



HAL
open science

Analyse, reconnaissance et réalisation des gestes pour l'entraînement en chirurgie laparoscopique robotisée

Fabien Despinoy

► **To cite this version:**

Fabien Despinoy. Analyse, reconnaissance et réalisation des gestes pour l'entraînement en chirurgie laparoscopique robotisée. Automatique / Robotique. Université Montpellier, 2015. Français. NNT : 2015MONT037 . tel-01261618v1

HAL Id: tel-01261618

<https://hal-lirmm.ccsd.cnrs.fr/tel-01261618v1>

Submitted on 25 Jan 2016 (v1), last revised 29 May 2017 (v2)

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.

THÈSE

Pour obtenir le grade de
Docteur

Délivré par l'Université de Montpellier

Préparée au sein de l'école doctorale
Information, Structures, Systèmes

Et de l'unité de recherche
**Laboratoire d'Informatique, de Robotique
et de Microélectronique de Montpellier**

Spécialité doctorale
Systèmes Automatiques et Microélectroniques

Présentée par **Fabien Despinoy**

**Analyse, reconnaissance et réalisation
des gestes pour l'entraînement en
chirurgie laparoscopique robotisée**

Soutenue le 14 décembre 2015 devant le jury composé de :

M. Jacques HUBERT	PU-PH	Université Lorraine	Président
Mme. Jocelyne TROCCAZ	DR CNRS	Université Joseph Fourier	Rapporteur
M. Jérôme SZEWCZYK	PU	Université Paris 6	Rapporteur
M. Philippe POIGNET	PU	Université Montpellier	Directeur
M. Pierre JANNIN	DR INSERM	Université Rennes 1	Co-Directeur
M. Nabil ZEMITI	MCU	Université Montpellier	Co-Encadrant



*Parce que certaines personnes sont irremplaçables ...
À Baptiste et à Hubert.*



Remerciements

Au travers de cette aventure scientifique, ces années ont été l'occasion de rencontrer des personnes que je souhaite remercier aujourd'hui. Je commencerai tout d'abord par remercier Jacques Hubert, qui m'a fait l'honneur de présider mon jury de thèse et d'apporter son expérience sur ces travaux. Je remercie également Jocelyne Troccaz et Jérôme Szewczyk pour avoir accepté de rapporter mon travail, ainsi que pour leurs regards critiques sur la qualité scientifique de ce manuscrit.

Ayant réalisé une grande partie de ma thèse au sein du LIRMM à Montpellier, je me dois de remercier les personnes avec qui j'ai passé d'excellents moments, aussi bien au sein du laboratoire qu'à l'extérieur. Merci à Nicolas, Divine et Alejandro, les anciens qui ont su nous montrer le chemin de la recherche. Merci également à Ederson, Jing, Julien, Mohamed et Moussab, d'excellents collègues avec le sourire. Nos discussions ont toujours été sympathiques. Merci également à David pour nos discussions le midi, un petit nouveau qui j'en suis sûr saura faire sa place dans la recherche et dans l'équipe robotique médicale.

Je souhaite également remercier des personnes qui m'ont beaucoup apporté au travers des échanges que nous avons pu avoir durant le travail ainsi qu'autour des verres. Merci à Andrea ("Colloooooo") pour sa joie et sa bonne humeur, un co-bureau à Montpellier des plus plaisant, un très bon ami et un très bon coéquipier aussi bien sur le terrain de football que derrière une manette de console. Je remercie également Johann pour sa bonne humeur, nos discussions scientifiques ainsi que le partage d'un plaisir en commun qu'est la moto. Tu fais un fabuleux père, j'en suis sûr ! Enfin, un très grand merci à Alonso, qui au fil du temps est devenu un vrai ami. Nous avons partagé beaucoup de choses ensemble, et j'espère que ce n'est qu'un début. Les échanges scientifiques autour d'un sandwich, d'une pizza, mais également les discussions personnelles m'ont permis de garder la tête sur les épaules, et pour cela je t'en remercie infiniment.

Cette aventure m'a également amené en Bretagne, et plus particulièrement à Rennes au

REMERCIEMENTS

sein du LTSI et de l'équipe MediCIS. J'y ai rencontré des collègues, des amis, et bien plus encore. Merci à Noémie pour ta joie et ta bonne humeur durant les pauses café et les repas du midi. Merci également à Frédo qui a su nous faire rigoler durant les repas ! Merci à Darko qui, même de passage, a su amener de la fraîcheur dans nos discussions.

J'ai également découvert deux personnes qui m'ont beaucoup appris durant mon séjour à Rennes, et sans qui ces travaux n'auraient été possibles. Merci à Cédric pour son aide, ses précieux conseils, ainsi que ses réponses à mes multiples questions. Également un grand merci à David, mon co-bureau de Rennes, avec qui j'ai passé de très bons moments au bureau (toujours en musique) mais aussi à l'extérieur. Ton aide et ton soutien m'ont permis d'avancer et d'arriver jusqu'ici. Je tiens également à remercier Germain, un collègue de Mulhouse, avec qui j'ai également beaucoup appris. J'espère que nous aurons l'occasion de travailler de nouveau ensemble.

Je souhaite très sincèrement remercier mes directeurs et mon encadrant de thèse, sans qui ces travaux n'auraient pu voir le jour. Je remercie Nabil pour son aide, sa gentillesse, ainsi que les nombreuses heures que nous avons passé ensemble durant les rédactions de projets et les relectures d'articles. Je remercie également Philippe pour m'avoir offert l'opportunité de réaliser cette thèse. Merci à vous pour vos conseils avisés et votre soutien. Enfin, je me dois de remercier Pierre qui m'a énormément apporté aussi bien d'un point de vue scientifique que personnel. Nos discussions ainsi que votre soutien m'ont été très précieux durant mon année rennaise. Merci également de vos relectures, de vos conseils et de votre encadrement qui m'ont été extrêmement bénéfiques. Ce fut un plaisir de travailler à vos côtés, et j'espère que cela n'est qu'un début. Je remercie également Salih, Étienne, Yacine et Bernard pour nos discussions scientifiques très intéressantes. Cela fait toujours plaisir d'échanger avec des personnes aussi compétentes que vous. Durant ces trois années, j'ai également eu l'occasion de réaliser des enseignements afin de découvrir la face cachée du monde académique. Et grâce à Alain, j'ai été conquis. J'ai pris un immense plaisir à réaliser ces travaux pratiques et je te remercie très sincèrement de m'en avoir offert l'opportunité. Cela m'a beaucoup aidé pour le choix de mon parcours professionnel.

Ces travaux n'auraient pu voir le jour sans le soutien de ma famille et de mes proches, qui me sont très chers et qui m'ont beaucoup apporté durant ces trois années. Merci à mes parents, ma sœur, mon (futur) beau-frère et également mon neveu et ma nièce qui sont des perles rares ! J'espère enfin pouvoir rattraper le temps perdu avec vous.

Enfin, cette aventure a également changé ma vie et m'a permis de rencontrer une personne rare et exceptionnelle avec qui j'ai l'honneur aujourd'hui de partager mon quotidien. Merci à toi Olga pour ton amour, ton soutien, ta joie et ton sourire qui m'aide à me sentir mieux chaque jour que je passe à tes côtés. Grâce à toi, je sais ce que le mot "Aimer" veut dire, et je t'en remercie du fond du cœur.



Liste des publications

Revue internationale

F. Despinoy, D. Bouget, G. Forestier, C. Penet, N. Zemiti, P. Poignet, P. Jannin. *Unsupervised Trajectory Segmentation for Surgical Gesture Recognition in Robotic Training*. IEEE Transactions on Biomedical Engineering, DOI 10.1109/TBME.2015.2493100 (early access), 2015.

Conférence internationale

F. Despinoy, A. Sanchez, N. Zemiti, P. Jannin, P. Poignet. *Comparative assessment of a novel optical human-machine interface for laparoscopic telesurgery*. International Conference on Information Processing in Computer-Assisted Interventions, vol. 8498, pages 21-30, 2014.

Conférence nationale

F. Despinoy, A. Sanchez, N. Zemiti, P. Jannin, P. Poignet. *A novel contactless human-machine interface for laparoscopic telesurgery*. Surgetica, du 3 au 5 décembre 2014.

Workshop national

F. Despinoy, A. Sanchez, N. Zemiti, P. Jannin, P. Poignet. *Comparative assessment of a novel optical human-machine interface for laparoscopic telesurgery*. GdR Robotique, Groupe de travail "Robotique et Santé", Thème F : Apprentissage et assistance aux gestes médico-chirurgicaux, Journée "Interfaces des assistants robotiques en chirurgie et en médecine interventionnelle", 21 mai 2014.



Table des matières

Table des matières	i
Table des figures	v
Liste des tableaux	ix
Introduction	1
I Analyse et compréhension du geste en chirurgie laparoscopique robotisée	3
1 Analyse du geste chirurgical : point de vue clinique	5
1.1 La chirurgie laparoscopique	6
1.1.1 Pratique chirurgicale	6
1.1.2 La chirurgie assistée par robot	8
1.1.3 Les problématiques persistantes	10
1.2 L'entraînement chirurgical	12
1.2.1 L'apprentissage des compétences	12
1.2.1.1 Les compétences non-techniques	12
1.2.1.2 Les compétences techniques	13
1.2.2 Les dispositifs d'entraînement aux gestes chirurgicaux	14
1.2.2.1 Les différentes techniques de simulation	14
1.2.2.2 Les simulateurs organiques	15
1.2.2.3 Les simulateurs non-organiques	16

1.2.2.4	Aperçu des simulateurs non-organiques existants	17
1.2.3	Les méthodes d'évaluation du geste chirurgical	19
1.2.3.1	Méthodes qualitatives basées sur un observateur externe	20
1.2.3.2	Méthodes quantitatives automatiques	21
1.3	Analyse et discussion	23
2	Analyse du geste chirurgical : point de vue scientifique	25
2.1	Compréhension quantitative des interventions	26
2.1.1	Analyse et modélisation de processus chirurgicaux	26
2.1.1.1	Définition des concepts	27
2.1.1.2	Niveaux de granularité	29
2.1.2	Vers une reconnaissance du geste chirurgical	31
2.1.2.1	Notions de surgeme et dexeme	31
2.1.2.2	Description du processus de reconnaissance	33
2.2	Etat de l'art des travaux pour la reconnaissance des gestes chirurgicaux	35
2.2.1	Conditions d'étude	36
2.2.1.1	Contexte médical	36
2.2.1.2	Acquisition des données	38
2.2.2	Processus de reconnaissance	42
2.2.2.1	Prétraitement des données	42
2.2.2.2	Extraction d'attributs	44
2.2.2.3	Méthodologie d'apprentissage	44
2.2.2.4	Spécification du processus	46
2.2.3	Processus de validation	46
2.2.3.1	Conditions expérimentales	47
2.2.3.2	Stratégie de validation	49
2.2.4	Analyse et discussion	51
2.2.4.1	Conditions d'étude	51
2.2.4.2	Processus de reconnaissance	52
2.2.4.3	Processus de validation	54
2.3	Conclusion et problématiques de la thèse	55
II	Méthodes pour la reconnaissance et la réalisation de gestes destinées à l'entraînement en chirurgie robotisée	57
3	Segmentation automatique et reconnaissance des gestes chirurgicaux	59
3.1	Introduction et description de la méthode	60
3.1.1	Segmentation non-supervisée	61
3.1.2	Classification et reconnaissance	63
3.2	Approche non-supervisée pour la segmentation de trajectoires	63

3.2.1	Prétraitement des données	63
3.2.1.1	Variables cinématiques	64
3.2.1.2	Filtrage	65
3.2.1.3	Normalisation	67
3.2.2	Sélection de délimiteurs temporels	68
3.2.2.1	Notion de persistance	69
3.2.2.2	Utilisation en pratique	70
3.2.3	Analyse par similarité spatio-temporelle	70
3.2.3.1	Métriques de similarité spatio-temporelle	71
3.2.3.2	Choix de la dimension du problème	74
3.2.4	Sélection automatique	76
3.2.5	Vérification de la méthode de segmentation	77
3.2.5.1	Protocole expérimental	77
3.2.5.2	Données	78
3.2.5.3	Résultats	78
3.3	Classification et reconnaissance des surgemes	79
3.3.1	Extraction d'attributs caractéristiques	81
3.3.1.1	Transformation des signaux	82
3.3.1.2	Approximation polynômiale	82
3.3.2	Algorithmes de classification	83
3.3.2.1	K-Plus Proches Voisins	84
3.3.2.2	Machines à Vecteurs de Support	85
3.3.3	Association de dexemes	86
3.3.4	Validation de la méthode de reconnaissance	86
3.3.4.1	Type de tâche et vocabulaire	87
3.3.4.2	Stratégie de validation	87
3.3.4.3	Données	90
3.3.4.4	Métriques d'évaluation	90
3.3.4.5	Résultats	91
3.3.5	Analyse et discussion	99
3.3.5.1	Segmentation non-supervisée	99
3.3.5.2	Classification et reconnaissance	101
3.4	Conclusion	104
4	Évaluation d'une nouvelle interface sans contact pour l'entraînement en chi- rurgie robotisée	107
4.1	État de l'art des interfaces homme-machine pour la commande des robots chirurgicaux	108
4.1.1	Commande par interface haptique	109
4.1.2	Commande par interface sans contact	112
4.1.3	Analyse et discussion	115

4.2	Étude d'une nouvelle interface sans contact pour l'entraînement en chirurgie robotisée	117
4.2.1	Plateforme expérimentale pour l'entraînement en chirurgie robotisée	117
4.2.1.1	Le robot Raven-II	117
4.2.1.2	Les interfaces homme-machine	119
4.2.1.3	La tâche d'entraînement chirurgical	121
4.2.2	Évaluation pour l'entraînement en chirurgie robotisée	123
4.2.2.1	Protocole expérimental	123
4.2.2.2	Données et métriques d'évaluation	123
4.2.2.3	Résultats	124
4.2.2.4	Analyse et discussion	126
4.3	Vers une analyse des trajectoires pour l'évaluation du niveau d'expertise . .	128
4.3.1	Amélioration de la commande sans contact	128
4.3.1.1	Modification du modèle	128
4.3.1.2	Filtrage des tremblements	129
4.3.2	Méthodologie d'analyse des trajectoires	130
4.3.2.1	Prétraitement des données	130
4.3.2.2	Approche non-supervisée pour la classification	130
4.3.3	Vérification de l'interface et du niveau d'expertise	133
4.3.3.1	Protocole expérimental	133
4.3.3.2	Données	133
4.3.3.3	Résultats	133
4.3.3.4	Analyse et discussion	136
4.4	Conclusion	137
	Conclusion et perspectives des travaux	139
	Bibliographie	147



