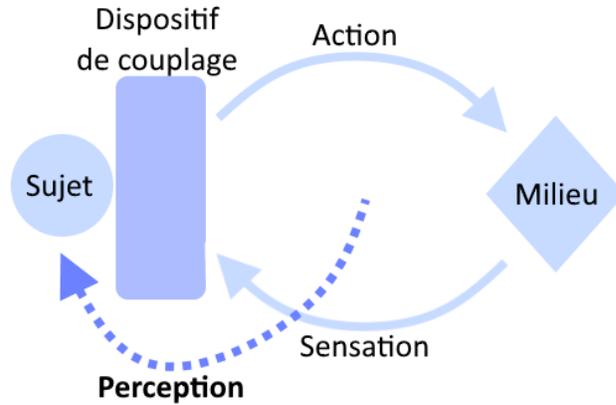


FIGURE 2.12 – Schéma général de la théorie sensorimotrice de la perception [Lenay, 2006]

sants, parce qu'elle est continuellement soumise à des perturbations externes, et constamment forcée de compenser ces perturbations. Ainsi, une machine autopoïétique est un système à relations stables dont l'invariant fondamental est sa propre organisation (le réseau de relations qui la définit).

[Varela et al., 1974, Varela et al., 1989]

Différents travaux en intelligence artificielle ont adopté ce paradigme. Winograd *et al.* par exemple [Winograd et al., 1989], proposent de concevoir leurs intelligences artificielles en utilisant la métaphore du vivant dans la représentation des connaissances, et s'éloignent ainsi de l'orientation rationaliste des modèles de l'époque. Dans le paradigme de l'énaction, l'humain qui interagit avec son environnement via un dispositif de couplage, perçoit des sensations en retour de ses actions, et va donc adapter son comportement et/ou ses gestes en conséquence. Ses actions sont ainsi guidées par sa perception qui, dans cette approche, ne peut exister sans le couplage action-sensation issu de la boucle sensorimotrice de la perception proposée par Lenay (figure 2.12). Ici la perception est au cœur du couplage et non dans le sujet [Lenay, 2006].

Thouvenin rappelle néanmoins que ce n'est pas l'interface qui est éactive, mais bien l'humain qui rend le monde éactif par le biais d'une interface [Thouvenin, 2009]. En outre, De Loor *et al.* soulignent que ce paradigme convient bien au domaine de la réalité virtuelle (et l'intelligence artificielle) puisqu'il y est possible de modéliser une boucle éactive comprenant le système et l'interaction entre celui-ci et l'humain [De Loor et al., 2008].

Nous avons ainsi présenté dans cette section les liens entre l'interaction gestuelle et les feedbacks, tout en nous plaçant dans le cadre de l'énaction qui consiste à renforcer le couplage entre l'utilisateur et l'environnement. Nous

allons donc maintenant nous intéresser aux moyens existants afin de guider les apprenants lors d'une interaction gestuelle en environnement virtuel, pour l'apprentissage de gestes précis. La section suivante explicite ainsi les différents types de feedbacks et expose différentes métaphores d'interaction utilisées pour le guidage, principalement des gestes.

2.2.3 Conclusion sur les interactions gestuelles

Concernant l'interaction gestuelle, nous avons évoqué la notion d'adaptation qui répond à la problématique de variabilité interpersonnelle et intrapersonnelle du geste. Un système adaptatif peut en effet tenir compte du contexte utilisateur et adapter son comportement en fonction de la personne qui interagit avec l'environnement et en fonction de ses gestes. Diversifier les feedbacks et proposer des retours différents selon ce contexte pourrait être une stratégie fructueuse.

Nous plaçons nos recherches dans le cadre théorique de l'énaction qui centre l'humain au sein d'une co-évolution avec l'environnement dans lequel il se trouve. Dans le prochain chapitre, nous étudions quelles sont les solutions possibles pour aider l'apprenant lors de l'apprentissage de gestes, notamment les méthodes de guidage.

2.3 Métaphores d'interaction pour le guidage

L'apprentissage moteur correspond à un changement durable des habilités motrices provoqué par la pratique et l'expérience. La formation aux gestes précis, comme des gestes techniques industriels, requiert de proposer différents scénarios pour apprendre les différentes gestuelles et acquérir le niveau de précision requis.

Dans cette section nous étudions plus en détail les feedbacks dans le cadre de l'apprentissage moteur, notamment lors du guidage visuel. Tout d'abord nous revenons sur quelques définitions et caractéristiques de ces feedbacks, puis dans un second temps nous aborderons les différents types de guidage en environnement virtuel : dynamique et non-dynamique.

2.3.1 Généralités sur les feedbacks

Feedback intrinsèque et feedback extrinsèque

La première distinction à faire concernant les feedbacks est leur provenance, ou plus simplement leur nature. Galligan différencie les feedbacks intrinsèques qui sont des informations reçues par l'apprenant comme résultat direct de la réalisation des gestes, des feedbacks extrinsèques qui sont non inhérents au mouvement mais qui améliorent les feedbacks intrinsèques [Galligan et al., 2000].

Le contrôle moteur est le processus par lequel les humains (ou les animaux par la même occasion) utilisent leurs fonctions cérébrales pour activer et contrôler les muscles et articulations nécessaires lors de la réalisation d'une activité

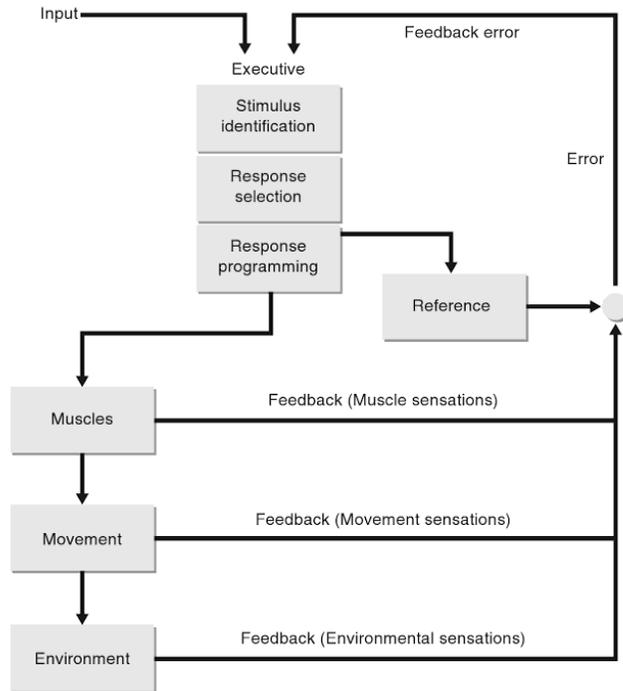


FIGURE 2.13 – Modèle étendu de la boucle fermée du contrôle moteur selon [Schmidt et al., 2005]

motrice (figure 2.13) [Schmidt and Lee, 1988, Schmidt et al., 2005]. Il utilise les différents sens et se base donc sur les feedbacks intrinsèques (aussi appelés *inherent feedback* par Schmidt) pour déterminer la force nécessaire pour faire une tâche. Ces informations proviennent principalement de :

- la vision : permet de déterminer l'emplacement et le mouvement d'objets (incluant nos propres membres) dans le monde extérieur (au corps) ;
- l'audition : nous informe de la nature et la spatialisation des mouvements dans notre environnement ;
- la proprioception : c'est la perception de la position des parties du corps et de la force déployée lors d'un mouvement.

En revanche, les feedbacks dits extrinsèques ou "augmentés" (*augmented feedback*) viennent compléter ces feedbacks intrinsèques. Ils comprennent en général des informations sur les résultats, on parle alors de *knowledge of results* (KR), et/ou des informations sur la qualité du mouvement ou de l'action réalisée, on parle dans ce cas-ci de *knowledge of performance* (KP). Par la suite, le terme feedback fera référence aux feedbacks extrinsèques.

Temporalité des feedbacks

La temporalité des feedbacks correspond au moment auquel les retours sont proposés à l'utilisateur, par rapport à la réalisation de l'action (du geste dans notre cas). Il existe ainsi trois types de feedbacks selon la temporalité : *a priori*, temps réel, *a posteriori*.

Les retours *a priori* sont plus communément appelés *feedforward* en Anglais. Ils fournissent de l'information à l'utilisateur avant que l'action ne soit réalisée. Cette information peut néanmoins persister durant la réalisation de la tâche ou disparaître au début de cette dernière. Dans le cadre du guidage par exemple, cela peut tout simplement être une forme indiquant où aller par la suite, ou bien le tracé du chemin à suivre.

En revanche, les feedbacks en temps réel (*concurrent*) apparaissent et sont proposés durant le mouvement, comme la flèche d'un GPS par exemple, tandis que les feedbacks *a posteriori* sont soumis à l'utilisateur après la fin du mouvement, comme le temps réalisé lors d'une course chronométré par exemple.

Impact des feedbacks selon la complexité de la tâche

Wulf & Shea définissent la complexité d'une tâche de la manière suivante :

We will judge tasks to be complex if they generally cannot be mastered in a single session, have several degrees of freedom, and perhaps tend to be ecologically valid. Tasks will be judged as simple if they have only one degree of freedom, can be mastered in a single practice session, and appear to be artificial

[Wulf and Shea, 2002]

Dans le cadre de l'apprentissage de mouvements simples, les feedbacks peuvent avoir soit un rôle descriptif en informant l'apprenant des erreurs commises, soit un rôle prescriptif en lui indiquant comment corriger ses erreurs [Tzetzis et al., 2008]. Si la tâche demandée est plus complexe et difficile, il est plutôt recommandé de proposer des feedbacks incluant des informations à la fois sur les erreurs et leur correction, afin d'améliorer les performances [Kernodle et al., 2001]. Cependant, les informations de correction d'erreur semblent avoir un rôle plus important pour des tâches complexes que pour des tâches simples [Tzetzis et al., 2008].

Dans leur *review*, Sigrist, Rauter, Riener & Wolf [Sigrist et al., 2013] exposent les résultats des travaux de différents chercheurs concernant l'apprentissage de tâches simples et de tâches complexes. Dans le cas des tâches simples, les feedbacks en temps réel ont tendance à être plutôt préjudiciables pour l'apprentissage puisqu'ils complexifient inutilement la tâche demandée. Cependant, comme le montre la méta-analyse de Marschall, Bund & Wiemeyer [Marschall et al., 2007], ce n'est pas nécessairement le cas pour les tâches complexes, notamment sportives : "*By tendency, it seems that the more complex the task, the more the trainee can profit from concurrent feedback*" [Sigrist et al., 2013]. Plusieurs raisons semblent l'expliquer. Tout

d'abord, les feedbacks en temps réel procurent à l'apprenant un autre centre d'attention durant l'apprentissage, qui est alors externe à la tâche réalisée [Shea and Wulf, 1999]. Or, Wulf montre dans sa *review* sur l'attention lors de l'apprentissage moteur [Wulf, 2007] que cela semble être bénéfique : "[...] *an external focus promotes automaticity in movement control*". Une autre raison, évoquée par Wulf & Shea [Wulf and Shea, 2002], est que les feedbacks en temps réel permettent de réduire la surcharge cognitive. En effet, lorsque des tâches complexes sont réalisées pour la première fois, la mémoire à court terme peut être surchargée, entraînant une diminution des performances de l'apprenant. La réduction de cette surcharge permettrait donc d'améliorer l'apprentissage de tâches complexes.

Pour le guidage de gestes précis dont la difficulté est la haute précision demandée, la diminution de la charge cognitive générée par l'apprentissage et de la complexité du geste permettrait aux apprenants d'appréhender plus facilement le geste à réaliser, afin de s'intéresser plus rapidement à l'amélioration de la précision du geste. Il semble donc intéressant d'utiliser des feedbacks en temps réel pour le guidage.

Par la suite, nous distinguons les métaphores dynamiques des métaphores non-dynamiques dans le cadre du guidage. Ces dernières sont principalement des techniques de *feedforward* et permettent de représenter le geste en affichant sa trace par exemple, qui est la manière la plus simple et la plus précise pour indiquer la trajectoire à suivre.

2.3.2 Guidage non-dynamique

Si l'on élargit notre étude au guidage global et non plus seulement au guidage du geste, il s'avère que la suggestion de la trajectoire idéale a été longuement étudiée dans le domaine du transport. Dans le cadre de la navigation maritime par exemple, cette trace du chemin à suivre peut être affichée à l'aide de multiples traces bordant ainsi la route à suivre. Bjorneseth *et al.*, ainsi que Benton et Walker par la suite, ont utilisé cette technique, appelée *Highway metaphor* (ou métaphore de l'autoroute) dans une application de réalité augmentée à la manière d'un HUD (figure 2.14) [Benton and Walker, 2004, Bjorneseth, 2003].

Pour le transport fluvial, Fricoteaux propose un environnement de formation en réalité virtuelle [Fricoteaux, 2012]. Ici, le module décisionnel a pour objectif de renforcer le couplage entre l'humain et l'environnement. Pour cela, il adapte les retours multimodaux envoyés à l'utilisateur en fonction du comportement de celui-ci. Ils peuvent prendre la forme d'aides visuelles, de messages sonores et de simplifications des conditions de navigation afin de l'aider à anticiper les manœuvres à effectuer. Il utilise notamment la métaphore de l'autoroute pour indiquer les différentes voies possibles (tracé noir sur la figure 2.15), mais également un système de surbrillance de la signalétique d'alignement (cercles bleus sur la figure 2.15) ou encore l'indication de la future position de la péniche (rectangle bleu positionné sur un autre rectangle vert - figure 2.15).

2.3. MÉTAPHORES D'INTERACTION POUR LE GUIDAGE

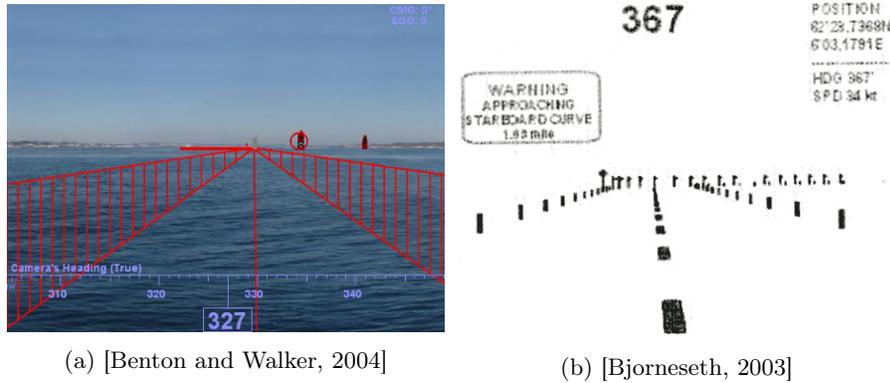


FIGURE 2.14 – Métaphore de l'autoroute

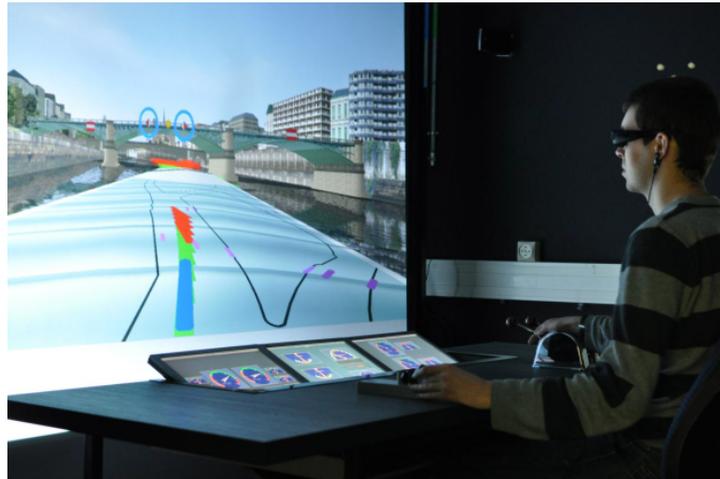


FIGURE 2.15 – Poste de l'apprenant sur le système OSE avec différentes métaphores de guidage [Fricoteaux et al., 2014]

En s'intéressant de nouveau au geste précis, il semble naturel de porter notre attention sur les gestes dans le domaine du médical. En chirurgie, plusieurs méthodes de réalité augmentée ont été utilisées pour aider les chirurgiens lors de leurs opérations. Wen *et al.* projettent sur le corps d'un mannequin des motifs représentant par exemple les modèles anatomiques [Wen et al., 2013] [Wen et al., 2014]. Le chirurgien peut grâce à certains gestes de la main définis au préalable, interagir avec le système et modifier les paramètres des modèles projetés comme la transparence ou l'orientation par exemple, et ainsi détecter certaines structures plus clairement. Dans le cas de l'exercice d'insertion d'aiguille de radio-fréquence, cela permet alors au chirurgien d'apprendre à la positionner correctement à l'aide d'un bras robotique, selon différentes struc-

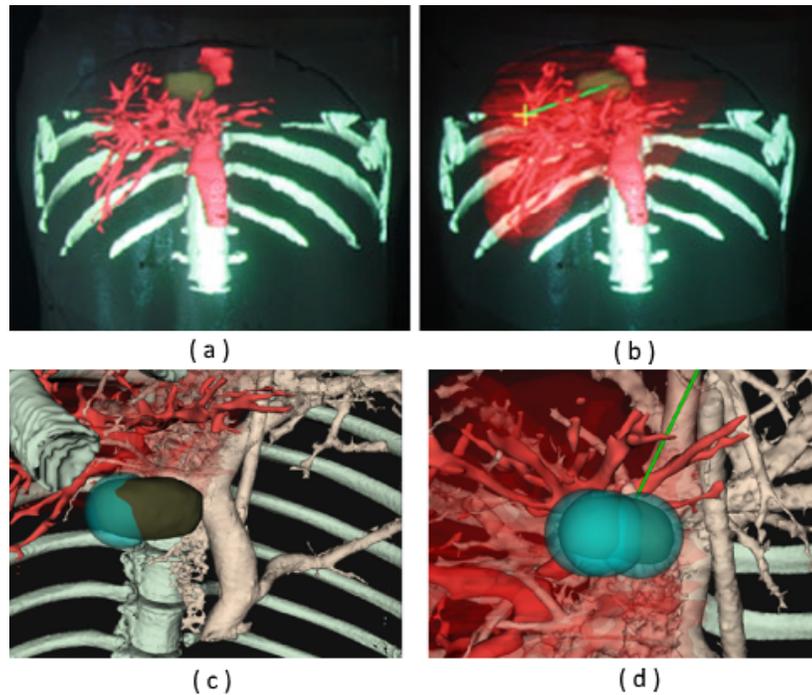


FIGURE 2.16 – Visualisation en réalité augmentée des zones cibles d'ablation, structures anatomiques critiques et voies d'insertion [Wen et al., 2013] [Wen et al., 2014]

tures anatomiques (figure 2.16a). Aussi, lorsque l'insertion de l'aiguille débute, le point d'insertion préétabli et la trajectoire actuelle d'insertion sont mis en évidence sur le mannequin pour guider le chirurgien (figure 2.16b).

Moreau *et al.* [Moreau et al., 2008], ont proposé une métaphore visuelle, appelée "*guide spheres*", dans le cadre de l'apprentissage de gestes obstétricaux avec des forceps. L'expérience se focalisait sur le placement des forceps sur la tête du fœtus, une étape cruciale pour minimiser les risques liés à leur utilisation. Pour avoir un geste de qualité, la formation ne se focalisait pas sur la durée du geste de l'apprenant dans un premier temps : "*the duration of the gesture should not be taken into account during the gesture analysis as precision is more important than time at the beginning of the training.*" [Moreau et al., 2008]. Lors des premières étapes de l'apprentissage, les apprenants réalisaient le geste sans aucune indication de l'expert. Celui-ci n'intervenait qu'au bout d'un certain temps, en leur montrant le geste correct, ainsi que l'interface du simulateur. Pour reproduire le geste de référence, les apprenants devaient traverser des sphères concentriques, de rayons différents (1cm, 2cm et 3cm), lors de l'apprentissage du placement des forceps (figure 2.17). Les sphères, au nombre de cinq, représentaient les positions de points particuliers du chemin de référence

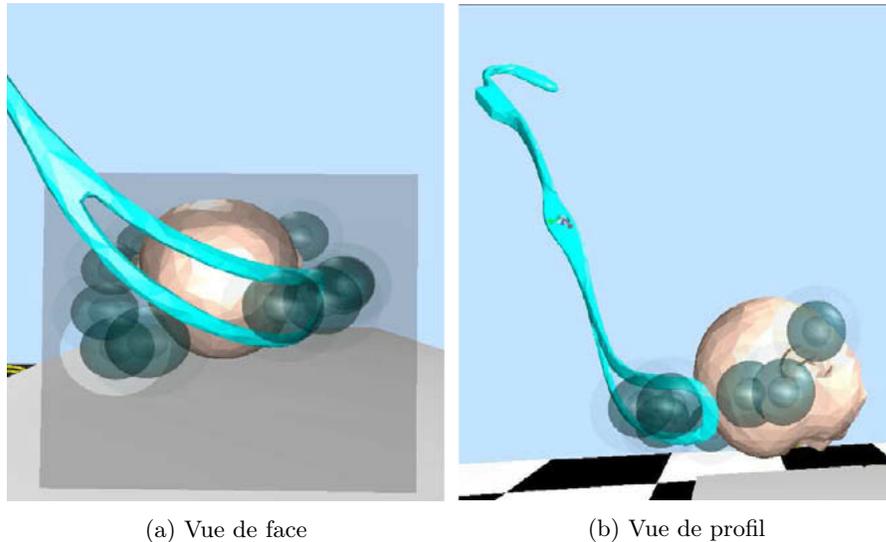


FIGURE 2.17 – *Guide spheres* pour l'apprentissage de gestes obstétriques [Moreau et al., 2008]

défini au préalable par l'expert. Le but pour les apprenants était de faire passer les lames des forceps le près possible des centres de ces sphères pour se rapprocher le plus possible de la performance de référence. Pour comparer les données des apprenants à ceux de l'expert, les auteurs comparaient les distances euclidiennes des trajectoires.

Concernant le sport et plus particulièrement le basket-ball, Covaci *et al.* ont proposé un simulateur de lancer franc⁴ [Covaci et al., 2014, Covaci et al., 2015]. Ils ont étudié l'impact de l'entraînement en environnement virtuel sur les performances selon trois types de conditions expérimentales : en vue à la première personne, en vue à la troisième personne et à la troisième personne avec des métaphores de guidage. De la même manière que [Moreau et al., 2008] précédemment, les auteurs proposent d'afficher des ellipses le long de la trajectoire idéale (figure 2.18). L'utilisateur doit alors faire passer le ballon de basket par le centre de chaque sphère pour avoir la performance optimale. De plus, ce système tient en partie compte du contexte utilisateur puisque utilise la taille de ce dernier pour déterminer la position de la première ellipse. Toutefois, il ne tient pas compte de la variabilité et des différents types de lancer possible.

Les métaphores de guidage non-dynamique permettent de visualiser directement la trace du geste à reproduire. Néanmoins ce manque de dynamisme ne semble pas être adéquat pour répondre à la variabilité du geste et globalement ne s'adapte donc pas à chaque utilisateur. Dans la prochaine partie, nous étu-

4. Pénalité accordée à un joueur ayant subi une faute lors d'un tir

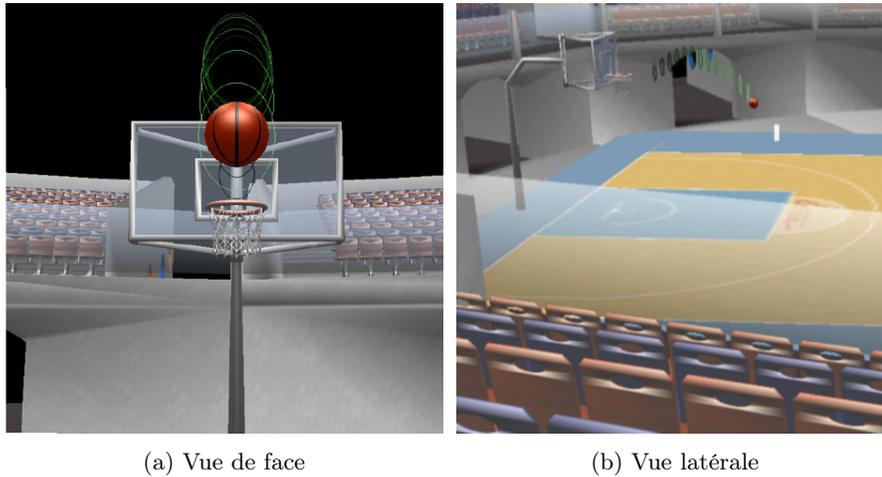


FIGURE 2.18 – Feedback pour le guidage, succession d'ellipses représentant la trajectoire idéale [Covaci et al., 2014]

dierons donc les métaphores d'interaction utilisées pour le guidage dynamique.

2.3.3 Guidage dynamique

Pour guider le geste des utilisateurs tout en répondant à nos problématiques, nous devons étudier l'impact d'un guidage dynamique. L'aspect dynamique semble en effet correspondre à cette notion d'adaptation. Bau et Mackay ont introduit dans [Bau and Mackay, 2008] les *dynamic guides* qui comprennent des mécanismes de feedforward et de feedback dynamiques pour l'apprentissage de gestes sur une interface de bureau en deux dimensions. La trace du geste à reproduire apparaît et évolue au cours de la réalisation de ce geste par exemple. Freeman *et al.* ont par la suite étendu leurs travaux et se sont intéressés aux gestes de la main pour interagir avec des tables interactives [Freeman et al., 2009].

Dans cette section nous étudions différents types de métaphores d'interaction pour le guidage dynamique. Tout d'abord, nous nous intéressons aux avatars qui permettent de représenter l'humain dans l'environnement virtuel, que ce soit l'utilisateur ou le formateur. Il est ainsi possible de lui montrer le geste à réaliser ou de lui indiquer les erreurs de positionnements de ses membres. Puis dans un second temps, nous aborderons les feedbacks indiquant la direction à suivre.

Avatars

Tout d'abord, il n'existe que très peu de systèmes évaluant d'un point de vue ergonomique les postures de travail d'un poste de production. Les plus



FIGURE 2.19 – Mesure de l'ergonomie d'un poste de travail en environnement virtuel

connues sont *ERGO-WIDE3* (développé par CLARTE⁵) et *ErgDesign* (développé par HRV⁶). *ERGO-WIDE3* est un progiciel utilisé sur une plateforme immersive de réalité virtuelle. Il permet l'aide à la conception et l'évaluation des postes de production par la réalité virtuelle (Figure 2.19a), en simulant la posture de travail. Cette posture est alors analysée afin d'anticiper les conséquences de celle-ci. *ERGO-WIDE3* est décomposé en deux modules :

- Ergo3 qui évalue en temps réel la posture de l'utilisateur en utilisant la méthode RULA présentée précédemment ;
- Wide3 qui permet une évaluation à distance par des ergonomes en utilisant un simple navigateur internet. Ils interagissent alors à distance avec l'utilisateur immergé.

ErgDesign est un environnement de simulation physique interactive avec une visualisation à l'échelle 1 : 1. Il permet de détecter les risques de troubles musculo-squelettiques en évaluant les gestes de l'utilisateur selon la méthode RULA. Ce système vise donc à réduire ces troubles musculo-squelettiques ainsi que le risque d'accidents en améliorant l'ergonomie du poste dès sa conception.

Ces deux environnements utilisent des avatars dont les différentes parties du corps sont colorées du vert (bonne posture) au rouge (posture très mauvaise) en passant par des teintes de jaune et d'orange (figure 2.19). Cela permet d'indiquer à l'apprenant si sa posture est correcte ou s'il y a des risques pour sa santé. En effet, ce type de système ne vise pas à guider le geste dans une optique de formation ou d'évaluation de la performance contrairement à l'apprentissage moteur, mais plutôt sur la qualité ergonomique du geste en détectant le risque de troubles musculo-squelettiques.

Pour des gestes libres (*mid-air*), Roveló *et al.* ont proposé un système de guidage appelé *Gestu-Wan* [Roveló *et al.*, 2015]. Les auteurs divisent le geste à reproduire en différentes sous-étapes qui sont rassemblées dans un arbre. Sur la figure 2.20, *Gestu-Wan* montre tout d'abord en vert l'étape du geste qui a

5. <http://www.clarte-lab.fr>

6. <http://www.hrv-simulation.com>

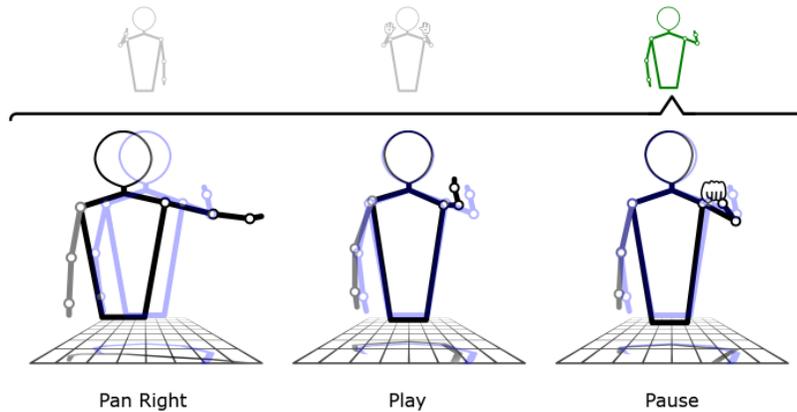


FIGURE 2.20 – *Gestu-Wan* montre les différentes étapes du geste à réaliser sous la forme d'avatars. L'utilisateur peut voir les étapes précédentes (en gris), les prochaines étapes à réaliser (en noir) et la posture actuelle (en bleu) [Rovelo et al., 2015]

été réalisée ainsi que les deux autres postures possibles en gris. Puis, sur la deuxième ligne, l'application propose en noir trois étapes possibles du geste, suite à la dernière étape réalisée, ainsi que la posture actuelle de l'utilisateur en bleu pour comparer. Cette technique est intéressante pour apprendre la succession d'étapes d'un geste libre assez simple et court mais s'avère difficilement transposable dans un exercice de gestes précis ou rapides.

YouMove est un système permettant aux utilisateurs d'enregistrer et d'apprendre à reproduire différents mouvements par le biais d'un *miroir augmenté* [Anderson et al., 2013]. En réalité, ce miroir est un simple écran sur lequel on vient projeter l'image de l'utilisateur à la manière d'un miroir (figure 2.21). Les apprenants réalisent alors leurs gestes devant cet écran où ils voient un avatar superposé sur leur reflet. Ils doivent alors suivre les mouvements de cet avatar pour reproduire le geste demandé. Eaves *et al.* avait précédemment proposé un système similaire, basé sur un *miroir augmenté*, pour de l'apprentissage moteur dans le cadre de la danse [Eaves et al., 2011]. Néanmoins, l'avatar n'était pas complet comme dans *YouMove*, mais se limitait à quelques points placés sur des articulations spécifiques. Les danseurs devaient alors prendre des poses de manière à faire correspondre les différentes parties de leur corps avec les points.

Toujours dans le cadre de la danse, Chan *et al.* ont proposé un système composé de deux avatars affichés sur un écran [Chan et al., 2011]. L'un des deux avatars représente le danseur et l'autre le professeur virtuel. L'avatar de l'élève est alors composé de cylindres représentant les différents segments du corps, entre chaque articulation. La couleur de chaque cylindre indique si la position du segment correspondant est correcte ou non. Le jaune indique une bonne position tandis que le rouge indique une erreur. L'apprenant peut ainsi

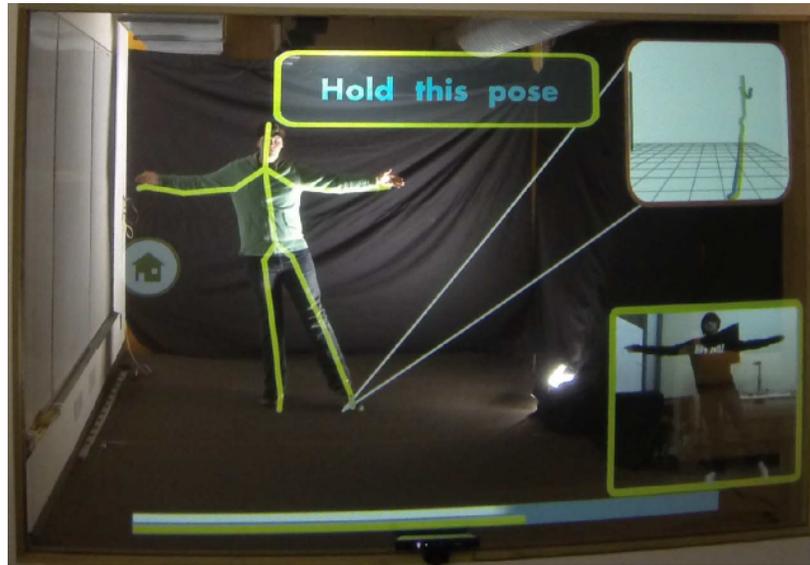


FIGURE 2.21 – *YouMove* comprend un *miroir augmenté* qui superposer l'image de l'utilisateur avec un avatar lui montre les gestes à reproduire [Anderson et al., 2013]

visualiser très rapidement si ses mouvements sont bons. Toutefois, une erreur indiquée par la couleur rouge peut être soit une erreur de posture, soit une erreur temporelle. L'utilisateur doit alors se référer au deuxième avatar, celui du professeur virtuel, et déterminer quel est le type d'erreur en analysant ses mouvements. Cela s'avère peu efficace pour de l'apprentissage de geste précis. L'apprenant doit d'une part se référer à son propre avatar pour déterminer s'il commet une erreur et d'autre part il doit déterminer à partir un deuxième avatar quel est le type d'erreur et doit déterminer par lui-même comment se corriger.