Table des figures

1.1	Illustration du principe de laparoscopie.	7
1.2	Le robot da Vinci commercialisé par la société Intuitive Surgical.	9
1.3	Répartition des interventions réalisées avec le robot da Vinci depuis 2011, en France.	10
1.4	Effets Indésirables Graves intervenants en chirurgie robotisée.	11
1.5	Les différentes techniques de simulation employées en France.	15
1.6	Les différents moyens de simulation pour le médical.	16
1.7	Les différents types de simulateurs laparoscopiques.	19
2.1	Décomposition temporelle et hiérarchique d'une fundoplicature par voie laparoscopique en utilisant une description à plusieurs niveaux.	28
2.2	Les différents niveaux de granularité utilisés pour décrire un processus chirurgical.	29
2.3	Répartition des niveaux de granularité utilisés pour la modélisation.	30
2.4	Décomposition en surgemes d'une tâche de transfert d'objet pour l'entraînement en chirurgie laparoscopique.	32
2.5	Processus généralisé pour la reconnaissance des gestes chirurgicaux.	33
3.1	Schéma du processus complet de segmentation et reconnaissance des gestes chirurgicaux.	62
3.2	Filtrage de la position cartésienne de l'instrument chirurgical selon l'axe X du repère du robot.	68
3.3	Normalisation du signal pour obtenir une représentation centrée réduite.	68
3.4	Illustration du principe de persistance topologique en 1D.	69

3.5	Évolution de la sélection des points caractéristiques en utilisant le seuil de persistance.	70
3.6	Extraction des points caractéristiques des signaux pour obtenir un vecteur de délimiteurs temporels.	71
3.7	Illustration de la distance de Hausdorff entre deux séries temporelles.	72
3.8	Illustration de la distance de Fréchet discrète entre deux séries temporelles.	73
3.9	Illustration de l'utilisation de l'algorithme de Déformation Temporelle Dynamique entre deux séries temporelles.	74
3.10	Illustration du calcul de dissimilarité suivant le choix de la dimension du problème.	75
3.11	Sélection automatique des délimiteurs temporels en utilisant l'algorithme de Suppression de Non-Maxima.	77
3.12	Analyse de la segmentation sur une trajectoire aléatoire. L'algorithme de sélection des scores permet une détection pertinente des points caractéristiques (en noirs), qui définissent les différents dexemes.	80
3.13	Trajectoires acquises durant la réalisation de la lettre "R" en utilisant la plateforme robotisée de téléopération.	81
3.14	Schéma de construction du vecteur caractéristique représentant un dexeme à partir de données multidimensionnelles.	83
3.15	Schéma de fonctionnement de l'algorithme des K-Plus Proches Voisins.	84
3.16	Schéma de fonctionnement de l'algorithme des Machines à Vecteurs de Support.	85
3.17	Illustration de la tâche d'entraînement chirurgical.	88
3.18	Graphique des performances de reconnaissance avec les deux classifieurs étudiés.	92
3.19	Matrices de confusion représentant les répartitions de classification par rapport aux surgemes de référence en se basant sur la vérité terrain.	94
3.20	Graphique des performances de reconnaissance avec les deux classifieurs étudiés ainsi que les différentes métriques de segmentation utilisées.	95
3.21	Matrices de confusion représentant les répartitions de classification par rapport aux surgemes de référence avec notre approche de reconnaissance.	97
3.22	Comparaison spatiale et temporelle des surgemes de références et des surgemes reconnus par notre approche de reconnaissance.	98
4.1	Les principales architectures cinématiques des manipulateurs haptiques.	110
4.2	Interfaces de commande séries de robots chirurgicaux.	111
4.3	Dispositifs mécaniques séries pour la commande de robots.	111
4.4	Dispositifs mécaniques parallèles pour la commande de robots.	112
4.5	Dispositifs à structures parallèles à câbles pour la commande de robots.	113
4.6	Robot ViKY porte endoscope contrôlé via une interface vocale.	114
4.7	Commande du robot EndoAssist par le suivi de la tête.	115
4.8	Interface sans contact pour la commande d'un instrument du da Vinci.	116
4.9	La plateforme robotisée pour l'entraînement aux gestes chirurgicaux.	118

4.10	Le robot Raven-II utilisé pour l'entraînement aux gestes en chirurgie laparoscopique robotisée.	119
4.11	Console de commande avec deux Sigma.7 pour le contrôle des instruments chirurgicaux.	120
4.12	Représentation des espaces de travail simplifiés des interfaces homme-machine.	121
4.13	Schéma d'acquisition des informations cartésiennes avec les interfaces pour la commande du robot.	122
4.14	Illustration de la tâche de transfert de pions.	122
4.15	Comparaison des interfaces en utilisant les métriques quantitatives.	125
4.16	Comparaison des interfaces en utilisant les métriques qualitatives.	125
4.17	Schéma d'acquisition des informations cartésiennes avec l'interface sans contact Leap Motion.	129
4.18	Illustration pour la construction de la matrice de similarité.	131
4.19	Schéma de fonctionnement de l'algorithme de classification ascendante hiérarchique.	132
4.20	Analyse de la similarité entre deux trajectoires d'un même instrument, provenant de deux opérateurs différents.	134
4.21	Regroupement des trajectoires basé sur la métrique 1D-DTD.	135
4.22	Regroupement des trajectoires basé sur la métrique MD-DTD.	135



Liste des tableaux

1.1	Description et répartition des compétences nécessaires en chirurgie.	13
1.2	Simulateurs destinés à l'apprentissage de la chirurgie laparoscopique manuelle ou robotisée.	18
1.3	Échelle d'évaluation GOALS. La notation est réalisée ici par l'attribution d'un score de 1 à 5 sur l'échelle de Linkert.	20
1.4	Extrait de la liste de vérification pour l'hémostate d'une plaie de veine cave inférieure (IVC) avec l'échelle d'évaluation OSATS.	21
2.1	Résumé des conditions d'études considérées dans les travaux de la littérature ciblant la reconnaissance de gestes médico-chirurgicaux.	37
2.2	Résumé des processus de reconnaissance utilisés dans la littérature ciblant la reconnaissance de gestes chirurgicaux.	43
2.3	Résumé des processus de validation employés la littérature pour qualifier et quantifier la reconnaissance des gestes chirurgicaux.	47
3.1	Vocabulaire de surgemes utilisé pour décrire la tâche de transfert d'objet.	88
3.2	Statistiques sur les durées des trajectoires pour la tâche d'entraînement.	89
3.3	Rappel, précision et score F^1 pour chaque surgeme en utilisant la vérité terrain.	93
3.4	Rappel, précision et score F^1 pour chaque surgeme en utilisant notre approche de reconnaissance.	96



Introduction

Les travaux de cette thèse s'inscrivent dans le cadre de la chirurgie assistée par ordinateur et ont été financés par le Labex CAMI¹. Ils ont été réalisés en cotutelle entre l'équipe DEXTER du LIRMM, dirigée par Philippe Poignet et l'équipe MédiCIS du LTSI, dirigée par Pierre Jannin.

Les avancées médicales ainsi que le développement des systèmes informatisés ont donné naissance à de nouvelles pratiques chirurgicales connues sous le nom de Gestes Médico-Chirurgicaux Assistés par Ordinateur (GMCAO). Dans ce contexte, l'ordinateur permet d'aider le praticien durant l'intervention en lui fournissant un ensemble de services. Dans le domaine de la chirurgie laparoscopique, ces services peuvent être informatifs telle que la visualisation d'imageries préopératoires permettant d'aider à la prise de décision. Ces services peuvent également être exploités pour l'assistance durant la réalisation de gestes chirurgicaux tels que les robots de téléopération et de comanipulation. Dans ce cas, l'utilisation d'un robot modifie la pratique chirurgicale car l'ensemble des gestes est réalisé généralement au travers d'une console et reproduit par le robot pour interagir avec le patient, interférant directement avec le déroulement de l'intervention. Une formation spécifique est alors nécessaire pour que le praticien acquiert les compétences techniques nécessaires afin de garantir le bon déroulement de la chirurgie.

Dans le contexte de la chirurgie laparoscopique, il existe plusieurs types de dispositifs pour l'entraînement aux gestes chirurgicaux. La formation clinique passe principalement par l'utilisation d'un banc d'essai où l'apprenti répète un ensemble de gestes spécifiques à l'accomplissement d'une tâche chirurgicale. L'évaluation est réalisée sous la supervision d'un expert qui fournit un ensemble d'indications verbales afin d'améliorer la gestuelle. Cependant, la chirurgie assistée par ordinateur a également permis le développement de dispositifs spécifiques pour l'entraînement, appelés simulateurs. Ces derniers permettent

1. Computer Assisted Medical Interventions : <http://cami-labex.fr/>

de reproduire virtuellement les mouvements des instruments ainsi que des tissus afin de recréer les conditions opératoires, apportant un réalisme supplémentaire pour l'apprentissage et l'évaluation de la gestuelle. Dans ce cadre, l'évaluation est réalisée par un ensemble de métriques calculées par l'ordinateur et retransmises ensuite à l'apprenti. De ce fait, ce dernier doit analyser les résultats et en déduire les possibilités d'améliorations, sans approbation par un expert. De plus, ces dispositifs restent coûteux à l'achat et souvent difficiles d'accès à l'hôpital. Il peut alors en résulter un manque de pratique, couplé à une évaluation perfectible.

Dans ces travaux, nous souhaitons améliorer les conditions d'entraînement et d'apprentissage ainsi que l'accès aux simulateurs nécessitant généralement une console robotisée spécifique et coûteuse. Pour ce faire, nous proposons de nouvelles approches pour l'analyse, l'évaluation et la réalisation des gestes durant l'entraînement en chirurgie laparoscopique robotisée.

Ce manuscrit aborde dans un premier temps le contexte clinique de ces travaux avec l'introduction des systèmes robotisés pour la chirurgie, ainsi que la nécessité d'une formation adaptée pour la manipulation sous assistance robotisée (chapitre 1).

Afin de proposer des méthodes d'évaluation plus adaptées à l'apprentissage des gestes dans le cadre de la chirurgie laparoscopique robotisée, nous réalisons un état de l'art des techniques existantes dans le domaine de la recherche pour la détection et la reconnaissance des gestes chirurgicaux (chapitre 2).

À partir de cet état de l'art, nous présentons notre contribution qui traite de la segmentation non-supervisée et la reconnaissance des gestes chirurgicaux en utilisant comme données les trajectoires des instruments chirurgicaux (chapitre 3).

Enfin, souhaitant également améliorer l'accès à la formation et à l'entraînement aux gestes chirurgicaux, nous présentons l'étude d'une nouvelle interface sans contact pour la commande des robots chirurgicaux (chapitre 4).

Dans le cadre du Labex CAMI, ces travaux contribuent au développement de briques méthodologiques et matérielles afin d'améliorer l'évaluation des gestes, de faciliter l'accès à la formation et d'offrir une plateforme destinée à l'entraînement aux gestes en chirurgie laparoscopique robotisée.

Première partie

Analyse et compréhension du geste en chirurgie laparoscopique robotisée

Analyse du geste chirurgical : point de vue clinique

“La connaissance s’acquiert par l’expérience, tout le reste n’est que de l’information”.

ALBERT EINSTEIN

Préambule

Ce chapitre aborde l’aspect clinique de la thèse. Nous présentons tout d’abord les fondamentaux de la chirurgie laparoscopique et traitons de l’apport de la robotique pour cette pratique. Puis nous évoquons les problématiques associées à cette chirurgie et attirons l’attention sur la nécessité d’un entraînement adapté aux gestes chirurgicaux. Enfin, nous présentons les dispositifs et méthodes actuellement utilisés en clinique pour réaliser et évaluer l’entraînement chirurgical.

Sommaire

1.1	La chirurgie laparoscopique	6
1.1.1	Pratique chirurgicale	6
1.1.2	La chirurgie assistée par robot	8
1.1.3	Les problématiques persistantes	10
1.2	L’entraînement chirurgical	12
1.2.1	L’apprentissage des compétences	12
1.2.1.1	Les compétences non-techniques	12
1.2.1.2	Les compétences techniques	13

1.2.2	Les dispositifs d'entraînement aux gestes chirurgicaux	14
1.2.2.1	Les différentes techniques de simulation	14
1.2.2.2	Les simulateurs organiques	15
1.2.2.3	Les simulateurs non-organiques	16
1.2.2.4	Aperçu des simulateurs non-organiques existants	17
1.2.3	Les méthodes d'évaluation du geste chirurgical	19
1.2.3.1	Méthodes qualitatives basées sur un observateur externe	20
1.2.3.2	Méthodes quantitatives automatiques	21
1.3	Analyse et discussion	23

1.1 La chirurgie laparoscopique

La chirurgie laparoscopique, également appelée chirurgie coelioscopique, est une pratique très ancienne qui date de plus d'un siècle. Les fondamentaux de cette chirurgie sont apparus dans les années 1850 avec l'urologue français Antonin Jean Desormeaux. Cependant, c'est uniquement en 1987 avec l'apparition des nouvelles technologies pour l'affichage du signal vidéo sur un écran déporté que le professeur Philippe Mouret réalisa la première cholécystectomie par voie laparoscopique [Spaner 1997]. Mais depuis ces années, la pratique de cette méthode a bien évolué, laissant place à de nouvelles perspectives et des voies d'accès de plus en plus étroites pour réduire le traumatisme subi par le patient.

1.1.1 Pratique chirurgicale

La laparoscopie est née d'une pratique ancienne appelée laparotomie. Cette dernière, également dénommée "chirurgie à ciel ouvert" consiste à réaliser une large incision de la paroi abdominale, le plus souvent partant du pubis jusqu'au sternum, afin d'avoir un accès direct aux organes abdominaux et pelviens. Néanmoins, cette pratique est devenue trop intrusive et a laissé place aux méthodes dites minimalement invasives, et notamment à la laparoscopie. Cette dernière consiste à aborder le site de l'intervention au travers de petites incisions dans la paroi abdominale. Des trocarts sont ensuite placés au niveau des incisions, par lesquels sont insérés de longs instruments à commande déportée afin d'exécuter le geste sous contrôle vidéo, comme présenté sur la figure 1.1. Cette méthode est aujourd'hui standardisée et utilisée en pratique pour les chirurgies viscérales, gynécologiques et urologiques.

Cette approche présente de nombreux bénéfices majoritairement en faveur du patient, dont la réduction du traumatisme pariétal et des saignements, la réduction des douleurs post-opératoires, la réduction de la durée d'hospitalisation ainsi que l'amélioration du résultat cosmétique. Néanmoins, elle implique également de nombreuses contraintes qui dégradent les conditions interventionnelles du chirurgien [Agha 2003] :

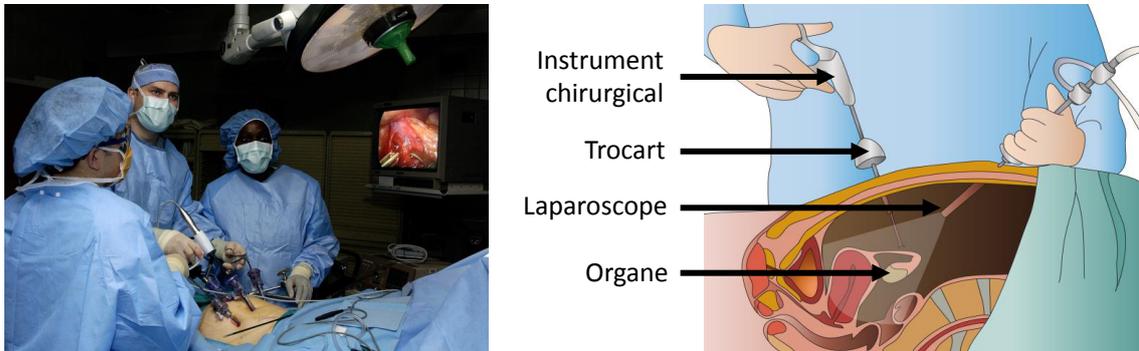


FIGURE 1.1: Illustration du principe de laparoscopie.

- La coordination main-œil est fortement altérée par le transfert de la scène observée sur moniteur déporté, ainsi que par l'inversion des mouvements entre la main et l'extrémité de l'outil. Ce phénomène, appelé point d'appui, est dû à la position du centre instantané de rotation qui se situe à l'intersection de la paroi abdominale et de l'axe du trocart [Breedveld 2001] ;
- L'utilisation d'une optique unique diminue la perception de la scène observée. Le champ opératoire en trois dimensions doit être reconstruit mentalement par le chirurgien à partir des images 2D provenant de la caméra [Breedveld 1999] ;
- Le retour haptique est détérioré par l'insertion de l'instrument à travers le trocart, produisant des frottements non négligeables. De plus, l'effet bras de levier généré par la conception des instruments laparoscopiques amplifie les efforts perçus par le chirurgien [Bholat 1999] ;
- L'ergonomie est largement réduite de par la posture adoptée par le chirurgien lors de l'intervention, impliquant des douleurs au niveau des membres supérieurs ainsi qu'au niveau du dos [Santos-Carreras 2012a]. Ces douleurs peuvent accélérer l'arrivée d'une fatigue physique et mentale.

Toutes ces complications ne sont pas sans conséquences puisqu'elles augmentent la charge cognitive du chirurgien, influant directement sur ses capacités motrices et sensorielles. Il en résulte une augmentation du temps d'apprentissage [Secin 2010], allongeant la durée de formation des chirurgiens. De plus, le modèle de compagnonnage "see one, do one, teach one" conceptualisé par Halsted [Tsuda 2009], qui veut que l'apprenti se forme aux côtés de son maître au bloc opératoire, en prenant peu à peu des responsabilités dans la réalisation des gestes, se retrouve limité à cause des problématiques induites par la chirurgie laparoscopique. Dans cette configuration, l'inversion des rôles entre le chirurgien qui tient les instruments et l'apprenti qui manipule l'endoscope est difficile. La gestion de l'endoscope permet de se familiariser avec la représentation mentale de l'espace de travail

en 2D, et permet également d'acquérir les notions théoriques concernant le déroulé de la procédure, mais n'autorise en aucun cas l'interne à s'entraîner à la réalisation de gestes techniques pourtant différents de ceux utilisés en laparotomie. Cependant, la robotique a maintenant conduit à faire évoluer les techniques interventionnelles, afin d'améliorer les conditions de travail des chirurgiens ainsi que la formation des apprentis [Perez 2012].

1.1.2 La chirurgie assistée par robot

À l'intersection des évolutions technologiques de la fin du XXe siècle, telles que le développement des systèmes d'imagerie et l'utilisation croissante des ordinateurs dans le domaine médical, et des contraintes médicales fortes où le bénéfice du patient devient une priorité, est née la chirurgie robotisée. Apparue dès 1985 dans le domaine de la neurochirurgie pour les biopsies des tumeurs cérébrales [Kwoh 1988], la robotique chirurgicale a depuis largement évolué touchant les domaines de la neurochirurgie, chirurgie otorhinolaryngologique, chirurgie orthopédique, radiologie interventionnelle ainsi que la chirurgie laparoscopique. Basés initialement sur des structures mécaniques anthropomorphiques permettant de reproduire les mouvements du membre supérieur, les robots ont ensuite évolué vers des structures destinées aux besoins médicaux. Dans la suite de ce manuscrit, nous concentrons notre analyse sur les dispositifs conçus pour l'assistance aux gestes en chirurgie laparoscopique. Pour plus d'information concernant d'autres domaines médicaux, un récent ouvrage retrace l'histoire de la robotique médicale ainsi que les différentes applications cliniques existantes [Troccaz 2012].

L'apport du robot en chirurgie laparoscopique a débuté dans les années 1990 avec l'apparition des systèmes téléopérés et du robot AESOP, par la société Computer Motion [Sackier 1994]. Dans les années 2000 est apparue la première version du robot da Vinci [Guthart 2000]. Initialement créée pour répondre aux besoins du DARPA (Defense Advanced Research Projects Agency) afin d'opérer les soldats blessés sur les champs de bataille et de réduire les coûts matériels et humains, la société Intuitive Surgical a développé le robot da Vinci. Certifié par la FDA (Food and Drug Administration) aux États-Unis, et ayant obtenu le marquage CE en Europe, ce robot est actuellement l'unique robot de téléopération commercialisé pour une utilisation au bloc opératoire, dans le cadre de chirurgies laparoscopiques. Ce système permet, grâce à une interface homme-machine (voir figure 1.2a) de piloter un robot distant équipé de plusieurs instruments chirurgicaux et d'une caméra (figure 1.2b). Son architecture de téléopération, adaptée aux blocs opératoires, permet d'améliorer l'espace et l'organisation du bloc tout en résolvant plusieurs problématiques liées à la laparoscopie. L'inversion des mouvements est gérée directement par le robot, tout comme la profondeur dans la scène qui est capturée par un système d'optique stéréoscopique et retransmise via une interface visuelle 3D. L'ergonomie du chirurgien se trouve largement améliorée grâce à une pose assise au niveau de la console : la tête est maintenue par un support sur la partie haute de la console, le poids du corps est réparti entre les jambes et les fessiers et un support central permet de soulager les bras et avant-bras



(a) L'interface homme-machine du da Vinci.

(b) Le robot da Vinci et ses instruments.



(c) Le robot da Vinci avec deux consoles au bloc opératoire.

FIGURE 1.2: Le robot da Vinci commercialisé par la société Intuitive Surgical.
Source: Intuitive Surgical.

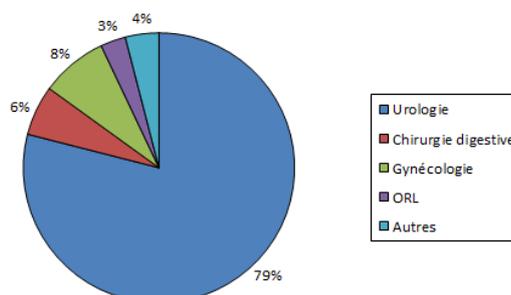


FIGURE 1.3: Répartition des interventions réalisées avec le robot da Vinci depuis 2011, en France.

Source: Rapport de la Haute Autorité de Santé [HAS 2014].

lors de la manipulation, réduisant les risques d'apparition de fatigue physique. Dotée de 7 degrés de liberté, l'interface mécanique du robot da Vinci permet de pleinement contrôler les instruments chirurgicaux ainsi que l'endoscope, et filtre également les mouvements parasites afin d'améliorer l'exécution des gestes et les conditions opératoires du chirurgien [Van Der Schatte Olivier 2009]. De plus, le robot peut être couplé à deux consoles simultanément pour partager la vision ainsi que la commande des instruments dans le cas de procédures complexes, favorisant le compagnonnage et la formation des apprentis (voir figure 1.2c).

Initialement destiné à assister le chirurgien en cardiologie, les capacités du robot da Vinci répondent finalement en plus grande partie aux besoins des chirurgiens laparoscopiques (figure 1.3), notamment grâce aux évolutions apportées sur le robot en termes de conception et fonctionnalités [Freschi 2013]. Cependant, certains aspects notamment liés au coût économique du dispositif, aux bénéfices apportés et à la formation des chirurgiens restent encore discutables.

1.1.3 Les problématiques persistantes

Bien que la robotique ait aujourd'hui apporté des systèmes perfectionnés répondant aux besoins des chirurgiens laparoscopiques, en termes d'ergonomie et de manipulation, le bénéfice médical reste encore discutable [Lepor 2009, Nix 2010] et certaines problématiques restent en suspens. Une de ces problématiques est le coût économique d'un tel dispositif, non négligeable pour un hôpital ou une clinique. Steinberg et al. [Steinberg 2008] ont déclaré que l'achat d'un robot da Vinci s'élève à plus d'un million et demi de dollars (US \$), ajouté à cela un budget d'entretien d'environ 10% du prix d'achat pour la maintenance annuelle (hors année d'achat). L'étude rapporte également qu'un surcoût de plus de 3000 dollars par intervention est nécessaire, dans le cadre de l'utilisation du robot et de l'achat des instruments chirurgicaux, à durée de vie limitée (environ 10 interventions). Le

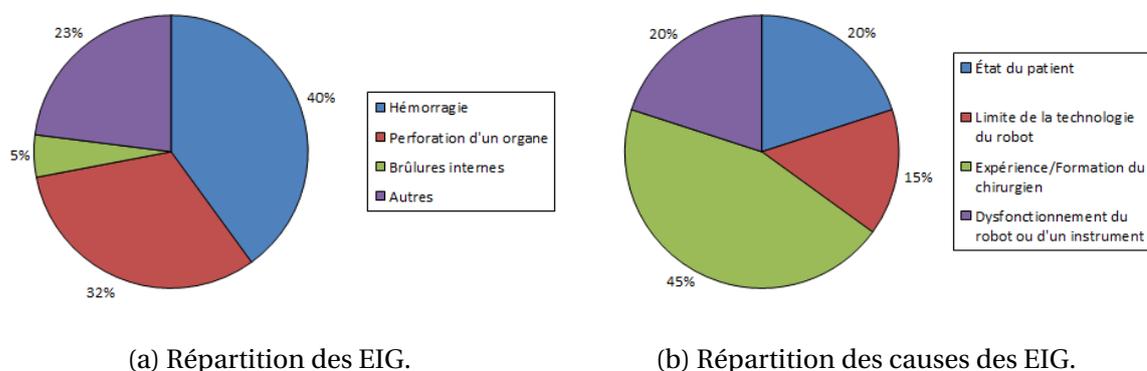


FIGURE 1.4: Effets Indésirables Graves intervenants en chirurgie robotisée.

Source: Rapport de la Haute Autorité de Santé [HAS 2014].

ratio coût économique/bénéfice médical devient alors largement discutable [Altieri 2015]. D'autre part, la principale lacune technologique du robot da Vinci est que ce dernier ne permet pas de retour d'efforts. Ce qui signifie que le chirurgien est dans l'incapacité de ressentir les efforts exercés par et sur les instruments. Ainsi, une pression excessive sur un tissu ou un organe ne peut être détectée que par la vision de ces déformations.

D'après deux récents rapports de la Haute Autorité de Santé [HAS 2014, HAS 2015], plus de 3000 robots à travers le monde, dont 80 répartis dans 69 établissements en France, ont permis de réaliser plus de 450000 chirurgies en 2012. Cependant d'après ces enquêtes menées dans les établissements français, plus de 30 EIG (Effets Indésirables Graves¹) sont apparus dans le cadre de 17000 interventions réalisées depuis l'acquisition du robot da Vinci, soit plus de 0,0017%. Il est néanmoins à signaler que ces anomalies ne sont pas toujours rapportées, ou parfois mal déclarées [Cooper 2013]. Dans l'étude de la Haute Autorité de Santé, les anomalies détectées sont réparties par nature (voir figure 1.4a) ainsi que par cause (figure 1.4b).

Ces graphiques montrent que 15% des événements indésirables intervenus durant une chirurgie robotisée sont dus à des limites de la technologie, et notamment à l'absence de retour d'effort pour la manipulation, mais également que 20% de ces événements sont apparus suite à un dysfonctionnement du robot ou d'un instrument robotisé. Toutefois, contre toute attente, cette enquête révèle que l'expérience et la formation du chirurgien est mise en cause dans 45% des incidents, le geste chirurgical se trouvant être la source du problème et causant également un décès durant l'enquête.

1. L'EIG est défini d'après l'article R.1123-39 du Code de la Santé Publique comme "l'évènement ou effet indésirable grave qui entraîne la mort, met en danger la vie de la personne, nécessite une hospitalisation ou prolongation d'hospitalisation, provoque une incapacité ou un handicap important ou durable, ou bien se traduit par une anomalie ou malformation congénitale".

Cette étude met en avant le développement croissant des systèmes robotisés destinés à la chirurgie laparoscopiques. L'assistance robotisée apporte un avantage indéniable en terme pratique pour le chirurgien. Cependant, cette analyse souligne également la nécessité de la formation et de la courbe d'apprentissage des opérateurs, dans ce contexte de forte expansion de l'utilisation du robot dans des domaines de chirurgie de plus en plus variés. Elle met également en avant le fait que les établissements de santé réalisant le plus grand nombre d'interventions robotisées ont déclaré le moins d'EIG, montrant l'influence de l'expérience du clinicien dans la pratique. Enfin, l'étude conclut sur l'importance que chaque opérateur puisse bénéficier d'une formation complète et adaptée avant la réalisation d'interventions chirurgicales robotisées, afin d'acquérir les compétences techniques nécessaires. Chaque utilisateur doit également pouvoir bénéficier d'un suivi régulier dans le cadre de la formation continue pour l'usage du robot. L'entraînement aux gestes médicaux devient alors un passage obligatoire et renforcé dans la formation chirurgicale, aussi bien en laparoscopie standard qu'avec une assistance robotisée.

1.2 L'entraînement chirurgical

L'entraînement en chirurgie laparoscopique passe par plusieurs étapes d'apprentissage telles que l'acquisition des connaissances anatomiques ou encore l'acquisition des connaissances liées au processus chirurgical. L'entraînement fait également appel à des facteurs et des facultés personnelles comme la prise de décision ou la communication, pouvant directement influencer sur le déroulement de l'intervention. Certaines recherches se concentrent alors sur l'analyse de ces compétences, ayant ainsi pour objectifs d'améliorer les conditions d'entraînement des chirurgiens mais également de comprendre les facteurs humains et sociaux qui interviennent durant l'apprentissage et la réalisation d'interventions chirurgicales.

1.2.1 L'apprentissage des compétences

L'acquisition des compétences médicales et chirurgicales est décomposée en plusieurs parties et structurée par le cursus médical, partant des compétences théoriques durant les premières années d'étude, jusqu'à la pratique vers la fin de la formation. Cependant, du point de vue de la recherche, la décomposition des compétences chirurgicales se fait plus directement en deux grandes catégories : les compétences dites "non-techniques" et les compétences dites "techniques" (voir tableau 1.1).

1.2.1.1 Les compétences non-techniques

Le terme de "compétences non-techniques" a été largement répandu et utilisé dans les années 1990 pour le domaine de l'aviation afin de qualifier les facteurs transversaux

Compétences chirurgicales	Non-techniques	Connaissance de la situation
		Prise de décision
		Leadership
		Travail en équipe
		Communication
		Gestion du stress et fatigue
	Techniques	Dextérité
		Connaissances conceptuelles
		Connaissances procédurales

Tableau 1.1: Description et répartition des compétences nécessaires en chirurgie.

ayant un impact sur les compétences des pilotes, hors prise en compte de la technique de pilotage. Des compagnies aériennes ont alors mis en place des programmes pour la gestion des "ressources humaines" (intitulés "Crew Resource Management") ainsi que la formation des "facteurs humains" pouvant intervenir au sein de l'équipage, et notamment à bord du cockpit. Mais plus largement, ces programmes ont ensuite été élargis à d'autres domaines afin de former et d'évaluer le personnel occupant des professions et postes à risques ou critiques, comme c'est le cas dans le domaine du médical, et plus particulièrement de la chirurgie [Reader 2006]. Les compétences non-techniques ont alors été définies comme les "compétences complémentaires aux compétences techniques faisant intervenir des facteurs cognitifs, sociaux et interpersonnels permettant de contribuer à des performances plus efficaces et sûrs" [Flin 2013]. En analysant la littérature récente afin de prendre en compte l'évolution des processus chirurgicaux qui ont eu lieu ces dernières années, la classification de ces facteurs présentée par Mitchell et al. [Mitchell 2013] est celle majoritairement retenue dans les travaux et également la plus adaptée au domaine médical et plus particulièrement à la description des compétences non-techniques intervenant au cours du processus chirurgical. Cette classification reprend également les termes décrits par Yule et al. [Yule 2006a, Yule 2006b] (voir tableau 1.1).