Ensuite, il existe les solutions d'avatar en vue à la première personne à l'aide d'un HMD. Yang *et al.* ont proposé la *Ghost Metaphor* (métaphore du fantôme) pour montrer le geste à réaliser à l'apprenant [Yang and Kim, 2002]. L'utilisateur se voit dans l'environnement virtuel à l'aide d'un avatar et ce geste de référence est affiché à l'aide d'un avatar semi-transparent, superposé à celui de l'utilisateur. Ce dernier essaye donc de réaliser le geste en reproduisant les mouvements de l'avatar fantôme. Cette technique semble intéressante pour le guidage, néanmoins elle ne permet pas d'indiquer à l'apprenant lorsqu'il commet des erreurs, ni comment se corriger.

De la même manière, Hoang *et al.* superpose l'avatar de l'apprenant et de l'instructeur dans le cadre de la formation à des gestes d'arts martiaux (figure 2.22) [Hoang et al., 2016]. Ils proposent également une vue à la première personne où les deux avatars ne sont plus superposés. L'avatar de l'instructeur

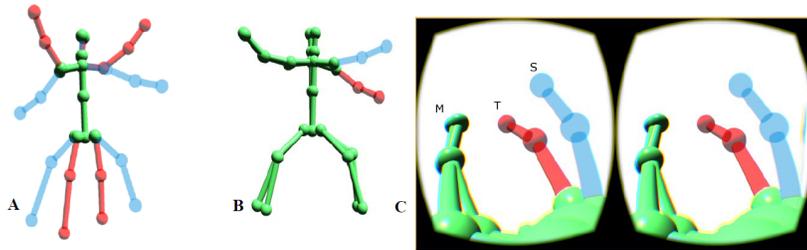


FIGURE 2.22 – A et B montrent la superposition des avatars. C montre la vue de l'étudiant à travers le HMD. S indique le bras de l'apprenant et T celui de l'instructeur. M les deux bras gauches (celui de l'apprenant et celui de l'instructeur) qui sont confondus, cela signifie que l'apprenant est à la bonne position [Hoang et al., 2016]

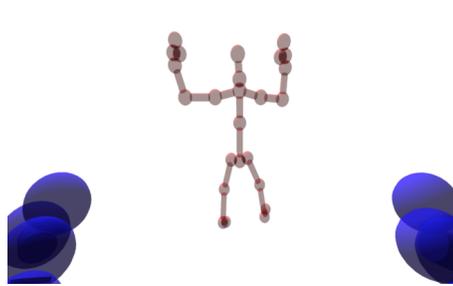


FIGURE 2.23 – Vue de l'apprenant. L'avatar de l'instructeur est situé en face de l'apprenant [Hoang et al., 2016]

est alors situé en face de l'apprenant (figure 2.23).

En résumé, la représentation complète du corps semble être une technique intéressante pour du guidage gestuel, mais n'est pas assez précise au niveau de l'origine des erreurs ou leur importance (erreur importante ou non). La plupart des approches présentées permettent d'observer son propre avatar ou corps, ainsi qu'un avatar reproduisant le geste de référence. Cela oblige ainsi les utilisateurs à s'intéresser aux deux avatars ce qui peut être coûteux en terme de charge cognitive et peut donc dégrader la qualité de l'apprentissage.

Direction et orientation

Intéressons-nous maintenant aux métaphores d'interaction destinées à l'orientation et à l'indication de la trajectoire à suivre. Tönnis et Klinker ont proposé dans leur *review* d'étudier les flèches 3D comme métaphore de guidage pour les environnements automobiles utilisant la réalité augmentée [Tönnis and Klinker, 2009]. Ils distinguent trois types de flèches (figure 2.24), dont les deux premiers ont été conçus afin de réduire l'inattention des conducteurs sans les distraire de la tâche de conduite. La première flèche indique la

2.3. MÉTAPHORES D'INTERACTION POUR LE GUIDAGE

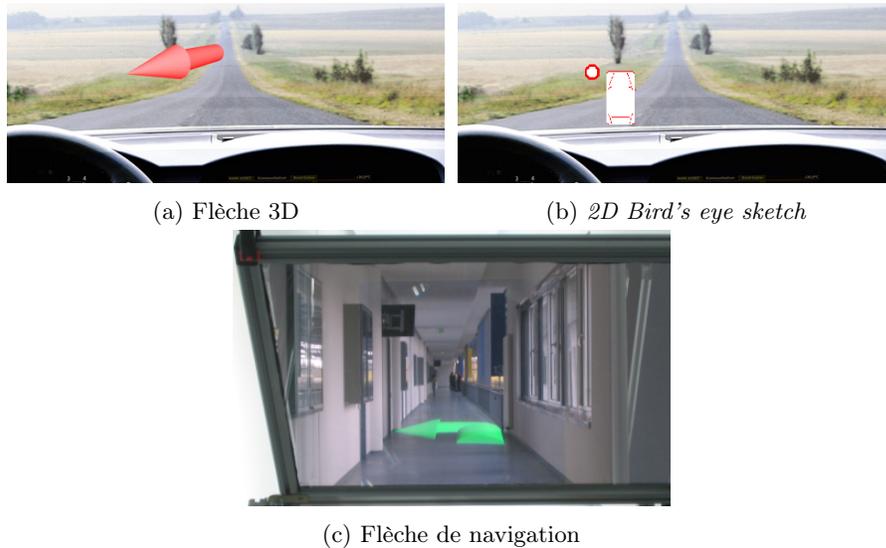


FIGURE 2.24 – Les trois types de flèches selon [Tönnis et al., 2009]

direction du danger le plus proche (figure 2.24a), le second type est un croquis d'une voiture avec une flèche en deux dimensions qui montre également la direction du danger le plus proche par rapport à la voiture (figure 2.24b). Ces deux types de flèches sont des feedbacks en temps réels qui permettent au conducteur de corriger sa trajectoire en fonction des dangers environnant. Enfin le troisième type est une flèche de navigation qui est directement projetée sur le chemin en face de soi ou sur la route (figure 2.24c). Néanmoins, dans le cadre de la navigation automobile le problème d'occlusion avec les autres véhicules réduit ses performances, notamment dans le cas d'un trafic assez dense. La métaphore devient peu visible, voire même difficile à identifier et à comprendre. Ces trois flèches permettent d'avoir une bonne orientation de la direction à suivre. George *et al.* ont proposé un feedback similaire sous la forme d'une girouette indiquant la direction d'un danger, sa dangerosité et sa criticité [George et al., 2012].

De même, Sodhi *et al.* ont proposé des aides visuelles en réalité augmentée pour le guidage des mouvements de la main (figure 2.25) [Sodhi et al., 2012]. Quatre types de feedbacks peuvent être utilisés et sont projetés directement sur la main de l'utilisateur. Toutefois seulement trois d'entre-eux s'intéressent à l'affiche de la direction à suivre, le quatrième sera donc présenté dans la prochaine section. La première de ces aides visuelles est une flèche noire en deux dimensions dans un cercle blanc (figure 2.25a). Elle indique à l'utilisateur s'il doit élever sa main ou l'abaisser, en changeant sa direction. La taille de la flèche indique la distance à la position cible. Ensuite, la deuxième flèche est cette fois-ci en trois dimensions et pointe dans la direction à suivre (figure 2.25b). Enfin, la dernière métaphore est appelée *3D Pathlet* et est composé

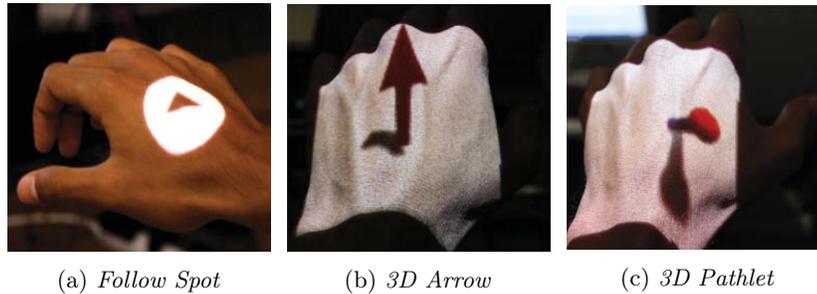


FIGURE 2.25 – Aides visuelles 3D directionnelles [Sodhi et al., 2012]

d'un point rouge représentant la position de la main de l'utilisateur, et d'un segment bleu indiquant la chemin à suivre (figure 2.25c). Cependant, lorsque la main n'est plus visible pas l'utilisateur ou orientée avec un angle important, les feedbacks ne sont également plus disponibles ce qui peut amoindrir la qualité du guidage et de l'entraînement aux gestes.

Enfin, Henderson et Feiner ont développé une application de réalité augmentée en utilisant un HMD afin d'assister les utilisateurs durant l'apprentissage d'une tâche de maintenance et/ou d'assemblage [Henderson and Feiner, 2011]. Des feedbacks en temps réel, comme des flèches 3D ou des labels, sont proposés aux apprenants pour les informer sur les prochains mouvements à faire. La couleur et la taille de chaque flèche change en fonction de l'activité de l'utilisateur.

Ces différentes techniques pour l'orientation permettent de guider les utilisateurs de manière intuitive puisqu'une flèche est la manière la plus simple pour indiquer une direction. Néanmoins, ce type d'aide visuelle ne permet pas d'indiquer l'écart de performance entre l'apprenant et l'expert. Seule la présence de ces feedbacks indique qu'une erreur est commise, ce qui oblige alors l'utilisateur à déduire que son geste est mauvais et qu'il doit se corriger.

2.3.4 Conclusion sur les métaphores d'interaction

Dans cette dernière partie de l'état de l'art, nous avons étudié différentes métaphores d'interaction pour le guidage, et plus particulièrement pour le guidage de gestes. Nous avons vu qu'il existait principalement deux types de métaphores : les métaphores dynamiques et les métaphores non-dynamiques. Les feedbacks dynamiques semblent plus pertinents pour le guidage de gestes précis puisqu'ils permettent de réaliser plus facilement des tâches complexes par rapport aux feedbacks non-dynamiques. La compréhension du geste à réaliser est alors plus rapide pour l'apprenant, ce qui lui permet de se concentrer d'avantage sur la précision de sa gestuelle.

Néanmoins, la représentation complète du corps par des avatars ne semble pas assez précise au niveau de la perception des erreurs ou trop coûteuse cognitivement, réduisant ainsi la qualité de l'apprentissage. Les feedbacks directionnels quant à eux sont plus compréhensibles pour l'apprenant mais n'indiquent

2.3. MÉTAPHORES D'INTERACTION POUR LE GUIDAGE

pas directement la nature des erreurs. L'utilisateur manque alors d'aide pour comprendre la nature de ses erreurs en temps réel. Nous positionnant dans une pédagogie basée sur l'apprentissage par l'erreur, l'approche que nous choisissons d'explorer est donc basée sur la visualisation des erreurs. Le chapitre suivant présente notre proposition. C'est une métaphore dynamique de guidage basée sur la représentation de l'erreur gestuelle.

Chapitre 3

EBAGG : une nouvelle métaphore de guidage

Nous présentons dans ce chapitre notre contribution et notre réponse à ces questions. Elle se présente sous la forme d'une métaphore de guidage nommée *EBAGG*. En nous basant sur l'état de l'art (chapitre 2), nous cherchons donc à proposer une métaphore dynamique de guidage en temps réel, répondant à la problématique soulevée en introduction de ce mémoire.

Lors de l'apprentissage de geste, nous proposons à l'apprenant de suivre une stratégie essai - erreur où il perçoit les conséquences de ses gestes sur l'environnement et modifie son comportement (et donc ses gestes) en conséquence. Cette approche basée sur la répétition et la compréhension des erreurs a déjà été appliquée à différents domaines, notamment au monde de la finance et du jeu [Young, 2009] où l'individu adapte son comportement en fonction de ses gains et pertes. Elle est également à la base de travaux sur la formation en environnement virtuel pour de la formation à des situations sociales et/ou techniques complexes [Barot et al., 2013, Lourdeaux et al., 2017].

Dans un premier temps, nous revenons sur les travaux connexes qui ont abordé des problématiques proches des nôtres. Puis nous présentons le concept de ce feedback et la méthode utilisée pour calculer le degré d'erreur commis par l'utilisateur. Enfin dans la dernière partie, nous exposons les différentes versions d'implémentation réalisées.

3.1 Travaux connexes

La représentation des erreurs pour le guidage est principalement utilisé dans des modalités autres que le visuel. L'haptique et le vibro-tactile s'y prêtent en effet assez bien. Par conséquent dans cette section nous ne nous focalisons pas uniquement sur le visuel, mais nous étudions également des méthodes de guidage issues de modalités diverses.

Tout d'abord, en ce qui concerne le vibro-tactile, Kapur *et al.* ont proposé un système pour la réhabilitation de patients victimes d'accident vasculaire

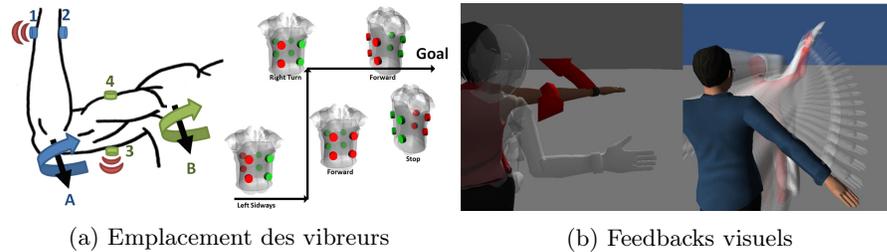


FIGURE 3.1 – Guidage multimodal pour le geste. La figure (a) présente l'équipement. La figure (b) présente les feedbacks visuels : le *ghost avatar* et les flèches directionnelles sur la gauche, la séquence de mouvement avec les *ghost avatars* sur la droite (l'avatar rouge correspond à un geste obligatoire) [Schönauer et al., 2012]

cérébral. Il comprend une manche recouvrant l'intégralité d'un des bras du patient. Elle permet de reconnaître les mouvements des membres supérieurs et donne des feedbacks vibro-tactiles pour aider les utilisateurs à réaliser des gestes simples du quotidien qu'ils ne sont plus capables de faire. Les vibreurs agissent de manière répulsive, forçant les utilisateurs à diriger leur mouvement dans la direction opposée à la vibration. Les patients souffrant d'apraxie idéomotrice¹ des membres ont souvent des difficultés à reproduire un geste qui leur est montré visuellement. L'usage de feedbacks vibro-tactiles est par conséquent avantageux.

Cette modalité ne se limite cependant pas qu'à la réhabilitation, elle peut également être utilisée pour souligner les erreurs commises par exemple. Pour améliorer l'apprentissage moteur humain, Lieberman et Breazeal ont développé un système appelé *TIKL* (*tactile interaction for kinesthetic learning*). C'est une combinaison comprenant différents vibreurs, qui analyse les mouvements des bras de l'utilisateur et lui retourne des feedbacks vibro-tactiles en temps réel lorsqu'il commet des erreurs de mouvement. Par exemple, si un apprenant doit abaisser son bras comme l'instructeur l'a fait avant lui, le vibreur situé sous l'avant bras vibrera s'il descend trop bas [Lieberman and Breazeal, 2007]. De plus, plus l'erreur est importante, plus l'intensité de la vibration sera forte. Sur le même principe, Schönauer *et al.* ont également étudié les effets des feedbacks vibro-tactiles pour le guidage de geste [Schönauer et al., 2012]. Toutefois, dans leur étude ils ont également ajouté des feedbacks visuels en plus des vibrations (figure 3.1). Les feedbacks vibro-tactiles proviennent des vibreurs placés sur les manches de la veste ainsi que sur le ventre, la poitrine et le dos (figure 3.1a). Le couplage de ces deux modalités leur permet de guider l'apprenant visuellement, à l'aide d'un avatar et de flèches directionnelles, pour lui montrer le geste à réaliser (figure 3.1b), et de corriger ses erreurs à l'aide des vibrations.

Wei *et al.* se sont quant à eux intéressés à la reproduction de gestes spécifiques pour atteindre une cible à l'aide d'un bras robotique [Wei et al., 2005].

1. Incapacité à exécuter un geste sur demande ou un geste mimé

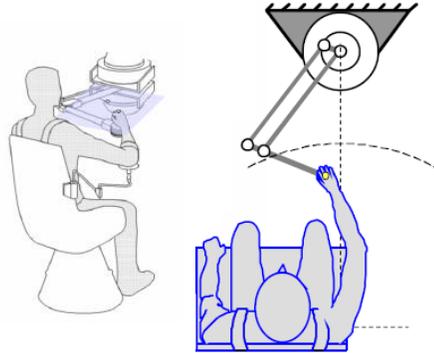


FIGURE 3.2 – Manipulation d’un bras robotique et guidage visuel projeté sur une surface cachant le bras de l’utilisateur (en bleu sur la figure) [Wei et al., 2005]

Des feedbacks visuels étaient affichés sur un plan de projection qui cachait le bras de l’utilisateur (figure 3.2). Le but était d’atteindre une cible à l’aide du bras robotique en s’aidant des indications visuelles. Cependant, les auteurs ont étudié l’effet de distorsions visuelles résultant en une augmentation de l’intensité des erreurs : les sujets ne voyaient donc pas leur trajectoire réelle, mais une version déformée où les écarts sont exagérés. Cette expérimentation a été réalisée sur des sujets sains et a montré qu’augmenter l’intensité de l’erreur perçue permettait d’améliorer le taux d’apprentissage, c’est-à-dire que les utilisateurs ayant l’augmentation de l’erreur avaient de meilleurs résultats en terme de performances, que ceux ne l’utilisant pas.

Concernant le guidage haptique, il est important de revenir sur quelques points de définition avant de poursuivre. Le guidage haptique est défini par Feygin *et al.* comme étant une assistance physique fournie par une interface haptique pour guider l’utilisateur à travers le mouvement idéal [Feygin et al., 2002]. En outre, les feedbacks haptiques élargissent cette définition et correspondent à n’importe quelle perception haptique qui guide vers, et non nécessairement à travers, le mouvement désiré [Sigrist et al., 2013]. Marchal-Crespo *et al.* ont étudié les effets des feedbacks haptiques² dans un simulateur de conduite automobile [Marchal-Crespo et al., 2010]. Les participants de l’expérimentation devaient conduire dans un environnement virtuel où des forces étaient appliquées sur leur volant lorsqu’ils commettaient des erreurs afin de les replacer sur la bonne trajectoire. La fréquence et l’intensité des feedbacks diminuaient lorsque les utilisateurs ne commettaient que quelques erreurs, et à l’inverse augmentaient dans le cas d’erreurs grossières. Les résultats de cette étude ont montré que les feedbacks haptiques avaient un effet bénéfique pour l’apprentissage à long-terme de la conduite.

2. Les auteurs parlent de guidage haptique, mais cela ne correspond pas aux définitions données précédemment



FIGURE 3.3 – *Hue Cue* : utilisation de couleurs pour représenter l’espace dans lequel se déplacer (bleu) et l’espace qu’il faut éviter (orange)) [Sodhi et al., 2012]

Enfin, comme mentionné dans la partie précédente, Sodhi *et al.* ont également proposé une métaphore d’interaction basée sur la visualisation des erreurs de trajectoire. *Hue Cue* est composée d’une flèche (*Follow Spot*) entourée par des zones colorées positivement (en bleu) et négativement (en orange), représentant un espace en deux dimensions (x et y) [Sodhi et al., 2012]. Pour réaliser le bon geste, les utilisateurs doivent déplacer leur main vers la zone bleue et s’éloigner de la zone orangée. La flèche *Follow Spot* permet de contrôler les erreurs sur la profondeur (z).

Les feedbacks basés sur l’erreur semblent ainsi être pertinents pour l’apprentissage de geste puisqu’ils s’axent sur la représentation de l’erreur et permettent notamment de montrer l’écart entre la performance de l’utilisateur (l’apprenant) et celle de référence (l’expert). De plus nous avons vu en introduction de ce mémoire, que la dépendance aux feedback semblait inhiber les sensations provenant de la proprioception (section 1.3.4). Or, ce type de feedback semble axer le ressenti de l’utilisateur d’avantage sur ses feedbacks intrinsèques plutôt que sur les feedbacks extrinsèques, ce qui pourrait être bénéfique pour répondre à cette problématique.

3.2 Modélisation de la métaphore

La métaphore EBAGG (*Error-based assistance for gesture guidance*) a été conçue pour le guidage de gestes précis en se basant sur le principe de l’énaction [Varela et al., 1992]. Au lieu d’afficher une traînée représentant le chemin idéal à suivre par exemple, la métaphore permet à l’utilisateur de suivre la bonne voie ou de revenir sur celle-ci si jamais il déviait en lui indiquant ses écarts. L’utilisateur apprend la bonne gestuelle en tenant compte des erreurs commises lors des expériences précédentes. Cela lui permet de se réguler et d’adapter ainsi son geste en fonction de ce que lui dit la métaphore de guidage. Étant donné

que EBAGG n'indique que les erreurs et ne montre pas la direction à prendre, l'apprenant n'a d'autre choix que de se faire sa propre trace mentale du geste en se basant sur les informations provenant de sa proprioception. Cette trace agit en tant que référence (figure 2.13) sur laquelle il se base par la suite pour réaliser son geste :

The reference of correctness represents the state of the feedback associated with the correct movements of the limbs during the intended action; it specifies the sensory qualities of a goal

[Schmidt et al., 2005]

En outre, lorsque l'apprenant réalise correctement le geste ou une partie du geste, EBAGG n'affichera aucune ou très peu de particules. Il peut alors mettre à jour sa trace mentale du geste à partir de ses sensations. Par la suite, nous émettons l'hypothèse que lorsque EBAGG n'est plus proposée à l'utilisateur, il se base sur sa trace mentale et sur les sensations liées pour reproduire le geste, limitant ainsi sa dépendance aux feedbacks durant l'apprentissage du geste.

EBAGG possède une forme primaire comme une sphère, une boule ou encore un cube, composée de particules pouvant vibrer entre elles. Elle voit son état modifié, notamment le niveau d'agitation de ses particules, en fonction de l'évaluation du critère du geste réalisé. Ce dernier peut être une position, une orientation ou tout autre caractéristique du geste susceptible d'être évaluée. La forme choisie peut s'avérer intéressante selon le critère d'erreur choisi, par exemple une sphère ou une boule sont des formes plus avantageuses pour représenter un espace en trois-dimensions, tandis qu'un cube est plus pertinent pour représenter l'orientation d'un objet. Toutefois, la forme de EBAGG peut être invisible et dans ce cas précis, seules les particules sont visibles par l'utilisateur.

Les particules symbolisent l'erreur commise. Lorsqu'il n'y a aucune erreur, ou qu'elle reste acceptable, aucune particule n'est présente sur EBAGG. Cependant, à partir d'un certain seuil où la valeur de l'erreur n'est plus acceptable, elles commencent à apparaître et plus l'erreur est importante, plus le nombre de particules augmente jusqu'à atteindre un nombre maximum défini au préalable. Cela peut être lorsque l'ensemble des particules recouvrent la totalité de la surface de la forme choisie ou la moitié par exemple. De plus, chaque particule émet une vibration afin d'attirer d'autant plus l'attention de l'utilisateur sur la présence d'une erreur.

Différents paramètres sont donc à prendre en compte et sont modifiables selon le cadre d'utilisation de EBAGG. Ils sont au nombre de cinq :

- **sa forme** : elle peut prendre n'importe quelle forme primaire (sphère, boule, cube, pyramide, cylindre, etc.), complexe ou peut n'avoir aucune forme visible. Le forme ainsi que sa taille dépendent principalement du critère d'erreur évalué ;
- **le critère d'erreur évalué** : une position, une vitesse, une orientation, une performance ;

- **ses particules** : chaque particule possède une forme, une taille, une couleur, une position, une visibilité et une intensité de vibration ;
- **sa couleur** : EBAGG possède une couleur de base mais celle-ci peut varier en fonction du critère d'erreur et même être dynamique (varier au cours du temps) selon l'utilisation ;
- **son opacité** : la métaphore est de base légèrement transparente afin de pouvoir visualiser les particules à travers sa structure.

Enfin, la précision de l'information retournée par la métaphore fluctue en fonction du nombre de particules qu'elle peut afficher. Plus ce nombre est grand, plus il sera possible de différencier la sévérité de l'erreur. Dans le cadre du guidage de geste, cette sévérité d'erreur est basée sur le calcul de l'écart entre une trajectoire idéale définie au préalable (geste expert) et la trajectoire actuelle de l'utilisateur. Notre méthode de calcul est décrite dans la section suivante (section 3.3).

3.3 Calcul de l'erreur

Pour déterminer les performances des utilisateurs, nous devons comparer leur geste à une référence enregistrée au préalable. Nous calculons la distance entre les deux gestes puis en fonction de l'écart obtenu, nous signalons l'erreur et l'importance de celle-ci via EBAGG. Nous choisissons donc d'utiliser l'algorithme de déformation temporelle dynamique (DTW) qui fournit des résultats interprétables comme l'écart entre deux séries temporelles (voir chapitre 2) [Berndt and Clifford, 1994, Keogh and Pazzani, 2001]. L'algorithme 1 reflète la version classique du DTW.

Contrairement aux méthodes d'apprentissages comme les HMM par exemple, un algorithme DTW traite directement des données labellisées et ne les modélise pas. L'écart obtenu n'est pas calculé entre une modélisation des données et les données réelles mais bien entre les données d'un geste et celles du geste de l'utilisateur.

Néanmoins les gestes sont généralement des séquences multidimensionnelles étant donné qu'ils sont représentés en tant que séries temporelles (signaux provenant de la capture de mouvements). Il est ainsi important de s'intéresser à la synchronisation de ces séries. Nous nous sommes donc tout d'abord porté vers l'extension multidimensionnelle du DTW (appelée MD-DTW) proposée par ten Holt *et al.* [Ten Holt et al., 2007]. Le calcul de la distance est alors plus précis qu'avec un DTW classique et l'écart obtenu est ainsi multidimensionnel. Toutefois, ce calcul est aussi plus coûteux en temps et en ressources. Nous nous sommes alors rabattus sur l'extension naturelle dont parle ten Holt *et al.* (appelée 1D-DTW) qu'ils comparent au MD-DTW. L'algorithme 2 présente notre version de cet algorithme, que nous avons appelé k-DTW.

Multi-dimensional (time) series are series in which multiple measurements are made simultaneously. Such series have an K -dimensional vector of feature values for each (time) instance of the series. They

Algorithme 1 : DTW

Entrées : Deux séries temporelles $X = (x_1, \dots, x_N)$ et $Q = (q_1, \dots, q_M)$
de longueurs respectives N et M
Sorties : La distance minimale entre les deux séries gap
La matrice D comprenant les distances
La matrice C des coûts cumulés

début

```

    Initialisation de la matrice des coûts cumulés
     $C \leftarrow \text{array}(N + 1, M + 1)$ 
    pour  $i \leftarrow 1$  à  $N$  faire
    |    $C(i, 0) \leftarrow \text{infinity}$ 
    pour  $i \leftarrow 1$  à  $M$  faire
    |    $C(0, i) \leftarrow \text{infinity}$ 
    Initialisation de la matrice des distances
     $D \leftarrow \text{array}(N, M)$ 
    pour  $i \leftarrow 1$  à  $N$  faire
    |   pour  $j \leftarrow 1$  à  $M$  faire
    |   |   Calcul des distances
    |   |    $D(i, j) \leftarrow \text{ValeurAbsolue}(X(0, i) - Q(0, j))$ 
    |   |   Calcul du coût cumulé
    |   |    $C(i, j) \leftarrow D(i, j) + \text{Minimum}(C(i - 1, j),$ 
    |   |   |    $C(i - 1, j - 1),$ 
    |   |   |    $C(i, j - 1))$ 
    |   |
    |    $gap \leftarrow C(n, m)/(n + m)$ 
    retourner ( $gap, D, C$ )

```

can be synchronised by simply picking one dimension to perform DTW with and warping the complete series according to the warp found in this dimension.

[Ten Holt et al., 2007]

Cette nouvelle version du DTW ne fonctionne cependant qu'en *offline*, c'est-à-dire une fois que le geste est terminé et que l'ensemble des données est enregistré. Il nous fallait donc adapter cette extension pour le temps réel et nous nous sommes donc inspirés des travaux de [Musillo et al., 2011, Sabinas et al., 2013] pour définir une version *early-recognition*. Pour déterminer la correspondance entre deux points des séquences temporelles, nous utilisons une fenêtre locale glissante en nous basant sur la correspondance précédente au lieu de la séquence complète du geste de référence. Enfin, nous avons également dû paralléliser les calculs de distance du DTW sur différents processus pour pouvoir en faire plusieurs en même temps, et afin que l'ensemble du système puisse fonctionner en temps réel.

Algorithme 2 : k-DTW

Entrées : Deux séries temporelles X et Q de longueurs respectives M et N (nombre de lignes) et de dimension K (nombre de colonnes)

Sorties : La distance minimale cumulée entre les deux séries gap
 La matrice D_{tot} comprenant les distances cumulées
 La matrice C_{tot} des coûts cumulés

début

Initialisation de la matrice des coûts cumulés

$C_k \leftarrow \text{array}(N + 1, M + 1)$

Initialisation de la matrice des distances

$D_k \leftarrow \text{array}(N, M)$

pour $k \leftarrow 1$ **à** K **faire**

x est un tableau de N valeurs

pour $i \leftarrow 0$ **à** $N - 1$ **faire**

$x(i) \leftarrow X(i, k)$

q est un tableau de M valeurs

pour $i \leftarrow 0$ **à** $M - 1$ **faire**

$q(i) \leftarrow Q(i, k)$

$gap, D, C = \text{DTW}(x, y)$ */* Appel du DTW classique */*

$C_{tot} \leftarrow C_{tot} + C$

$D_{tot} \leftarrow D_{tot} + D$

$gap_{tot} \leftarrow gap_{tot} + gap$

retourner $(gap_{tot}, D_{tot}, C_{tot})$

3.4 Implémentations

Ayant pour objectif de guider la gestuelle de l'utilisateur, le critère d'erreur choisi est donc la position de l'objet dont les mouvements sont suivis. Comme dit précédemment, une sphère se prête mieux à la représentation de l'espace en trois dimensions, puisque chacun de ses points est à une distance égale de son centre. Associée aux particules, elle nous semble donc être la forme la plus simple et la plus efficace pour différencier les directions de l'erreur. Les particules permettent d'indiquer les erreurs de trajectoires selon le geste de chaque apprenant. Chacun de ces gestes est différent, donc les erreurs le sont également. Ainsi, à chaque tentative de réalisation du geste, l'utilisateur n'aura les mêmes retours provenant de la métaphore qu'à l'essai précédent. Ceci est également valable entre chaque utilisateur puisqu'ils réalisent tous un geste différent. EBAGG s'intéresse ainsi à la performance de l'utilisateur dans le cadre de l'apprentissage moteur..