1.2.1.2 Les compétences techniques

Les compétences techniques sont, pour leur part, plus largement étudiées de par leur impact direct et évident sur le déroulement de la chirurgie [Reznick 1993, Hamdorf 2000]. Ces compétences font intervenir des facteurs définis comme techniques, impliquant d'avoir reçu au préalable une formation spécifique, ainsi que pratique, pour l'acquisition des connaissances associées. D'après la littérature, il existe trois catégories de compétences techniques (voir tableau 1.1) : la dextérité [Moorthy 2003], les compétences conceptuelles et les compétences procédurales [McCormick 1997]. Les compétences

conceptuelles et procédurales sont décrites par McCormick comme les compétences se référant aux connaissances acquises durant la formation médicale, ainsi qu'aux connaissances dites "stratégiques" qui se ramènent à de la résolution de problème, à de la planification, à de l'optimisation et à de la modélisation des concepts et processus. D'un autre côté, la dextérité implique quant à elle des facteurs dits "manuels" comme la gestuelle, la précision, l'habileté physique et le ressenti ainsi que la gestion des efforts. Ces facteurs, plus concrets, font l'objet de recherches avancées pour la formation des apprentis et l'évaluation de leurs compétences techniques [Datta 2002, Ahmed 2011, Reiley 2011].

Les compétences techniques, ainsi que la formation des praticiens en chirurgie robotisée, sont les facteurs majoritairement remis en cause par les études de la Haute Autorité de Santé présentées précédemment. Parmi ces compétences techniques, une part non négligeable, à hauteur de 25%, traite de la dextérité lors de la manipulation [Barnes 1987]. Dans ces travaux, nous souhaitons contribuer à la résolution de cette problématique afin de faire bénéficier aux apprentis, mais également aux chirurgiens, de dispositifs d'entraînement évolués. Ces dispositifs peuvent permettre d'apprendre mais aussi d'entretenir la dextérité, notamment lors de l'utilisation d'une assistance robotisée. Dans cet objectif, nous focalisons ces travaux de recherche sur le développement de méthodes et systèmes permettant l'apprentissage et l'entraînement aux gestes dans le cadre de la chirurgie robotisée.

1.2.2 Les dispositifs d'entraînement aux gestes chirurgicaux

Actuellement, l'entraînement pour l'acquisition des compétences nécessaires à la réalisation des gestes chirurgicaux, aussi bien en laparoscopie manuelle qu'avec assistance robotisée, se fait au travers de la simulation. Béguin et al. [Béguin 1997] ont défini la simulation comme "une méthode d'enseignement de savoir-faire et d'habiletés utilisés dans des tâches pour lesquelles un enseignement direct s'avère impossible pour des raisons déontologiques (sécurité et sûreté), économiques (coût du matériel) ou techniques (très faible probabilité d'occurrence des incidents ou accidents)". La simulation dépend principalement des techniques employées et également du dispositif mis en jeu au cours de l'exercice.

1.2.2.1 Les différentes techniques de simulation

Il existe différentes techniques de simulation utilisées dans les écoles et les établissements destinés à la formation médicale, et plus particulièrement chirurgicale. Un précédent rapport de la Haute Autorité de Santé [HAS 2012] a mis en avant cinq techniques de simulation :

- Haute-fidélité avec matériel (exemple : Mannequin piloté par informatique),
- Haute-fidélité avec environnement réaliste (exemple : Situation de consultation simulée),

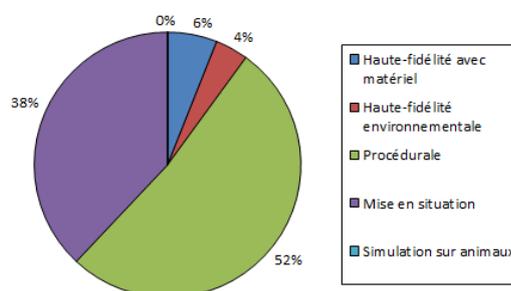


FIGURE 1.5: Les différentes techniques de simulation employées en France.
 Source: Rapport de la Haute Autorité de Santé [HAS 2012].

- Simulation procédurale (exemple : Apprentissage de technique : pose de perfusion, intubation, etc.),
- Mise en situation (exemple : Jeux de rôles, études de cas),
- Expérimentation animale (exemple : Exercice de suture sur animal).

Dans ce rapport, un bilan détaillé des techniques de simulation employées dans les établissements et écoles médicales en France a été établi. En analysant les résultats de cette enquête, la simulation dite "procédurale" (l'apprentissage ciblé de techniques) et la "mise en situation" (études de cas ou jeux de rôles avec peu d'investissement matériel) sont les méthodes les plus régulièrement employées, comme le montre la figure 1.5, là où la simulation sur animaux n'est que faiblement employée de par le caractère éthique de ce type d'exercice. Ces différentes techniques sont adaptables suivant la nature de la simulation et le type de simulateur employé.

Dans le cadre de l'enseignement des compétences, aussi bien techniques que non-techniques, la simulation peut se faire au travers de différents moyens, comme illustré sur la figure 1.6. On constate deux grandes familles de simulateurs : les simulateurs organiques et les simulateurs non-organiques.

1.2.2.2 Les simulateurs organiques

Les simulateurs dits organiques, ou *in-vivo*, concernent le domaine de la simulation procédurale faite sur cadavre humain ou animal. La simulation humaine du vivant fait quant à elle référence au concept de patients standardisés. L'avantage de ces types de simulateurs est la capacité d'immersion en conditions opératoires réelles, et l'utilisation de techniques et dispositifs proches ou identiques à ceux utilisés sur patient. Néanmoins, les entraînements sur animaux ou cadavres ont un coût économique non négligeable, et demandent une organisation complexe pour la mise en place d'entraînements répétés [van Velthoven 2006]. De plus, le nombre de répétitions est limité par la détérioration des

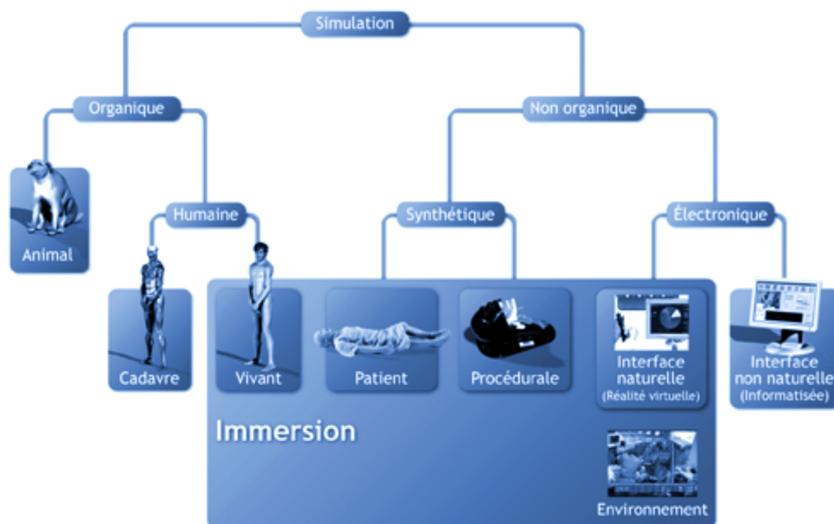


FIGURE 1.6: Les différents moyens de simulation pour le médical.
 Source: Inspiré par le Dr Amitai Ziv et adoptée par l'Université de PennState
 [Chiniara 2007]

tissus et organes manipulés. Ces types de simulateur ne sont que peu accessibles, par principes éthiques et moraux, et ne servent que principalement à la validation de dispositifs et procédés médicaux afin de reproduire les conditions et contraintes opératoires réelles.

1.2.2.3 Les simulateurs non-organiques

L'autre grande famille des simulateurs est celle des simulateurs non organiques. Ce groupe peut se diviser en deux sous-catégories que sont les simulateurs "électroniques" et les simulateurs "synthétiques", selon que les informations générées par l'ordinateur y jouent un rôle prépondérant ou non. Dans le cas de la simulation synthétique, on retrouve les mannequins à haute-fidélité capables de reproduire plusieurs caractéristiques humaines, ainsi que la simulation procédurale permettant de reproduire uniquement certaines techniques spécifiques, ou de réaliser des scénarios médicaux. De l'autre côté, les simulateurs électroniques se décomposent en deux groupes dits "interface naturelle" et "interface non-naturelle", selon leur degré d'immersion et de reproduction et d'immersion en environnement réel. Les principaux avantages des simulateurs non-organiques sont notamment leur coût, ainsi que leur accès plus largement ouvert pour l'entraînement à la pratique médicale et chirurgicale [Tsuda 2009]. Ils permettent en outre, grâce au développement de la réalité virtuelle ainsi que de la réalité augmentée, de reproduire des conditions opératoires proches du réel, tout en garantissant une reproductibilité sans limite des gestes [Botden 2007], sans détérioration des organes ou tissus simulés, et montrent

néanmoins une forte corrélation avec l'évaluation réelle sur patient [Matsuda 2012].

Pour les raisons évidentes présentées ci-dessus, les travaux de ce manuscrit vont se concentrer sur l'entraînement aux gestes chirurgicaux, et l'application en simulation non-organique qui permet notamment d'améliorer l'apprentissage et l'efficacité des techniques nécessaires en chirurgie laparoscopique [Aggarwal 2007]. Visant cet objectif, nous souhaitons présenter plus en détail quelques simulateurs non-organiques, et plus particulièrement électroniques et procéduraux, actuellement utilisés pour l'entraînement aux gestes chirurgicaux.

1.2.2.4 Aperçu des simulateurs non-organiques existants

Une liste aussi exhaustive que possible des différents simulateurs non-organiques existant actuellement sur le marché est présentée dans le tableau 1.2. Ces simulateurs sont classés suivant différents critères que sont : le type d'entraînement permis par le système, le support employé pour la simulation, la capacité de retour haptique du dispositif, les types de tâches chirurgicales autorisées et le type de métrique employé pour l'évaluation. Il est à noter que la très grande majorité des dispositifs présentés dans ce tableau ne permet pas l'accès aux informations quant à la nature des métriques d'évaluation, telles que la vitesse ou le nombre de collisions. De plus, les prix de ces systèmes ne sont pas directement accessibles par le grand public.

Nous constatons que ces simulateurs sont répartis en deux grandes catégories : les simulateurs avec un support de type "physique" (voir figure 1.7a) où le dispositif permet l'entraînement dans un environnement réel (fil, aiguille, pièces en plastique, tissus déformables), et les simulateurs avec un support de type "réalité virtuelle" où le dispositif permet, grâce à un logiciel, l'entraînement dans un environnement simulé proche de la réalité sur patient (voir figures 1.7b et 1.7c). Les simulateurs de réalité virtuelle deviennent prépondérants grâce à leur facilité d'utilisation et leur simplicité d'évolution, avec la mise à jour du système et des modules d'activité, mais la gestion du retour haptique devient plus complexe obligeant à motoriser la plateforme afin de reproduire les efforts d'interaction avec l'environnement. Apportant finalement un meilleur réalisme, ils sont toutefois plus onéreux qu'un simulateur physique, le prix étant justifié par le développement logiciel ainsi que l'architecture mécanique et électronique utilisée. Nous constatons que parmi les simulateurs de réalité virtuelle certains sont spécifiquement conçus pour l'apprentissage des techniques dans le cadre de chirurgies robotisées. Ces derniers étant plus complets en termes de matériel, ils sont néanmoins les plus onéreux sur le marché du fait de la technologie employée pour reproduire les mouvements du praticien, réalisés au travers de l'interface homme-machine. Soulignons enfin que la plupart de ces dispositifs permettent de réaliser des tâches d'entraînement dites "basiques" afin d'acquérir les compétences minimales en termes de dextérité (mouvements des instruments et de la caméra, entraînement aux sutures, etc.). Néanmoins, une faible majorité d'entre eux permettent de s'en-

Nom	Société / Institution	Type d'entraînement	Support	Retour haptique	Tâches		Méthodes d'évaluation	Référence
					Basiques	Procédures		
LAP Mentor III	Simbionix	Manuel	Réalité virtuelle	Oui	Oui	Oui	Automatique	[LapMentor 2015]
LAP Mentor Express	Simbionix	Manuel	Réalité virtuelle	Non	Oui	Oui	Automatique	[LapMentor 2015]
LapSim Haptic System	Surgical Science	Manuel	Réalité virtuelle	Oui	Oui	Oui	Automatique	[LapSim 2015]
LapSim Non-Haptic System	Surgical Science	Manuel	Réalité virtuelle	Non	Oui	Oui	Automatique	[LapSim 2015]
Simendo Pro2	Simendo	Manuel	Réalité virtuelle	Non	Oui	Non	Automatique	[SimPro 2015]
SEP Basic	SimSurgery	Manuel	Réalité virtuelle	Non	Oui	Non	Automatique	[SEBasic 2015]
LapVR	CAE Healthcare	Manuel	Réalité virtuelle	Non	Oui	Oui	Automatique	[LapVR 2015]
MIST VR	Manchester Royal Infirmary	Manuel	Réalité virtuelle	Non	Oui	Non	Automatique	[Wilson 1997]
Xperience Team Trainer	Mimic Simulation	Manuel	Réalité virtuelle	Non	Oui	Oui	Automatique	[Xperience 2015]
FLS Trainer	Limbs & Things	Manuel	Physique	-	Oui	Non	Manuelle	[FLSTrainer 2015]
LaproTrain	Endosim	Manuel	Physique	Oui	Oui	Non	Manuelle	[LaproTrain 2015]
PoP Trainer	Optimist	Manuel	Physique	Oui	Oui	Oui	Manuelle	[PoPTrainer 2015]
HUESAD	Hiroshima University	Manuel	Physique	Oui	Oui	Non	Automatique	[Egi 2008]
Surgical Workstation	Israël Center for Med. Sim.	Manuel	Physique	Oui	Oui	-	-	[SurgWork 2015]
oeSim	eoSurgical	Manuel	Physique	Oui	Oui	Non	Manuelle	[OeSim 2015]
Red Dragon	Washington University	Manuel	Physique	Oui	Oui	Non	Automatique	[RedDragon 2015]
Blue Dragon	Washington University	Manuel	Physique	Oui	Oui	Non	Automatique	[BlueDragon 2015]
ADEPT	University of Dundee	Manuel	Physique	Oui	Oui	Non	Automatique	[Francis 2001]
T-serie	3-Dmed	Manuel	Physique	-	Oui	Non	Manuelle	[Tserie 2015]
MLabs	Mimic Simulation	Manuel	Physique	-	Oui	Non	Manuelle	[MLabs 2015]
ICSAD	Imperial College London	Manuel	Physique	Oui	Non	Non	Automatique	[Hayter 2009]
Robotix Mentor	Simbionix	Robotisé	Réalité virtuelle	Non	Oui	Oui	Automatique	[RobMentor 2015]
Maestro AR	Mimic Simulation	Robotisé	Réalité virtuelle	Non	Non	Oui	Automatique	[MaestroAR 2015]
SEP-Robot	SimSurgery	Robotisé	Réalité virtuelle	Non	Oui	Non	Automatique	[SEPRobot 2015]
RoSS II	Simulated Surgical Systems	Robotisé	Réalité virtuelle	Non	Oui	Non	Automatique	[RoSSII 2015]
dv-Trainee	Mimic Simulation	Robotisé	Réalité virtuelle	Oui	Oui	Oui	Automatique	[DVTrainer 2015]
da Vinci Skills Simulator	Intuitive Surgical	Robotisé	Réalité virtuelle	-	Oui	Oui	Automatique	[DVSS 2015]
Xron	BBZ Medical Technologies	Robotisé	Réalité virtuelle	-	Oui	Oui	Automatique	[Xron 2015]

Tableau 1.2: Simulateurs destinés à l'apprentissage de la chirurgie laparoscopique manuelle ou robotisée.



(a) PoP Trainer.
Source: Optimist.

(b) LAP Mentor III.
Source: Simbionix.

(c) Robotix Mentor.
Source: Simbionix.

FIGURE 1.7: Les différents types de simulateurs laparoscopiques.

traîner sur des procédures simulées, permettant alors d'étendre les notions basiques dans un contexte chirurgical plus réaliste.

Ce tableau met également en lumière le développement évident de techniques destinées à l'évaluation de la dextérité. Alors que peu de dispositifs permettent une évaluation manuelle basée sur un observateur externe, impliquant la présence d'un expert pour évaluer et qualifier les performances de l'opérateur, la plupart des dispositifs comportent des méthodes avancées pour l'analyse et la quantification du geste chirurgical basées sur différentes métriques calculées de manière automatique.

1.2.3 Les méthodes d'évaluation du geste chirurgical

On distingue actuellement deux grandes méthodes d'évaluation pour qualifier et quantifier la qualité du geste chirurgical :

- Les méthodes dites "qualitatives", basées sur un observateur externe et régulièrement utilisées en clinique. Des études complètes de leur fiabilité, validité, faisabilité et reproductibilité ont été réalisées [Ahmed 2011, Wolf 2013], permettant de juger de la qualité du système de notation au regard de différents critères. Bien que ces méthodes de notation aient été initialement conçues pour l'évaluation de compétences au bloc opératoire, leurs utilisations dans des conditions expérimentales ont également été validées au travers différentes analyses ;
- Les méthodes automatiques dites "quantitatives", moins répandues en clinique mais de plus en plus utilisées par les simulateurs, où les métriques et grandeurs physiques employées pour quantifier la qualité du geste montrent une forte corrélation avec l'expertise de l'opérateur.

	1	2	3	4	5
Depth perception	Constantly overshoots target, wide swings, slow to correct		Some overshooting or missing of target, but quick to correct		Accurately directs instruments in the correct plane to target
Bimanual dexterity	Uses only one hand, ignores nondominant hand, poor coordination between hands		Uses both hands, but does not optimize interaction between hands		Expertly uses both hands in a complimentary manner to provide optimal exposure
Efficiency	Uncertain, inefficient efforts; many tentative movements; constantly changing focus or persisting without progress		Slow, but planned movements are reasonably organized		Confident, efficient and safe conduct, maintains focus on task until it is better performed by way of an alternative approach
Tissue handling	Rough movements, tears tissue, injures adjacent structures, poor grasper control, grasper frequently slips		Handles tissues reasonably well, minor trauma to adjacent tissue (ie, occasional unnecessary bleeding or slipping of the grasper)		Handles tissues well, applies appropriate traction, negligible injury to adjacent structures
Autonomy	Unable to complete entire task, even with verbal guidance		Able to complete task safely with moderate guidance		Able to complete task independently without prompting

Tableau 1.3: Échelle d'évaluation GOALS. La notation est réalisée ici par l'attribution d'un score de 1 à 5 sur l'échelle de Linkert.

Source: *Extrait de Vassiliou et al. [Vassiliou 2005].*

1.2.3.1 Méthodes qualitatives basées sur un observateur externe

Dans le cadre de l'évaluation qualitative, l'apprenti qui réalise l'entraînement chirurgical est supervisé par un chirurgien expert. L'expert cherche à analyser le geste aux moyens de méthodes qualitatives réparties en trois groupes : les échelles de notation générique, les échelles de notation spécifique, liée à une procédure, et enfin la combinaison d'outils génériques et spécifiques [Ahmed 2011].

Les échelles de notation générique permettent de structurer l'évaluation suivant une grille de critères génériques prédéterminés. L'expert cherche à qualifier la qualité des gestes réalisés en attribuant un score pour chaque critère demandé. Pour ces outils nous pouvons notamment citer les méthodes GRS (Global Rating Scale) et GOALS (Global Operative Assessment of Laparoscopic Skills) [Vassiliou 2005] ainsi que leurs variantes. Un exemple de grille d'évaluation du score GOALS est présenté dans le tableau 1.3.

Les échelles dites de notation spécifique permettent de fournir plus de précision sur la qualité du geste, dans un contexte de procédure défini. Le système de notation décompose (de manière manuelle) la procédure en un ensemble de tâches prédéterminées et dépendantes de la chirurgie ciblée. Dans ce cadre, nous pouvons citer les méthodes OSATS (Objective Structured Assessment of Technical Skills) [Martin 1997] et le score OCHRA (Observational Clinical Human Reliability Analysis) [Tang 2006], ainsi que leurs variantes. Un exemple d'évaluation utilisant le score OSATS est présenté dans le tableau 1.4.

Enfin, il existe des outils permettant de combiner les avantages des échelles de notation spécifique liées à des interventions chirurgicales et des analyses plus génériques. Ces outils

ITEM	Not Done/ Done Incorrectly	Done Correctly
<u>CONTROL OF HEMORRHAGE</u>		
1. Applies pressure to stop bleeding <u>first</u>	0	1
2. Asks assistant to suction field	0	1
3. Inspects injury by carefully releasing the IVC	0	1
4. Ensures all equipment needed for repair is at hand before starting	0	1
5. Control of bleeding point (use deBakey forceps /Satinsky clamp or prox/distal pressure)	0	1
<u>REPAIR</u>		
6. Select appropriate suture (4.0/5.0/6.0 polypropylene)	0	1
7. Select appropriate needle driver (vascular)	0	1
8. Select appropriate forceps (deBakey)	0	1
9. Needle loaded 1/2-2/3 from tip 90% of time	0	1

Tableau 1.4: Extrait de la liste de vérification pour l'hémostate d'une plaie de veine cave inférieure (IVC) avec l'échelle d'évaluation OSATS.

Source: Extrait de Martin et al. [Martin 1997].

peuvent être mélangés pour fournir un score global, ou rester indépendants afin d'analyser séparément les performances des apprentis. Pour ce genre de méthodes d'évaluation, citons notamment le ICLEPS (Imperial College London Evaluation of Procedural Skills) [Pandey 2006] et le STSAF (Structured Technical Skills Assessment Form) [Chipman 2009]. De plus, Naik et al. [Naik 2007] ont prouvé que ces types d'outils fournissent des méthodes d'évaluation efficaces pour le transfert de compétences entre l'environnement expérimental et l'environnement clinique sur patient.

L'ensemble des méthodes d'évaluation proposées ci-dessus présente les mêmes limites. D'une part, le coût économique important de l'entraînement associé à la présence d'un observateur expert externe, rendant de plus l'utilisation des compétences expertes impossibles au même moment. De ce fait, la gestion des ressources humaines se voit également complexifiée. D'autre part, dans chaque cas, quel que soit l'environnement considéré (expérimental ou clinique) ou bien l'outil de notation adopté, l'évaluation se fait au travers de l'observateur créant alors une notation dépendante de l'évaluateur, de par la nature subjective de son jugement. De nombreuses équipes ont pris en compte ces problématiques et s'intéressent de plus en plus à l'évaluation de compétences basée sur des méthodes automatiques, utilisant des mesures physiques pour en déduire le niveau d'expertise de l'opérateur.

1.2.3.2 Méthodes quantitatives automatiques

Ces méthodes, à l'inverse des méthodes manuelles, reposent sur une évaluation automatique de la qualité du geste et sont directement intégrées au sein de la plupart des

simulateurs actuels. Pour ce faire, le système acquiert des mesures physiques afin de pouvoir classer et évaluer le niveau d'expertise de l'utilisateur. Nous pouvons distinguer deux grandes catégories pour ces méthodes d'évaluation automatique : les méthodes prenant en compte l'analyse du résultat final, et les méthodes basées sur l'analyse des mouvements.

L'analyse du résultat est principalement axée sur le calcul d'un score global, en prenant en compte plusieurs paramètres durant l'exécution du geste chirurgical. La principale méthode d'analyse du geste basée sur l'étude du résultat est le test FLS (Fundamentals of Laparoscopic Skills), provenant du simulateur MISTELS (McGill Inanimate System for Training and Evaluation of Laparoscopic Skills) [Derossis 1998]. Le banc d'entraînement MISTELS permet d'évaluer l'apprenti au travers de cinq tâches basiques de laparoscopie que sont le transfert d'objet, la découpe d'un cercle, la ligature, la suture extracorporelle et la suture intracorporelle. Le score qui résulte de chaque tâche est alors calculé à partir du temps d'exécution qui est comparé avec un temps de référence, minoré par un score de pénalité associé aux erreurs durant l'exécution. Cette méthode, peu coûteuse, a été adoptée par la SAGES (Society of American Gastrointestinal and Endoscopic Surgeons) et par l'ACS (American College of Surgeons) afin d'évaluer les compétences techniques des apprentis chirurgiens digestifs. Dans le même objectif, certains simulateurs de réalité virtuelle ont implémenté une détection automatique du niveau d'expertise en prenant en compte automatiquement le nombre d'erreurs (cible manquée ou perte d'objet avec le da Vinci Skills Simulator) [Lyons 2013].

D'un autre côté, l'analyse du mouvement se base sur un ensemble de métriques, qui peuvent se combiner suivant les facteurs techniques que l'on souhaite étudier. Pour ces métriques, nous retrouvons majoritairement :

- Le temps, qui peut s'utiliser comme une métrique globale [Moorthy 2004], ou plus locale en prenant en compte le temps d'approche [Tokunaga 2012], le temps passé sur la zone de travail et le temps d'immobilité [Oropesa 2013] ;
- La longueur de chemin parcouru, le nombre de mouvements et la vitesse moyenne de ces mouvements, qui permettent de caractériser dans une certaine mesure l'économie de mouvement globale, notamment dans le cadre de la chirurgie robotisée [Judkins 2009] ;
- La fluidité du geste mesurée à partir du jerk (dérivée troisième de la position), donne une bonne approximation du score de confiance dans le geste réalisé [Chmarra 2010] ;
- Les directions de déplacements permettent d'étudier les composantes principales du mouvement [Cotin 2002] ;
- Les forces, qui permettent de quantifier les efforts exercés par les instruments sur les tissus [Horeman 2014].
- Les positions des instruments au cours du temps, qui fournissent une information de trajectoire permettant de définir le geste en trois dimensions [Uemura 2014].

Une étude détaillée des différentes métriques utilisées pour l'analyse du mouvement, dans le cadre de l'évaluation de la qualité des gestes chirurgicaux, est réalisée dans les travaux de Wolf [Wolf 2013]. Cette étude prend en compte le niveau d'analyse, le contexte d'utilisation et décrit la validité des métriques explorées.

1.3 Analyse et discussion

Ce chapitre présente le contexte clinique de la chirurgie laparoscopique et l'introduction des robots chirurgicaux. L'évolution des systèmes manuels vers des systèmes de plus en plus perfectionnés et robotisés a conduit à améliorer les conditions opératoires des chirurgiens. Cependant, l'importance de l'entraînement chirurgical s'accroît avec la complexification des interventions, l'objectif étant d'améliorer le service médical rendu au patient. Les études de la Haute Autorité de Santé présentées précédemment montrent que ce sont les compétences techniques des chirurgiens qui sont majoritairement mises en jeu. Ces compétences peuvent être apprises au travers de différents dispositifs d'entraînement pour l'apprentissage des gestes chirurgicaux. De plus, en analysant les dispositifs existants sur le marché, on constate que leur évolution est sans précédent, permettant ainsi à l'apprenti chirurgien d'acquérir les compétences techniques nécessaires à partir d'activités basiques, puis d'évaluer ses capacités dans des situations simulées proches de la réalité. Différentes méthodes d'évaluation sont alors utilisées pour juger de la qualité de réalisation.

En étudiant les méthodes basées sur un observateur externe pour l'évaluation du geste en clinique, on constate que la présence d'un expert est obligatoire afin de qualifier la qualité des gestes réalisés. De plus, l'aspect subjectif de la notation pose alors la question de la reproductibilité inter-observateur, limitant la portabilité du protocole d'évaluation d'un centre à un autre. En outre, on constate également que suivant la méthode employée, l'évaluation du geste se fait à différentes échelles. Alors que certaines méthodes se focalisent sur les activités, d'autres essaient de décrire plus précisément les gestes. À l'inverse, les méthodes d'évaluation automatiques utilisées afin de quantifier le geste possèdent des avantages indéniables, dont notamment l'aspect objectif dans la notation. Une fois les métriques normalisées, ces dernières peuvent permettre de comparer objectivement et directement un opérateur à une référence experte, sans souci de fiabilité et de reproductibilité interpersonnelle. Bien que l'acquisition de dispositifs équipés de la sorte soit plus coûteuse pour les écoles et établissements, l'évaluation est plus détaillée, permettant d'évaluer les points forts ainsi que les points faibles de l'apprenti au travers d'un ensemble de métriques. Ces méthodes d'évaluation sont actuellement implémentées sur la grande majorité des dispositifs de réalité virtuelle employés pour l'entraînement aux gestes chirurgicaux.

Toutefois, le paradigme conceptualisé par Halsted "see one, do one, teach one" ne semble plus en vigueur dans le cadre d'évaluation automatique, dérivant vers une représentation

du type “perfect practice make perfect” [Tsuda 2009]. L’entraînement par démonstration est alors remplacé par une évaluation basée sur un ensemble de métriques calculées automatiquement, où l’opérateur doit maximiser les scores de chacune d’elles. Ce type d’analyse cherche à corrélérer des facteurs discriminants avec les compétences techniques de l’opérateur [Uemura 2014], mais perd toutes notions procédurales, ne cherchant ainsi qu’à quantifier la tâche au travers de métriques et non plus de gestes. Une alternative entre l’évaluation d’une tâche chirurgicale basée observateur qui est capable de reconnaître les gestes mais évalue de manière subjective, et l’évaluation automatique basée sur un ensemble de métriques, pourrait être une évaluation automatique des gestes. Pour ce faire, le système doit être capable dans un premier temps de détecter les différents gestes de l’opérateur afin, dans un second temps, de les quantifier indépendamment les uns des autres. L’avantage de ce type d’approche est d’associer une évaluation objective se basant sur des métriques avec la capacité de discernement des gestes que possède un observateur humain. Le résultat serait une analyse intelligente de la tâche pour une évaluation locale, au niveau des gestes, mais également globale en prenant en compte l’enchaînement de ces gestes. Ce type d’évaluation aurait donc la propriété d’évaluer à plusieurs niveaux la tâche réalisée par l’opérateur, d’être reproductible, mais également de lui fournir des informations complémentaires quant aux améliorations à réaliser sur chaque geste pour parfaire l’apprentissage.