Deux prototypes ont ainsi vu le jour au cours de nos travaux. Tout deux développés sur le moteur graphique Unity³, ils reprennent les choix de conceptions

3. <http://unity3d.com/>

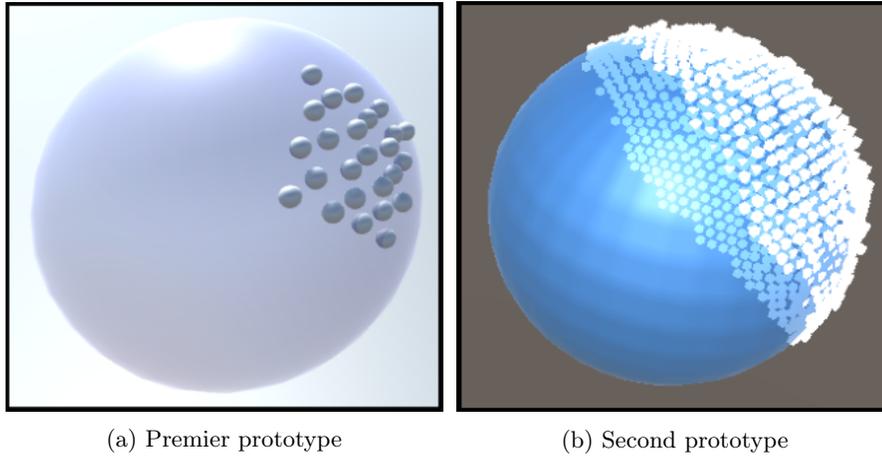


FIGURE 3.4 – La métaphore EBAGG

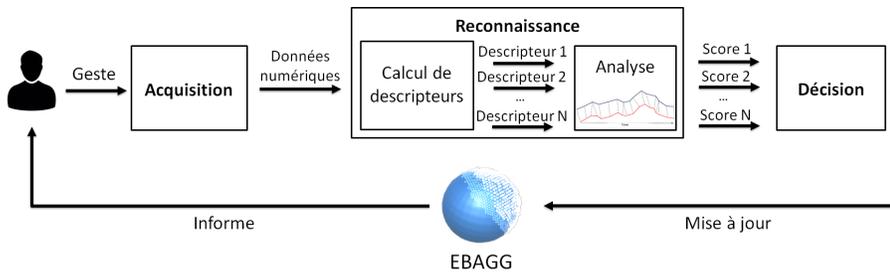


FIGURE 3.5 – Schéma de fonctionnement de la métaphore EBAGG

présentés précédemment. La figure 3.4 montre les deux versions de EBAGG. La première version fut développée dans le cadre d’une première expérimentation (chapitre 4). Elle ne comprenait qu’un maximum de 25 particules et fut modélisée avec les éléments proposés par défaut dans Unity, à savoir l’objet 3D de la sphère. Un second prototype fut développé en partenariat avec la société Reviatech⁴, membre du projet KIVA et en charge du développement de l’environnement virtuel final. Ce second prototype était cette fois-ci généré sous Unity à partir d’un *shader*⁵ développé spécifiquement pour EBAGG.

La performance étant évaluée par notre algorithme DTW, les résultats sont retournés et interprétés par EBAGG qui met alors à jour le nombre et la position des particules (figure 3.5). Le DTW se base sur les données du geste de l’utilisateur pour calculer les descripteurs (positions) nécessaires pour l’analyse et la mesure de similarité avec le geste de référence. La figure 3.6 quant à elle

4. <http://reviatech.com/>

5. Programme destiné à paramétrer le processus du rendu réalisé par le GPU (ombres, lumière, texture, etc.)

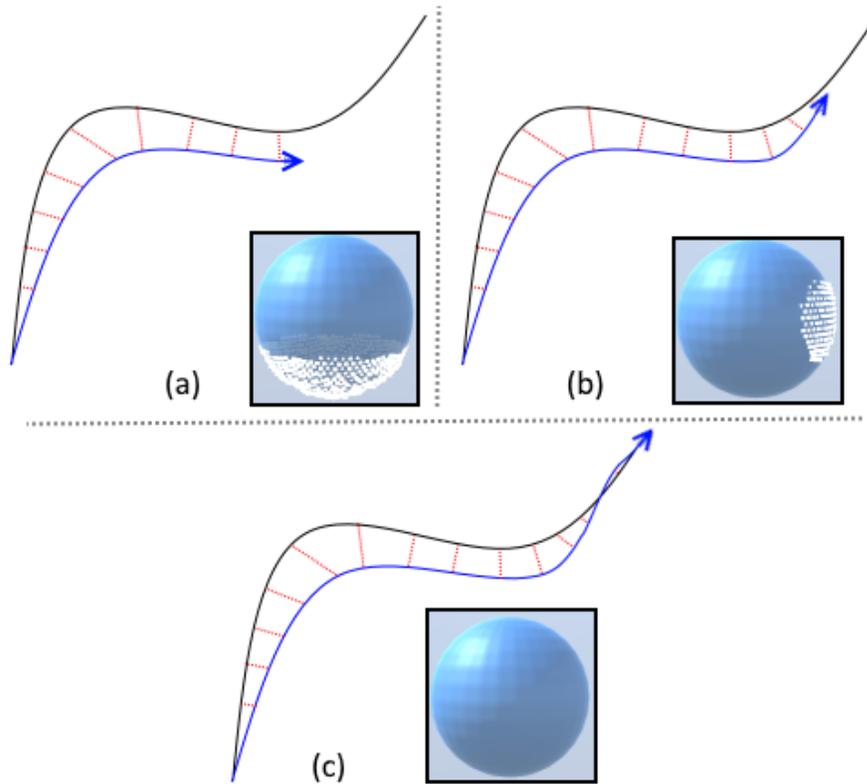


FIGURE 3.6 – Comportement de la métaphore EBAGG. La trajectoire du geste de référence est représentée en noir, tandis que celle du geste utilisateur est en bleu. La distance les séparant est en rouge

représente le comportement de EBAGG lors de la présence d'erreur. En (a) l'écart est plutôt important, donc EBAGG affiche un nombre important de particules dans la direction de l'erreur. Toutefois en (b) cet écart a diminué, réduisant ainsi le nombre de particules. Enfin en (c), la trajectoire de l'utilisateur est semblable à la référence, il n'y a donc plus aucune particule à la surface.

Conclusion

Les deux prochains chapitres présentent les différentes expérimentations que nous avons menées au cours de nos travaux. Elles nous ont permis de tester les performances de la métaphore EBAGG pour le guidage lors d'exercices de reproduction de gestes précis. Toutefois nous n'avons sélectionné qu'un certain nombre de paramètres de la métaphore. Pour ces expérimentations, la couleur de la sphère a été choisie arbitrairement et était fixe, et enfin seule la position

CHAPITRE 3. EBAGG : UNE NOUVELLE MÉTAPHORE DE GUIDAGE

a été testée comme critère d'erreur.

Chapitre 4

Evaluation de la métaphore pour le guidage

4.1 Introduction

Afin de répondre aux problématiques soulevées, nous avons proposé dans le chapitre précédent la métaphore EBAGG. Nous conduisons donc ici une première expérimentation visant à évaluer la capacité de notre métaphore à guider l'apprenant. Nous cherchons ainsi à évaluer son impact sur les performances lors d'un exercice de reproduction de gestes précis. Les apprenant doivent reproduire deux gestes en étant partiellement aidés par EBAGG au cours de la session d'entraînement. La difficulté des deux gestes étant différente, nous évaluons également son impact sur les performances selon cette complexité. L'objectif est donc d'une part de s'assurer que EBAGG est adéquate pour le guidage de geste en environnement virtuel, et d'autre part d'évaluer sa capacité à corriger efficacement le geste en fonction de la difficulté de celui-ci.

Ce chapitre se décompose de la manière suivante. Tout d'abord nous revenons sur les hypothèses que nous avons établies au préalable. Dans la partie 4.3, nous détaillons le protocole expérimental. Nous revenons sur la tâche (reproduction de deux gestes), sur les conditions et le dispositif expérimental. Puis dans la partie 4.4, nous exposons et explicitons les résultats obtenus. Nous discutons de l'ensemble de cette expérimentation dans la partie 4.5, avant de conclure.

4.2 Hypothèses

Le premier enjeu de cette expérimentation est de vérifier que EBAGG permet effectivement de guider le geste de l'apprenant. La vision est souvent perçue comme étant le sens le plus important en raison de sa forte interaction avec l'environnement dans notre vie de tous les jours. Elle prédomine notamment sur les autres sens pour ce qui concerne les informations de spatialisation [Nesbitt et al., 2003]. Le choix de la modalité visuelle semble donc pertinent.

De plus, nous avons vu dans la partie 3.1 que la visualisation de l’erreur permettait le guidage. Il semblerait que les informations relatives à la correction des erreurs soient plus important pour les tâches complexes comme les gestes précis, que les mouvements simples [Tzetzis et al., 2008]. Plus la difficulté est élevée, plus l’apprenant aurait besoin d’une combinaison de feedbacks descriptifs et de feedbacks normatifs [Williams and Hodges, 2005, Wulf et al., 1998]. Dans notre cas, EBAGG est bien une combinaison des deux. En effet, la simple présence des particules joue le rôle du feedback descriptif puisque cela signifie qu’une erreur est commise, tandis que leur position sur la sphère joue le rôle du feedback normatif en indiquant la direction de l’erreur et donc implicitement la correction dans le sens opposé. Pour déterminer si EBAGG permet donc bien de guider l’apprenant lors de la réalisation de ses gestes, nous évaluons ses performances et notamment la présence d’une amélioration notable ou non. Les performances sont obtenues en évaluant les scores retournés par le DTW à chaque instant du geste. En comparant les scores des premières itérations du geste avec les dernières itérations, nous obtenons une tendance sur l’amélioration des performances.

Ensuite, comme indiqué dans l’introduction de cette partie, nous supposons que l’influence de EBAGG sur les performances est étroitement liée à la difficulté du geste. Pour cette expérimentation, nous avons choisi deux gestes complexes à réaliser. Pour rappel, selon [Wulf and Shea, 2002] un geste est défini comme étant complexe lorsqu’il n’est pas possible de l’acquérir en une seule session d’entraînement. Nous proposons donc aux participants de reproduire deux gestes, le premier étant plus facile à réaliser que le deuxième, tout en gardant une complexité importante en jouant sur la précision de la trajectoire du geste. L’écart entre un geste de l’utilisateur et le geste de référence correspondant doit en effet être relativement faible (de l’ordre du centimètre). EBAGG est ainsi présente durant la moitié des itérations afin de corriger les erreurs de trajectoire. Le second enjeu de cette expérimentation est donc d’observer l’influence de la difficulté du geste.

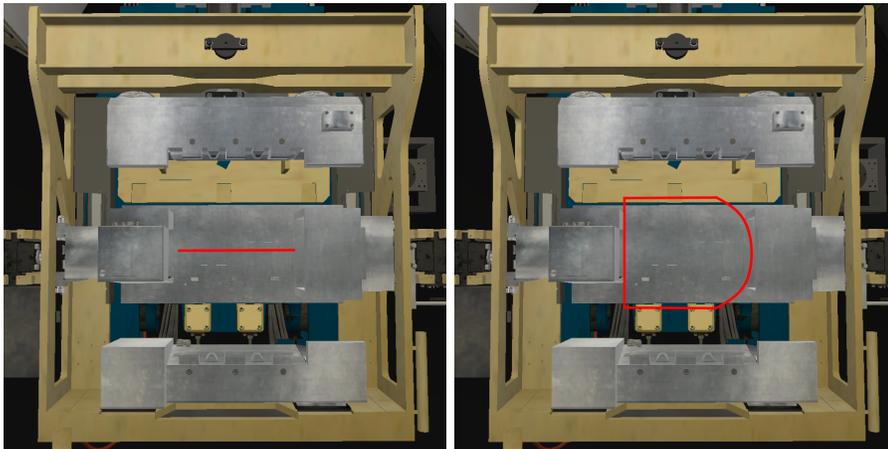
L’expérimentation doit donc être en mesure de vérifier les hypothèses suivantes :

- hypothèse 1 (**H1**) : l’utilisation de EBAGG durant l’entraînement permet aux apprenants d’obtenir de meilleures performances lors des dernières itérations du geste que lors des premières itérations ;
- hypothèse 2 (**H2**) : lors de la phase d’entraînement avec EBAGG, on observe une amélioration significative des performances lors des dernières itérations du geste par rapport aux premières itérations.

4.3 Expérimentation

4.3.1 Tâche

Les participants doivent reproduire deux gestes avec une haute précision en face d’un moule virtuel projeté sur le mur situé devant eux. Ce moule fait partie



(a) Premier geste : linéaire (b) Second geste : en forme de lettre "D"

FIGURE 4.1 – Visualisation des deux gestes à réaliser. Le tracé rouge symbolise la forme du geste.

d'une usine virtuelle développée par Reviatech dans le cadre du projet KIVA. Il est représentatif du véritable poste de moulage de Montupet. Les deux gestes étaient réalisés à l'aide d'une soufflette à air comprimé qui est habituellement sur ce poste.

Le premier geste à reproduire est un tracé linéaire en trois dimensions (figure 4.1a). La difficulté est de réaliser le mouvement d'un point à un autre en ne se déplaçant que selon une seule dimension de l'espace (la largeur x), soit une translation, tout en restant invariant sur les deux autres dimensions (la hauteur y et la profondeur z).

Le second geste est un carré avec un arc de cercle à la place de l'un de ses côtés (figure 4.1b), prenant ainsi la forme d'une lettre "D" élargie. Cette fois-ci, le geste doit être réalisé dans le plan vertical et n'est plus une simple translation. Les participants doivent néanmoins rester invariant sur la profondeur z , la soufflette devant rester à une distance constante du moule.

La principale difficulté est de réaliser les deux gestes à la bonne profondeur.

4.3.2 Protocole expérimental

Avant de commencer, le contexte des recherches est décrit à chaque participant afin leur permettre de mieux comprendre certains aspects de la tâche à réaliser, comme l'utilisation d'une véritable soufflette à air comprimé par exemple. L'expérimentation se déroule suivant le protocole suivant. Tout d'abord, une première phase de familiarisation avec l'environnement est proposée aux apprenants, principalement avec les lunettes stéréoscopiques et le *head tracking* (suivi des mouvements de la tête). Lorsqu'ils se sentent suffisamment familiarisés, le principe de EBAGG leur est exposé. Pour l'illustrer, un cube virtuel est

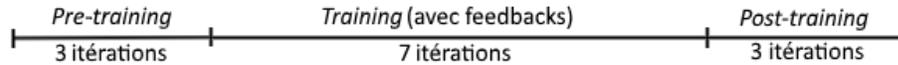


FIGURE 4.2 – Différentes phases et nombre d'itérations correspondant pour la réalisation d'un geste

placé au centre du moule et la position de la soufflette est alors comparée avec la sienne. Les particules apparaissent sur EBAGG lorsque la soufflette est au centre du cube et il n'y en a pas lorsqu'elle est à l'intérieur de celui-ci. Les participants peuvent alors déplacer eux-mêmes la soufflette afin de bien assimiler le comportement de EBAGG.

Une fois la phase de familiarisation terminée, le premier geste est alors expliqué au participant. Aucun repère visuel n'est donné lors de son explication hormis les points de départ et d'arrivée. Le geste est montré deux fois, puis l'expérimentation peut commencer. Cette dernière est divisée en trois phases successives pour les deux gestes (figure 4.2). Tout d'abord, il y a la phase de *pre-training* comprenant trois itérations du geste, où les apprenants ne recevaient aucun feedback sur le geste. A ce moment-là, EBAGG apparaît pour sept itérations. C'est la phase de *training* avec feedback, où EBAGG indique aux apprenants lorsqu'ils commettent une erreur et la direction de celle-ci afin qu'ils puissent se corriger. Enfin, il y a une dernière phase de trois itérations sans feedback, la phase de *post-training*. Une fois les treize itérations du premier geste effectuées, le second geste est montré aux apprenants qui doivent alors reproduire le même processus, mais cette fois-ci avec le nouveau geste.

Lorsque toutes les itérations du deuxième geste ont été réalisées, il est demandé aux participants de remplir un questionnaire afin d'avoir une évaluation qualitative de l'expérimentation et de EBAGG. Il comprend douze questions adaptées du questionnaire *Presence* de Witmer [Witmer and Singer, 1998] servant à mesurer la *présence* dans les environnements virtuels. La présence est définie par Bouvier comme étant "le sentiment authentique d'exister dans un monde autre que le monde physique où notre corps se trouve" [Bouvier, 2009]. N'étant pas le *focus* de notre expérimentation, nous nous sommes concentrés sur les questions portant sur la qualité de l'interaction et l'auto-évaluation des performances. Les participants doivent alors répondre aux questions par un score allant de 0 ("Pas du tout d'accord") à 10 ("Parfaitement en accord").

4.3.3 Participants

Vingt-deux participants ont pris part à cette première expérimentation (cinq femmes et dix-sept hommes). Ils étaient âgés de 21 à 42 ans (*mean* : 25.3, *std* : 4.6) et trois d'entre eux étaient gauchers. La majorité de ces participants étaient des étudiants de l'université ou des membres du laboratoire. Aucun d'entre eux n'avait utilisé la métaphore EBAGG avant cette expérience. De plus, très peu d'entre-eux avaient déjà eu une expérience de réalité virtuelle en dehors des visiocasques.



FIGURE 4.3 – Caméra OptiTrack Prime 13W

4.3.4 Dispositif expérimental

L'expérimentation a eu lieu dans une salle équipée de caméras de capture de mouvement : cinq caméras OptiTrack¹ Prime 13W. Chacune d'entre elles possède une fréquence ajustable de 30 fps (*frames per second* - images par seconde) à 240 fps. Pour cette expérience, leur fréquence d'échantillonnage était de 120 fps, avec une latence de 8 ms. Comme on peut le voir sur la figure 4.3, elles possèdent dix diodes électroluminescentes émettant des rayons infrarouges qui réfléchissent sur les marqueurs placés sur les objets dont on souhaite capturer les mouvements. Chaque caméra détecte alors ces marqueurs dont les positions sont envoyées à l'environnement virtuel via le logiciel Motive², développé par OptiTrack.

L'environnement virtuel était projeté sur le mur à l'aide d'un vidéo-projecteur à ultra courte focale Optoma³ EH319USTi (figure 4.4). Sa résolution était 1080p (1920 x 1080 pixels) à 120 Hz, permettant d'avoir un rendu stéréoscopique actif. Chaque participant portait donc des lunettes stéréoscopiques Volfoni⁴ EDGE RF sur lesquelles étaient placés des marqueurs réfléchissants afin d'assurer le *head tracking*. Enfin, les participants étaient placés à environ 150 cm du mur à l'aide d'une marque au sol.

Le rendu de cet environnement virtuel était réalisé à l'aide d'un ordinateur ayant un processeur Intel⁵ i7 6700 possédant quatre cœurs à 3.40 - 4.00 GHz, ainsi qu'une carte graphique Nvidia⁶ GTX 970 possédant 4 Go de mémoire DDR5.

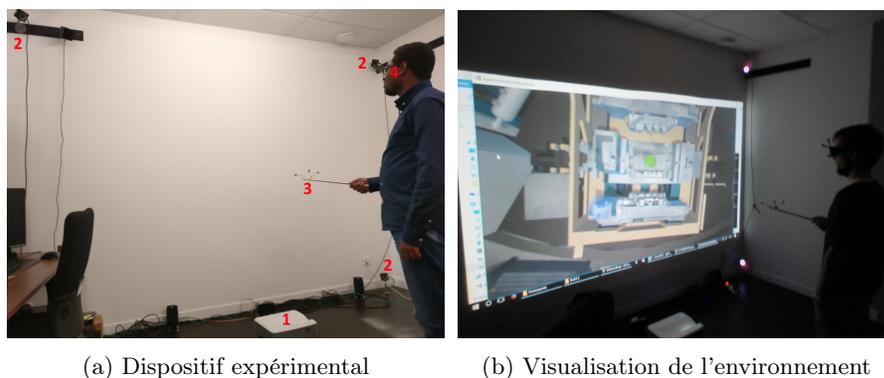


FIGURE 4.4 – Plateforme de l’expérimentation : (1) vidéoprojecteur, (2) caméras infrarouge, (3) soufflette couplée aux marqueurs réfléchissants, (4) lunettes stéréoscopiques

4.3.5 Evaluation

Pour évaluer les performances des participants, nous comparons leur geste à un geste de référence enregistré au préalable. Nous cherchons à calculer la distance entre ces deux gestes et en fonction de la valeur de l’écart obtenu, des particules apparaissent ou non sur la métaphore EBAGG dans la direction de l’erreur (figure 4.5). Pour mesurer ces distances à chaque instant, nous utilisons l’algorithme DTW présenté dans le chapitre 3 et qui permet d’évaluer la similarité entre deux séquences temporelles. Il nous fournit des résultats interprétables tels que la correspondance de la dernière position de l’outil tenu par l’utilisateur lors de la réalisation du geste avec la position de référence et la distance euclidienne les séparant (l’ensemble de ces trois données représente ce que l’on appellera par la suite *matching*), ainsi que le coût cumulé calculé sur le geste partiel (geste en cours de réalisation). Comme dit précédemment, la distance de chaque *matching* sert de donnée d’entrée pour notre métaphore, tandis que le coût cumulé agit en tant que score pour évaluer la performance globale de l’utilisateur. Plus sa valeur est proche de zéro, plus la performance de l’utilisateur est proche de celle de la référence.

4.4 Résultats

Avant d’analyser les résultats de l’expérimentation, nous appliquons aux données le test statistique de Shapiro-Wilk qui teste l’hypothèse nulle selon

1. <http://optitrack.com/>
2. <http://optitrack.com/products/motive/>
3. <http://www.optoma.fr/>
4. <http://volfon.com>
5. <http://intel.com/>
6. <http://www.nvidia.com/>

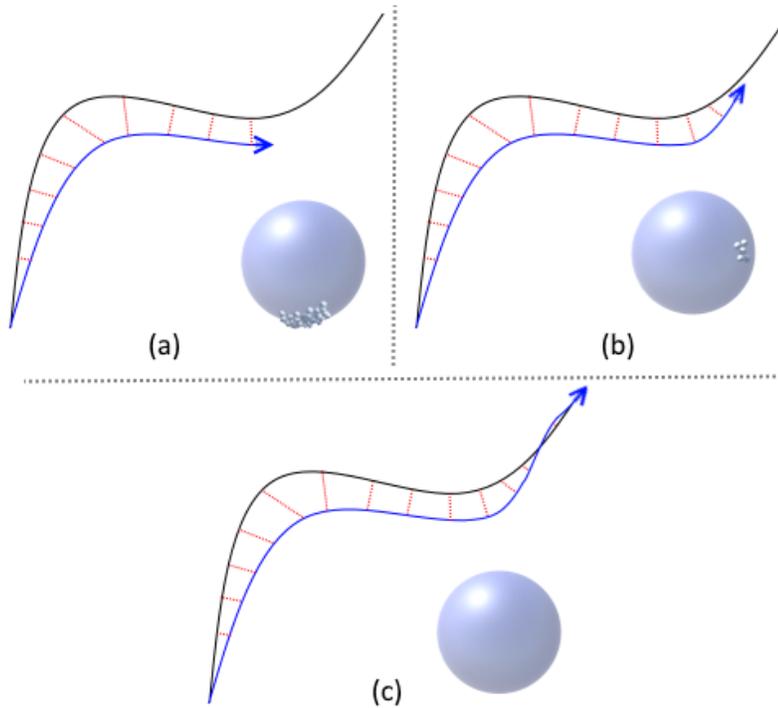


FIGURE 4.5 – Comportement de la métaphore EBAGG en fonction de l'écart de performance entre l'utilisateur et la référence. Le geste de référence est représenté par la trajectoire noire, tandis que celui de l'utilisateur est représenté en bleu. Les distances calculées entre deux points de chaque séquence sont symbolisées par les segments en pointillés rouges. Cette figure est présentée plus en détail au chapitre 3 et sert ici de rappel.

laquelle l'échantillon x_1, \dots, x_n provient d'une population normalement distribuée. Dans notre cas, puisque la valeur de la p -value étant inférieure à 0.05 nous rejetons l'hypothèse nulle, ce qui signifie que nos données ne suivent pas une loi normale. Par conséquent, nous ne pouvons pas utiliser de tests paramétriques pour évaluer nos résultats. Nous utilisons ainsi le test non-paramétrique des rangs signés de Wilcoxon pour comparer les valeurs moyennes des résultats selon les phases de l'expérimentation. Nous comparons ainsi les valeurs de la phase de *pre-training* avec le *post-training*, mais également les premières itérations de la phase de *training* avec les dernières de cette même phase. Pour compléter cette analyse quantitative, nous avons également proposé aux participants le questionnaire, présenté précédemment, à la fin de l'expérimentation.

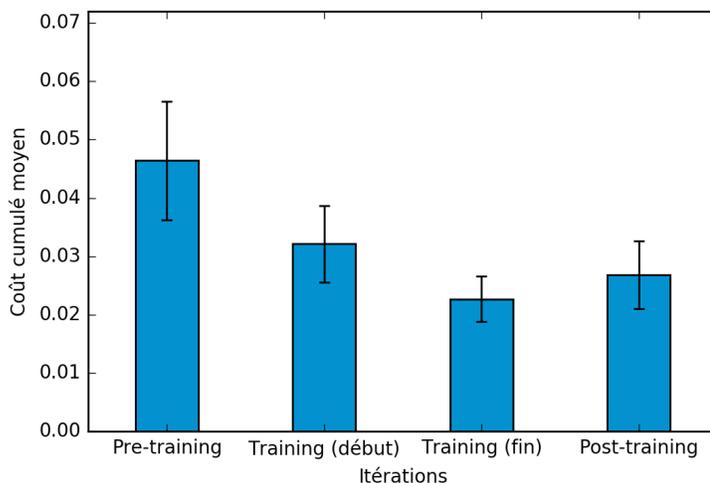


FIGURE 4.6 – Diagramme en barres illustrant la performance des participants au cours du temps lors de la réalisation du premier geste. La phase de *training* est divisée en deux sous-parties.

4.4.1 Résultats pour le premier geste

La figure 4.6 illustre les résultats et l'évolution des performances pour les phases de *pre-training*, *training* et *post-training* lors de la réalisation du premier geste (tracé linéaire en trois dimensions). Nous pouvons dès lors observer une amélioration des performances au cours du temps. La comparaison quantitative des phases de *pre-training* et de *post-training*, qui sont les deux phases où la métaphore EBAGG n'est pas visible, indique une amélioration significative avec une *p-value* égale à 0.014. La valeur moyenne du coût cumulé retourné par l'algorithme DTW décroît de 0.046 (*standard deviation* (std) : 0.024) à 0.026 (std : 0.014). Nous pouvons également remarquer une diminution de la variance, indiquant que les participants tendent vers les mêmes performances en fin d'exercice. Ces résultats nous permettent de valider notre première hypothèse **H1** pour le premier geste.

Concernant notre seconde hypothèse, l'analyse statistiques révèle une hausse significative des performances lors de la phase de *training* avec la métaphore EBAGG. En effet, lors des trois premières itérations la valeur moyenne du coût cumulé est à 0.032 (std : 0.016), tandis qu'elle est à 0.023 (std : 0.009) pour les trois dernières itérations, et la *p-value* issue du test de Wilcoxon est égale à 0.0011. Notre hypothèse **H2** semble donc elle aussi être vérifiée pour ce premier geste.

Néanmoins, nous pouvons remarquer sur la figure 4.6, que le score moyen croit légèrement en *post-training* par rapport à la fin de la phase de *training*,

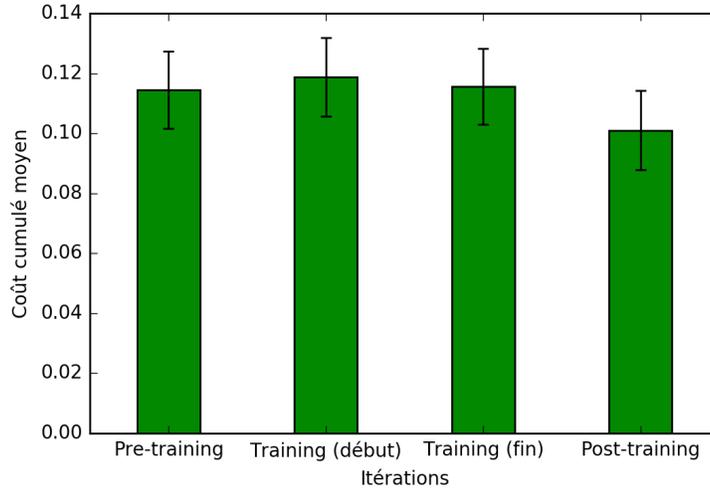


FIGURE 4.7 – Diagramme en barres illustrant la performance des participants au cours du temps lors de la réalisation du second geste. La phase de *training* est divisée en deux sous-parties.

ainsi que la variance. Ceci peut être expliqué par la disparition de la métaphore EBAGG en *post-training*, perturbant légèrement les apprenants. Cette différence n'est cependant pas significative pour un seuil à 0.05, avec une *p-value* égale à 0.098.

4.4.2 Résultats pour le second geste

La figure 4.7 illustre également une amélioration pour le second geste (forme carrée en trois dimensions). L'analyse statistique réalisée avec le test de Wilcoxon révèle en effet un effet significatif entre la phase de *pre-training* et la phase de *post-training*. La valeur moyenne du coût cumulée étant passée de 0.114 (std : 0.032) à 0.101 (std : 0.032) et celle de la *p-value* est en effet égale à 0.01. Ces résultats semblent donc valider notre première hypothèse **H1** pour le second geste.

Nous réalisons une nouvelle fois l'analyse statistique pour la phase de *training* en comparant les résultats des trois premières itérations avec les trois dernières. On remarque sur la figure 4.7 une légère amélioration de 0.119 (std : 0.032) à 0.116 (std : 0.03). Cependant, cette amélioration n'est pas significative pour un seuil à 0.05, la valeur de la *p-value* étant égale à 0.23. Notre seconde hypothèse **H2** n'est donc pas vérifiée pour le second geste. Ces résultats peuvent s'expliquer par la complexité du geste. Ce deuxième geste est plus difficile à réaliser que le deuxième et semble donc nécessiter plus d'itérations d'entraînement pour corriger entièrement le geste. La comparaison des phases de *pre-training*

et de *post-training* semble également aller dans ce sens puisque lorsque nous retirons la métaphore EBAGG, les résultats s'améliorent immédiatement.

4.4.3 Résultats qualitatifs

L'objectif du questionnaire proposé est d'obtenir une évaluation qualitative de l'expérimentation et de la métaphore EBAGG par les participants. L'ensemble des questions et des réponses fournies se trouve en annexe B. Dans cette section nous ne traiterons que certaines questions qui semblent expliquer en partie les résultats précédents. Pour rappel, les notes des réponses vont de 0 à 10.

La figure 4.8 comprend deux diagrammes présentant les résultats des questions portant sur la compréhension du comportement de la métaphore EBAGG. Le diagramme 4.8a montre que les participants se sont rapidement adaptés à l'aide visuelle, avec une moyenne de 6.1 sur 10 (std : 1.96). Néanmoins, même s'ils comprennent vite la manière dont cette métaphore fonctionne, ces résultats sont à nuancer avec le diagramme 4.8b. Celui-ci traite de la gêne provoquée par EBAGG lors de l'entraînement. Les résultats sont d'ailleurs assez disparates avec une moyenne de 5 sur 10 et une variance élevée (std : 2.49). Pour la majorité des participants, lorsqu'ils répondaient s'être rapidement habitué à EBAGG, ils n'étaient pas ou très peu perturbé par la métaphore et au contraire, lorsqu'ils avaient mis du temps à s'y habituer, l'aide visuelle a eu tendance à les perturber durant l'entraînement.

Cette compréhension de la métaphore EBAGG impacte directement la qualité de l'interaction puisqu'elle interfère sur la concentration de l'utilisateur. En effet, si celui-ci ne parvient pas à comprendre le comportement de l'aide visuelle, il aura tendance à focaliser son attention sur le feedback afin de saisir les informations retournées, plutôt que de se concentrer sur la réalisation du geste. Les diagrammes de la figure 4.9 présentent les résultats sur la concentration des participants au début de la phase de *training* (trois premières itérations) et à la fin de cette même phase (trois dernières itérations). Le premier diagramme 4.9a montre que les apprenants portaient d'avantage leur regard sur EBAGG et sur les erreurs qu'elle indiquait plutôt que sur le mouvement, la moyenne des participants étant de 5.82 sur 10 (std : 2.28). Tandis qu'à la fin de la phase de *training* ils étaient plutôt concentrés sur la tâche (diagramme 4.9b), à savoir le geste à réaliser, la moyenne des participants étant alors à 6.73 (std : 2.12).

Enfin, nous avons questionné les participants sur leur ressenti sur leur performance et nous leur avons demandé s'ils avaient eu le sentiment de s'être amélioré dans la réalisation de leurs gestes au cours de l'expérimentation. Ils ont majoritairement eu l'impression de s'améliorer au cours du temps malgré le peu d'itérations de geste, avec une moyenne de 6.6 sur 10 (std : 2.04) comme le montre le diagramme de la figure 4.10.