Afin de répondre à cette problématique, une thématique de recherche émergente tente de résoudre ces problèmes, faisant appel à des méthodes scientifiques avancées pour modéliser l’ensemble du processus chirurgical. L’idée est alors de modéliser les gestes chirurgicaux afin de les reconnaître, puis de les évaluer. Dans la suite de ce manuscrit, notre première contribution aborde le développement d’une nouvelle méthode pour la reconnaissance des gestes dans le cadre de l’entraînement en chirurgie robotisée. Pour ce faire, nous présentons dans le chapitre suivant un état de l’art des méthodes existantes dans ce domaine afin de positionner notre contribution par rapport aux travaux de la littérature.

Analyse du geste chirurgical : point de vue scientifique

Préambule

Ce chapitre présente dans un premier temps le concept de modélisation des processus chirurgicaux. Pour ce faire, nous détaillons le cadre applicatif de ce type d'outil, puis nous concentrons notre analyse sur le processus destiné à la reconnaissance du geste chirurgical où un état de l'art des travaux existants permet de mettre en lumière la nécessité d'aborder une nouvelle approche. Nous terminons ce chapitre en introduisant les problématiques et le cadre de ces travaux de recherche.

Sommaire

2.1	Compréhension quantitative des interventions	26
2.1.1	Analyse et modélisation de processus chirurgicaux	26
2.1.1.1	Définition des concepts	27
2.1.1.2	Niveaux de granularité	29
2.1.2	Vers une reconnaissance du geste chirurgical	31
2.1.2.1	Notions de surgeme et dexeme	31
2.1.2.2	Description du processus de reconnaissance	33
2.2	Etat de l'art des travaux pour la reconnaissance des gestes chirurgicaux . .	35
2.2.1	Conditions d'étude	36
2.2.1.1	Contexte médical	36
2.2.1.2	Acquisition des données	38
2.2.2	Processus de reconnaissance	42
2.2.2.1	Prétraitement des données	42
2.2.2.2	Extraction d'attributs	44

2.2.2.3	Méthodologie d'apprentissage	44
2.2.2.4	Spécification du processus	46
2.2.3	Processus de validation	46
2.2.3.1	Conditions expérimentales	47
2.2.3.2	Stratégie de validation	49
2.2.4	Analyse et discussion	51
2.2.4.1	Conditions d'étude	51
2.2.4.2	Processus de reconnaissance	52
2.2.4.3	Processus de validation	54
2.3	Conclusion et problématiques de la thèse	55

2.1 Compréhension quantitative des interventions

Suite aux avancées médicales, notamment avec l'apparition de la laparoscopie et de la chirurgie minimalement invasive, et à l'émergence de nouvelles technologies au sein du bloc opératoire avec les systèmes d'imagerie et les robots de téléopération, le déroulement des interventions s'est complexifié. Il en résulte une surcharge d'informations accessibles par le personnel médical et le chirurgien. De ce fait, la chirurgie devient plus longue, ce qui peut augmenter les risques d'erreurs et de complications au cours de l'intervention.

Afin de réduire ces erreurs, certains établissements ont mis en place des formations spécifiques pour l'entraînement en chirurgie robotisée, comme c'est le cas à Nancy avec la création d'un DIU (diplôme interuniversitaire) de chirurgie robotisée. Ces formations permettent notamment d'apprendre la gestuelle, mais l'évaluation qui en résulte peut être perfectible, plus particulièrement dans le cadre de l'entraînement au travers de méthodes d'évaluation automatiques comme nous l'avons abordé précédemment. Dans ces travaux, nous souhaitons développer une nouvelle approche pour la reconnaissance des gestes chirurgicaux, à des fins d'évaluation des performances mais également d'amélioration du processus d'apprentissage. Pour ce faire, nous présentons dans un premier temps la modélisation des processus chirurgicaux permettant d'introduire le contexte de recherche scientifique de nos travaux.

2.1.1 Analyse et modélisation de processus chirurgicaux

Les recherches abordant la reconnaissance des gestes chirurgicaux emploient certains concepts et outils qui émanent initialement du domaine de la modélisation de processus chirurgicaux. Dans la suite de ce chapitre nous étudions et définissons uniquement les concepts nécessaires à la compréhension de nos travaux. Pour plus d'informations sur ce domaine de recherche, nous nous référerons aux travaux de Lalys et al. [Lalys 2014] où un état de l'art de la littérature introduit les différents concepts et outils.

2.1.1.1 Définition des concepts

L'analyse et la modélisation du processus chirurgical¹ est une thématique de recherche qui a émergé depuis une quinzaine d'années. Une définition formelle du concept de processus chirurgical a été donnée par Neumuth et al. [Neumuth 2006b], mais nous préférons la définition plus complète fournie par Jannin et al. [Jannin 2007]. Le terme de processus chirurgical y est décrit comme "l'automatisation du processus médical de gestion du patient, en partie ou dans sa totalité, durant lequel documents, informations, images ou tâches sont transférés d'un participant à un autre pour la réalisation d'actions, régies par un ensemble de règles procédurales". Ce processus chirurgical décrit principalement les relations temporelles existantes entre les différentes actions intervenant au cours dudit processus. Cependant, l'arrivée de la chirurgie assistée par ordinateur ainsi que des systèmes d'assistance en chirurgie ont complexifié ces processus médicaux. En conséquence, des problèmes de communication, d'accès et d'échanges d'informations sont apparus. Les travaux du domaine se sont alors intensifiés afin d'analyser ces processus pour fournir des possibilités d'optimisation, permettant ainsi d'améliorer le service médical rendu au patient.

L'acquisition de données, aussi bien pré- que peropératoire, est la première étape dans la description du processus chirurgical. Il existe deux méthodes pour acquérir ces données. La première méthode, basée sur un observateur externe, vise à décrire manuellement l'ensemble des actions, parfois au travers d'un logiciel dédié [Neumuth 2006a]. L'acquisition peut se faire durant l'intervention, ou bien en utilisant les enregistrements vidéo pour l'annotation [Padoy 2012]. Une seconde méthode consiste à bénéficier des avancées technologiques pour réaliser ces acquisitions de manière automatique au travers de différents capteurs. Ces capteurs peuvent être des caméras [Lalys 2011], des capteurs RFID (radio-identification) ainsi que des capteurs de position et d'état des systèmes [Kranzfelder 2014, Stauder 2014], des capteurs magnétiques placés sur les instruments [Uemura 2014] ou bien directement l'utilisation des informations provenant du robot chirurgical [Lin 2006]. Ces données servent ensuite à la création d'un modèle, qui peut être mathématique, représentant le processus observé. La création de ce modèle nécessite néanmoins d'avoir une description structurée et un vocabulaire adapté au processus analysé.

Parmi les premiers travaux du domaine, nous retrouvons principalement ceux de Mackenzie et al. [Mackenzie 2001], dont l'idée est de créer un modèle de procédure à partir d'une décomposition structurée à plusieurs niveaux décrivant l'ensemble des actions chirurgicales réalisées au cours de l'intervention. Pour ce faire, un diagramme sous forme d'arbre est construit (figure 2.1). La description des actions chirurgicales est alors effectuée en utilisant un vocabulaire propre à l'intervention. Jannin et al. [Jannin 2001] ont également cherché à modéliser une procédure neurochirurgicale en utilisant une décomposition similaire, basée sur l'utilisation du langage UML (Unified Markup Language). Suite

1. "Surgical Process Modelling" (SPM) en anglais.

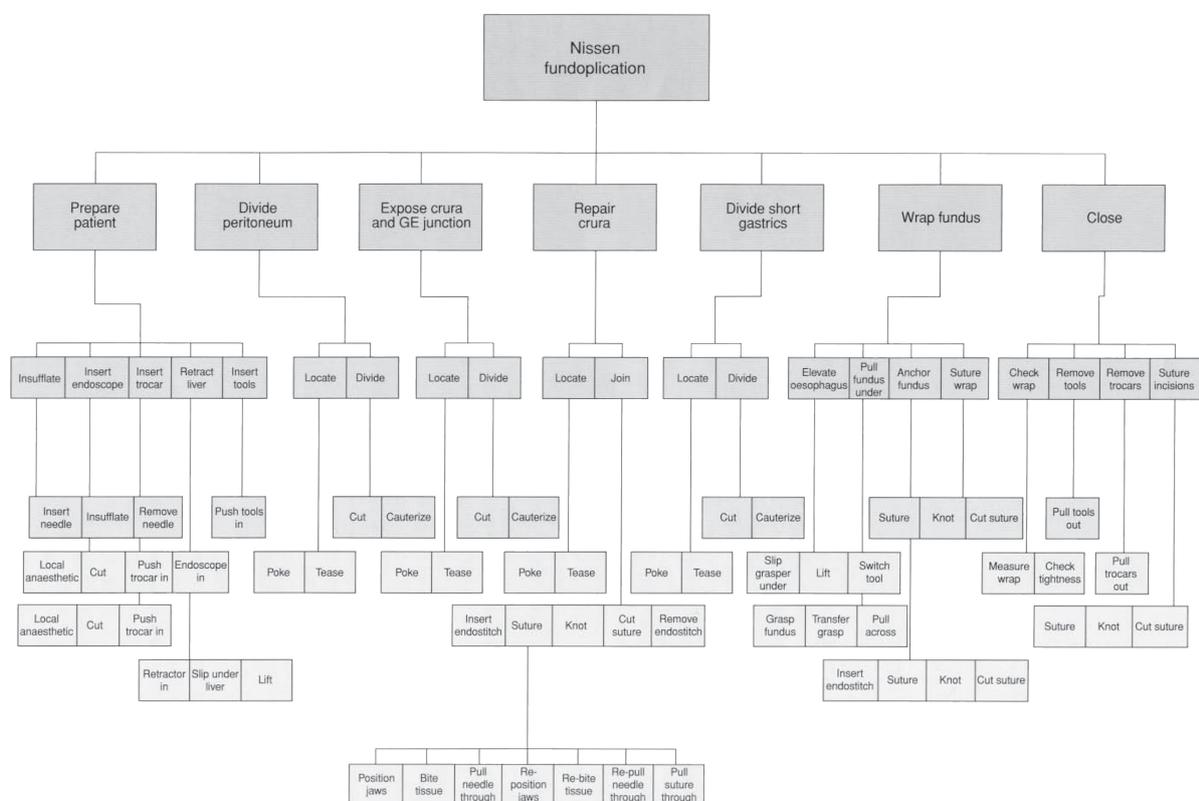


FIGURE 2.1: Décomposition temporelle et hiérarchique d'une fundoplication par voie laparoscopique en utilisant une description à plusieurs niveaux.
 Source: Schéma provenant des travaux de Mackenzie et al. [Mackenzie 2001]

à l'acquisition de plusieurs interventions, ce modèle a été amélioré et validé dans le cadre d'une assistance pour la planification préopératoire [Jannin 2007]. Le développement des travaux et des applications dans le domaine ont permis de créer un formalisme utilisé pour décrire les processus chirurgicaux. Dans cet objectif, Neumuth et al. [Neumuth 2006b] utilisent une représentation à deux dimensions. La première, qui est horizontale, correspond aux classifications des actions suivant cinq critères complémentaires : l'organisation (qui ou quoi effectue la tâche?), la fonction (l'activité et la structure sur laquelle elle s'applique), l'information (les données requises et/ou produites), l'opération (les instruments chirurgicaux employés) et le comportement (l'enchaînement temporel des tâches et leur causalité). La seconde dimension, verticale, correspond à l'organisation. Cette dimension se décompose de manière hiérarchique et se réfère à la notion de granularité.

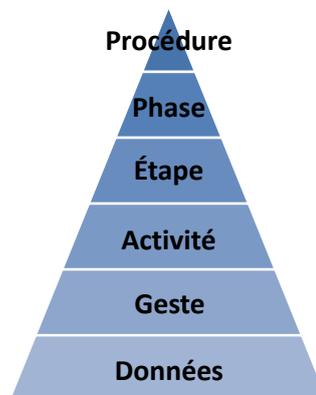


FIGURE 2.2: Les différents niveaux de granularité utilisés pour décrire un processus chirurgical.

Source: Schéma inspiré et adapté de Lalys et al. [Lalys 2014].

2.1.1.2 Niveaux de granularité

La notion de granularité pour la modélisation des processus chirurgicaux a été décrite par [Neumuth 2006b, Trelhu 2011]. Cependant l'unique et complète analyse de l'état de l'art traitant la modélisation de processus chirurgicaux a été réalisée par Lalys et al. [Lalys 2014]. Ils y reprennent la définition de granularité et y associent un vocabulaire adapté permettant d'organiser leur analyse des travaux du domaine en utilisant un référentiel commun. La granularité y est définie comme le niveau d'abstraction utilisé pour décrire le processus chirurgical, qui correspond en réalité au point de vue de la description (macro ou micro). Dans ce manuscrit, nous emploierons une partie du formalisme présenté ci-après pour distinguer chaque niveau de granularité du processus chirurgical (voir figure 2.2). De plus, afin d'illustrer nos propos, le paragraphe suivant fera référence à certaines informations contenues dans la figure 2.1.

Le niveau le plus élevé pour la description d'un processus chirurgical est la procédure elle-même. Mackenzie et al. [Mackenzie 2001] se sont appuyés sur l'exemple de la fundoplicature qui est une intervention chirurgicale dont le but est d'empêcher le reflux gastrique dans l'œsophage. Pour ce faire, il est nécessaire d'enrouler la partie supérieure de l'estomac autour de la partie inférieure de l'œsophage. Cette intervention requiert différentes manipulations de la part du chirurgien, définissant un ensemble de phases. Ces phases sont généralement caractérisées par une durée temporelle importante, et ne fournissent que peu de précision quant aux événements intervenant au sein du bloc opératoire. Des exemples de phases dans le cadre d'une fundoplicature sont la séparation du péritoine, ou bien la déformation du fundus (partie haute de l'estomac). Chaque phase se compose ensuite d'un ensemble d'étapes, fournissant généralement plus de précision sur la description du processus chirurgical. Ces étapes peuvent être l'insertion d'un trocart, la loca-

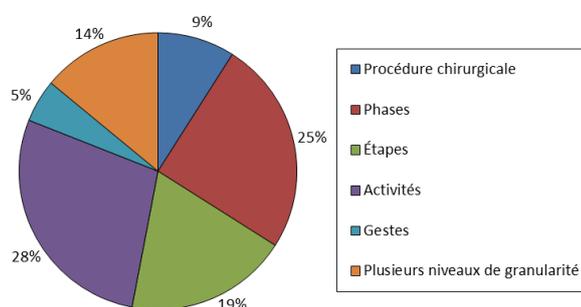


FIGURE 2.3: Répartition des niveaux de granularité utilisés pour la modélisation.
Source: Schéma adapté de Lalys et al. [Lalys 2014].

lisation ou bien le déplacement d'un organe. Chaque phase contient alors une ou plusieurs activités. On citera ici la découpe de tissus, la réalisation d'une suture ou bien la cautérisation. L'avant dernier niveau de granularité est la description des gestes. Dans ce cas, nous pouvons donner comme exemple la réalisation d'une suture. De plus, un sous ensemble de gestes nécessaires à la réalisation d'une suture peuvent être : attraper l'aiguille, insérer l'aiguille dans le tissu, tirer l'aiguille. Enfin, le dernier niveau est la description du processus par des données dites bas-niveaux. Le terme de données bas-niveaux fait référence à l'utilisation d'informations telles que des images médicales, des vidéos (du bloc opératoire ou du champ opératoire), des trajectoires (des instruments chirurgicaux ou bien du personnel médical) ainsi que des données binaires (comme l'état d'utilisation d'un instrument chirurgical à un moment précis). Ces données, pour la plupart numériques, permettent de décrire l'ensemble des gestes, activités, étapes et phases intervenant au cours d'une procédure chirurgicale.