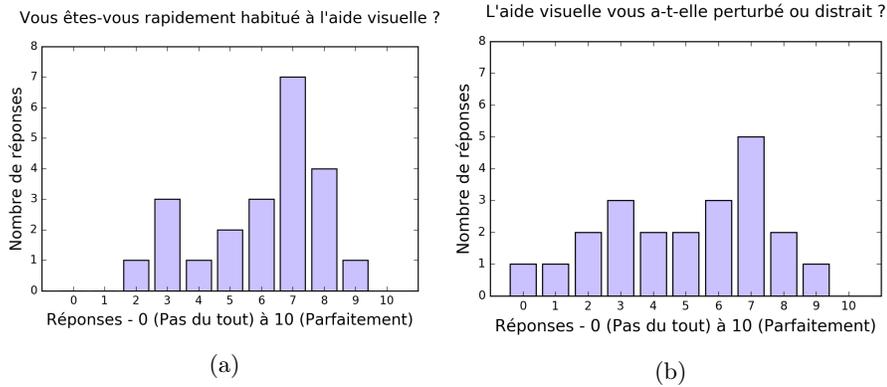


FIGURE 4.8 – Réponses des participants aux questions sur le comportement de la métaphore EBAGG.

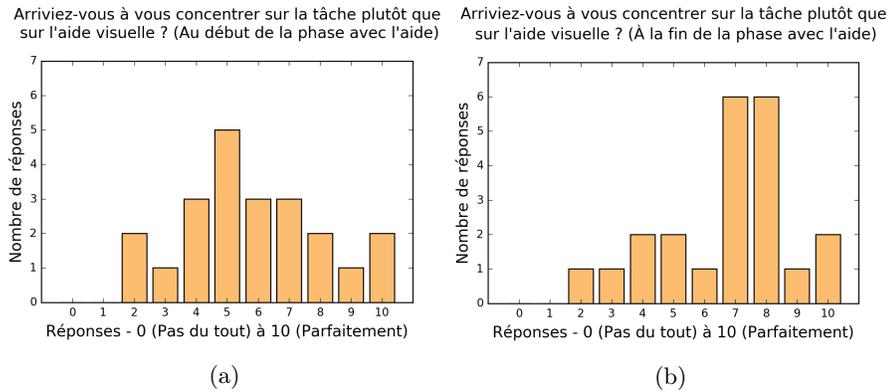


FIGURE 4.9 – Réponses des participants à la question : "Arrivez-vous à vous concentrer sur la tâche plutôt que sur l'aide visuelle?"

4.5 Discussion

Les résultats du premier geste sont relativement bons dès la phase de *pre-training*, ce qui indique que ce geste était assez simple à réaliser. La principale difficulté était de maîtriser la trajectoire de l'outil afin qu'il réalise une droite parfaite. Néanmoins, nous pouvons remarquer qu'ils s'améliorent rapidement dès que la métaphore EBAGG apparaît pour atteindre en moyenne 0.026 en fin de phase de *training*, ce qui est un coût cumulé assez faible et donc une bonne performance. De plus, la variance diminue tout au long de cet entraînement, surtout lors de la formation avec feedback. Les participants ont eu tendance à avoir des résultats assez similaires et à tendre vers le même niveau de performance. De plus, les corrections du geste apprises durant la phase avec feedback semblent perdurer à court terme en *post-training*. Les performances diminuent

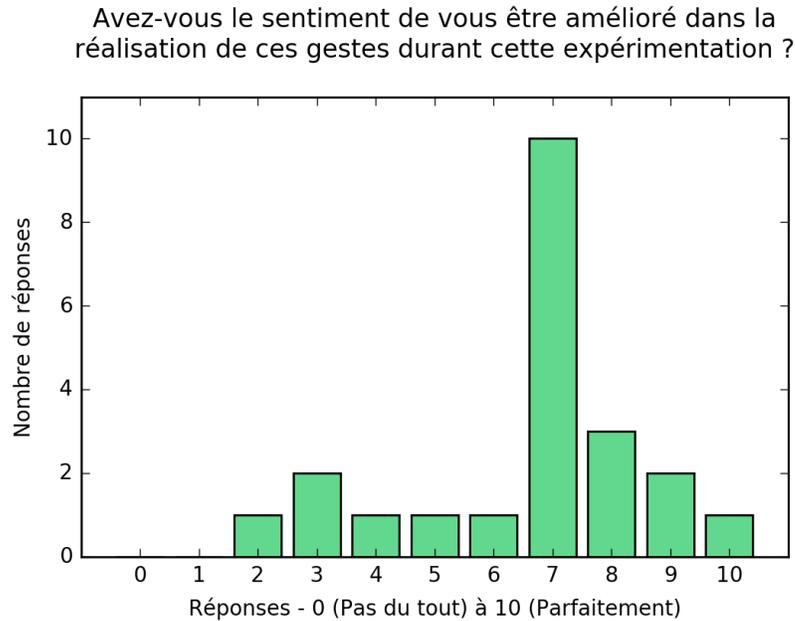


FIGURE 4.10 – Réponses des participants à la question : "Avez-vous le sentiment de vous être amélioré dans la réalisation de ces gestes durant cette expérimentation ?"

quelque peu et la variance augmente, néanmoins cette différence n'étant pas significative, cela n'influe pas ou très peu sur la qualité du geste final. Cette régression peut s'expliquer par le retrait de la métaphore qui a pu déstabiliser quelque peu les participants qui s'étaient habitués à sa présence. Toutefois, les performances étant similaires à la phase de *training*, cela pourrait signifier que les apprenants ne sont que très peu dépendants à la métaphore EBAGG à court terme et qu'ils arrivent à reproduire rapidement un geste assez simple, avec très peu d'itérations d'entraînement.

Cependant, les résultats obtenus pour le second geste nuancent ces conclusions. Pour un geste plus complexe à reproduire, l'amélioration des performances n'est pas aussi explicite. Aux vu des résultats du questionnaire et du diagramme 4.7, la métaphore semble avoir perturbé les participants au début de la phase de *training*, ce qui pourrait expliquer la baisse de performance. La difficulté du geste est aussi à prendre en compte. Il est en effet assez difficile de réaliser le geste tout en restant dans un plan vertical, c'est-à-dire en restant constant selon l'axe z . Leurs des premières itérations avec feedback, les apprenants avaient pour la majorité d'entre eux de nombreuses erreurs de trajectoire à corriger, principalement à cause de la profondeur. Une autre cause de ces résultats provient du paramétrage de la métaphore. En effet, le seuil de tolérance pour l'affichage des particules, soit le signalement de la présence

d'erreur dans le geste, était trop faible. Par conséquent, un faible écart par rapport à la trajectoire de référence impliquait un trop grand nombre de particules sur EBAGG, comme s'il s'agissait d'une erreur grossière. Ce paramétrage trop stricte a donc contribué à perturber les participants lors de la réalisation de leurs gestes et explique en partie la non-validation de notre seconde hypothèse pour ce deuxième geste. De plus, dès que la métaphore n'est plus visible pendant la phase de *post-training*, les performances s'améliorent immédiatement, ce qui conforte notre idée que son comportement gênait les apprenants, sans pour autant les empêcher de s'améliorer au vu de l'amélioration significative entre la phase de *pre-training* et celle de *post-training*.

Enfin, *a posteriori* de cette expérimentation, il s'avère qu'il aurait été plus judicieux de faire un questionnaire pour chaque geste plutôt que d'en faire un global à la fin de l'expérimentation. Nous aurions pu avoir une plus grande précision sur les résultats qualitatifs obtenus afin de déceler plus facilement les incompréhensions ou les gênes ressentis par les apprenants lors de la réalisation des gestes.

4.6 Conclusion

Dans ce chapitre, nous avons présenté une première expérimentation sur l'utilisation de la métaphore EBAGG dans le cadre de l'interaction gestuelle. L'objectif était d'observer sa capacité à guider l'apprenant lors de la reproduction de gestes précis et d'évaluer son impact. Pour cela, nous avons proposé aux participants de reproduire deux gestes en étant partiellement aidés par notre feedback visuel. Leurs gestes étaient comparés à un geste de référence. EBAGG leur indiquait alors la présence ou non d'erreurs, ainsi que la direction de ces dernières. Les particules présentes à la surface de la métaphore devenaient alors visibles. Plus l'erreur était importante, plus le nombre de particules apparentes étaient grand.

Les résultats présentés ont montré que notre feedback permettait d'améliorer les performances des utilisateurs lors de la reproduction de gestes. Les participants ont amélioré leurs performances après seulement quelques itérations. Néanmoins, l'utilisation de EBAGG pourrait présenter une charge cognitive qui influencerait sur le nombre d'itérations suffisantes pour l'entraînement, en fonction de la complexité du geste. Dans notre expérimentation, les sept itérations de la phase de *training* étaient suffisantes pour le premier geste et insuffisantes pour le second. En effet, ce geste était plus difficile à reproduire. Toutefois, l'étude qualitative a montré que les utilisateurs comprenaient rapidement les retours de la métaphore EBAGG, ce qui facilitait l'entraînement. Les apprenants arrivaient à se focaliser sur la tâche plutôt que sur l'aide visuelle après seulement quelques itérations.

Il serait donc intéressant de s'intéresser à cette charge cognitive et de comparer EBAGG à d'autres métaphores de guidage comme la visualisation de la trajectoire par exemple, afin de déterminer si notre métaphore a de meilleures performances. De plus, dans cette expérimentation, EBAGG avait une position

fixe en raison du matériel utilisé, notamment l'utilisation d'un seul écran pour le rendu du feedback. La position de la métaphore ainsi que sa mobilité influent peut-être sur la compréhension de son comportement et sur la perception des retours qu'elle propose. Les résultats de ces travaux ont également donné lieu à une publication à la conférence IEEE ICALT 2017 [Jeanne et al., 2017a]. Dans le chapitre suivant, nous présentons une expérimentation ayant lieu dans un CAVE⁷. Nous y abordons la question de l'évolution des performances lorsque l'on retire le feedback et donc de la dépendance lors de la reproduction d'un geste précis.

7. Acronyme pour *cave automatic virtual environment*

Chapitre 5

Evaluation de l'impact de EBAGG sur les performances des utilisateurs

5.1 Introduction

Dans le chapitre précédent, nous avons présenté une expérimentation visant à évaluer la capacité de la métaphore EBAGG à guider l'apprenant lors d'un exercice de reproduction de gestes précis. Nous nous sommes intéressés à l'évaluation de sa capacité à guider l'apprenant en lui indiquant ses erreurs de trajectoire. Nous avons ainsi constaté que notre feedback permettait effectivement de guider l'apprenant dans la réalisation de ses gestes et de l'aider à corriger ses erreurs afin de se rapprocher le plus possible du geste de référence. De plus, pour un geste difficile à réaliser, quelques itérations d'entraînement ont suffi pour améliorer significativement les performances des apprenants. L'étude qualitative a également montré que les utilisateurs comprenaient rapidement le comportement de la métaphore et les informations qu'elle leur retournait. Toutefois, le paramétrage trop exigeant ainsi que la charge cognitive générée par la métaphore ont quelque peu perturbé les participants de l'expérimentation lors de la reproduction d'un geste difficile.

Suite à ces résultats et conclusions, nous devons toutefois nous interroger sur leur signification et les mettre en perspective. Nous n'avons testé jusque là que l'utilisation de EBAGG lors d'un exercice de reproduction de gestes. Certes, l'amélioration des performances est visible pour les deux gestes malgré un paramétrage trop exigeant de la métaphore, néanmoins peut-on pour autant affirmer que cette hausse provient de la présence du feedback ? La simple répétition des gestes ne serait-elle pas la cause de cette amélioration ?

Nous proposons donc une nouvelle expérimentation visant à comparer les performances obtenues à l'aide de EBAGG, à un groupe de contrôle ne recevant aucune aide lors de la réalisation de leur geste. Il est également intéressant de faire cette comparaison avec des métaphores usuelles de guidage gestuel en nous

basant sur les différentes aides présentées dans le chapitre 2. Cette comparaison nous permet de déterminer si notre métaphore permet d'apprendre plus efficacement le geste proposé, c'est-à-dire qu'après le même nombre d'itérations d'entraînement avec feedback, les utilisateurs ayant eu EBAGG auraient de meilleurs résultats une fois l'aide retirée.

Ce chapitre se décompose de la manière suivante. Tout d'abord nous revenons sur le choix des différentes aides usuelles de guidage gestuel, puis dans la section 5.3 sur les hypothèses que nous avons établies au préalable de l'expérimentation. Dans la partie 5.4 nous détaillons le protocole expérimental, l'explicitation du geste à reproduire, puis les conditions et le dispositif expérimental. Ensuite, dans la partie 5.5, nous présentons les résultats obtenus. Nous discutons de l'ensemble de cette expérimentation dans la partie 5.6, avant de conclure.

5.2 Métaphores de guidage choisies

La première expérimentation ne concernait que notre métaphore EBAGG. Ici nous comparons ses performances à deux autres feedbacks. A partir des travaux étudiés dans l'état de l'art de ce mémoire (section 2.3), nous avons déterminé quelles étaient les techniques usuelles de guidage visuel les plus appropriées pour notre expérimentation.

Tout d'abord, différentes études ont été menées sur les avatars en vue à la troisième personne [Rovelo et al., 2015, Velloso et al., 2013], directement superposés sur l'utilisateur [Anderson et al., 2013] ou même sur une combinaison des deux superposant l'avatar de l'expert sur l'avatar de l'utilisateur [Hoang et al., 2016, Schönauer et al., 2012, Yang and Kim, 2002]. Néanmoins, cette expérimentation vise à comparer EBAGG à d'autres méthodes de guidage visuel du geste, il nous semble donc plus pertinent de sélectionner des techniques assez proches de notre feedback dans leur fonctionnement. Les avatars représentent partiellement ou entier, le corps de l'utilisateur ou de l'expert, tandis que EBAGG agit comme un élément extérieur venant potentiellement s'ajouter aux conseils de l'expert supervisant l'entraînement. De plus d'un point de vue technique, l'utilisation d'un avatar signifiait de capturer les positions et mouvements des différentes parties du corps de l'utilisateur, ce qui est notablement différent de techniques plus classiques comme EBAGG où nous ne suivons que les mouvements d'un outil ou d'une seule partie du corps comme la main par exemple. Nous faisons donc le choix de ne pas retenir cette solution.

Les techniques de feedforward sont en revanche très utiles pour guider les apprenants selon la trajectoire idéale durant la réalisation de leurs gestes, puisqu'ils renseignent sur le geste à reproduire avant que les utilisateurs ne s'y essaient. Contrairement aux **feedbacks** qui sont fournis durant (c'est-à-dire de l'ordre de millièmes ou centièmes de secondes après) ou *a posteriori*, donc après la réalisation de l'action, les **feedforward** sont en effet proposés *a priori*, c'est-à-dire avant que le geste ne soit réalisé. Pour le guidage gestuel, différentes techniques de feedforward sont utilisées mais restent néanmoins assez

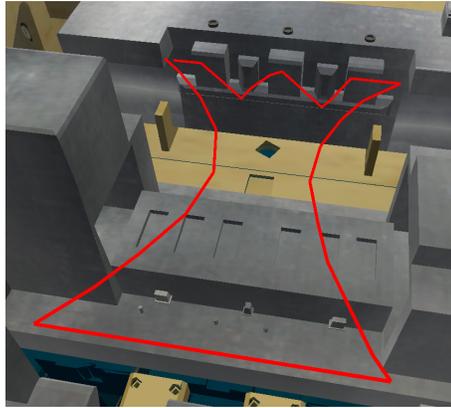


FIGURE 5.1 – Tracé de la trajectoire du geste de référence à reproduire

similaires. Nous avons donc choisi comme première métaphore de représenter la trajectoire du geste de référence (figure 5.1). Nous nous inspirons des travaux de Covaci *et al.* [Covaci et al., 2015] et de Moreau [Moreau et al., 2007], qui positionnent des points de passage le long de la trajectoire idéale et par lesquels les utilisateurs doivent passer. Nous avons cependant décidé de relier ces points afin d'avoir un tracé continu qui nous semble plus facile à suivre avec un outil que des points espacés.

La seconde métaphore choisie est un feedback se basant sur l'indication de l'orientation, c'est-à-dire sur la visualisation de la direction à prendre pour suivre la trajectoire idéale. Nous nous basons sur les travaux de Henderson et Feiner [Henderson and Feiner, 2011], Tönnis et Klinker [Tönnis and Klinker, 2009], et surtout sur ceux de Sodhi *et al.* [Sodhi et al., 2012] (*3D Arrow*) pour modéliser une flèche en trois dimensions qui pointe dans la direction dans laquelle aller (figure 5.2). Dans son implémentation, nous avons décidé de n'afficher la flèche uniquement lorsque des erreurs sont commises. Ainsi, si l'utilisateur est assez proche de la trajectoire idéale, la flèche est invisible. L'apparition de la flèche lui indique alors une erreur de trajectoire et son orientation montre la direction qu'il doit prendre. De plus, le degré de l'erreur est symbolisé par la taille de la flèche qui grossit ou rétrécit en fonction de l'écart à la performance de référence.

Enfin, nous utilisons la deuxième version de EBAGG (figure 5.3) qui gère mieux son comportement vis-à-vis de la valeur de la distance retournée par l'algorithme DTW et qui possède un plus grand nombre de particules à sa surface.

5.3 Hypothèses

La première expérimentation visait à étudier les performances de EBAGG uniquement. L'ensemble des participants recevait une aide visuelle en prove-



FIGURE 5.2 – Flèche rouge en trois dimensions montrant la direction dans laquelle se diriger pour suivre à nouveau la trajectoire idéale du geste de référence à reproduire

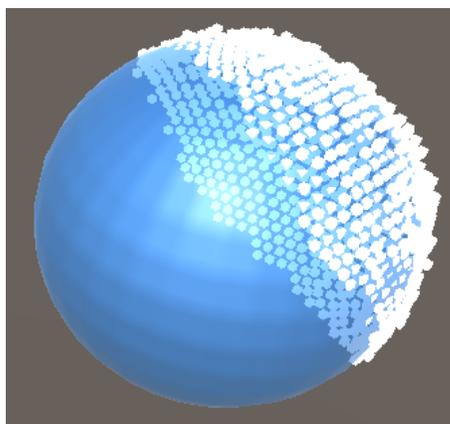


FIGURE 5.3 – EBAGG et ses particules indiquant la direction et l'intensité de l'erreur

nance de la métaphore EBAGG qui indiquait les erreurs de trajectoires. La seconde étape de notre recherche est donc de comparer les performances de notre feedback à des métaphores couramment utilisées pour guider la gestuelle des utilisateurs. EBAGG permet-elle en effet d'obtenir de meilleures performances lors d'un exercice de reproduction de gestes? Toutefois, cette étude ne s'intéresse évidemment qu'à l'amélioration des performances de gestes précis lors de la réalisation d'une tâche industrielle. Les résultats et conclusions obtenus ne s'appliquent que dans ce cadre-là.

De plus, cette étude cherche à comparer les résultats obtenus par les groupes ayant comme aide les différentes métaphores de guidage à un groupe de contrôle qui ne reçoit aucune feedback. L'exercice de celui-ci ne comprend qu'une ré-

pétition de gestes. L'objectif ici est de vérifier que ce sont bien les aides de guidage qui améliorent significativement les performances et non la répétition du geste.

Suite à la précédente expérimentation, nous avons également décelé une probable charge cognitive importante pour EBAGG lors de l'entraînement aux gestes. Pour un geste plus difficile à reproduire, un nombre supérieur d'itérations d'apprentissage était requis pour obtenir une amélioration importante des performances. Il semble donc intéressant d'évaluer à la fois les performances des différents groupes d'apprentissage, mais également la charge cognitive développée lors de l'entraînement. EBAGG semble en effet avoir une charge cognitive importante, sera-t-elle tout de même inférieure aux autres ? Il y a-t-il une différence au niveau de la charge entre des aides dynamiques, des aides statiques et l'absence d'aide de guidage ?

Enfin, le problème de dépendance soulevé précédemment dans ce manuscrit devrait être visible dans cette étude. Pour rappel, l'hypothèse de guidage stipule que l'utilisation de de feedbacks en continu pendant la phase d'apprentissage génère une dépendance de l'apprenant vis-à-vis de ces feedbacks [Salmoni et al., 1984, Schmidt, 1991, Sigrist et al., 2013]. EBAGG a été pensée et conçue dans le but de réduire au maximum cette dépendance. Cette dernière devrait donc être moins importante pour notre feedback que pour les autres groupes d'apprentissage.

L'expérimentation vise donc à vérifier les hypothèses suivantes :

- hypothèse 1 (**H1**) : les sujets des trois groupes recevant l'aide d'une métaphore de guidage durant la phase d'entraînement avec feedback auront de meilleures performances à la fin de l'exercice que les sujets du groupe de contrôle ;
- hypothèse 2 (**H2**) : les sujets du groupe ayant EBAGG comme métaphore de guidage auront de meilleures performances que les autres groupes ;
- hypothèse 3 (**H3**) : à la fin de l'exercice, les sujets du groupe ayant EBAGG comme métaphore de guidage auront des performances similaires à celles qu'ils avaient la fin de l'entraînement avec feedback.

5.4 Expérimentation

5.4.1 Tâche

Lors de cette expérimentation, les participants doivent également reproduire un geste avec une haute précision en face d'un moule virtuel, mais cette fois-ci dans un CAVE (notre plateforme est décrite dans la section 5.4.4). Un CAVE (acronyme de *Cave automatic virtual environment*) est une salle immersive, généralement de forme généralement cubique, qui fut conçue à l'origine par Cruz-Neira *et al.* (*The CAVE*) [Cruz-Neira et al., 1992]. Cette plateforme comprend plusieurs murs (de deux à six), pouvant inclure le sol et le plafond, sur lesquels sont projetées les images de l'environnement virtuel. La stéréoscopie permet d'avoir l'image et donc les objets de l'environnement en trois

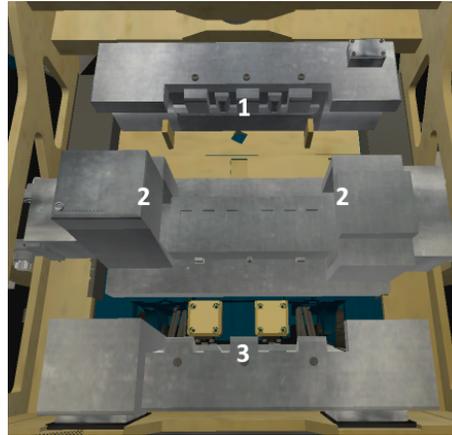


FIGURE 5.4 – Moule : (1) chape injecteur, (2) appuis noyaux, (3) chape collecteur

dimensions. De plus, le *tracking* des mouvements de l'utilisateur est réalisé à l'aide de caméras infrarouges. L'environnement virtuel utilisé est le même que pour la première expérimentation. Néanmoins, cette fois-ci le moule n'est plus projeté en vue de dessus puisque comme dans la première expérimentation, mais en vue réelle à l'échelle 1 :1 dans le CAVE. Les utilisateurs peuvent donc se déplacer autour du moule et visualiser l'ensemble des éléments de l'environnement.

Le geste à reproduire est une modélisation des actions couramment réalisées sur le poste réel de moulage chez Montupet. Il comprend le nettoyage de certaines parties du moules : les appuis noyaux de toit sur le pont de coulée et à l'opposé, la chape injecteur et la chape collecteur (figure 5.4). Le geste modélisé est représenté par sa trace sur la figure 5.5.

5.4.2 Participants

Quarante-huit sujets ont participé à cette expérimentation (douze femmes et trente-six hommes) et ils étaient âgés de 18 à 60 ans (*mean* : 27, *std* : 9.36). Seulement six participants étaient gauchers, mais ils ont eu les mêmes conditions expérimentales puisque les feedbacks dynamiques (la flèche et EBAGG) suivaient les mouvements du *wand*¹. La majorité de ces personnes étaient des étudiants de l'université ou des membres du laboratoire. Ils avaient des expériences diverses avec les équipements de réalité virtuelle, notamment le CAVE, certains (moins de 5%) n'ayant jamais utilisé de lunettes stéréoscopiques. Les

1. En réalité virtuelle, un wand est un dispositif d'interaction. Il comprend généralement un *joystick* pour se déplacer et des boutons pour interagir avec l'environnement virtuel. Sa position dans l'espace peut également être obtenue via un capteur ou en utilisant une combinaison de marqueurs et de caméras infrarouges.

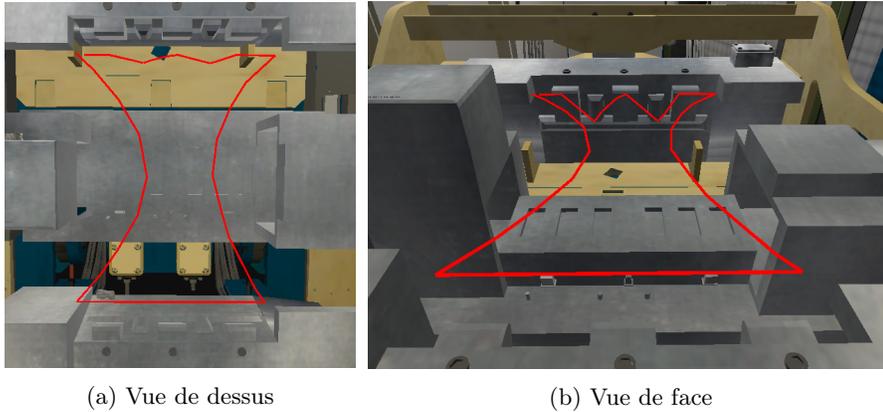


FIGURE 5.5 – Tracé de la trajectoire du geste de référence à reproduire

sujets de cette expérimentation n'ont reçu aucune rémunération pour leur participation.

5.4.3 Protocole expérimental

Les participants sont au préalable aléatoirement partagés en quatre groupes :

- le groupe de contrôle qui n'a aucune aide de guidage durant cet exercice ;
- le groupe ayant l'aide de guidage "trace" ;
- le groupe ayant l'aide de guidage "flèche 3D" ;
- le groupe ayant l'aide de guidage "EBAGG".

Ils doivent au préalable s'équiper de surchaussures jetables afin de ne pas salir le revêtement du CAVE (voir section 5.4.4), puis le déroulement de l'exercice leur est expliqué. Tout d'abord, les participants sont immergés dans un environnement d'entraînement basique où ils peuvent s'accommoder à l'équipement de réalité virtuelle et explorer l'environnement virtuel qui est l'usine virtuelle utilisée dans la précédente expérimentation. Durant cette phase de familiarisation, le superviseur de l'expérimentation leur explique la tâche qu'ils doivent réaliser. Dans notre première étude, ils devaient reproduire le geste treize fois mais cela était insuffisant pour un geste complexe. Dans celle-ci, ils doivent donc reproduire vingt fois le geste présenté dans la section 5.4.1. En déplaçant la soufflette à air comprimé virtuelle, les sujets doivent passer par des points spécifiques du moule tout en se déplaçant selon les trois dimensions (voir figure 5.5). Le geste leur était montré deux fois sans qu'ils ne soient autorisés à s'entraîner à le reproduire afin que tous les participants aient les mêmes conditions expérimentales.

Ensuite, les sujets des groupes ayant une métaphore de guidage durant cet entraînement de reproduction d'un geste précis se voient expliquer le concept et

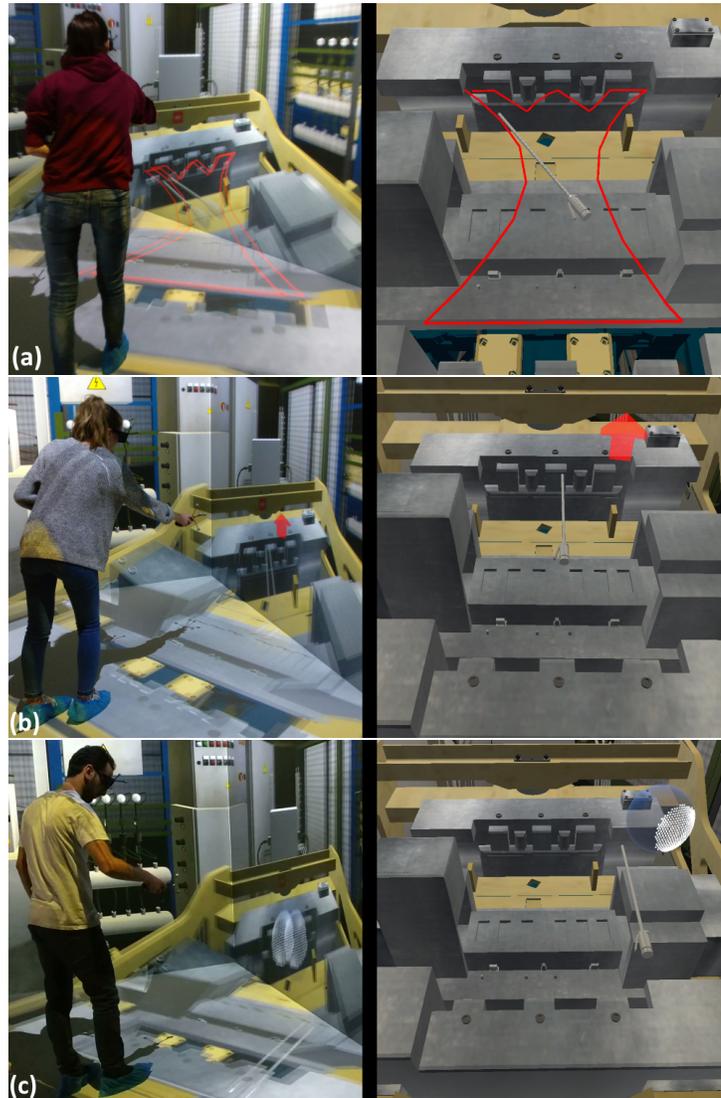


FIGURE 5.6 – Chaque figure illustre la vue d'un participant de l'expérimentation utilisant une métaphore de guidage et la visualisation de cette aide dans l'environnement virtuel

le fonctionnement de celle-ci. Pour illustrer le principe de la flèche 3D et de la métaphore EBAGG, un cube vert est placé au centre du moule et la position de la soufflette virtuelle est alors comparée avec celle du cube. Cette comparaison assez simpliste permet ainsi d'expliquer assez facilement le comportement des particules de EBAGG ainsi que l'orientation et la variation de la taille de la

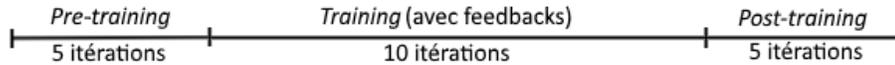


FIGURE 5.7 – Différentes phases de l'expérimentation

flèche 3D.

L'exercice comprend trois phases d'entraînement pour les groupes recevant l'aide d'une métaphore de guidage (figure 5.7) : une phase de *pre-training* de cinq itérations du geste où les sujets ne reçoivent l'aide d'aucun feedback sur leurs performances ; une phase de *training* de dix itérations où les participants sont respectivement aidés par la trace du geste (figure 5.6(a), la flèche 3D (figure 5.6(b)) ou la métaphore EBAGG (figure 5.6(c)) selon le groupe auquel ils appartiennent ; une phase de *post-training* de cinq itérations où les aides de guidages n'étaient plus affichées. Le groupe de contrôle quant à lui ne doit que reproduire le geste vingt fois sans aucune aide provenant d'une métaphore de guidage.

Une fois les vingt itérations du geste réalisées, les participants sont soumis à un questionnaire où leurs sont posées des questions sur l'expérimentation et sur la métaphore de guidage utilisée (à l'exception du groupe de contrôle).

5.4.4 Dispositif expérimental

L'étude se déroule dans la plateforme TRANSLIFE² du laboratoire Heudiasyc. Elle comprend principalement un CAVE qui a la particularité d'avoir une face mobile. En effet, traditionnellement en forme de "U" (figure 5.8a), ici un des murs peut pivoter avoir une configuration en "L" (figure 5.8b). Nous avons fait le choix de faire cette expérimentation en utilisant la configuration en "L" du CAVE puisqu'elle permet tout simplement d'avoir une plus grande superficie au sol et d'ainsi avoir plus d'espace autour de l'utilisateur (figure 5.9). Cela permet également au superviseur de l'expérimentation de venir à côté de l'apprenant pour observer sans être trop proche de celui-ci pour ne pas le perturber. En outre, contrairement à la plateforme de la première expérimentation, l'utilisation du CAVE permet à l'utilisateur de voir son propre corps au sein de l'environnement 3D, ce qui favorise le sentiment d'immersion.

La plateforme est composée de trois écrans en verre de 3.4m x 2.5m, ainsi qu'une superficie au sol de 7m x 3.4m (pour la configuration en "L"). Ce dernier comprend une peinture réfléchissante blanche, ce qui nous permet d'avoir la projection d'une image stéréoscopique de bonne qualité. Chaque écran est apparié à un projecteur stéréoscopique (*Christie Mirage 3D Projector*³) d'une résolution de 1920 x 1200 pixels. Les participants portent donc des lunettes stéréoscopiques Volfoni EDGE RF sur lesquelles sont placés des marqueurs réfléchissants afin d'assurer le *head tracking*.