L'analyse de l'état de l'art pour la modélisation de processus chirurgicaux réalisé par Lalys et al. [Lalys 2014] a permis de quantifier la répartition des travaux de la communauté suivant le niveau de granularité utilisé pour la modélisation (voir figure 2.3). En analysant le diagramme, on peut voir que la majorité des travaux (28%) traite de l'analyse et la modélisation de processus chirurgicaux au travers des activités. À l'inverse, une faible proportion (seulement 5%) aborde l'analyse et la modélisation de ces processus par l'étude des gestes médico-chirurgicaux, qui sont pourtant directement liés aux compétences techniques du chirurgien.

Ces travaux de thèse abordent donc une problématique qui n'a été que très peu explorée, en comparaison d'autres approches basées sur des niveaux de granularité supérieurs. Néanmoins, nous verrons qu'analyser et reconnaître le geste chirurgical apporte

un nombre conséquent d'informations, qui peuvent être utiles à l'évaluation des compétences techniques des praticiens.

2.1.2 Vers une reconnaissance du geste chirurgical

Lorsque la modélisation est réalisée, l'un des objectifs est d'être capable de reconnaître les événements de processus chirurgicaux similaires. Pour ce faire, des méthodes de reconnaissance automatique connues sous le nom d'apprentissage automatique peuvent être employées. Ces méthodes, majoritairement utilisées pour reconnaître les phases et activités au sein des processus chirurgicaux [Lalys 2012, Quéllec 2014], peuvent également s'appliquer dans le cadre de la reconnaissance des gestes chirurgicaux. Afin de mieux comprendre leur fonctionnement et leur utilisation, nous introduisons dans le paragraphe suivant le vocabulaire employé dans ces travaux nécessaire à la compréhension du geste chirurgical. Puis, nous présentons les différentes étapes intervenant dans le processus de reconnaissance du geste chirurgical.

2.1.2.1 Notions de *surgeme* et *dexeme*

Afin de comprendre la réalisation d'un geste nous souhaitons introduire deux notions. Ces notions viennent en complément du schéma de la figure 2.3, et sont spécifiques à l'analyse du geste chirurgical. Ainsi, de la même manière qu'une procédure comporte un ensemble de primitives telles que des phases, des étapes, des activités et des gestes, nous décomposons le geste en deux catégories.

Dans un premier temps, nous définissons la notion de *surgeme*, présentée précédemment dans les travaux de Lin et al. [Lin 2006] ainsi que de Reiley et al. [Reiley 2008]. Un *surgeme*² représente un geste chirurgical réalisé à une main, avec un unique instrument, comportant une action et un but chirurgical précis et agrémenté d'une sémantique explicite. Identique à la notion de geste retenue par Lalys et al. [Lalys 2014], nous préférons l'utilisation de ce terme qui, en complément de la prochaine notion, fournit plus de sens aux travaux de ce manuscrit. De plus, ce terme a maintenant été adopté par la communauté, permettant ainsi d'utiliser un vocabulaire commun pour représenter des notions similaires. À titre d'exemple, nous illustrons nos propos avec la figure 2.4 qui représente la trajectoire d'un instrument robotisé de chirurgie laparoscopique durant la réalisation d'une tâche d'entraînement pour le transfert d'objet. Ce type de tâche est régulièrement employé en laparoscopie pour évaluer les compétences psychomotrices des apprentis [Derossis 1998]. Dans cet entraînement, l'activité chirurgicale principale à réaliser est le transfert d'un pion en silicone, d'un point de départ jusqu'à une cible prédéfinie. Cette activité se répète alors plusieurs fois pour évaluer les compétences techniques du chirurgien. Afin d'accomplir cette tâche, l'un des *surgemes* à réaliser est l'insertion du pion dans la cible en utilisant le

2. Le mot "surgeme" provient de la composition des mots anglais "surgical" et "gesteme", où ce dernier fait référence à la notion de primitive de geste.

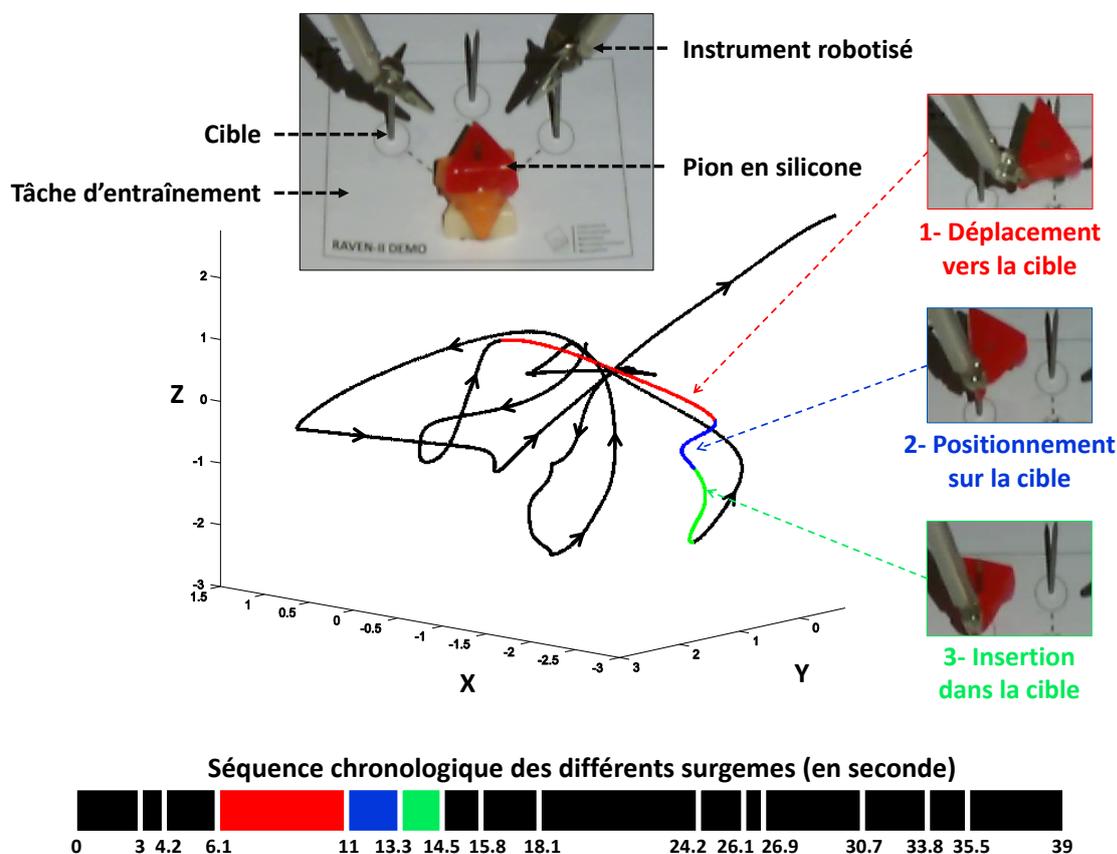


FIGURE 2.4: Décomposition en surgemes d'une tâche de transfert d'objet pour l'entraînement en chirurgie laparoscopique.

porte-aiguille de la main gauche.

Dans un second temps, nous définissons la notion de *dexeme*, présentée initialement dans les travaux de Reiley et al. [Reiley 2009b] ainsi que de Varadarajan [Varadarajan 2009]. Un *dexeme*³ représente une partie d'un surgeme, réalisé à une main avec un instrument, et dénué de sens médical. Ces *dexemes* ne sont pas spécifiques aux surgemes mais permettent, une fois la séquence de *dexeme* exécutée, d'accomplir un surgeme. En reprenant l'exemple précédent du transfert d'objet présenté avec la figure 2.4, là où le surgeme correspond à l'insertion du pion dans la cible en utilisant l'instrument gauche, un premier *dexeme* sera de se déplacer vers le bas pour atteindre la cible. Les *dexemes* représentent des gestes primitifs, simples, sans objectif médical prédéterminé, traduisant une intention directe de l'opérateur.

3. Le mot "dexeme" provient de la composition des mots anglais "dexterous" et "gesteme".

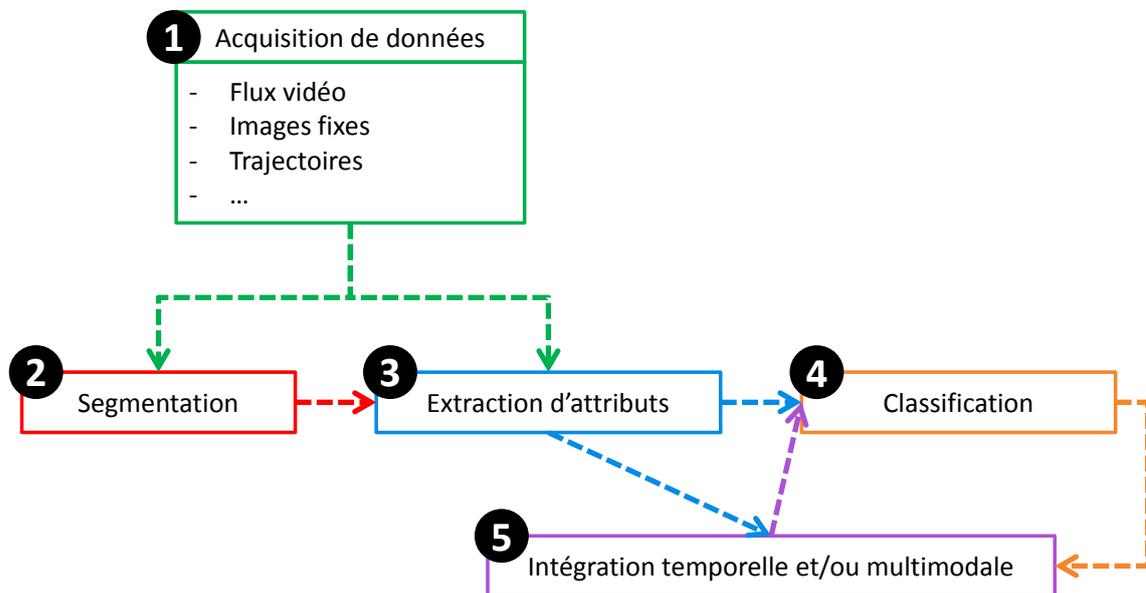


FIGURE 2.5: Processus généralisé pour la reconnaissance des gestes chirurgicaux.
Source: Schéma inspiré et adapté de Penet [Penet 2013].

Dans la suite de ce manuscrit, nous emploierons les notions de *surgeme* et *dexeme* pour l'analyse et la reconnaissance des gestes chirurgicaux.

2.1.2.2 Description du processus de reconnaissance

Les principes de reconnaissance et d'apprentissage automatiques ont été appliqués à une multitude de domaines tels que la segmentation des phonèmes pour la reconnaissance de mots [Andre-Obrecht 1988, Kamakshi Prasad 2004], la détection de piéton [Benenson 2012, Benenson 2014], en passant par la reconnaissance d'évènement dans les vidéos pour la détection de violence [Penet 2013], ainsi que la reconnaissance de concepts et d'objets divers pour la compréhension de la situation et de l'environnement [Fang 2014]. À des fins de compréhension, ces principes sont rappelés dans les paragraphes suivants.

Schématisation du processus

Pour clarifier l'ensemble des étapes intervenant dans le processus de reconnaissance des gestes chirurgicaux, nous appuyons notre description sur le schéma de la figure 2.5. La première étape du processus concerne l'acquisition de données. Cette étape est primordiale car elle permet d'injecter des données à traiter dans le système de reconnaissance. De plus, il est nécessaire que les spécifications techniques pour l'acquisition soient connues (type de données, caractéristiques intrinsèques, fréquence d'acquisition) afin d'assurer le

bon conditionnement des données.

La seconde étape permet de segmenter les données. Cette segmentation assure le découpage des données d'entrée en plusieurs sous-ensembles, dont les tailles peuvent varier suivant les besoins définis par l'utilisateur. Ce découpage peut être structuré ou non fait appel à des algorithmes basés sur l'apprentissage, suivant la méthode employée. À noter que cette étape est facultative et dépend des objectifs de l'utilisateur.

La troisième étape est l'extraction d'attributs. Elle calcule des caractéristiques permettant de représenter les données dans un espace de dimension fixe, tout en maximisant leur degré de discrimination. Dans le cas où le processus de reconnaissance contient une étape de segmentation et une étape d'extraction d'attributs, ces deux étapes deviennent indissociables dans le sens où les attributs extraits sont dépendant de la segmentation effectuée en amont, caractérisant alors les segments retenus.

La quatrième étape du processus réalise la classification, que l'on appelle également caractérisation. Cette étape clé du processus permet d'associer aux données d'entrée des étiquettes, ou labels, suivant leur appartenance réelle. La classification est une étape d'apprentissage dont les différents modes sont décrits ci-dessous.

Additionnellement, la cinquième étape du processus de reconnaissance est l'intégration temporelle et/ou multimodale. Cette étape permet de prendre en compte l'aspect temporel des données afin d'améliorer les capacités de reconnaissance du système. Cette prise en compte peut être précoce (après l'extraction d'attributs) ou tardive (en aval de la classification). De plus, le système peut également prendre en compte différentes modalités. Dans ce cas, une synchronisation des données est requise.

Il est à noter que ce schéma est une généralisation du processus de reconnaissance. Dans certains cas, des méthodes hybrides peuvent être employées pour segmenter et classer simultanément.

Modes d'apprentissage

Il existe quatre grands modes d'apprentissage qui se caractérisent par le retour dont on dispose sur les données :

- L'apprentissage supervisé : à partir des couples d'entrée-sortie disponibles, le système apprend une fonction afin de pouvoir, avec un jeu de données, reproduire les sorties associées. Dans ce cas, le système cherche à minimiser l'erreur en généralisation ;
- L'apprentissage semi-supervisé : à partir des données, dont une partie des couples d'entrée-sortie est connue et l'autre non, le système a pour objectif de généraliser un modèle appris sur les couples connus. Cependant, ces couples ne sont pas représentatifs d'une vérité générale puisqu'une partie seulement est connue. Le système cherche alors à modifier le modèle généralisé et à réorganiser les hypothèses effectuées sur le modèle de départ afin de minimiser l'erreur moyenne sur la base de tests (couples dont on ne connaît que les entrées) ;

- L'apprentissage non-supervisé : à partir des données, ne connaissant pas les couples d'entrée-sortie, le système cherche à apprendre une structure permettant d'expliquer leurs relations. Une méthode de classification courante dans ce cas est la détection de groupes⁴ ;
- L'apprentissage par renforcement : à partir des données, le système apprend un modèle évolutif par une série de récompenses ou de punitions. Les punitions et récompenses indiquent les points de renforcement (positifs ou négatifs) à prendre en compte dans le modèle pour répondre au problème donné.