2. <https://translife.hds.utc.fr/>

3. <https://www.christiedigital.com>

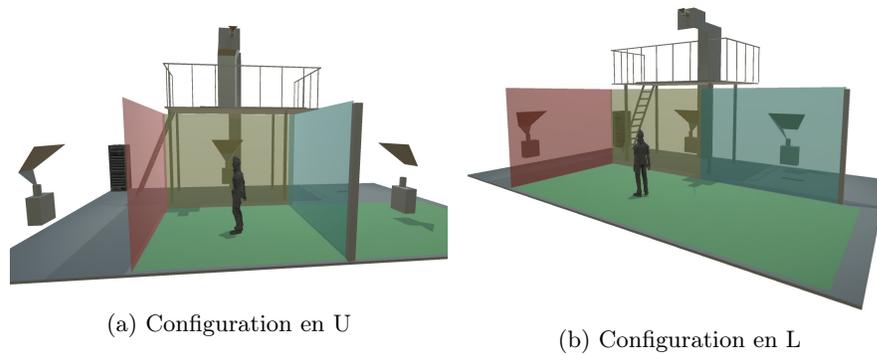


FIGURE 5.8 – CAVE TRANSLIFE



FIGURE 5.9 – Apprenant réalisant un geste dans le CAVE TRANSLIFE

Le *tracking* est assuré par le système de capture de mouvements du CAVE (figure 5.9). Il comprend un total de dix caméras infrarouges incluant huit OptiTrack Prime 13W et deux Optitrack Prime 13 qui ont les mêmes capteurs d'images présentés dans le chapitre précédent. La différence entre ces deux types de caméras se situe au niveau des types de diodes électroluminescentes utilisées (dix diodes ultra-puissantes pour les Prime 13W contre soixante-deux diodes classiques pour les Prime 13) et des lentilles. Le modèle Prime 13W permet en effet d'avoir un FOV horizontal (*field of view* ou champ de vue) de 86° , qui est plus large que les 56° du modèle Prime 13. Les caméras Prime 13W étant des caméras à grand angle sont ainsi plus adaptées pour un CAVE.

Ensuite, chaque caméra détecte les marqueurs situés dans la scène dont les positions sont envoyées à l'environnement virtuel via le logiciel Motive. Ici,

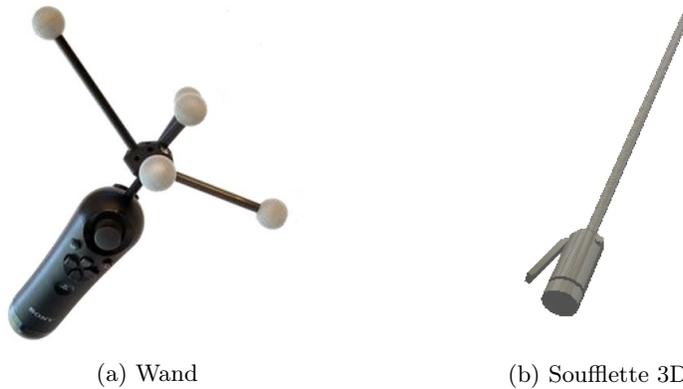


FIGURE 5.10 – Wand utilisé pour l'expérimentation associé au modèle 3D de la soufflette

nous utilisons une manette PlayStation Move⁴ comme *wand* sur laquelle nous plaçons des marqueurs réfléchissants (figure 5.10a). Le modèle 3D de la soufflette (figure 5.10b) était alors associé au wand (figure 5.9). Les déplacements du modèle 3D correspondaient alors aux mouvements du wand.

Enfin, le rendu de cet environnement virtuel est réalisé à l'aide de la station de travail HP Z840⁵, un ordinateur destiné à cet effet, notamment grâce aux deux cartes graphiques Nvidia M6000 ayant 24 Go de mémoire chacune. Il possède également deux processeurs Intel Xeon E5-2667 possédant chacun six cœurs à 2.90 - 3.50 GHz.

5.4.5 Evaluation des performances

EBAGG et la flèche fournissent un feedback en temps réel. Ces deux aides visuelles utilisent toutes les deux l'algorithme DTW que nous avons présenté dans la partie 3 de ce mémoire. De la même manière que dans la première expérimentation, nous calculons la distance entre le geste de l'utilisateur et le geste de référence enregistré au préalable à l'aide cet algorithme. Pour rappel, il nous fournit des résultats interprétables tels que la correspondance de la dernière position de l'outil tenu par l'utilisateur lors de la réalisation du geste avec la position de référence et la distance euclidienne les séparant (l'ensemble de ces trois données représente le *matching*), ainsi que le coût cumulé calculé sur le geste partiel (geste en cours de réalisation). Ici, la distance de chaque *matching* sert de donnée d'entrée pour notre métaphore EBAGG mais également pour la flèche 3D, qui modifient leur comportement en fonction de sa valeur. De même, le coût cumulé agit en tant que score pour évaluer la performance globale de l'utilisateur. Plus sa valeur est proche de zéro, plus la performance de l'utilisateur est proche de celle de la référence.

4. <https://www.playstation.com>

5. <http://www.hp.com/>

Les données des gestes d'un participant sont enregistrées au moment il débute son geste en appuyant sur l'un des boutons du *wand*, jusqu'au nouvel appui du bouton signifiant la fin du geste. Durant cette phase d'acquisition des données, la position du *wand* est sauvegardée et transmise à l'algorithme DTW avec une fréquence de cinq messages par seconde.

En plus de cette évaluation quantitative, un questionnaire est également utilisé pour obtenir une évaluation qualitative de l'expérimentation, mais aussi des différentes métaphores de guidage utilisées.

5.5 Résultats

Cette section décrit les résultats obtenus pour l'expérimentation décrite précédemment. Les résultats sont analysés pour chaque groupe ; nous réalisons une régression linéaire sur les résultats des quinze premières itérations du geste, ce qui correspond aux phases de *pre-training* et de *training*, afin de déterminer si l'entraînement avec ou sans métaphore de guidage a un impact sur les performances. Nous obtenons ainsi des droites selon l'équation 5.1, dont les valeurs des coefficients sont regroupées dans le tableau 5.1. Nous analysons par la suite les résultats de la phase de *post-training* pour chaque groupe avant de nous intéresser à l'étude qualitative via le questionnaire que nous avons proposé aux participants.

$$y = \alpha * x + \beta \tag{5.1}$$

TABLE 5.1 – Valeurs de la pente et de l'ordonnée à l'origine 5.1 pour chaque groupe

	Groupes			
	Contrôle	Trace 3D	Flèche 3D	EBAGG
valeurs α	-0.0004408	-0.002269	-0.000056	-0.0008376
valeurs β	0.0588256	0.059406	0.04546	0.0480623

5.5.1 Impact de l'entraînement avec et sans métaphore de guidage

Avant d'analyser les résultats de l'expérimentation, nous appliquons comme lors de la première expérimentation le test statistique de Shapiro-Wilk aux données de chaque groupe afin de vérifier si notre échantillon provient d'une population normalement distribuée.

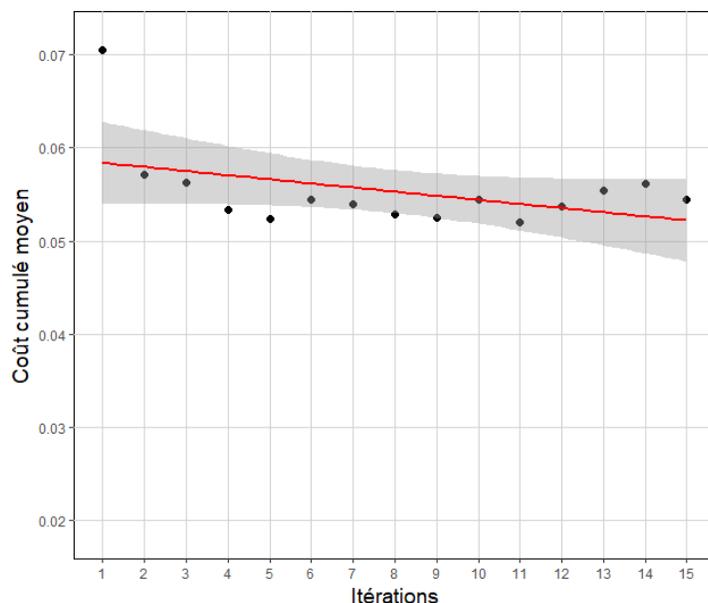


FIGURE 5.11 – Régression linéaire sur le coût cumulé moyen des quinze premières itérations pour le groupe de contrôle

Groupe contrôle

La figure 5.11 présente la régression linéaire pour les quinze premières itérations du groupe de contrôle n'ayant reçu l'aide d'aucun feedback.

La pente (équation 5.1 et tableau 5.1) de cette droite a une faible valeur ($\alpha = -0.0004408$). En effet, la valeur du coût cumulé moyen étant de $\bar{y}_{pre} = 0.0579$ durant la phase de *pre-training*, diminue pour atteindre $\bar{y}_{tbegin} = 0.0536$ puis $\bar{y}_{tend} = 0.0543$ pour respectivement les cinq premières itérations et les cinq dernières itérations de la phase de *training*. Il y a donc bien une légère amélioration.

Le test de Shapiro-Wilk appliqué aux données montre qu'elles sont normalement distribuées. Nous effectuons donc une analyse de la variance sur ces résultats. Nous observons alors qu'il n'y a aucun effet significatif ($F(1, 13) = 3.1549, p\text{-value} = 0.09909$). Les niveaux de performances restent relativement constants et nous pouvons donc en conclure que les participants du groupe contrôle ne se sont pas significativement améliorés. De plus, le coefficient de détermination a pour valeur $R^2 = 0.1953$, ce qui signifie que cette régression linéaire n'est pas convenablement adaptée pour décrire la distribution des données. Ces résultats nous montre qu'en l'absence de métaphore de guidage, les participants n'ont pas réussi à s'améliorer. La simple répétition du geste ne semble donc pas être suffisante comme stratégie d'entraînement.

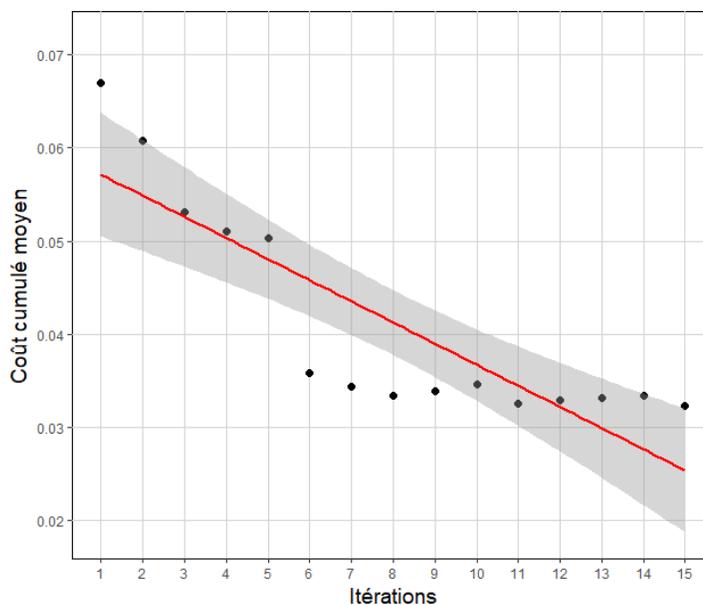


FIGURE 5.12 – Régression linéaire sur le coût cumulé moyen des quinze premières itérations pour le groupe utilisant la trace comme métaphore de guidage

Groupe trace

Ici, la figure 5.12 présente la régression linéaire pour les quinze premières itérations du groupe ayant eu la trace du geste comme métaphore de guidage.

La pente (équation 5.1 et tableau 5.1) de la droite est une nouvelle fois négative, néanmoins sa valeur absolue est dix fois supérieure à celle du groupe précédent ($\alpha = -0.002269$). Cela indique une amélioration nettement supérieure pour le groupe ayant eu la trace comme aide durant l'entraînement. De plus, on remarque que la valeur moyenne du coût cumulé diminue considérablement entre la phase de *pre-training* ($\bar{y}_{pre} = 0.0564$) et la phase de *training* ($\bar{y}_{tbegin} = 0.0344$ et $\bar{y}_{tend} = 0.0329$). Une forte amélioration des performances est notamment visible sur la figure 5.12 dès la sixième itération. C'est en effet à cette itération que la trace apparaissait.

Le test de Shapiro-Wilk appliqué aux données a ici aussi montré qu'elles étaient normalement distribuées. Néanmoins, cette fois l'analyse de la valeur a montré qu'il y a avait un effet significatif de la métaphore de la trace sur l'amélioration des performances ($F(1, 13) = 36.88, p\text{-value} = 0.00004$). De plus, le coefficient de détermination est égal à $R^2 = 0.7393$, ce qui indique que notre modèle correspond convenablement aux données des quinze premières itérations. La trace comme métaphore de guidage semble donc avoir amélioré les performances des participants. L'affichage de la trace du geste de référence permet ainsi d'obtenir facilement de bon résultats et rapidement, ce qui montre

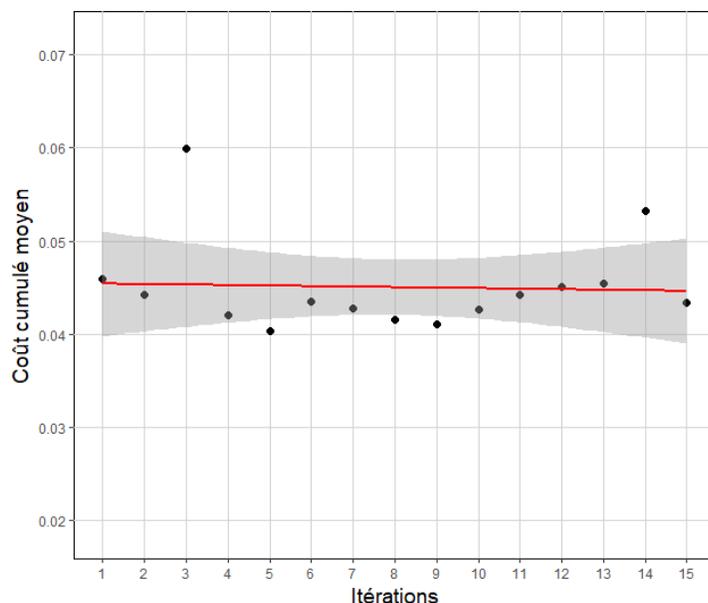


FIGURE 5.13 – Régression linéaire sur le coût cumulé moyen des quinze premières itérations pour le groupe utilisant la flèche 3D comme métaphore de guidage

la facilité que les utilisateurs ont eu pour s’y habituer et pour l’utiliser.

Groupe flèche 3D

La figure 5.13 présente la régression linéaire pour les quinze premières itérations du groupe ayant eu la flèche 3D comme métaphore de guidage.

Cette fois-ci la pente (équation 5.1 et tableau 5.1) est nettement inférieure aux deux premières ($\alpha = -0.000056$). On peut notamment voir sur la figure 5.13 que la droite est pratiquement à l’horizontale. La valeur moyenne du score diminue effectivement entre la phase de *pre-training* ($\bar{y}_{pre} = 0.0465$) et la phase de *training* ($\bar{y}_{tbegin} = 0.0423$ et $\bar{y}_{tend} = 0.0463$). Toutefois, cette amélioration est minime et on peut remarquer que les performances des apprenants se détériorent durant la phase de *training*.

Le test de Shapiro-Wilk appliqué aux données nous montre qu’elles suivent une loi normale. Néanmoins, l’analyse de la variance nous montre bien que cette métaphore de guidage n’a aucun effet significatif sur l’amélioration des performances ($F(1, 13) = 0.0312, p = 0.8625$). De plus, la valeur du coefficient de détermination est égale à $R^2 = 0.0024$, impliquant que la régression linéaire n’est pas convenablement adaptée pour décrire la distribution des données. Cette métaphore ne semble donc pas adéquate pour le guidage de geste selon les conditions de l’expérimentation.

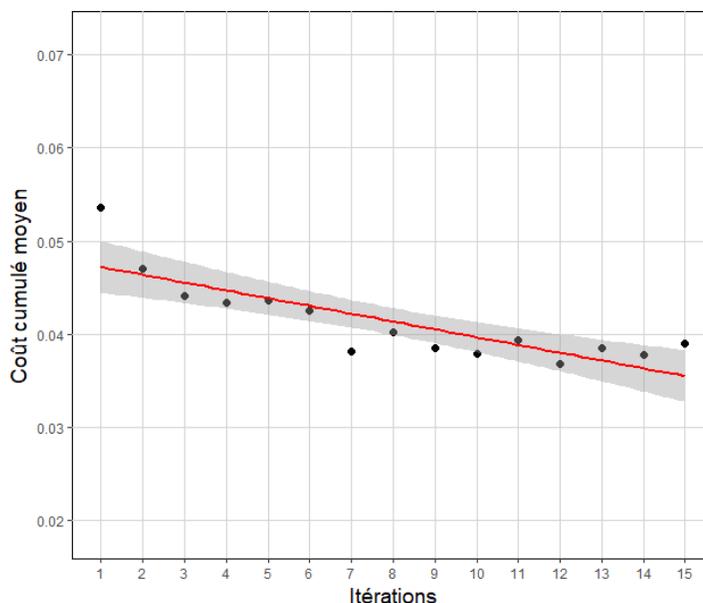


FIGURE 5.14 – Régression linéaire sur le coût cumulé moyen des quinze premières itérations pour le groupe utilisant EBAGG comme métaphore de guidage

Groupe EBAGG

Enfin, la figure 5.14 présente la régression linéaire pour les quinze premières itérations du groupe ayant eu la métaphore EBAGG comme aide pour guidage du geste.

La pente (équation 5.1 et tableau 5.1) est inférieure à celle du groupe ayant eu la trace, néanmoins elle reste supérieure à celles des deux autres groupes ($\alpha = -0.000056$). On peut remarquer sur la figure 5.14 que les participants se sont améliorés au cours de ces quinze premières itérations. Toutefois, il cette hausse des performances n'est pas aussi accentuée que sur la figure 5.12 où l'écart était important entre les itérations 5 et 6. Ici, l'amélioration est progressive. La valeur moyenne passe en effet de $\bar{y}_{pre} = 0.0464$ pour la phase de *pre-training*, à $\bar{y}_{tbegin} = 0.0395$ pour les cinq premières itérations de la phase de *training*, puis $\bar{y}_{tend} = 0.0383$ pour les cinq dernières itérations de cette même phase.

Une nouvelle fois, le test de Shapiro-Wilk appliqué aux données à montré que ces dernières suivaient une loi normale. L'analyse de la variance révèle que la métaphore EBAGG a un impact significatif sur les performances des apprenants sur ces quinze itérations du geste ($F(1, 13) = 28.74, p = 0.00013$). De plus, la valeur du coefficient de détermination est de $R^2 = 0.6885$, ce qui implique que le modèle linéaire proposé est convenablement adapté pour décrire

la distribution des données. Ainsi l'utilisation de la métaphore EBAGG comme aide pour le guidage semble permettre une amélioration des performances lors de l'entraînement à la reproduction de geste précis.

5.5.2 Analyse de la phase de *post-training*

Lors de cette dernière phase de cinq itérations, les métaphores de guidage sont retirées aux trois groupes. L'objectif ici est de déterminer si ces aides visuelles ont un quelconque impact sur l'apprentissage à très court terme du geste.

La figure 5.15 présente les résultats de toutes les phases d'entraînement pour l'ensemble des groupes. Contrairement à ce que l'on pouvait prévoir, les quatre groupes n'ont pas des scores du même ordre de grandeur pour la phase de *pre-training* alors qu'ils ont tous réalisé ces cinq itérations du geste selon les mêmes conditions expérimentales. En effet, pour le groupe de contrôle et pour le groupe de la trace du geste, la valeur moyenne du coût cumulé est d'environ $\bar{y}_{pre} \simeq 0.057$, tandis que pour les groupes de la flèche 3D et de EBAGG la valeur est de l'ordre de $\bar{y}_{pre} \simeq 0.046$. Par conséquent, les résultats de la phase de *post-training* des différents groupes ne sont pas comparables entre eux, puisque l'écart entre ces derniers pourrait être expliqué par une composante autre que l'entraînement. Néanmoins, l'impact des données aberrantes aurait pu être réduit en ayant un plus grand nombre de participants. Nous choisissons donc de faire une comparaison **intragroupe** plutôt que intergroupe. Nous analysons ainsi la différence de performance, si elle existe, entre la phase de *pre-training* et celle de *post-training*. Nous utilisons une nouvelle fois le test de Shapiro-Wilk afin de déterminer si les données sont normalement distribuées. Si elles suivent bien une distribution normale, nous utilisons alors le test t de Student pour données appariées.

Pour le groupe de contrôle, les performances des participants durant cette phase de *post-training* sont du même ordre de grandeur que celles des phases précédentes : $\bar{y}_{pre} = 0.0579$, $\bar{y}_{tbegin} = 0.0537$, $\bar{y}_{tend} = 0.0543$ et $\bar{y}_{post} = 0.0538$. Ces résultats rejoignent nos conclusions précédentes et confirment la tendance de notre régression linéaire, le niveau de performance des participants reste constant tout au long de l'exercice de reproduction de geste, il n'y a pas d'amélioration significative entre les phases de *pre-training* et de *post-training* ($t = 1.24$, $p - value = 0.28$).

Néanmoins, les résultats obtenus par le groupe ayant eu la trace ne correspondent pas à la tendance établie par la régression linéaire. Les performances diminuent fortement durant la phase de *post-training* lorsque le feedback n'est plus présent pour les guider (figure 5.15). La valeur moyenne du coût cumulé augmente (donc les performances diminuent) par rapport à la fin de la phase de *training*, de $\bar{y}_{tend} = 0.0329$ à $\bar{y}_{post} = 0.0441$ ($t = -33.7$, $p - value < 0.001$), mais est toutefois inférieure à la valeur moyenne du score de la phase de *pre-training* qui est de $\bar{y}_{pre} = 0.0564$ ($t = 3.81$, $p - value < 0.05$). Cela suggère qu'une part importante des informations reçues et assimilées durant la phase d'entraînement avec la trace du geste sont perdues lorsque la métaphore est

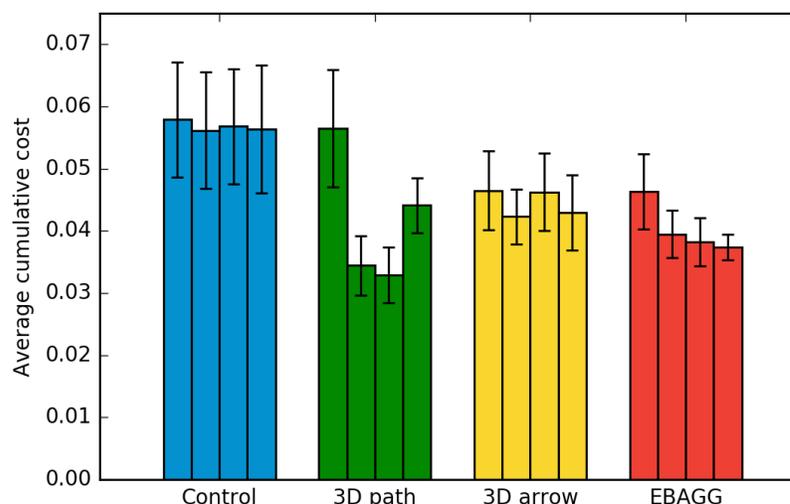


FIGURE 5.15 – Coûts cumulatifs moyens des vingt itérations répartis selon les trois phases d’entraînement et par groupes respectifs. Sur ce diagramme, pour chaque groupe la première barre correspond à la phase de *pre-training*, la seconde et la troisième à la phase de *training* et enfin la quatrième représente la phase de *post-training*. Chaque barre représente donc cinq itérations du geste et les barres d’erreurs correspondent les intervalles de confiance à 95%.

retirée. Cependant, l’amélioration des résultats entre la phase de *pre-training* et celle de *post-training* montre que cette aide semble avoir été bénéfique pour l’amélioration des performances des participants.

Puis, pour le groupe de la flèche 3D, il n’y a pas d’amélioration significative entre les phases de *pre-training* et de *post-training* ($t = 1.02, p\text{-value} = 0.37$), comme suggéré par la régression linéaire.

Si nous comparons les données de la phase de *pre-training* ($\bar{y}_{pre} = 0.0464$) avec celles du *post-training* ($\bar{y}_{post} = 0.0374$) pour le groupe EBAGG, nous pouvons observer une amélioration significative ($t = 4.57, p\text{-value} < 0.01$). La figure 5.15 nous montre également que les performances obtenues lors de la phase de *post-training* sont relativement proches des performances de la fin de la phase de *training*, voire même légèrement meilleures. En effet, le score diminue de $\bar{y}_{tend} = 0.0383$ à $\bar{y}_{post} = 0.0374$. Cette différence n’est cependant pas significative ($t = 1.64, p\text{-value} = 0.15$), ce qui veut dire que lorsque EBAGG n’était plus affichée, les participants arrivaient à se remémorer les informations fournies par la métaphore et à en réutiliser une grande partie. La légère amélioration pourrait être expliquée par une surcorrection de la part des apprenants lors de la phase de *training*, qui se seraient alors davantage focaliser

CHAPITRE 5. EVALUATION DE L'IMPACT DE EBAGG SUR LES PERFORMANCES DES UTILISATEURS

sur la métaphore que sur la réalisation du geste. De plus, on peut observer sur la figure 5.15 que l'intervalle de confiance à 95% est minime pour la phase de *post-training* du groupe EBAGG. Cela nous indique que les performances des différents participants de ce groupe sont assez similaires.

Pour terminer avec cette analyse des résultats de la phase de *post-training*, nous effectuons une analyse plus fine de l'impact sur l'apprentissage de la métaphore de la trace et de EBAGG. Nous calculons l'erreur quadratique moyenne pour chaque groupe et nous comparons alors les cinq dernières valeurs moyennes de la phase de *post-training* obtenues lors de l'expérimentation, avec les cinq valeurs moyennes prédites par notre modèle linéaire. Ces valeurs sont reportées dans le tableau 5.2 pour la trace et dans le tableau 5.3 pour EBAGG. Nous n'étudions pas les résultats du groupe de contrôle, ni ceux du groupe de la flèche 3D puisqu'il n'y a eu aucun impact significatif sur l'amélioration des performances lors des quinze premières itérations du geste.

TABLE 5.2 – Coûts cumulés moyens estimés par le modèle et valeurs empiriques obtenues lors de l'expérimentation pour le groupe de la trace du geste

	Itérations				
	16	17	18	19	20
Valeurs estimées y	0.02310	0.02083	0.01856	0.01630	0.01403
Valeurs empiriques y	0.04381	0.04322	0.04434	0.04453	0.04471

TABLE 5.3 – Coûts cumulés moyens estimés par le modèle et valeurs empiriques obtenues lors de l'expérimentation pour le groupe EBAGG

	Itérations				
	16	17	18	19	20
Valeurs estimées y	0.03466	0.03382	0.03299	0.03215	0.03131
Valeurs empiriques y	0.03806	0.03744	0.03638	0.03735	0.03768

L'erreur quadratique moyenne est calculée avec l'équation suivante :

$$MSE = \frac{1}{n} \sum_{i=j}^k (y_i - \hat{y}_i)^2 \quad (5.2)$$

Ici, n représente le nombre d'itérations ($n = 5$), j est la première itération étudiée ($j = 16$), et k la dernière itération ($k = 20$).

A partir de la table 5.2, nous obtenons une valeur de l'erreur quadratique moyenne égale à $MSE_t = 6.66 \times 10^{-4}$, tandis qu'à partir de la table 5.3, nous obtenons $MSE_e = 2.08 \times 10^{-5}$.

Dans les deux cas, nous pouvons remarquer qu'il y a une détérioration des niveaux de performances par rapport aux valeurs estimées du modèle linéaire,

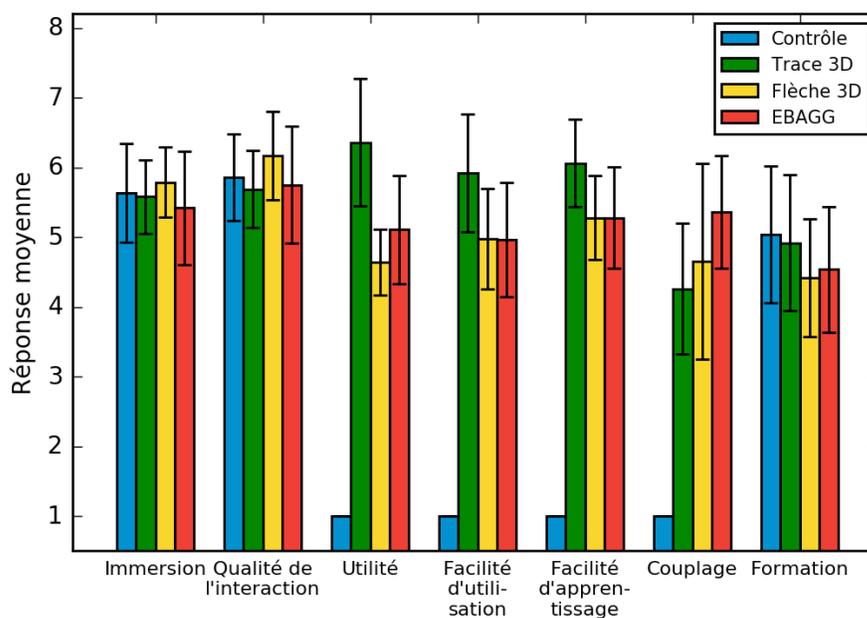


FIGURE 5.16 – Résultats du questionnaire utilisant une échelle de Likert à sept points

c'est-à-dire que les valeurs empiriques sont plus élevées que les valeurs estimées par le modèle. Pour la trace du geste, les valeurs empiriques ont pour valeur moyennes $\bar{y}_{emp} = 0.0441$ contre $\bar{y}_{est} = 0.0186$ pour les valeurs estimées. Cela pourrait être expliqué par l'absence de feedback durant la phase de *post-training*; les participants n'ont pas eu des performances aussi bonnes que celles obtenues lors de la phase de *training*. Néanmoins, l'écart est beaucoup plus faible pour EBAGG, comme on peut le voir sur la figure 5.15. Pour le groupe EBAGG, les performances en *post-training* étaient légèrement meilleures, tandis que pour la trace du geste, elles étaient significativement inférieures.

5.5.3 Analyse qualitative

Après la réalisation des vingt itérations du geste, les participants doivent répondre à un questionnaire. Celui-ci comprend vingt-six questions, auxquelles il faut ajouter l'évaluation NASA-TLX (*NASA Task Load Index*) que nous présenterons par la suite. Les questions sont adaptées du questionnaire Presence de Witmer [Witmer and Singer, 1998] et du questionnaire USE de Lund [Lund, 2001] qui évalue l'attitude de l'utilisateur vis-à-vis de l'utilisabilité d'un produit ou d'une application. Notre questionnaire propose donc à l'utilisateur de noter sur une échelle de Likert à sept points leur accord avec les affirmations

CHAPITRE 5. EVALUATION DE L'IMPACT DE EBAGG SUR LES PERFORMANCES DES UTILISATEURS

proposées (1 : Fort désaccord, 7 : Tout à fait d'accord). Ces affirmations sont réparties selon sept catégories : l'immersion, la qualité de l'interaction, *l'utilité de la métaphore*, *la facilité d'utilisation de la métaphore*, *la facilité d'apprentissage de la métaphore*, *le couplage entre l'utilisateur et la métaphore* et enfin la formation (les catégories en italiques ne sont pas proposées au groupe de contrôle puisque ses membres ne reçoivent aucune aide durant l'exercice). L'objectif de cette étude qualitative est d'obtenir une évaluation subjective de la part des participants sur l'expérimentation et sur les différentes métaphores de guidage, selon les sept catégories mentionnées.

Questionnaire

La figure 5.16 présente les résultats du questionnaire selon les sept catégories. L'ensemble des affirmations et les réponses correspondantes se trouvent dans l'annexe C. La première catégorie concernait donc la qualité de l'immersion, plus précisément si les participants se sont sentis immergés dans l'environnement virtuel et si l'équipement ne procurait aucune gêne. Globalement, tous les participants étaient satisfaits de la qualité de l'immersion (*mean* : 5.79, *std* : 1.13). Seulement huit des quarante-huit participants ont ressenti une gêne vis-à-vis des lunettes stéréoscopiques qui étaient trop grandes pour eux. De même, la majorité d'entre-eux ont eu le sentiment que leurs interactions étaient naturelles (*mean* : 5.87, *std* : 1.18).

Les trois catégories suivantes n'étaient pas proposées au groupe de contrôle. En effet, ces affirmations, issues du questionnaire USE, concernaient les métaphores de guidage qui aidaient les apprenants durant la phase de *training*. Ici, c'est la métaphore de la trace qui a obtenu les meilleurs résultats. Pour les participants, elle était la plus utile avec un score moyen de 6.36 sur 7 (*std* : 0.8), la plus facile à utiliser (*mean* : 5.92, *std* : 1.26) et celle dont l'apprentissage était le plus simple (*mean* : 6.06, *std* : 1.07). Cette évaluation entièrement favorable à la trace 3D peut très certainement s'expliquer par la simplicité de compréhension de cette métaphore et de son utilisation. La manière la plus facile et la plus rapide pour informer quelqu'un sur la direction à prendre est d'en afficher le chemin, contrairement aux deux autres feedbacks qui sont focalisés sur la correction de l'erreur. On peut notamment remarquer sur la figure 5.16 que la flèche 3D et EBAGG ont des résultats assez similaires. Les feedbacks dynamiques semblent donc être plus difficile à appréhender et à utiliser.

Ensuite, les affirmations concernent le couplage, qui dans notre cas est défini comme étant une influence mutuelle et continue entre un utilisateur humain et une métaphore de guidage. Ils se régulent entre-eux et co-évoluent [Varela et al., 1992]. Ainsi, les métaphores dynamiques adaptent leur comportement en fonction des actions des utilisateurs, et ces derniers modifient la manière dont il réalise leur geste selon les retours qu'ils obtiennent de la métaphore. Par conséquent, ce n'est plus la trace 3D du geste de référence qui obtient les meilleurs résultats, mais bien notre métaphore EBAGG avec un score de 5.36 sur 7 (*std* : 1.42). Cela suggère donc que EBAGG aurait l'impact le plus important sur l'entraînement. La trace 3D obtient notamment les plus

mauvais résultats avec une moyenne de 4.26 sur 7 (*std* : 2.48) alors qu'elle avait les meilleures résultats pour les trois catégories précédentes. Une des raisons possibles est que c'est une métaphore statique alors que les deux autres sont dynamiques. Les utilisateurs semblent s'appropriier plus facilement des feedbacks dynamiques en tant que partie d'eux-mêmes, comme une extension de leur bras, puisque ces métaphores suivaient les mouvements de l'outil, tandis qu'une métaphore statique semble plutôt appartenir au moule virtuel, et donc à l'environnement virtuel.

Enfin, la dernière catégorie s'intéresse à l'entraînement en lui-même, s'ils ont le sentiment d'avoir progressé au cours des vingt itérations, s'ils ont ressenti le besoin d'avoir une métaphore de guidage pour les aider à faire le geste, que ce soit durant la phase de *training* ou lors des autres phases. Les résultats sont similaires pour les quatre groupes, toutefois le groupe de contrôle a le meilleur score avec une moyenne de 5.04 (*std* : 1.73), ce qui n'est pas si surprenant étant donné qu'ils n'ont reçu aucun feedback durant l'exercice.

NASA-TLX

Une fois ce questionnaire terminé, nous procédons à une évaluation de la charge de travail, qui peut être définie comme étant l'exigence cognitive d'une tâche [Galy et al., 2012, Miyake, 2001]. L'exigence cognitive (*cognitive demand*) et la charge cognitive (*cognitive load*) sont deux notions distinctes. L'exigence cognitive correspond au niveau de mémoire et aux ressources attentionnelles nécessaires pour réaliser une tâche [Smith and Stein, 1998], tandis que la charge cognitive est la quantité d'information que la mémoire de travail peut contenir lors de la réalisation d'une tâche [Sweller, 1988, Sweller, 1994]. Ainsi, une exigence cognitive faible impliquerait une charge cognitive plus faible.

NASA-TLX est une méthode d'évaluation multidimensionnelle qui permet de quantifier subjectivement la charge de travail totale, c'est-à-dire la charge physique et la charge mentale de travail. Elle a été proposée par Hart et Staveland [Hart and Staveland, 1988] du *Human Performance Group* appartenant au *Ames Research Center* de la NASA. Ce questionnaire se divise en deux étapes. Tout d'abord, les participants doivent répondre par une note de 1 (note la plus faible) à 20 (note la plus élevée) à six questions portant sur six critères d'évaluations : exigence mentale, exigence physique, exigence temporelle, performance globale, niveau de frustration et l'effort requis. Ensuite dans un second temps, ils doivent pondérer les différents critères en les comparant deux à deux et sélectionner à chaque fois lequel des deux critères était le plus important durant l'expérimentation. Cette pondération vient par la suite se coupler aux résultats de la première étape pour obtenir finalement une note sur 100 qui nous donne la charge de travail totale.

La figure 5.17 présente les moyennes obtenues pour la charge de travail totale et pour chaque groupe. On remarque immédiatement que les métaphores dynamiques obtiennent un score plus faible que les deux autres groupes, EBAGG ayant la charge de travail totale la plus faible avec un score de 55.14

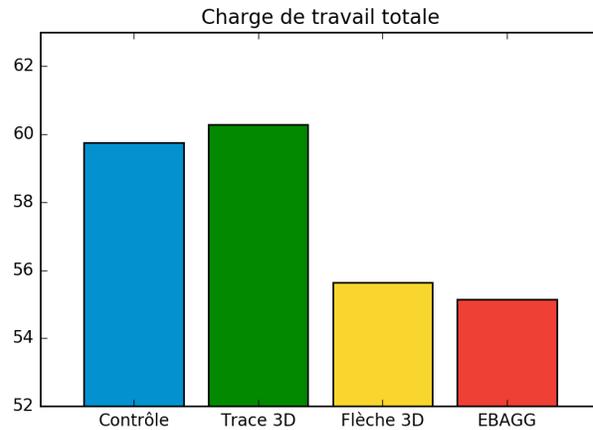


FIGURE 5.17 – Charge de travail totale pour les différents groupes selon l'évaluation NASA-TLX

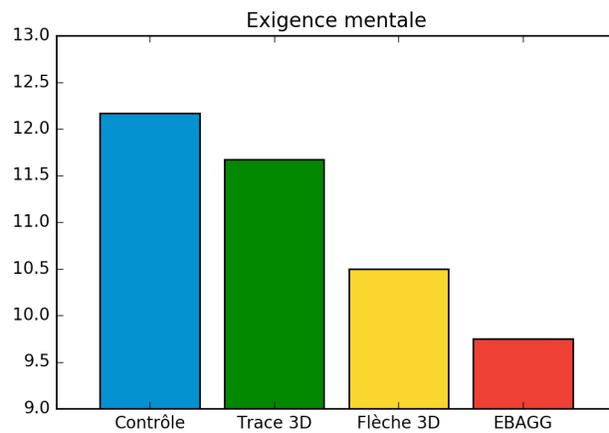


FIGURE 5.18 – Scores moyen des différents groupes pour la question sur l'exigence mentale de l'étape 1 de l'évaluation NASA-TLX

sur 100 (*std* : 12.86), ce qui rejoint nos remarques précédentes. De plus, si l'on étudie la figure 5.18 qui présente les résultats de l'évaluation NASA-TLX pour le critère de l'exigence mentale (non pondérée par les résultats de l'étape 2), EBAGG possède une nouvelle fois le score moyen le plus faible avec 9.75 (*std* : 4.88).

5.6 Discussion

5.6.1 Groupe contrôle

Les résultats du groupe de contrôle indiquent que la simple répétition des gestes n'a eu aucun ou très peu d'impact sur l'entraînement. Les niveaux de performances des participants de ce groupe ne varient que très peu. La majorité des utilisateurs ont par ailleurs répondu dans le questionnaire qu'ils avaient besoin d'avoir des feedbacks durant l'exercice afin d'être au courant de leurs erreurs et de pouvoir les corriger. De plus, l'ensemble des sujets des trois autres groupes ont eu de meilleurs résultats, ce qui suggère la validation de notre première hypothèse **H1**.

5.6.2 Groupe EBAGG

La métaphore EBAGG a quant à elle présenté des résultats intéressants et a permis aux participants du groupe correspondant d'avoir une amélioration significative de leurs performances. Il y avait un écart important entre les résultats de la phase de *pre-training* et celle de *post-training*. Les sujets de ce groupe se sont en effet entraînés efficacement, afin de reproduire correctement le geste de référence. Ils se sont améliorés de manière linéaire et continue tout au long des vingt itérations et ont ainsi obtenu de meilleurs résultats que les autres groupes lors de la phase de *post-training*. Le questionnaire nous a également montré que le couplage entre les utilisateurs et la métaphore était plus important avec EBAGG. Cela validerait donc notre seconde hypothèse **H2**. Toutefois, comme précisé précédemment, nous ne pouvons déterminer si cette différence est significative ou non. Nous avons besoin de plus données et donc de plus de participants pour comparer les résultats quantitatifs de la phase de *post-training*.

Nous pouvons également observer des similarités dans les niveaux de performance entre les dernières itérations de la phase de *training* et la phase de *post-training*, voire même une légère amélioration, bien qu'elle ne soit pas significative. Cela suggère que les informations sur le geste que les participants ont assimilées durant la phase de *training*, persistent une fois que l'on retire la métaphore EBAGG de l'exercice, elles restent ancrées dans leur mémoire. Ils se souviennent alors des erreurs que EBAGG avait soulignées et les corrections qu'ils ont dû opérer pour ajuster leur geste. La disparition de EBAGG ne les a pas perturbés dans leur réalisation du geste et leur a permis de le faire correctement. Cela validerait donc notre hypothèse **H3**. Ainsi, cette métaphore pourrait donc être une solution au problème de dépendance présenté dans l'introduction de ce mémoire.

5.6.3 Groupe trace

Les performances du groupe ayant eu la trace 3D du geste de référence comme métaphore de guidage se sont significativement améliorées durant la

phase de *training* où la métaphore était affichée. Cela pourrait être expliqué par son efficacité et sa simplicité d'utilisation, comme le suggère les résultats du questionnaire. Afficher la trace du geste à reproduire semble être plus naturel pour les participants, dont la tâche ne se réduisait plus qu'à suivre le tracé affiché devant eux. Néanmoins, la forte chute de performance durant la phase de *post-training* est à souligner. Ceci rejoint également le problème de dépendance mentionné précédemment. Les participants semblent avoir été d'avantage concentrés à suivre correctement la trace 3D qu'à reproduire le geste. Par conséquent, lorsque la trace 3D n'était plus affichée, ils avaient des difficultés à se remémorer le geste et à le réaliser correctement. De plus, la comparaison des erreurs quadratiques moyennes montre que l'écart entre les valeurs empiriques et les valeurs prédites par le modèle était plus important pour ce groupe. Cela renforce notre idée d'une dépendance générée par la métaphore de la trace 3D. En revanche, EBAGG semble ne générer qu'une faible dépendance. Les utilisateurs utilisant la trace 3D s'intéresseraient donc d'avantage aux retours visuels proposés par l'environnement, tandis que ceux qui utilisent EBAGG se concentrerait plus sur leurs feedbacks intrinsèques et donc sur leur proprioception. Cependant, cette expérimentation était une analyse de l'amélioration des performances lors d'une session d'entraînement, à très court terme donc. Il serait donc intéressant d'effectuer la même expérience en étudiant cette fois-ci l'impact des différentes métaphores sur l'apprentissage à long terme et vérifier si les résultats sont les mêmes.

5.6.4 Groupe flèche 3D

Les résultats du groupe de la flèche 3D nous indique que cette métaphore n'a eu aucun effet significatif sur l'entraînement. Nous remarquons également que les performances se sont détériorées durant la phase de *training*. Néanmoins, les réponses du questionnaires nous permettent d'expliquer ces résultats. En effet, malgré la facilité de compréhension de ce feedback, les participants ont eu du mal à comprendre les informations que la métaphore indiquait durant l'entraînement. Sans repère agissant comme référence, les utilisateurs n'étaient pas capables de comprendre avec précision dans quelle direction la flèche pointait. Ils ont notamment signalé que ce feedback était plus un obstacle qu'une aide, dans le cadre du guidage de gestes précis. Toutefois, la charge de travail générée par la flèche 3D est assez proche de celle de EBAGG et reste relativement éloignée de celle de la trace 3D. Néanmoins, son exigence mentale est plus élevée par rapport à EBAGG. Il semblerait donc que les feedbacks dynamiques permettent d'avoir une charge de travail plus faible lors d'un apprentissage de gestes précis, et donc une charge cognitive plus faible durant l'exercice par rapport à une métaphore statique ou une absence de feedback. Wulf et Shea décrivaient dans [Wulf and Shea, 2002] que lors d'un apprentissage de tâches complexes, un feedback visuel concurrent pourrait prévenir d'une surcharge cognitive. Ils permettraient ainsi de rendre les informations plus accessibles et aideraient les apprenants à se construire un chemin mental du geste.

5.7 Limitations

Certaines limitations concernant EBAGG sont néanmoins apparues durant cette expérimentation. Certains participants ont déclaré que lorsque le *wand* était trop proche de de leur corps, la sphère devenait gigantesque et masquait une grande partie de la scène. En effet, sa position devenait trop proche de celle des caméras virtuelles situées dans l’environnement virtuel et qui permettent le rendu. EBAGG n’était également parfois plus visible dans cette situation, puisqu’elle se déplaçait derrière ces mêmes caméras. Les participants devaient donc faire un pas en arrière pour ne plus être trop proche du moule virtuel et afin de réaliser leur geste sans aucune gêne.

Le taux de calcul réalisé par l’algorithme DTW était également l’une des limites de la métaphore. En effet, il est crucial qu’il soit élevé afin que le comportement de EBAGG soit cohérent. Plus ce taux est élevé, plus la position des particules est mise à jour sur la métaphore. Ce taux de calcul dépend de différents facteurs. D’une part, il prend en considération le nombre de positions successives enregistrées du geste de référence. Notre algorithme parcourt en réalité entièrement la série temporelle à chaque nouveau calcul afin de trouver la meilleure correspondance entre les points. D’autre part, il calcule également une nouvelle fois les précédents *matchings* afin de les actualiser. Nous étions donc limités par le nombre de points que le geste de référence pouvait contenir, afin que notre algorithme fonctionne en temps réel. À certaines parties du geste comme les deux courbes (pour rappel, la trace du geste est visible sur la figure 5.1), l’écart entre deux positions était trop important. L’algorithme ayant retourné un *matching* avec le premier point, il le gardait comme référence pour évaluer la performance de l’utilisateur, puisque le point suivant était trop éloigné. Par conséquent, lorsque les participants commettaient des erreurs de trajectoire, EBAGG leur indiquait de manière erronée qu’ils étaient également trop en avance ou en retard au niveau de leur geste, comme s’il y avait une certaine vitesse à avoir alors que la temporalité n’était pas évaluée dans l’expérimentation.

5.8 Conclusion

Dans ce chapitre, nous avons présenté notre seconde expérimentation qui était une comparaison de EBAGG avec des métaphores de guidage usuelles, ainsi qu’un groupe de contrôle. L’objectif était d’analyser les performances des participants de chaque groupe lors d’une session d’entraînement à la reproduction d’un geste précis, afin de déterminer quelle métaphore avait le plus d’impact sur l’amélioration de ces performances. Les trois groupes avec feedback étaient donc aidés par ces métaphores de guidage durant dix des vingt itérations d’entraînement, tandis que le groupe de contrôle ne recevait aucun feedback. Pour comparer leurs performances, nous nous sommes basés sur une évaluation quantitative en utilisant les résultats de notre algorithme DTW, ainsi que sur un questionnaire.

CHAPITRE 5. EVALUATION DE L'IMPACT DE EBAGG SUR LES PERFORMANCES DES UTILISATEURS

Notre métaphore EBAGG a eu de meilleurs résultats en ce qui concerne les performances des utilisateurs, comparés aux trois autres groupes. Les participants se sont améliorés de manière linéaire et continue au cours des vingt itérations de l'exercice, bien que EBAGG était un peu plus difficile à prendre en main et à utiliser que la trace 3D du geste de référence. D'autre part, la flèche 3D a quelque peu perturbé les participants du groupe correspondant, tandis que la trace 3D semble avoir généré une dépendance chez les sujets, ce qui a fortement détérioré les performances des participants une fois la métaphore retirée. Pour EBAGG, au contraire, les performances des participants sont restées semblables lors de la phase de *post-training*, au même niveau que lors de la phase de *training* avec feedback. De plus, EBAGG a obtenu les meilleurs résultats concernant le couplage, résultant sur un impact plus important lors de l'entraînement. La charge de travail de la reproduction de geste est également plus faible en utilisant la métaphore EBAGG.

Suite aux deux expérimentations, présentées dans les chapitres 4 et 5, nous constatons que la métaphore EBAGG permet de guider efficacement l'utilisateur pour un geste précis et d'obtenir des résultats satisfaisants en terme d'apprentissage du geste à très court terme, lors d'un exercice de reproduction d'un geste. De plus, la comparaison avec les métaphores usuelles a mis en évidence la supériorité de EBAGG concernant les performances des utilisateurs, notamment en ce qui concerne le syndrome de la dépendance soulevé par l'hypothèse de guidage, dans le cadre de l'apprentissage moteur.

De plus, la problématique du paramétrage de la métaphore a été corrigée dans cette expérimentation. Si l'on compare les tableaux B.2 et C.3 en annexe, on peut s'apercevoir que EBAGG provoque moins de gêne chez les participants dans cette expérimentation, par rapport à la précédente. La gêne restante est principalement due à la prise en main de la métaphore.

Une possible amélioration des limitations de EBAGG serait d'utiliser un ordinateur dédié aux calculs, mais également de faire du calcul GPU (*graphics processing unit*). Cela nous permettrait d'avoir un geste de référence beaucoup plus précis, avec plus de points. Toutefois, ce nouveau scénario soulève de nouvelles questions. EBAGG surpasserait-elle toujours les autres métaphores de guidage, comme la trace 3D par exemple? La trace étant une métaphore statique, l'ajout de point ne changerait que très peu son comportement, tandis que EBAGG serait plus précise dans ses retours à l'utilisateur et son taux de rafraîchissement serait plus important. Il semblerait donc que EBAGG surpasserait toujours la trace 3D. Il serait également intéressant de continuer cette expérimentation avec de nouveaux participants pour atteindre un nombre total d'au moins cent sujets. Cela nous permettrait de réduire les données aberrantes afin de pouvoir réaliser une comparaison intergroupe. Les résultats de ces travaux ont également donné lieu à une publication à la conférence ACM VRST 2017 [Jeanne et al., 2017b].

Dans le prochain et dernier chapitre, nous dressons un bilan de nos travaux et nous proposons différentes perspectives d'évolution de notre approche.

Chapitre 6

Conclusion et perspectives

Ce chapitre présente un résumé de nos travaux ainsi que notre contribution en réponse aux problématiques soulevées lors de l'introduction de ce mémoire de thèse. Nous évoquons par la suite les limites de notre modèle et enfin nous proposons des pistes de recherches futures.

6.1 Retours sur notre contribution

Nous avons proposé une nouvelle métaphore dynamique de guidage qui vise à répondre aux problèmes soulevés par l'apprentissage de gestes. En effet, il convient tout d'abord de caractériser le geste de l'utilisateur, c'est-à-dire de déterminer en quoi le geste réalisé est un bon geste ou un mauvais geste. Dans le cadre du guidage, nous nous intéressons à la trajectoire empruntée pour réaliser le geste demandé. L'utilisateur doit passer par des positions précises. La vitesse d'exécution n'est pas prise en compte, il n'y a donc aucune exigence temporelle. Nous ne nous intéressons pas non plus à l'orientation de l'objet éventuellement utilisé, comme par exemple un outil dans un cadre industriel, ni à l'orientation ou à la posture du corps de l'utilisateur. En résumé, nous qualifions un bon geste dans le cadre de l'utilisation de la métaphore EBAGG comme étant une succession de points au cours du temps dont les positions sont relativement proches du geste de référence défini au préalable par un expert. Pour déterminer la distance entre les deux gestes, nous utilisons un algorithme de déformation temporelle dynamique (DTW en Anglais). En plus de cet écart, cet algorithme nous retourne un coût cumulé agissant comme un score : plus ce coût est faible, plus l'écart entre les deux gestes sera petit et donc plus le geste de l'utilisateur sera bon.

L'apprentissage de gestes aborde également les questions du guidage. L'apprenant doit savoir comment réaliser son geste et surtout comment se corriger lorsqu'il commet des erreurs. Nous proposons donc un feedback en temps réel afin de retourner des informations à l'utilisateur pendant la réalisation de son geste. En nous basant sur un apprentissage par l'erreur, les informations reçues par l'apprenant concernent ainsi la présence ou non d'erreurs de trajectoire. Il

sait alors à chaque instant s'il est sur la bonne trajectoire ou s'il doit se corriger pour l'être de nouveau. Pour cela, nous avons adapté l'algorithme classique du DTW afin qu'il puisse effectuer à la fois des calculs et retourner les résultats en temps réel. De plus, ne nous intéressant qu'à la trajectoire du geste, il nous fallait un feedback capable d'indiquer les erreurs, c'est-à-dire les écarts de trajectoire. Pour les représenter équitablement selon les trois dimensions de l'espace, nous avons décidé donner la forme d'une sphère à EBAGG.

En outre, la présence d'erreur est indiquée par l'affichage de particules sur la surface de EBAGG. Lorsque des particules apparaissent sur la sphère, l'utilisateur doit corriger son geste dans la direction opposée. Leur nombre varie en fonction de la gravité de l'erreur : plus elle est importante, plus le nombre de particules sera grand. Il dépend entièrement de la distance calculée et retournée par l'algorithme DTW. L'aspect dynamique de la métaphore EBAGG permet ainsi de retourner des informations aux utilisateurs en fonction de leur geste, et ces informations sont donc différentes d'un utilisateurs à l'autre, ou même d'un geste à un autre pour un même utilisateur. Ceci nous permet de tenir compte de la grande variabilité des utilisateurs à la fois sur le plan interpersonnel et sur le plan intrapersonnel.

Enfin, le dernier problème soulevé était la dépendance générée par l'utilisation de feedbacks en continu durant l'apprentissage. L'étude proposée dans le chapitre 5 a montré que lors d'un exercice de reproduction de gestes précis, les utilisateurs ayant eu EBAGG comme métaphore de guidage avaient, lorsqu'elle n'était plus affichée, des performances similaires à celles obtenues lors de l'entraînement avec feedback, voire même légèrement meilleures. N'ayant aucune représentation visuelle du geste, les apprenants se créent ainsi une image mentale du geste à partir des sensations qu'ils ont eu lors des itérations précédentes du geste, en se concentrant d'avantage sur leur proprioception (théorie de Schmidt sur l'apprentissage et le contrôle moteur (section 2.3.1)). Cela leur permet de reproduire le geste aisément sans l'aide de feedback.

6.2 Limitations et Perspectives d'évolution

Nous utilisons des méthodes de capture de mouvement pour obtenir la position en temps réel de l'utilisateur. Contrairement à d'autres techniques qui suivent l'intégralité des mouvements de l'utilisateur [Anderson et al., 2013, Anderson and Bischof, 2013, Han et al., 2016, Rovelo et al., 2015, Velloso et al., 2013], EBAGG n'évalue que les positions du wand ou de la partie du corps dont les mouvements sont suivis. Notre algorithme DTW ne prend en entrée que les positions successives d'une seule entité et les compare aux positions de la référence. Il serait techniquement possible d'afficher les erreurs de plusieurs parties du corps sur la métaphore, en ajoutant un ensemble de particules d'une couleur différentes des premières par exemple. Dans le cadre du projet KIVA par exemple, il serait également intéressant de pouvoir fournir à l'apprenant des informations sur ses erreurs de positions mais également sur l'orientation de l'outil. Néanmoins, cela risquerait d'être incom-

préhensible pour l'utilisateur qui devrait alors se concentrer sur EBAGG pour en comprendre les informations et arriver à distinguer les deux types d'erreurs.

Il est également possible d'enrichir EBAGG en couplant la modalité visuelle à de l'haptique [Feygin et al., 2002, Marchal-Crespo et al., 2013, Sigrist et al., 2013], du vibro-tactile [Bloomfield and Badler, 2007, Louison et al., 2017, Lieberman and Breazeal, 2007, Lindeman et al., 2004] ou du sonore [Helmer et al., 2010, Hummel et al., 2010, Sigrist et al., 2015]. De plus, les travaux de Ronsse *et al.* [Ronsse et al., 2010] montre que la modalité sonore générerait moins de dépendance vis-à-vis du feedback que la modalité visuelle, et focaliserait l'attention de l'utilisateur sur sa proprioception. Toutefois, l'utilisation de ces modalités peut être perturbée dans des environnements de travail où les conditions sont difficiles (températures élevées ou très basses et niveau sonore important par exemple). L'usage de la multimodalité pourrait donc être bénéfique pour des milieux moins contraignants. Il serait donc également possible d'utiliser EBAGG pour indiquer la présence d'erreur selon d'autres critères d'erreur, comme la vitesse ou l'orientation par exemple.

L'expérimentation du chapitre 5 a permis d'obtenir des résultats intéressants. En effet, les utilisateurs ayant eu EBAGG comme métaphore de guidage semblent n'avoir qu'une faible dépendance au feedback lors de cette expérimentation. Toutefois, cet exercice de reproduction d'un geste précis n'a été réalisé qu'avec un nombre relativement limité de participants, à savoir douze personnes par groupe, pour un total de quarante-huit. Il serait donc intéressant de poursuivre ou de reproduire cette expérimentation en augmentant le nombre de participants à au moins une centaine de personnes afin d'avoir au minimum vingt sujets par groupe. Cela permettrait de réduire l'impact des données aberrantes et nous permettrait notamment de réaliser une comparaison intergroupe. Cette nouvelle expérimentation serait également l'occasion de déterminer si les effets sur l'apprentissage de EBAGG à très court terme se généralisent et sont également présents à long terme. Pour cela nous pourrions proposer aux participants un nouvel exercice de reproduction de gestes quelques semaines ou mois après la première session afin de vérifier si le geste a correctement été appris ou non.

Enfin, l'une des limitations principales sur l'utilisation de la métaphore EBAGG est d'ordre technologique comme mentionné précédemment. L'algorithme DTW utilisé en temps réel est assez coûteux en ressources pour l'ordinateur qui est également utilisé pour le rendu de l'environnement en trois dimensions dans le CAVE. Une des solutions proposées est donc de transférer les calculs sur un ordinateur dédié qui communique les résultats à l'ordinateur de rendu et donc à EBAGG. Cela permettrait également de faire du calcul sur un processeur graphique qui serait alors libre de toute utilisation et qui accélérerait nettement le temps de calcul.

Concernant le problème de dépendance, l'une des améliorations possible serait de reprendre le principe de *fading feedback* pour EBAGG, qui réduit la fréquence des feedbacks au cours du temps [Sigrist et al., 2013]. Des études ont montré que cette méthode semblait être efficace [Crowell and Davis, 2011, Kovacs and Shea, 2011] sans que l'on sache exactement quelle est la fréquence

optimale de disparition de la métaphore. Ainsi, EBAGG pourrait être rendue invisible lorsque les performances réalisées par un utilisateur sont bonnes et que le degré de l'erreur ne dépasse un certain seuil durant plusieurs itérations consécutives, afin de ne pas l'habituer à la présence de la métaphore et donc de réduire la dépendance qu'elle peut générer. En cas de légères erreurs, EBAGG ne serait pas affichée directement, mais on attendrait plutôt l'itération suivante pour l'afficher de nouveau. Toutefois, si une erreur importante était commise, la métaphore serait alors visible directement.

6.3 Applications futures

Dans le cadre du projet KIVA, une phase de tests va être réalisée au cours de l'année 2018 sur une période de six mois à un an, chez l'industriel. L'objectif est d'évaluer l'environnement de formation, notamment l'impact de EBAGG sur l'apprentissage des opérateurs de Montupet. Étant destinée à des gestes lents, EBAGG interviendra dans les scénarios pour les opérateurs novices qui débutent leur formation.

Cette évaluation nous permettra de tester notre métaphore en conditions réelles et d'obtenir des premiers résultats concernant l'apprentissage à long terme.

Annexe A

Taxonomies

A.1 Textes en environnement virtuel

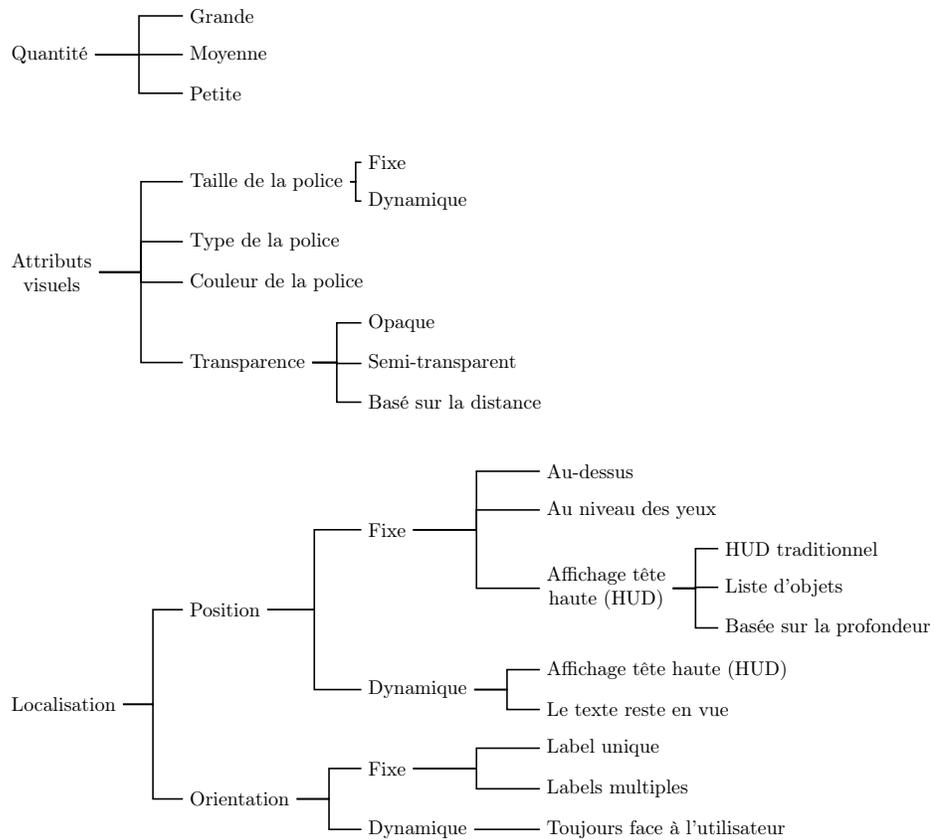


FIGURE A.1 – Taxonomie de Bowman [Bowman, 1999]

A.2 Mouvement en environnement virtuel

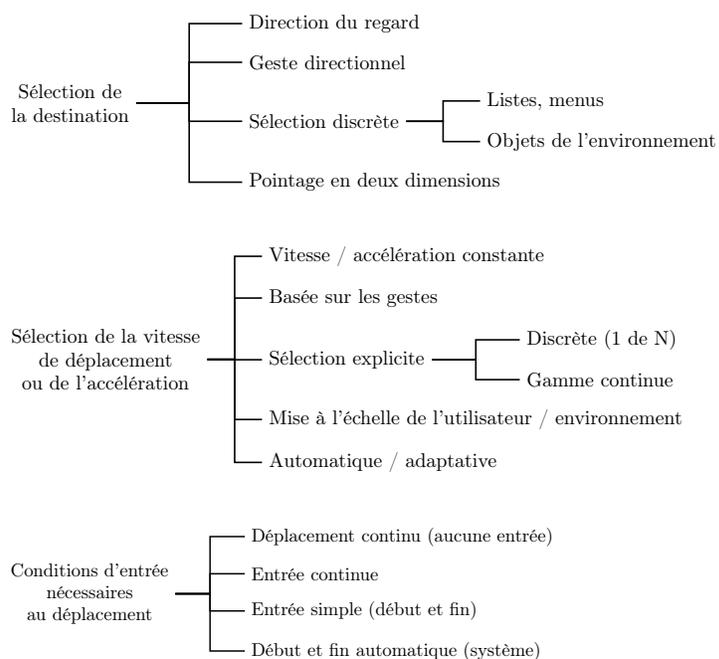


FIGURE A.2 – Taxonomie de Bowman [Bowman, 1999]

A.3 Sélection et manipulation d'objets en environnement virtuel

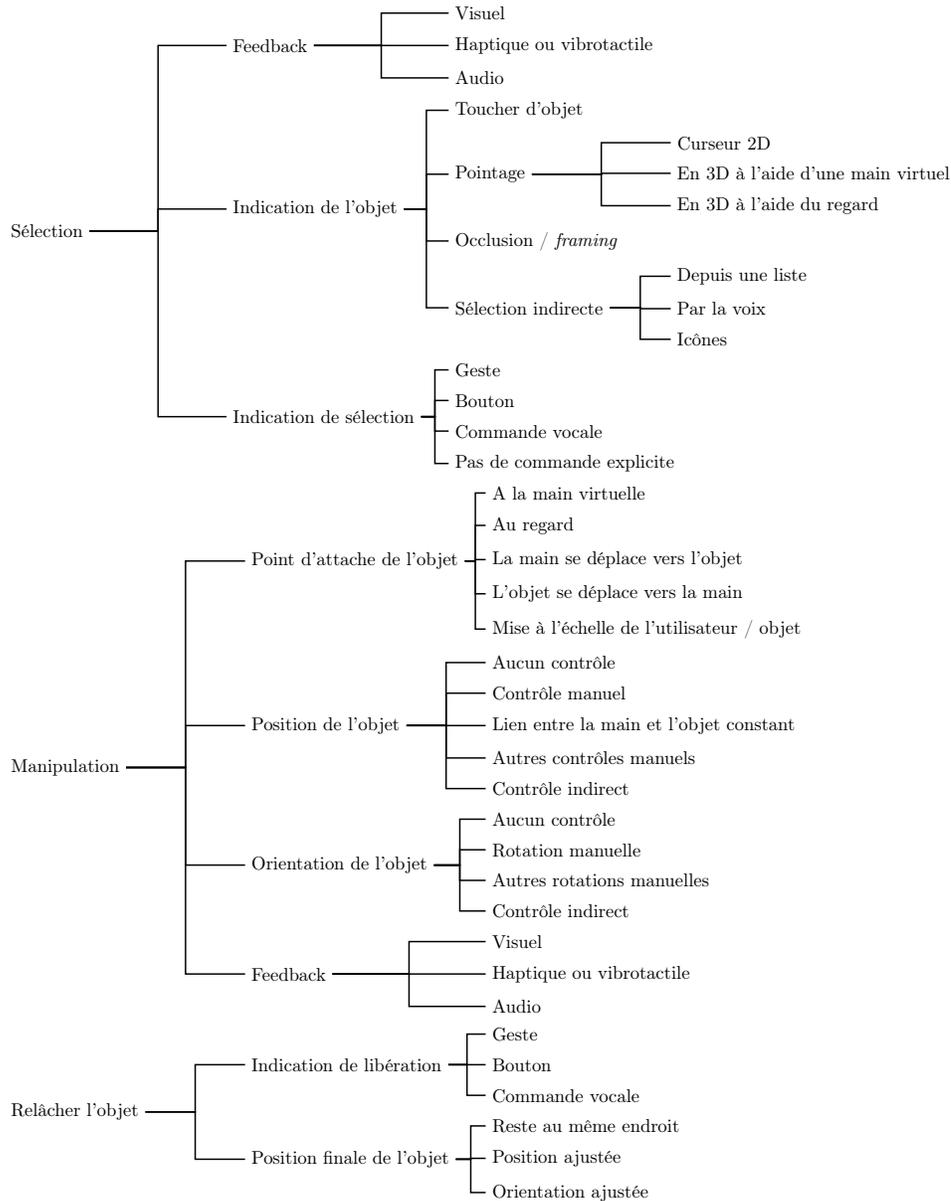


FIGURE A.3 – Taxonomie de Bowman [Bowman, 1999]

A.3. SÉLECTION ET MANIPULATION D'OBJETS EN ENVIRONNEMENT VIRTUEL

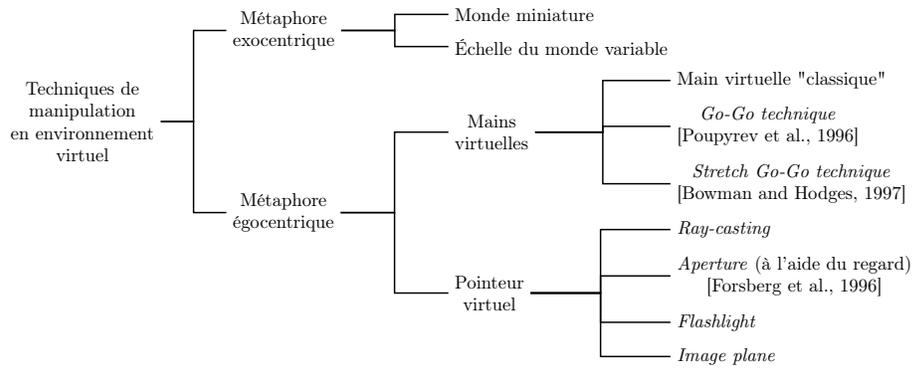


FIGURE A.4 – Taxonomie de Poupyrev [Poupyrev and Ichikawa, 1999]

Annexe B

Questionnaire et résultats de la première expérimentation

B.1 Questionnaire

1. L'environnement virtuel réagissait-il correctement aux actions que vous avez initiées ?

0 1 2 3 4 5 6 7 8 9 10

Pas du tout Parfaitement

2. Vos interactions avec l'environnement virtuel vous semblaient-elles naturelles ?

0 1 2 3 4 5 6 7 8 9 10

Pas du tout Parfaitement

3. Votre ressenti du déplacement d'objets dans l'espace était-il convaincant ?

0 1 2 3 4 5 6 7 8 9 10

Pas du tout Parfaitement

B.1. QUESTIONNAIRE

4. Vous êtes-vous rapidement adapté à cette expérience en environnement virtuel ?

0 1 2 3 4 5 6 7 8 9 10

Pas du tout Parfaitement

5. Avez-vous eu l'impression de bien maîtriser vos mouvements et interactions avec l'environnement virtuel à la fin de l'expérience ?

0 1 2 3 4 5 6 7 8 9 10

Pas du tout Parfaitement

6. Votre expérience en environnement virtuel vous a-t-elle semblé cohérente avec vos expériences du monde réel ?

0 1 2 3 4 5 6 7 8 9 10

Pas du tout Parfaitement

7. Vous êtes-vous rapidement habitué à l'aide visuelle proposée ?

0 1 2 3 4 5 6 7 8 9 10

Pas du tout Parfaitement

8. L'aide visuelle vous a-t-elle perturbé ou distrait pendant la réalisation de vos gestes ?

0 1 2 3 4 5 6 7 8 9 10

Pas du tout Complètement

ANNEXE B. QUESTIONNAIRE ET RÉSULTATS DE LA PREMIÈRE
EXPÉRIMENTATION

9. Si vous avez ressenti une gêne vis-à-vis de l'aide, à quel moment était-ce
(plusieurs réponses possibles) :

Lors des premières itérations de formation avec l'aide

Après quelques itérations

Lors des dernières itérations

10. Avez-vous ressenti un délai entre vos actions et les informations fournies
par l'aide ?

0 1 2 3 4 5 6 7 8 9 10

Pas du tout Enormément

11. Arriviez-vous à vous concentrer sur la tâche plutôt que sur l'aide visuelle ?
(Au début de la phase avec l'aide)

0 1 2 3 4 5 6 7 8 9 10

Pas du tout Parfaitement

12. Arriviez-vous à vous concentrer sur la tâche plutôt que sur l'aide visuelle ?
(À la fin de la phase avec l'aide)

0 1 2 3 4 5 6 7 8 9 10

Pas du tout Parfaitement

13. Avez-vous le sentiment de vous être amélioré dans la réalisation de ces
gestes durant cette expérimentation ?

0 1 2 3 4 5 6 7 8 9 10

Pas du tout Parfaitement

B.2 Résultats

TABLE B.1 – Moyennes et écarts-types du questionnaire de la première expérimentation (* voir tableau B.2)

		<i>Mean</i>	<i>Std</i>
Question	1	7.86	1.70
	2	6.55	1.87
	3	7.18	1.84
	4	6.59	1.74
	5	6.09	1.63
	6	6.32	1.59
	7	6.05	1.96
	8	5.00	2.49
	9	*	*
	10	2.09	1.9
	11	5.82	2.28
	12	6.73	2.12
	13	6.59	2.04

TABLE B.2 – Résultats de la question 9

		Réponse			Nombre de réponses
		Premières itérations	Après quelques itérations	Dernières itérations	
Question 9					3
	X				7
		X			3
				X	3
	X	X			1
	X			X	0
		X		X	3
	X	X		X	2

Annexe C

Résultats de la seconde expérimentation

C.1 Questionnaire

C.1.1 Immersion

1. Je me suis rapidement adapté(e) à cette expérience en environnement virtuel

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

2. Avez-vous ressenti la moindre gêne vis-à-vis de l'équipement de réalité virtuelle? (plusieurs réponses possibles, cochez la case pour oui, laissez-la vide pour non)

- CAVE (bâti)
- Stéréoscopie (vision 3D)
- Lunettes stéréoscopiques
- Wand (manette)
- Aucune gêne

3. L'expérience en environnement virtuel m'a semblé cohérente avec mes expériences du monde réel

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

C.1.2 Interaction

4. Mes interactions avec l'environnement virtuel me semblaient naturelles

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

5. L'environnement virtuel réagissait correctement aux actions que j'ai initiées

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

6. Mon ressenti du déplacement d'objets dans l'espace était convaincant

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

C.1.3 Feedback

Cette section du questionnaire n'était pas proposée au groupe de contrôle.

Utilité

7. La métaphore m'a aidé à être plus efficace

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

ANNEXE C. RÉSULTATS DE LA SECONDE EXPÉRIMENTATION

8. La métaphore m'a été utile

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

9. La métaphore m'a donné plus de contrôle sur mes gestes

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

Facilité d'utilisation

10. La métaphore est facile à utiliser (dans son fonctionnement)

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

11. On l'utilise sans le moindre effort physique

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

12. On l'utilise sans le moindre effort mental

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

13. La métaphore permet d'accomplir le geste plus facilement

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

14. Grâce à la métaphore, on corrige ses erreurs plus rapidement et plus facilement

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

Facilité d'apprentissage

15. On comprend facilement comment l'utiliser

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

16. J'ai appris à l'utiliser rapidement

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

17. Je me suis facilement remémoré comment l'utiliser (après les premières itérations sans feedback)

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

Couplage

18. J'ai eu l'impression que mes mouvements étaient suivis par la métaphore

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

ANNEXE C. RÉSULTATS DE LA SECONDE EXPÉRIMENTATION

19. J'ai eu l'impression qu'il y avait un délai entre mes actions et les informations fournies par la métaphore

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

20. Je me suis rapidement habitué(e) à l'aide proposée

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

21. La métaphore me proposait de nouvelles directions au cours du geste

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

22. L'aide visuelle m'a perturbé(e) ou distrait(e) pendant la réalisation de mes gestes

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

23. Si vous avez ressenti une gêne vis-à-vis de l'aide, à quel moment était-ce (plusieurs réponses possibles, cochez la case pour oui, laissez-la vide pour non) :

- Lors des premières itérations de formation avec l'aide
- Après quelques itérations
- Lors des dernières itérations

24. Le comportement de la métaphore évoluait en fonction de mes performances

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

C.1.4 Formation

25. J'ai le sentiment de m'être amélioré(e) dans la réalisation des gestes durant cette expérimentation

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

26. J'ai ressenti le besoin d'avoir une aide visuelle pour m'aider dans la réalisation des gestes au cours des dernières itérations de l'expérience

	1	2	3	4	5	6	7	
Désaccord	<input type="radio"/>	Accord						

C.1.5 Charge de travail

Evaluation de la charge

Pour les réponses suivantes, veuillez indiquer une valeur non-décimale comprise entre 1 (faible coût) et 20 (coût élevé) inclus.

27. Jusqu'à quel point les activités mentales et perceptives étaient requises pour réaliser le geste demandé (ex : penser, décider, calculer, se remémorer, regarder, chercher, etc.) ?
28. Jusqu'à quel point les activités physiques étaient requises pour réaliser le geste demandé (ex : pousser, tirer, tourner, contrôler, activer, etc.) ?
29. Jusqu'à quel point avez-vous ressenti la pression du temps (rythme, vitesse, etc.) ?
30. Jusqu'à quel point pensez-vous avoir réussi l'objectif de la tâche demandée (apprentissage du geste) ?
31. Jusqu'à quel point avez-vous eu à travailler, en terme d'effort, pour atteindre votre niveau de performance (mentalement et physiquement) ?
32. Jusqu'à quel point étiez-vous confiant durant cette expérimentation (1 : découragé et 20 : très confiant) ?

Poids dimensionnel de la charge cognitive

Pour chacune des paires suivantes, sélectionnez la dimension qui vous a semblé la plus importante en terme de charge de travail.

33. Effort - Performance
34. Exigence temporelle - Frustration
35. Exigence temporelle - Effort
36. Exigence physique - Frustration
37. Performance - Frustration
38. Exigence physique - Exigence temporelle
39. Exigence physique - Performance
40. Exigence temporelle - Exigence mentale
41. Frustration - Effort
42. Performance - Exigence mentale
43. Performance - Exigence temporelle
44. Exigence mentale - Effort
45. Exigence mentale - Exigence physique
46. Effort - Exigence physique
47. Frustration - Exigence mentale

C.2 Résultats

TABLE C.1 – Moyennes et écarts-types du questionnaire de la seconde expérimentation (* voir tableau C.2, ** voir tableau C.3, # question non proposée)

	Contrôle		Trace 3D		Flèche 3D		EBAGG	
	<i>mean</i>	<i>std</i>	<i>mean</i>	<i>std</i>	<i>mean</i>	<i>std</i>	<i>mean</i>	<i>std</i>
Q1	5.75	1.49	5.91	0.67	6.00	0.95	5.50	1.68
Q2	*	*	*	*	*	*	*	*
Q3	5.50	1.00	5.25	0.97	5.58	0.90	5.33	1.23
Q4	5.58	1.38	5.33	1.15	6.00	0.95	5.00	2.22
Q5	6.17	0.58	5.83	1.03	6.42	1.17	6.25	0.87
Q6	5.83	1.19	5.92	1.17	6.08	0.79	6.00	0.60
Q7	#	#	6.42	0.79	4.42	1.56	5.08	1.56
Q8	#	#	6.33	0.65	4.50	1.51	5.17	1.19
Q9	#	#	6.33	1.07	5.00	1.81	5.01	1.44
Q10	#	#	6.58	0.52	4.58	1.38	4.92	1.51
Q11	#	#	5.17	1.47	5.83	1.19	5.92	1.44
Q12	#	#	5.17	1.47	4.58	1.56	4.25	1.06
Q13	#	#	5.92	1.17	4.83	1.59	4.83	1.53
Q14	#	#	6.75	0.45	5.08	1.56	4.92	1.38
Q15	#	#	6.25	0.97	5.33	0.99	5.17	1.19
Q16	#	#	6.50	0.67	4.92	1.24	5.00	1.48
Q17	#	#	5.42	1.24	5.58	1.08	5.67	1.55
Q18	#	#	4.25	2.22	5.08	1.51	5.92	1.17
Q19	#	#	2.58	2.07	3.17	1.85	2.75	1.87
Q20	#	#	6.25	0.87	4.58	1.17	4.92	1.38
Q21	#	#	1.83	1.53	5.08	1.78	4.83	1.70
Q22	#	#	1.58	0.90	3.67	2.19	2.33	1.44
Q23	#	#	**	**	**	**	**	**
Q24	#	#	1.42	0.67	4.00	1.35	4.92	1.51
Q25	5.33	1.44	5.42	1.17	5.17	1.40	5.33	1.56
Q26	4.75	2.01	4.42	1.68	3.67	1.72	3.75	1.60

ANNEXE C. RÉSULTATS DE LA SECONDE EXPÉRIMENTATION

TABLE C.2 – Résultats de la question 2

		Nombre de réponses				
		Contrôle	Trace 3D	Flèche 3D	EBAGG	Total
Q2	CAVE (bâti)	0	0	0	1	1
	Stéréoscopie (vision 3D)	0	0	0	1	1
	Lunettes 3D	1	5	1	1	8
	Wand	2	0	0	3	5
	Aucune gêne	9	7	11	7	34

TABLE C.3 – Résultats de la question 23

Réponse			Nombre de réponses		
			Trace 3D	Flèche 3D	EBAGG
Premières itérations	Après quelques itérations	Dernières itérations			
			10	5	6
X			1	3	1
	X		0	2	1
		X	1	1	1
X	X		0	0	2
X		X	0	0	0
	X	X	0	0	0
X	X	X	0	1	1

TABLE C.4 – NASA-TLX : Réponses des questions 27 à 32 pour les participants des groupes Contrôle et Trace 3D

	Part.	Q27	Q28	Q29	Q30	Q31	Q32
Contrôle	1	3	12	7	13	13	14
	2	12	12	7	13	13	14
	3	4	1	1	15	7	15
	4	12	8	1	7	12	17
	5	16	12	9	15	15	14
	6	13	9	8	14	12	16
	7	15	18	10	15	16	14
	8	17	17	5	12	16	12
	9	12	12	12	10	10	18
	10	13	7	11	15	12	14
	11	15	18	8	15	17	14
	12	14	15	11	10	15	12
Trace 3D	13	13	15	16	14	14	18
	14	15	9	5	15	15	13
	15	7	4	3	13	7	17
	16	8	18	15	15	17	16
	17	11	12	10	17	11	18
	18	12	16	1	17	6	18
	19	15	16	15	12	9	12
	20	10	5	12	15	15	16
	21	13	4	2	14	11	12
	22	15	11	4	17	13	16
	23	8	10	6	15	11	14
	24	13	5	5	9	10	10

ANNEXE C. RÉSULTATS DE LA SECONDE EXPÉRIMENTATION

TABLE C.5 – NASA-TLX : Réponses des questions 27 à 32 pour les participants des groupes Flèche 3D et EBAGG

	Part.	Q27	Q28	Q29	Q30	Q31	Q32
Flèche 3D	25	5	1	1	15	1	15
	26	10	2	2	10	3	18
	27	11	8	3	13	11	15
	28	14	3	1	14	10	15
	29	13	18	5	12	16	11
	30	4	2	4	9	8	12
	31	10	8	7	9	11	13
	32	15	14	5	14	15	16
	33	12	14	13	14	10	17
	34	4	6	2	18	8	14
	35	15	13	10	11	15	12
	36	13	10	3	17	12	19
EBBAG	37	16	14	15	10	15	15
	38	2	16	1	13	8	4
	39	15	10	16	7	10	8
	40	5	10	1	13	10	15
	41	15	3	10	8	9	10
	42	7	12	2	16	4	16
	43	15	18	12	18	18	14
	44	6	2	2	15	6	15
	45	7	3	4	10	7	10
	46	7	12	5	14	5	16
	47	8	13	5	14	15	16
	48	14	16	8	14	11	15

TABLE C.6 – NASA-TLX : nombre d'occurrences de chaque critère (questions 33 à 47) pour les participants des groupes Contrôle et Trace 3D

	Part.	E.Men.	E.Phy.	E.Tem.	Perf.	Eff.	Fru.
Contrôle	1	2	3	0	5	3	2
	2	3	3	4	3	2	0
	3	1	3	4	5	1	1
	4	3	3	5	3	1	0
	5	5	3	0	2	4	1
	6	4	2	1	5	3	0
	7	3	3	2	4	2	1
	8	4	3	2	0	5	1
	9	1	4	2	5	2	1
	10	5	4	3	2	1	0
	11	4	0	2	3	5	1
	12	1	4	0	4	3	3
Trace 3D	13	0	2	5	4	3	1
	14	5	2	0	3	3	2
	15	4	2	1	5	3	0
	16	1	5	3	3	3	0
	17	2	5	1	3	4	0
	18	2	2	3	5	3	0
	19	5	3	2	4	1	0
	20	3	2	4	1	5	0
	21	1	3	1	5	4	1
	22	5	4	1	2	3	0
	23	4	3	1	4	3	0
	24	4	5	2	3	1	0

ANNEXE C. RÉSULTATS DE LA SECONDE EXPÉRIMENTATION

TABLE C.7 – NASA-TLX : nombre d’occurrences de chaque critère (questions 33 à 47) pour les participants des groupes Flèche 3D et EBAGG

	Part.	E.Men.	E.Phy.	E.Tem.	Perf.	Eff.	Fru.
Flèche 3D	25	0	0	0	5	4	2
	26	4	3	0	5	2	1
	27	4	2	0	5	3	1
	28	4	0	2	4	4	1
	29	5	3	1	2	4	0
	30	3	0	1	5	2	4
	31	4	1	1	5	1	3
	32	4	3	1	2	4	1
	33	1	4	4	4	1	1
	34	2	3	3	2	0	5
	35	4	3	0	1	3	4
	36	2	4	1	5	3	0
EBBAG	37	5	3	1	4	2	0
	38	1	3	0	5	4	2
	39	3	1	0	3	3	5
	40	2	1	4	5	3	0
	41	3	1	0	5	2	4
	42	1	4	0	5	2	3
	43	0	4	2	5	1	3
	44	2	0	5	4	1	3
	45	3	0	2	5	2	3
	46	3	4	1	4	3	0
	47	4	5	1	2	3	0
	48	5	2	1	4	3	0

TABLE C.8 – NASA-TLX : charge de travail moyenne du groupe de Contrôle

P.	E.Men.	E.Phy.	E.Tem.	Perf.	Eff.	Fru.	Total	<i>Mean (Std)</i>
1	2,00	11,70	0,00	21,33	13,20	9,73	57,97	59,76 (14,03)
2	12,00	11,70	10,13	12,80	8,80	0,00	55,43	
3	1,33	1,00	1,33	25,00	2,33	5,00	36,00	
4	12,00	8,00	1,67	7,00	4,00	0,00	32,67	
5	26,67	12,00	0,00	10,00	20,00	4,67	73,33	
6	17,33	6,00	2,67	23,33	12,00	0,00	61,33	
7	15,00	18,00	6,67	20,00	10,67	4,67	75,00	
8	22,67	17,00	3,33	0,00	26,67	4,00	73,67	
9	4,00	16,00	8,00	16,67	6,67	6,00	57,33	
10	21,67	9,33	11,00	10,00	4,00	0,00	56,00	
11	20,00	0,00	5,33	15,00	28,33	4,67	73,33	
12	4,67	20,00	0,00	13,33	15,00	12,00	65,00	

TABLE C.9 – NASA-TLX : charge de travail moyenne du groupe Trace 3D

P.	E.Men.	E.Phy.	E.Tem.	Perf.	Eff.	Fru.	Total	<i>Mean (Std)</i>
13	0,00	10,00	26,67	18,67	14,00	6,00	75,33	60,28 (12,40)
14	25,00	6,00	0,00	15,00	15,00	8,67	69,67	
15	9,33	2,67	1,00	21,67	7,00	0,00	41,67	
16	2,67	30,00	15,00	15,00	17,00	0,00	79,67	
17	7,33	20,00	3,33	17,00	14,67	0,00	62,33	
18	8,00	10,67	1,00	28,33	6,00	0,00	54,00	
19	25,00	16,00	10,00	16,00	3,00	0,00	70,00	
20	10,00	3,33	16,00	5,00	25,00	0,00	59,33	
21	4,33	4,00	0,67	23,33	14,67	4,00	51,00	
22	25,00	14,67	1,33	11,33	13,00	0,00	65,33	
23	10,67	10,00	2,00	20,00	11,00	0,00	53,67	
24	17,33	8,33	3,33	9,00	3,33	0,00	41,33	

ANNEXE C. RÉSULTATS DE LA SECONDE EXPÉRIMENTATION

TABLE C.10 – NASA-TLX : charge de travail moyenne du groupe Flèche 3D

P.	E.Men.	E.Phy.	E.Tem.	Perf.	Eff.	Fru.	Total	<i>Mean (Std)</i>
25	0,00	0,00	0,00	25,00	1,33	10,00	36,33	55.64 (12.55)
26	13,33	2,00	0,00	16,67	2,00	6,00	40,00	
27	14,67	5,33	0,00	21,67	11,00	5,00	57,67	
28	18,67	0,00	0,67	18,67	13,33	5,00	56,33	
29	21,67	18,00	1,67	8,00	21,33	0,00	70,67	
30	4,00	0,00	1,33	15,00	5,33	16,00	41,67	
31	13,33	2,67	2,33	15,00	3,67	13,00	50,00	
32	20,00	14,00	1,67	9,33	20,00	5,33	70,33	
33	4,00	18,67	17,33	18,67	3,33	5,67	67,67	
34	2,67	6,00	2,00	12,00	0,00	23,33	46,00	
35	20,00	13,00	0,00	3,67	15,00	16,00	67,67	
36	8,67	13,33	1,00	28,33	12,00	0,00	63,33	

TABLE C.11 – NASA-TLX : charge de travail moyenne du groupe EBAGG

P.	E.Men.	E.Phy.	E.Tem.	Perf.	Eff.	Fru.	Total	<i>Mean (Std)</i>
37	26,67	14,00	5,00	13,33	10,00	0,00	69,00	55.14 (12.86)
38	0,67	16,00	0,00	21,67	10,67	2,67	51,67	
39	15,00	3,33	0,00	7,00	10,00	13,33	48,67	
40	3,33	3,33	1,33	21,67	10,00	0,00	39,67	
41	15,00	1,00	0,00	13,33	6,00	13,33	48,67	
42	2,33	16,00	0,00	26,67	2,67	16,00	63,67	
43	0,00	24,00	8,00	30,00	6,00	14,00	82,00	
44	4,00	0,00	3,33	20,00	2,00	15,00	44,33	
45	7,00	0,00	2,67	16,67	4,67	10,00	41,00	
46	7,00	16,00	1,67	18,67	5,00	0,00	48,33	
47	10,67	21,67	1,67	9,33	15,00	0,00	58,33	
48	23,33	10,67	2,67	18,67	11,00	0,00	66,33	

Bibliographie

- [Ahmed, 2016] Ahmed, S. (2016). Diminution de la fatigue physique lors de l'interaction gestuelle : Approche méthodologique et comparaison virtuel et réel.
- [Aigner et al., 2012] Aigner, R., Wigdor, D., Benko, H., Haller, M., Lindbauer, D., Ion, A., Zhao, S., and Koh, J. (2012). Understanding mid-air hand gestures : A study of human preferences in usage of gesture types for hci. *Microsoft Research TechReport MSR-TR-2012-111*, 2.
- [Alon et al., 2009] Alon, J., Athitsos, V., Yuan, Q., and Sclaroff, S. (2009). A unified framework for gesture recognition and spatiotemporal gesture segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 31(9) :1685–1699.
- [Amokrane and Lourdeaux, 2009] Amokrane, K. and Lourdeaux, D. (2009). Virtual reality contribution to training and risk prevention. In *IC-AI*, pages 726–732.
- [Anderson and Bischof, 2013] Anderson, F. and Bischof, W. F. (2013). Learning and performance with gesture guides. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1109–1118. ACM.
- [Anderson et al., 2013] Anderson, F., Grossman, T., Matejka, J., and Fitzmaurice, G. (2013). Youmove : enhancing movement training with an augmented reality mirror. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, pages 311–320. ACM.
- [Aylett and Luck, 2000] Aylett, R. and Luck, M. (2000). Applying artificial intelligence to virtual reality : Intelligent virtual environments. *Applied Artificial Intelligence*, 14(1) :3–32.
- [Bailador et al., 2007] Bailador, G., Roggen, D., Tröster, G., and Triviño, G. (2007). Real time gesture recognition using continuous time recurrent neural networks. In *Proceedings of the ICST 2nd international conference on Body area networks*, page 15. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).
- [Barot et al., 2013] Barot, C., Lourdeaux, D., Burkhardt, J.-M., Amokrane, K., and Lenne, D. (2013). V3s : A virtual environment for risk-management

- training based on human-activity models. *Presence : Teleoperators and Virtual Environments*, 22(1) :1–19.
- [Bau and Mackay, 2008] Bau, O. and Mackay, W. E. (2008). Octopocus : a dynamic guide for learning gesture-based command sets. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*, pages 37–46. ACM.
- [Baum and Eagon, 1967] Baum, L. E. and Eagon, J. A. (1967). An inequality with applications to statistical estimation for probabilistic functions of markov processes and to a model for ecology. *Bulletin of the American Mathematical Society*, 73(3) :360–363.
- [Bayart et al., 2005] Bayart, B., Pocheville, A., and Kheddar, A. (2005). An adaptive haptic guidance software module for i-touch : example through a handwriting teaching simulation and a 3d maze. In *Haptic Audio Visual Environments and their Applications, 2005. IEEE International Workshop on*, pages 6–pp. IEEE.
- [Benesh et al., 1956] Benesh, R., Benesh, J., and Haskell, A. L. (1956). *An introduction to Benesh dance notation*. Black.
- [Bennes et al., 2012] Bennes, L., Bazzaro, F., and Sagot, J.-C. (2012). Virtual reality as a support tool for ergonomic-style convergence : multidisciplinary interaction design methodology and case study. In *Proceedings of the 2012 Virtual Reality International Conference*, page 24. ACM.
- [Benton and Walker, 2004] Benton, C. and Walker, R. (2004). Augmented reality for maritime navigation : The next generation of electronic navigational aids. In *Proc. 7th Marine Transportation System Research & Technology Coordination Conf., Washington, DC*.
- [Berndt and Clifford, 1994] Berndt, D. J. and Clifford, J. (1994). Using dynamic time warping to find patterns in time series. In *KDD workshop*, volume 10, pages 359–370. Seattle, WA.
- [Bernstein, 1967] Bernstein, N. (1967). The co-ordination and regulation of movements. *The co-ordination and regulation of movements*.
- [Bevilacqua et al., 2007] Bevilacqua, F., Guédy, F., Schnell, N., Fléty, E., and Leroy, N. (2007). Wireless sensor interface and gesture-follower for music pedagogy. In *Proceedings of the 7th international conference on New interfaces for musical expression*, pages 124–129. ACM.
- [Bevilacqua et al., 2010] Bevilacqua, F., Zamborlin, B., Sypniewski, A., Schnell, N., Guédy, F., and Rasamimanana, N. (2010). Continuous realtime gesture following and recognition. In *Gesture in embodied communication and human-computer interaction*, pages 73–84. Springer.
- [Bideau et al., 2010] Bideau, B., Kulpa, R., Vignais, N., Brault, S., Multon, F., and Craig, C. (2010). Using virtual reality to analyze sports performance. *IEEE Computer Graphics and Applications*, 30(2) :14–21.

BIBLIOGRAPHIE

- [Billinghurst and Buxton, 2011] Billinghurst, M. and Buxton, B. (2011). Gesture based interaction. *Haptic input*, 24.
- [Biswas and Basu, 2011] Biswas, K. K. and Basu, S. K. (2011). Gesture recognition using microsoft kinect®. In *Automation, Robotics and Applications (ICARA), 2011 5th International Conference on*, pages 100–103. IEEE.
- [Bjorneseth, 2003] Bjorneseth, O. (2003). Hots (highway on the sea), a new approach to lowvisibility navigation. In *Proceeding of International Conference on Maritime Simulation and Ship Manoeuvrability*, pages pp–RB23.
- [Bloomfield and Badler, 2007] Bloomfield, A. and Badler, N. I. (2007). Collision awareness using vibrotactile arrays. In *Virtual Reality Conference, 2007. VR'07. IEEE*, pages 163–170. IEEE.
- [Bobillier-Chaumon et al., 2005] Bobillier-Chaumon, M.-E., Carvallo, S., Tarpin-Bernard, F., and Vacherand-Revel, J. (2005). Adapter ou uniformiser les interactions personnes-systèmes? *Revue d'Interaction Homme-Machine Vol*, 6(2).
- [Bolt and Herranz, 1992] Bolt, R. A. and Herranz, E. (1992). Two-handed gesture in multi-modal natural dialog. In *Proceedings of the 5th annual ACM symposium on User interface software and technology*, pages 7–14. ACM.
- [Boutet, 2008] Boutet, D. (2008). Une morphologie de la gestualité : structuration articulaire. *Cahiers de linguistique analogique*, (5) :81–115.
- [Bouvier, 2009] Bouvier, P. (2009). *La présence en réalité virtuelle, une approche centrée utilisateur*. PhD thesis, Université Paris-Est.
- [Bowman, 1999] Bowman, D. A. (1999). Interaction techniques for common tasks in immersive virtual environments. *Georgia Institute of Technology*.
- [Bowman and Hodges, 1997] Bowman, D. A. and Hodges, L. F. (1997). An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Proceedings of the 1997 symposium on Interactive 3D graphics*, pages 35–ff. ACM.
- [Bowman et al., 1999] Bowman, D. A., Hodges, L. F., Allison, D., and Wine- man, J. (1999). The educational value of an information-rich virtual environment. *Presence : teleoperators and virtual environments*, 8(3) :317–331.
- [Bowman et al., 2012] Bowman, D. A., McMahan, R. P., and Ragan, E. D. (2012). Questioning naturalism in 3d user interfaces. *Communications of the ACM*, 55(9) :78–88.
- [Bowman et al., 2003] Bowman, D. A., North, C., Chen, J., Polys, N. F., Pyla, P. S., and Yilmaz, U. (2003). Information-rich virtual environments : theory, tools, and research agenda. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 81–90. ACM.
- [Bril and Roux, 2002] Bril, B. and Roux, V. (2002). Le geste technique. réflexions méthodologiques et anthropologiques.

- [Cadoz, 1994] Cadoz, C. (1994). Le geste canal de communication homme/machine : la communication" instrumentale". *Technique et science informatiques*, 13(1) :31–61.
- [Cadoz, 2006] Cadoz, C. (2006). Retour au réel : le sens du feedback rétroaction–interaction–énaction. *Le feedback dans la création musicale, Rencontres musicales pluridisciplinaires*.
- [Cao and Balakrishnan, 2003] Cao, X. and Balakrishnan, R. (2003). Visionwand : interaction techniques for large displays using a passive wand tracked in 3d. In *Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 173–182. ACM.
- [Carmona and Climent, 2012] Carmona, J. M. and Climent, J. (2012). A performance evaluation of hmm and dtw for gesture recognition. In *Iberoamerican Congress on Pattern Recognition*, pages 236–243. Springer.
- [Carpentier, 2015] Carpentier, K. (2015). *Scénarisation personnalisée dynamique dans les environnements virtuels pour la formation*. PhD thesis, Compiègne.
- [Chan et al., 2011] Chan, J. C., Leung, H., Tang, J. K., and Komura, T. (2011). A virtual reality dance training system using motion capture technology. *IEEE Transactions on Learning Technologies*, 4(2) :187–195.
- [Chen and Tseng, 2007] Chen, Y.-T. and Tseng, K.-T. (2007). Multiple-angle hand gesture recognition by fusing svm classifiers. In *Automation Science and Engineering, 2007. CASE 2007. IEEE International Conference on*, pages 527–530. IEEE.
- [Cole et al., 2002] Cole, J., Gallagher, S., and McNeill, D. (2002). Gesture following deafferentation : A phenomenologically informed experimental study. *Phenomenology and the Cognitive Sciences*, 1(1) :49–67.
- [Conté, 1954] Conté, P. (1954). *Technique générale du mouvement*. Arts et mouvement.
- [Covaci et al., 2014] Covaci, A., Olivier, A.-H., and Multon, F. (2014). Third person view and guidance for more natural motor behaviour in immersive basketball playing. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, pages 55–64. ACM.
- [Covaci et al., 2015] Covaci, A., Olivier, A.-H., and Multon, F. (2015). Visual perspective and feedback guidance for vr free-throw training. *Computer Graphics and Applications, IEEE*, 35(5) :55–65.
- [Crowell and Davis, 2011] Crowell, H. P. and Davis, I. S. (2011). Gait retraining to reduce lower extremity loading in runners. *Clinical biomechanics*, 26(1) :78–83.
- [Cruz-Neira et al., 1992] Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., and Hart, J. C. (1992). The cave : audio visual experience automatic virtual environment. *Communications of the ACM*, 35(6) :64–73.

BIBLIOGRAPHIE

- [David, 2005] David, G. (2005). Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders. *Occupational medicine*, 55(3) :190–199.
- [Davis and Shah, 1994] Davis, J. and Shah, M. (1994). Visual gesture recognition. *IEE Proceedings-Vision, Image and Signal Processing*, 141(2) :101–106.
- [De Loor et al., 2008] De Loor, P., Manac’h, K., Fronville, A., and Tisseau, J. (2008). Requirement for an enactive machine : Ontogenesis, interaction and human in the loop. In *International Conference on Enactive Interfaces*.
- [Demuth et al., 2014] Demuth, H. B., Beale, M. H., De Jess, O., and Hagan, M. T. (2014). *Neural network design*. Martin Hagan.
- [Donikian, 2004] Donikian, S. (2004). Modélisation, contrôle et animation d’agents virtuels autonomes évoluant dans des environnements informés et structurés. *Hdr, IRISA*.
- [Eaves et al., 2011] Eaves, D. L., Breslin, G., and Van Schaik, P. (2011). The short-term effects of real-time virtual reality feedback on motor learning in dance. *Presence*, 20(1) :62–77.
- [Eshkol and Wachmann, 1958] Eshkol, N. and Wachmann, A. (1958). *Movement notation*. Weidenfeld and Nicolson.
- [Falzon, 2004] Falzon, P. (2004). *Ergonomie*. Presses Universitaires de France.
- [Feygin et al., 2002] Feygin, D., Keehner, M., and Tendick, R. (2002). Haptic guidance : Experimental evaluation of a haptic training method for a perceptual motor skill. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2002. HAPTICS 2002. Proceedings. 10th Symposium on*, pages 40–47. IEEE.
- [Fisher et al., 1987] Fisher, S. S., McGreevy, M., Humphries, J., and Robinett, W. (1987). Virtual environment display system. In *Proceedings of the 1986 workshop on Interactive 3D graphics*, pages 77–87. ACM.
- [Fitts and Posner, 1967] Fitts, P. M. and Posner, M. I. (1967). Human performance.
- [Forsberg et al., 1996] Forsberg, A., Herndon, K., and Zeleznik, R. (1996). Aperture based selection for immersive virtual environments. In *Proceedings of the 9th annual ACM symposium on User interface software and technology*, pages 95–96. ACM.
- [Franklin and Graesser, 1996] Franklin, S. and Graesser, A. (1996). Is it an agent, or just a program? : A taxonomy for autonomous agents. In *International Workshop on Agent Theories, Architectures, and Languages*, pages 21–35. Springer.
- [Freeman et al., 2009] Freeman, D., Benko, H., Morris, M. R., and Wigdor, D. (2009). Shadowguides : visualizations for in-situ learning of multi-touch and whole-hand gestures. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, pages 165–172. ACM.

- [Frenoy, 2016] Frenoy, R. (2016). *Modélisation de l'activité gestuelle et sélection automatique de feedback pour des environnements interactifs d'apprentissage : application à la calligraphie*. PhD thesis, Compiègne.
- [Fricoteaux, 2012] Fricoteaux, L. (2012). *Interaction informée par un système décisionnel pour un retour multimodal en environnement virtuel : application à la formation au transport fluvial*. PhD thesis, Compiègne.
- [Fricoteaux et al., 2014] Fricoteaux, L., Thouvenin, I., and Mestre, D. (2014). Gulliver : A decision-making system based on user observation for an adaptive training in informed virtual environments. *Engineering Applications of Artificial Intelligence*, 33 :47–57.
- [Fuchs et al., 2006] Fuchs, P., Moreau, G., and Berthoz, A. (2006). *Le traité de la réalité virtuelle : L'Homme et l'environnement virtuel*, volume 1. Presse des Mines.
- [Galligan et al., 2000] Galligan, F. et al. (2000). Acquiring skill. *Advanced PE for Edexcel*, 1 :102–108.
- [Galy et al., 2012] Galy, E., Cariou, M., and Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types? *International Journal of Psychophysiology*, 83(3) :269–275.
- [George et al., 2012] George, P., Thouvenin, I., Fremont, V., and Cherfaoui, V. (2012). Daaria : Driver assistance by augmented reality for intelligent automobile. In *Intelligent Vehicles Symposium (IV), 2012 IEEE*, pages 1043–1048. IEEE.
- [Grossman et al., 2004] Grossman, T., Wigdor, D., and Balakrishnan, R. (2004). Multi-finger gestural interaction with 3d volumetric displays. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 61–70. ACM.
- [Han et al., 2017] Han, F., Reily, B., Hoff, W., and Zhang, H. (2017). Space-time representation of people based on 3d skeletal data : A review. *Computer Vision and Image Understanding*, 158 :85–105.
- [Han et al., 2016] Han, P.-H., Chen, K.-W., Hsieh, C.-H., Huang, Y.-J., and Hung, Y.-P. (2016). Ar-arm : Augmented visualization for guiding arm movement in the first-person perspective. In *Proceedings of the 7th Augmented Human International Conference 2016*, page 31. ACM.
- [Hart and Staveland, 1988] Hart, S. G. and Staveland, L. E. (1988). Development of nasa-tlx (task load index) : Results of empirical and theoretical research. *Advances in psychology*, 52 :139–183.
- [Hatze, 1986] Hatze, H. (1986). Motion variability—its definition, quantification, and origin. *Journal of Motor Behavior*, 18(1) :5–16.
- [Helmer et al., 2010] Helmer, R., Farrow, D., Lucas, S., Higginson, G., and Blanchonette, I. (2010). Can interactive textiles influence a novice's throwing technique? *Procedia Engineering*, 2(2) :2985–2990.

BIBLIOGRAPHIE

- [Henderson and Feiner, 2011] Henderson, S. J. and Feiner, S. K. (2011). Augmented reality in the psychomotor phase of a procedural task. In *Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on*, pages 191–200. IEEE.
- [Hijazi et al., 2015] Hijazi, S., Kumar, R., and Rowen, C. (2015). Using convolutional neural networks for image recognition. Technical report, Tech. Rep., 2015.[Online]. Available : <http://ip.cadence.com/uploads/901/cnn-wp-pdf>.
- [Hoai and De la Torre, 2014] Hoai, M. and De la Torre, F. (2014). Max-margin early event detectors. *International Journal of Computer Vision*, 107(2) :191–202.
- [Hoang et al., 2016] Hoang, T. N., Reinoso, M., Vetere, F., and Tanin, E. (2016). Onebody : Remote posture guidance system using first person view in virtual environment. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, page 25. ACM.
- [Hong et al., 2000] Hong, P., Turk, M., and Huang, T. S. (2000). Gesture modeling and recognition using finite state machines. In *Automatic face and gesture recognition, 2000. proceedings. fourth ieee international conference on*, pages 410–415. IEEE.
- [Hummel et al., 2010] Hummel, J., Hermann, T., Frauenberger, C., and Stockman, T. (2010). Interactive sonification of german wheel sports. In *Proceedings of ISON 2010-Interactive Sonification Workshop : Human Interaction with Auditory Displays*.
- [Jauregui et al., 2014] Jauregui, D. A. G., Argelaguet, F., Olivier, A.-H., Marchal, M., Multon, F., and Lecuyer, A. (2014). 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*, 20(4) :654–661.
- [Jeanne et al., 2017a] Jeanne, F., Soullard, Y., Oker, A., and Thouvenin, I. (2017a). EBAGG : Error-Based Assistance for Gesture Guidance in Virtual Environments. In *IEEE ICALT 2017*, Timisoara, Romania.
- [Jeanne et al., 2017b] Jeanne, F., Thouvenin, I. M., and Alban, L. (2017b). A study on improving performance in gesture training through visual guidance based on learners' errors. In *ACM Symposium on Virtual Reality Software and Technology (VRST)*.
- [Jego, 2013] Jego, J.-F. (2013). *Interaction basée sur des gestes définis par l'utilisateur : Application à la réalité virtuelle*. PhD thesis, Ecole Nationale Supérieure des Mines de Paris.
- [Karam and m. c. schraefel, 2005] Karam, M. and m. c. schraefel (2005). A taxonomy of gestures in human computer interactions.
- [Ke et al., 2013] Ke, S.-R., Thuc, H. L. U., Lee, Y.-J., Hwang, J.-N., Yoo, J.-H., and Choi, K.-H. (2013). A review on video-based human activity recognition. *Computers*, 2(2) :88–131.

- [Kelso et al., 1980] Kelso, J. S., Holt, K., Kugler, P., and Turvey, M. (1980). 2 on the concept of coordinative structures as dissipative structures : Ii. empirical lines of convergence. *Advances in Psychology*, 1 :49–70.
- [Kendon, 1988] Kendon, A. (1988). How gestures can become like words. *Cross-cultural perspectives in nonverbal communication*, 1 :131–141.
- [Keogh and Pazzani, 2001] Keogh, E. J. and Pazzani, M. J. (2001). Derivative dynamic time warping. In *Sdm*, volume 1, pages 5–7. SIAM.
- [Kernodle et al., 2001] Kernodle, M. W., Johnson, R., and Arnold, D. R. (2001). Verbal instruction for correcting errors versus such instructions plus videotape replay on learning the overhand throw. *Perceptual and Motor Skills*, 92(3c) :1039–1051.
- [Kluger and DeNisi, 1996] Kluger, A. N. and DeNisi, A. (1996). The effects of feedback interventions on performance : a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological bulletin*, 119(2) :254.
- [Kovacs and Shea, 2011] Kovacs, A. J. and Shea, C. H. (2011). The learning of 90 continuous relative phase with and without lissajous feedback : external and internally generated bimanual coordination. *Acta psychologica*, 136(3) :311–320.
- [Kugler et al., 1980] Kugler, P. N., Kelso, J. S., and Turvey, M. T. (1980). 1 on the concept of coordinative structures as dissipative structures : I. theoretical lines of convergence. *Advances in Psychology*, 1 :3–47.
- [Kurtenbach, 1990] Kurtenbach, G. (1990). Gestures in human-computer communication. *The Art of Human-Computer Interface Design, B*.
- [Laban and Ullmann, 1971] Laban, R. and Ullmann, L. (1971). *The mastery of movement*. Plays, Inc.
- [Lara and Labrador, 2013] Lara, O. D. and Labrador, M. A. (2013). A survey on human activity recognition using wearable sensors. *IEEE Communications Surveys and Tutorials*, 15(3) :1192–1209.
- [LaViola Jr et al., 2001] LaViola Jr, J. J., Feliz, D. A., Keefe, D. F., and Zeleznik, R. C. (2001). Hands-free multi-scale navigation in virtual environments. In *Proceedings of the 2001 symposium on Interactive 3D graphics*, pages 9–15. ACM.
- [Le Bellu et al., 2010] Le Bellu, S., Lahlou, S., and Nosulenko, V. (2010). Capter et transférer le savoir incorporé dans un geste professionnel. *Social science information*, 49(3) :371–413.
- [LeCun et al., 1990] LeCun, Y., Boser, B. E., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. E., and Jackel, L. D. (1990). Handwritten digit recognition with a back-propagation network. In *Advances in neural information processing systems*, pages 396–404.

BIBLIOGRAPHIE

- [Lenay, 2006] Lenay, C. (2006). Enaction, externalisme et suppléance perceptivo. *Intellectica*, 43 :27–52.
- [Levy and McNeill, 1992] Levy, E. T. and McNeill, D. (1992). Speech, gesture, and discourse. *Discourse processes*, 15(3) :277–301.
- [Li and Jarvis, 2009] Li, Z. and Jarvis, R. (2009). Real time hand gesture recognition using a range camera. In *Australasian Conference on Robotics and Automation*, pages 21–27.
- [Lieberman and Breazeal, 2007] Lieberman, J. and Breazeal, C. (2007). Tikl : Development of a wearable vibrotactile feedback suit for improved human motor learning. *IEEE Transactions on Robotics*, 23(5) :919–926.
- [Lindeman et al., 2004] Lindeman, R. W., Page, R., Yanagida, Y., and Sibert, J. L. (2004). Towards full-body haptic feedback : the design and deployment of a spatialized vibrotactile feedback system. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 146–149. ACM.
- [Loke et al., 2005] Loke, L., Larssen, A. T., and Robertson, T. (2005). Labanotation for design of movement-based interaction. In *Proceedings of the second Australasian conference on Interactive entertainment*, pages 113–120. Creativity & Cognition Studios Press.
- [Louison et al., 2017] Louison, C., Ferlay, F., Keller, D., and Mestre, D. R. (2017). Operators’ accessibility studies for assembly and maintenance scenarios using virtual reality. *Fusion Engineering and Design*.
- [Lourdeaux et al., 2017] Lourdeaux, D., Benabbou, R., Huguet, R., and Lacaze-Labadie, R. (2017). Humans : suite logicielle pour la scénarisation d’environnements virtuels pour la formation à des situations socio-techniques complexes. In *Conférence Nationale sur les Applications Pratiques de l’Intelligence Artificielle*.
- [Lund, 2001] Lund, A. M. (2001). Measuring usability with the use questionnaire. *Usability interface*, 8(2) :3–6.
- [Marchal-Crespo et al., 2010] Marchal-Crespo, L., McHughen, S., Cramer, S. C., and Reinkensmeyer, D. J. (2010). The effect of haptic guidance, aging, and initial skill level on motor learning of a steering task. *Experimental brain research*, 201(2) :209–220.
- [Marchal-Crespo et al., 2013] Marchal-Crespo, L., van Raaij, M., Rauter, G., Wolf, P., and Riener, R. (2013). The effect of haptic guidance and visual feedback on learning a complex tennis task. *Experimental brain research*, 231(3) :277–291.
- [Marschall et al., 2007] Marschall, F., Bund, A., and Wiemeyer, J. (2007). Does frequent augmented feedback really degrade learning? a meta-analysis. *Bewegung und Training*, 1 :75–86.

- [McAtamney and Corlett, 1993] McAtamney, L. and Corlett, E. N. (1993). Rula : a survey method for the investigation of work-related upper limb disorders. *Applied ergonomics*, 24(2) :91–99.
- [McGuinness-Scott, 1983] McGuinness-Scott, J. (1983). *Movement study and Benesh movement notation : an introduction to applications in dance, medicine, anthropology, and other studies*. Oxford Univ Pr.
- [McMahan, 2011] McMahan, R. P. (2011). *Exploring the effects of higher-fidelity display and interaction for virtual reality games*. PhD thesis.
- [McNeill, 1992] McNeill, D. (1992). *Hand and mind : What gestures reveal about thought*. University of Chicago Press.
- [McNeill and Levy, 1980] McNeill, D. and Levy, E. (1980). *Conceptual representations in language activity and gesture*. ERIC Clearinghouse.
- [Miyake, 2001] Miyake, S. (2001). Multivariate workload evaluation combining physiological and subjective measures. *International journal of psychophysiology*, 40(3) :233–238.
- [Moreau et al., 2008] Moreau, R., Ochoa, V., Pham, M. T., Boulanger, P., Redarce, T., and Dupuis, O. (2008). A method to evaluate skill transfer and acquisition of obstetric gestures based on the curvatures analysis of the position and the orientation. *Journal of biomedical informatics*, 41(6) :991–1000.
- [Moreau et al., 2007] Moreau, R., Pham, M. T., Redarce, T., and Dupuis, O. (2007). A new learning method for obstetric gestures using the birthsim simulator. In *Robotics and Automation, 2007 IEEE International Conference on*, pages 2279–2284. IEEE.
- [Mori et al., 2006] Mori, A., Uchida, S., Kurazume, R., Taniguchi, R.-i., Hasegawa, T., and Sakoe, H. (2006). Early recognition and prediction of gestures. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 3, pages 560–563. IEEE.
- [Murakami and Taguchi, 1991] Murakami, K. and Taguchi, H. (1991). Gesture recognition using recurrent neural networks. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 237–242. ACM.
- [Murthy and Jadon, 2009] Murthy, G. and Jadon, R. (2009). A review of vision based hand gestures recognition. *International Journal of Information Technology and Knowledge Management*, 2(2) :405–410.
- [Muscillo et al., 2011] Muscillo, R., Schmid, M., Conforto, S., and D’Alessio, T. (2011). Early recognition of upper limb motor tasks through accelerometers : real-time implementation of a dtw-based algorithm. *Computers in biology and medicine*, 41(3) :164–172.
- [Myers and Rabiner, 1981] Myers, C. S. and Rabiner, L. R. (1981). A comparative study of several dynamic time-warping algorithms for connected-word recognition. *Bell Labs Technical Journal*, 60(7) :1389–1409.

BIBLIOGRAPHIE

- [Nesbitt et al., 2003] Nesbitt, K. et al. (2003). Designing multi-sensory displays for abstract data.
- [Newell, 1986] Newell, K. M. (1986). Constraints on the development of coordination. *Motor development in children : Aspects of coordination and control*, 34 :341–360.
- [Oppermann, 1994] Oppermann, R. (1994). *Adaptive user support : ergonomic design of manually and automatically adaptable software*. CRC Press.
- [Oppermann and Rashev, 1997] Oppermann, R. and Rashev, R. (1997). Adaptability and adaptivity in learning systems. *Knowledge transfer*, 2 :173–179.
- [Ou et al., 2003] Ou, J., Fussell, S. R., Chen, X., Setlock, L. D., and Yang, J. (2003). Gestural communication over video stream : supporting multimodal interaction for remote collaborative physical tasks. In *Proceedings of the 5th international conference on Multimodal interfaces*, pages 242–249. ACM.
- [Poupyrev et al., 1996] Poupyrev, I., Billingham, M., Weghorst, S., and Ichikawa, T. (1996). The go-go interaction technique : non-linear mapping for direct manipulation in vr. In *Proceedings of the 9th annual ACM symposium on User interface software and technology*, pages 79–80. ACM.
- [Poupyrev and Ichikawa, 1999] Poupyrev, I. and Ichikawa, T. (1999). Manipulating objects in virtual worlds : Categorization and empirical evaluation of interaction techniques. *Journal of Visual Languages & Computing*, 10(1) :19–35.
- [Quek et al., 2002] Quek, F., McNeill, D., Bryll, R., Duncan, S., Ma, X.-F., Kirbas, C., McCullough, K. E., and Ansari, R. (2002). Multimodal human discourse : gesture and speech. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 9(3) :171–193.
- [Rabiner, 1989] Rabiner, L. R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2) :257–286.
- [Ramey et al., 2011] Ramey, A., González-Pacheco, V., and Salichs, M. A. (2011). Integration of a low-cost rgb-d sensor in a social robot for gesture recognition. In *Proceedings of the 6th international conference on Human-robot interaction*, pages 229–230. ACM.
- [Rautaray and Agrawal, 2015] Rautaray, S. S. and Agrawal, A. (2015). Vision based hand gesture recognition for human computer interaction : a survey. *Artificial Intelligence Review*, 43(1) :1–54.
- [Rekimoto, 1997] Rekimoto, J. (1997). Pick-and-drop : a direct manipulation technique for multiple computer environments. In *Proceedings of the 10th annual ACM symposium on User interface software and technology*, pages 31–39. ACM.

-
- [Rekimoto, 2002] Rekimoto, J. (2002). Smartskin : an infrastructure for free-hand manipulation on interactive surfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 113–120. ACM.
- [Rich, 1989] Rich, E. (1989). Stereotypes and user modeling. In *User models in dialog systems*, pages 35–51. Springer.
- [Rimé and Schiaratura, 1991] Rimé, B. and Schiaratura, L. (1991). Gesture and speech.
- [Ronsse et al., 2010] Ronsse, R., Puttemans, V., Coxon, J. P., Goble, D. J., Wagemans, J., Wenderoth, N., and Swinnen, S. P. (2010). Motor learning with augmented feedback : modality-dependent behavioral and neural consequences. *Cerebral cortex*, 21(6) :1283–1294.
- [Rovelo et al., 2015] Rovelo, G., Degraen, D., Vanacken, D., Luyten, K., and Coninx, K. (2015). Gestu-wan-an intelligible mid-air gesture guidance system for walk-up-and-use displays. In *Human-Computer Interaction*, pages 368–386. Springer.
- [Sabinas et al., 2013] Sabinas, Y., Morales, E. F., and Escalante, H. J. (2013). A one-shot dtw-based method for early gesture recognition. In *Iberoamerican Congress on Pattern Recognition*, pages 439–446. Springer.
- [Salmoni et al., 1984] Salmoni, A. W., Schmidt, R. A., and Walter, C. B. (1984). Knowledge of results and motor learning : a review and critical reappraisal. *Psychological bulletin*, 95(3) :355.
- [Schmidt, 1991] Schmidt, R. A. (1991). Frequent augmented feedback can degrade learning : Evidence and interpretations. In *Tutorials in motor neuroscience*, pages 59–75. Springer.
- [Schmidt and Lee, 1988] Schmidt, R. A. and Lee, T. (1988). *Motor control and learning*. Human kinetics.
- [Schmidt et al., 2005] Schmidt, R. A., Lee, T. D., et al. (2005). *Motor control and learning : A behavioral emphasis*, volume 4. Human kinetics Champaign, IL.
- [Schmidt and Wrisberg, 2008] Schmidt, R. A. and Wrisberg, C. A. (2008). *Motor learning and performance : A situation-based learning approach*. Human Kinetics.
- [Schmidt and Wulf, 1997] Schmidt, R. A. and Wulf, G. (1997). Continuous concurrent feedback degrades skill learning : Implications for training and simulation. *Human factors*, 39(4) :509.
- [Schönauer et al., 2012] Schönauer, C., Fukushi, K., Olwal, A., Kaufmann, H., and Raskar, R. (2012). Multimodal motion guidance : techniques for adaptive and dynamic feedback. In *Proceedings of the 14th ACM international conference on Multimodal interaction*, pages 133–140. ACM.

BIBLIOGRAPHIE

- [Schöner et al., 1992] Schöner, G., Zanone, P. G., and Kelso, J. (1992). Learning as change of coordination dynamics : Theory and experiment. *Journal of motor behavior*, 24(1) :29–48.
- [Serre, 1984] Serre, J.-C. (1984). La danse parmi les autres formes de la motricité. *La recherche en danse*, 3 :135–156.
- [Shea and Wulf, 1999] Shea, C. H. and Wulf, G. (1999). Enhancing motor learning through external-focus instructions and feedback. *Human Movement Science*, 18(4) :553–571.
- [Shmuelof et al., 2012] Shmuelof, L., Krakauer, J. W., and Mazzoni, P. (2012). How is a motor skill learned? change and invariance at the levels of task success and trajectory control. *Journal of neurophysiology*, 108(2) :578–594.
- [Shneiderman, 1993] Shneiderman, B. (1993). Direct manipulation : a step beyond programming languages. *Sparks of innovation in human-computer interaction*, page 17.
- [Sigrist et al., 2015] Sigrist, R., Rauter, G., Marchal-Crespo, L., Riener, R., and Wolf, P. (2015). Sonification and haptic feedback in addition to visual feedback enhances complex motor task learning. *Experimental brain research*, 233(3) :909–925.
- [Sigrist et al., 2013] Sigrist, R., Rauter, G., Riener, R., and Wolf, P. (2013). Augmented visual, auditory, haptic, and multimodal feedback in motor learning : a review. *Psychonomic bulletin & review*, 20(1) :21–53.
- [Simoneau et al., 1996] Simoneau, S., St-Vincent, M., and Chicoine, D. (1996). Les lésions musculosquelettiques associées au travail répétitif, mieux les comprendre pour mieux les prévenir.
- [Smisek et al., 2013] Smisek, J., Jancosek, M., and Pajdla, T. (2013). 3d with kinect. In *Consumer depth cameras for computer vision*, pages 3–25. Springer.
- [Smith and Stein, 1998] Smith, M. S. and Stein, M. K. (1998). Selecting and creating mathematical tasks : From research to practice. *Mathematics teaching in the middle school*, 3(5) :344–50.
- [Sodhi et al., 2012] Sodhi, R., Benko, H., and Wilson, A. (2012). Lightguide : projected visualizations for hand movement guidance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 179–188. ACM.
- [Suma et al., 2011] Suma, E. A., Lange, B., Rizzo, A. S., Krum, D. M., and Bolas, M. (2011). Faast : The flexible action and articulated skeleton toolkit. In *Virtual Reality Conference (VR), 2011 IEEE*, pages 247–248. IEEE.
- [Sutton, 1981] Sutton, V. (1981). Sutton movement writing & shorthand. *Dance Research Journal*, pages 78–85.

-
- [Sutton, 1982] Sutton, V. (1982). *Dance Writing Shorthand for Modern & Jazz Dance*. Center for Sutton Movement Writing.
- [Sweller, 1988] Sweller, J. (1988). Cognitive load during problem solving : Effects on learning. *Cognitive science*, 12(2) :257–285.
- [Sweller, 1994] Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, 4(4) :295–312.
- [Teitelbaum et al., 2004] Teitelbaum, O., Benton, T., Shah, P. K., Prince, A., Kelly, J. L., and Teitelbaum, P. (2004). Eshkol–wachman movement notation in diagnosis : The early detection of asperger’s syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, 101(32) :11909–11914.
- [Ten Holt et al., 2007] Ten Holt, G., Reinders, M., and Hendriks, E. (2007). Multi-dimensional dynamic time warping for gesture recognition.
- [Thouvenin, 2009] Thouvenin, I. M. (2009). *Interaction et connaissance : construction d’une expérience dans le monde virtuel*. Hdr, Université de Technologie de Compiègne.
- [Tomás, 2008] Tomás, J.-L. (2008). La transmission des gestes de métier en chirurgie cardiaque. *Centre de Recherche sur le Travail et le Développement (EA 4132), CNAM, équipe clinique de l’activité. Site de l’Institut de recherche Robert-Sauvé en santé et sécurité du travail (en ligne)*. <http://www.irsst.qc.ca>.
- [Tönnis and Klinker, 2009] Tönnis, M. and Klinker, G. (2009). Augmented 3d arrows reach their limits in automotive environments. In *Mixed Reality In Architecture, Design And Construction*, pages 185–202. Springer.
- [Tönnis et al., 2009] Tönnis, M., Klinker, G., and Plavšić, M. (2009). Survey and classification of head-up display presentation principles. *Proceedings of the International Ergonomics Association (IEA)*.
- [Tzetzis et al., 2008] Tzetzis, G., Votsis, E., and Kourtessis, T. (2008). Back issues. *Journal of Sports Science and Medicine*, 7 :371–378.
- [Van den Oord et al., 2013] Van den Oord, A., Dieleman, S., and Schrauwen, B. (2013). Deep content-based music recommendation. In *Advances in neural information processing systems*, pages 2643–2651.
- [Varela et al., 1974] Varela, F. G., Maturana, H. R., and Uribe, R. (1974). Autopoiesis : the organization of living systems, its characterization and a model. *Biosystems*, 5(4) :187–196.
- [Varela et al., 1989] Varela, F. J., Dumouchel, P., and Bourguine, P. (1989). *Autonomie et connaissance*. Seuil.
- [Varela et al., 1992] Varela, F. J., Thompson, E., and Rosch, E. (1992). *The embodied mind : Cognitive science and human experience*. MIT press.

BIBLIOGRAPHIE

- [Vatavu et al., 2005] Vatavu, R.-D., Pentiuc, Ş.-G., and Chaillou, C. (2005). On natural gestures for interacting in virtual environments. *Advances in electrical and computer engineering*, 24(5) :72–79.
- [Velloso et al., 2013] Velloso, E., Bulling, A., and Gellersen, H. (2013). Motionma : motion modelling and analysis by demonstration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1309–1318. ACM.
- [Vignais et al., 2013] Vignais, N., Miezal, M., Bleser, G., Mura, K., Gorecky, D., and Marin, F. (2013). Innovative system for real-time ergonomic feedback in industrial manufacturing. *Applied ergonomics*, 44(4) :566–574.
- [Wei et al., 2005] Wei, Y., Patton, J., Bajaj, P., and Scheidt, R. (2005). A real-time haptic/graphic demonstration of how error augmentation can enhance learning. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pages 4406–4411. IEEE.
- [Wen et al., 2013] Wen, R., Chui, C.-K., Ong, S.-H., Lim, K.-B., and Chang, S. K.-Y. (2013). Projection-based visual guidance for robot-aided rf needle insertion. *International journal of computer assisted radiology and surgery*, 8(6) :1015–1025.
- [Wen et al., 2014] Wen, R., Tay, W.-L., Nguyen, B. P., Chng, C.-B., and Chui, C.-K. (2014). Hand gesture guided robot-assisted surgery based on a direct augmented reality interface. *Computer methods and programs in biomedicine*, 116(2) :68–80.
- [Williams and Hodges, 2005] Williams, A. M. and Hodges, N. J. (2005). Practice, instruction and skill acquisition in soccer : Challenging tradition. *Journal of sports sciences*, 23(6) :637–650.
- [Winograd et al., 1989] Winograd, T., Flores, F., and Peytavin, J.-L. (1989). *L’intelligence artificielle en question*. Presses universitaires de France.
- [Witmer and Singer, 1998] Witmer, B. G. and Singer, M. J. (1998). Measuring presence in virtual environments : A presence questionnaire. *Presence : Teleoperators and virtual environments*, 7(3) :225–240.
- [Wu and Balakrishnan, 2003] Wu, M. and Balakrishnan, R. (2003). Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 193–202. ACM.
- [Wulf, 2007] Wulf, G. (2007). Attentional focus and motor learning : A review of 10 years of research. *E-journal Bewegung und Training*, 1(2-3) :1–11.
- [Wulf and Shea, 2002] Wulf, G. and Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic bulletin & review*, 9(2) :185–211.

- [Wulf et al., 1998] Wulf, G., Shea, C. H., and Matschiner, S. (1998). Frequent feedback enhances complex motor skill learning. *Journal of motor behavior*, 30(2) :180–192.
- [Yang and Kim, 2002] Yang, U. and Kim, G. J. (2002). Implementation and evaluation of "just follow me" : an immersive, vr-based, motion-training system. *Presence : Teleoperators and Virtual Environments*, 11(3) :304–323.
- [Young, 2009] Young, H. P. (2009). Learning by trial and error. *Games and economic behavior*, 65(2) :626–643.