Parmi ces modes, l'apprentissage supervisé est le plus couramment utilisé dans la littérature. Son intérêt réside dans une meilleure maîtrise du modèle dans le sens où la fonction dépend majoritairement du jeu de données employé pour l'apprentissage. Néanmoins, les capacités de généralisation sont limitées et dépendent de la variabilité des données. Dans ce cas, l'apprentissage semi- ou non-supervisé se prête mieux à la généralisation, mais demande un effort plus important à la mise en place, et ne peut généralement pas se révéler aussi fiable. Dans le contexte de reconnaissance des gestes chirurgicaux, l'apprentissage par renforcement n'a jamais été employé à notre connaissance.

Les notions d'apprentissage introduites vont nous permettre de présenter, avec le niveau de détail nécessaire à la compréhension, l'état de l'art des travaux s'intéressant à la problématique de reconnaissance des gestes dans le cadre de l'entraînement en chirurgie.

2.2 Etat de l'art des travaux pour la reconnaissance des gestes chirurgicaux

La reconnaissance des gestes chirurgicaux est une thématique très récente, mais actuellement en pleine expansion notamment avec l'apparition des robots au bloc opératoire destinés à l'assistance. Néanmoins, la frontière entre la reconnaissance d'activités et de gestes chirurgicaux est parfois ambiguë suivant l'interprétation des définitions et le vocabulaire employé par chaque équipe. Dans cet état de l'art, nous analysons uniquement les travaux abordant la reconnaissance de surgemes et dexemes. De ce fait, nous avons exclus les travaux abordant l'apprentissage des compétences techniques chirurgicales utilisant une modélisation statistique du niveau d'expertise, sans prendre en compte la détection du geste chirurgical. Excluant la reconnaissance sémantique, ce type d'approche infère généralement un niveau de compétence au travers de facteurs physiques [Megali 2006, Reiley 2009b, Malpani 2014, Uemura 2014].

Les publications présentes dans cet état de l'art introduisent des travaux réalisés entre Janvier 2000 et Juin 2015. Avec initialement 54 articles, nous avons affiné notre champ d'analyse en sélectionnant les travaux traitant explicitement de la reconnaissance des gestes chirurgicaux. Nous présentons dans cet état de l'art 25 travaux en lien avec le développement

4. "Clustering" en anglais

de méthodes pour la segmentation et la détection de gestes chirurgicaux. De plus, afin de faciliter l'analyse de ces travaux, nous avons décrit ces derniers à travers trois grandes catégories que sont les conditions de l'étude, le processus de reconnaissance et le processus de validation. Chaque catégorie est représentée par plusieurs descripteurs, eux-mêmes caractérisés par plusieurs critères permettant de décrire de manière structurée le contenu des travaux. Afin d'améliorer la lisibilité, nous avons également choisi de présenter cette analyse à travers trois tableaux distincts reprenant les catégories citées précédemment. À noter que dans certains cas, suivant l'avancement des travaux et le niveau de détail fourni dans les publications, certaines données peuvent être absentes ou manquer de précision. Nous renvoyons le lecteur vers la publication correspondante pour plus d'information.

2.2.1 Conditions d'étude

Afin de décrire les conditions des études considérées dans cet état de l'art, nous utilisons deux descripteurs que sont le contexte médical et la méthodologie d'acquisition des données, incluant une partie de la description des données. Le tableau 2.1 fournit un résumé des différentes conditions d'études considérées dans les travaux de la littérature.

2.2.1.1 Contexte médical

Le contexte médical décrit le scénario de l'étude d'un point de vue clinique. Afin de le caractériser, nous utilisons trois critères : le domaine médical ciblé, l'application médicale envisagée par les travaux et le type de données recueilli durant les expériences. Cette description permet au lecteur d'interpréter la méthodologie utilisée suivant le contexte expérimental des travaux.

Domaines d'application

La reconnaissance des gestes chirurgicaux a été appliquée à plusieurs domaines médicaux. Parmi ces travaux, 80% d'entre eux ont été réalisés uniquement dans le cadre de la chirurgie laparoscopique. Nous retiendrons notamment ceux de Rosen et al. [Rosen 2001, Rosen 2002, Rosen 2006] qui ont instauré des pistes méthodologiques intéressantes tout en étant les précurseurs de la plupart des travaux.

Concernant la chirurgie ophtalmique, impliquant également les travaux de microchirurgie, seuls 12% des travaux ont été appliqués à ce domaine. On citera ici les travaux de Hundtofte et al. [Hundtofte 2002], de Kragic et al. [Kragic 2003] ainsi que ceux de Novi et al. [Novi 2012].

Dans cet état de l'art, seule une publication, celle de Ahmadi et al. [Ahmadi 2009], s'attaque à deux domaines médicaux différents que sont la chirurgie laparoscopique et la neurochirurgie. Enfin, Holden et al. [Holden 2014] sont les seuls à appliquer leur méthodologie dans le cadre de la radiologie interventionnelle.

2.2. ETAT DE L'ART DES TRAVAUX POUR LA RECONNAISSANCE DES GESTES CHIRURGICAUX

	Contexte médical		Méthodologie d'acquisition des données										Spécification														
	Domaine médical	Application médicale	Type de données	Système d'acquisition										Nature des données	Référentiel d'acquisition												
	Laparoscopie	Radiologie interv.	Neurochirurgie	Ophtalmologie		Robot de téléopération	Robot de manipulation	Simulateur	Gant instrumenté	Outil Instrumenté	Capteur	Trajectoire	Vidéo	Force	Pression	Instrument	Interface	Tâche	Corps	Bras	Poignet	Main	Doigt	Fréquence d'acquisition	Nombre de signaux	Résolution	
[Ahmadi2009]	x		x								x	x				x								100Hz	12		
[Ahmadi2013]	x		x									x					x							30Hz	39		
[BejarHaro2012]	x		x										x											30Hz		320x240	
[Bodenstedt2012]	x		x									x															
[Dosis2005]	x		x																					5Hz, 20Hz	4		
[Holden2014]	x		x																								
[Hundtofte2002]		x																									
[Kragic2003]		x																									
[Lea2015]	x		x																					30Hz	39	640x480	
[Lin2005], [Lin2006]	x		x																					10Hz	39		
[Normm2008]	x		x																					60Hz	9		
[Novi2012]		x																									
[Paday2011]	x		x																					40Hz	13		
[Reiley2008]	x		x																					23Hz	36		
[Reiley2010]	x		x																					30Hz	41		
[Rosen2001], [Rosen2002]	x		x																					30Hz	7		
[Rosen2006]	x		x																					30Hz	13		
[Saggio2011]	x		x																								
[Tao2012]	x		x																					30Hz	39		
[Tao2013]	x		x																					30Hz	39	320x240	
[VanDenBerg2010]	x		x																					30Hz	39		
[Varadarajan2009]	x		x																					30Hz	39		
[Zappella2013]	x		x																					30Hz	39	320x240	

Tableau 2.1: Résumé des conditions d'études considérées dans les travaux de la littérature ciblant la reconnaissance de gestes médico-chirurgicaux.

Applications envisagées

Le critère d'application envisagée désigne l'objectif de ces travaux. En d'autres termes, il décrit leur bénéfice du point de vue médical et leur intérêt dans un futur proche ou lointain tel qu'il est envisagé par les auteurs. Parmi ces travaux, 60% d'entre eux ont pour objectif l'analyse de compétences. De plus, 93% d'entre eux sont appliqués au domaine de la laparoscopie, montrant l'intérêt grandissant pour cette thématique et l'importance de l'évaluation de compétences dans le cadre de la chirurgie robotisée. Nous citerons notamment les travaux de Rosen et al. [Rosen 2001, Rosen 2002, Rosen 2006] ainsi que ceux de Reiley et al. [Reiley 2008] et de Lea et al. [Lea 2015].

D'après les auteurs, 36% des travaux sont destinés à fournir une assistance robotisée. Des applications très intéressantes sont présentées par Nomm et al. [Nomm 2008], ainsi que par Padoy et al. [Padoy 2011]. Enfin, Ahmadi et al. [Ahmadi 2009] ne prônent aucune application concrète et décrivent simplement le bénéfice de leur méthodologie dans le cadre de la reconnaissance du contexte chirurgical à des fins techniques et administratives.

Dans cet état de l'art, seuls les travaux de Ahmadi et al. [Ahmadi 2009], Lea et al. [Lea 2015] ainsi que Varadarajan et al. [Varadarajan 2009] ont développé des méthodes de reconnaissance et évalué ces dernières dans le cadre applicatif envisagé par leur étude.

Types de données

Nous utilisons ce critère pour représenter la viabilité des données par rapport à un scénario clinique, positionnant ainsi le réalisme des données par au contexte médical correspondant [Jannin 2008]. Là où les études sur fantôme permettent de gérer l'environnement et d'améliorer la qualité des données acquises, les études en clinique intègrent plus de variabilité et de complexité du fait de l'acquisition des données dans un environnement non contraint. Ainsi, 80% des travaux publiés ont acquis des données sur fantôme uniquement, impliquant un contexte d'étude maîtrisé, manquant de réalisme et induisant des biais par rapport à la réalité. À l'inverse, l'acquisition de données dans un environnement réel admet d'importantes contraintes en termes de disposition et d'hygiène, mais permet d'évaluer les travaux dans un environnement réaliste. Dans ce cadre d'acquisition, 16% des travaux présentent des acquisitions dans un environnement in-vivo, aussi bien sur animaux [Rosen 2001, Rosen 2002, Rosen 2006] que sur patient [Nomm 2008]. Enfin, seuls Ahmadi et al. [Ahmadi 2009] (4%) présentent des acquisitions à la fois sur fantôme et en environnement in-vivo.

2.2.1.2 Acquisition des données

Dans cette paragraphe, nous décrivons les travaux du domaine suivant la méthodologie d'acquisition des données employée. Pour ce faire, nous utilisons plusieurs critères que sont le système d'acquisition utilisé, sa nature, la nature des données acquises, le référentiel d'acquisition considéré et enfin les spécifications techniques des données.

Systèmes d'acquisition

Parmi les systèmes d'acquisition recensés dans la littérature, le robot da Vinci est employé dans 56% des travaux. Dans le même esprit, Rosen et al. [Rosen 2006] utilisent le Blue DRAGON développé en interne afin d'acquérir des données lors d'interventions laparoscopiques sur animaux. De manière similaire mais plus expérimentale, Van Den Berg et al. [Van Den Berg 2010] utilisent des Berkeley Surgical Robots, tentant de reproduire les capacités obtenues avec le robot da Vinci. Nous noterons que seuls Dosis et al. [Dosis 2005] utilisent deux systèmes d'acquisition, que sont le robot da Vinci et le système ICSAD destiné à l'acquisition de données pour l'entraînement en laparoscopie. Abordant également l'entraînement en chirurgie laparoscopique, les travaux de Saggio et al. [Saggio 2011] se concentrent exclusivement sur la description d'un système d'acquisition pour la reconnaissance de gestes basée sur le gant HITEG. Novi et al. [Novi 2012] s'attaquent de la même manière à la notion d'entraînement aux gestes mais appliquée à la chirurgie oculaire grâce l'utilisation de l'Ocular Trauma Simulator. En restant dans le cadre de la chirurgie oculaire et de la microchirurgie, Hundtofte et al. [Hundtofte 2002] ainsi que Kragic et al. [Kragic 2003] ont utilisé le JHU Steady Hand Robot pour l'acquisition de données. Dans littérature, certaines publications ciblent l'acquisition de données à partir d'instruments chirurgicaux [Rosen 2001, Rosen 2002] et d'aiguilles [Holden 2014], en se basant sur l'intégration et l'adaptation de capteurs pour la mesure d'informations. De manière moins intrusive, Ahmadi et al. [Ahmadi 2009] emploient des accéléromètres pour capturer les mouvements du chirurgien. Enfin, seuls Nomm et al. [Nomm 2008] utilisent des caméras afin de capturer des informations permettant, à l'usage, de remplacer l'infirmière du bloc opératoire par un robot d'assistance.

Nature des systèmes

Les systèmes présentés précédemment peuvent se classer en différentes catégories, selon leur utilisation. Ainsi, 56% des travaux emploient uniquement un robot de téléopération (le robot da Vinci ou les Berkeley Surgical Robots) pour réaliser l'acquisition de données. En effet, l'utilisation de ce type de système est fort intéressante car elle permet d'améliorer le confort d'utilisation grâce à une console séparée, améliorant également la précision manuelle du chirurgien durant l'intervention. Nous remarquons que parmi ces travaux, seulement une publication décrit l'utilisation d'un robot de téléopération qui n'est pas le système da Vinci de chez Intuitive Surgical.

Néanmoins, il existe certaines alternatives, et notamment des robots conçus pour rester à proximité du patient. C'est le cas des robots de comanipulation utilisés dans 12% des travaux de la littérature. Ces robots permettent d'augmenter la dextérité du clinicien grâce à l'utilisation d'un dispositif actif, généralement sous la forme d'un instrument robotisé. Ce type de système s'utilise en parallèle ou dans le prolongement du membre supérieur, permettant d'assister la manipulation des instruments chirurgicaux.

L'acquisition de données peut se faire également aux moyens de systèmes moins perfec-

tionnés qu'un robot. C'est le cas dans 20% des travaux où les auteurs ont décidé d'utiliser différents systèmes instrumentés tels que des gants et outils médicaux, ainsi que des capteurs de type accéléromètre, afin d'acquérir des données sans encombrer le bloc opératoire. Enfin, dans le cadre de l'entraînement chirurgical, seulement 12% des travaux emploient un système d'acquisition de type simulateur. À noter qu'un seul des travaux de la littérature aborde l'acquisition de données sur deux systèmes que sont un robot de téléopération et un simulateur [Dosis 2005].

Nature des données

La nature des données correspond à leur représentation physique. Dans les travaux de la littérature étudiant la reconnaissance des gestes chirurgicaux, une majorité atteignant 88% recueille une partie des données liées aux trajectoires d'instruments ou de parties anatomiques humaines mises en jeu au cours de l'intervention. Nous entendons ici par trajectoire l'ensemble ou un sous-ensemble des variables géométriques acquises durant un mouvement dans l'espace cartésien. Ainsi les informations de position, rotation, vitesse (linéaire et/ou angulaire), accélération (normale et/ou tangentielle), jerk, et toutes autres variables se rapportant au déplacement d'un objet dans l'espace sont entendus.

Complémentaire aux trajectoires, la mesure de forces fournit des informations importantes sur l'exécution du geste. Ainsi, deux des publications retenues emploient uniquement des capteurs d'efforts pour analyser les forces et couples exercés au cours de la procédure [Rosen 2001, Rosen 2002]. Très proche de la mesure d'efforts, Saggio et al. [Saggio 2011] présentent leur gant instrumenté, équipé de capteurs permettant de mesurer la pression exercée par chaque doigt durant la manipulation. Enfin, 16% des travaux utilisent des données vidéo pour effectuer la reconnaissance des gestes en se basant sur l'analyse d'images peropératoires.

La technologie embarquée dans certains des systèmes présentés précédemment permet également d'effectuer l'acquisition de données multimodales, afin d'améliorer la perception de l'environnement. Ainsi, 16% des travaux combinent l'acquisition de trajectoire et d'efforts afin de reconnaître les gestes chirurgicaux, et 75% de ceux qui utilisent la vidéo l'agrément d'informations de trajectoire. C'est notamment le cas de Tao et al. [Tao 2013], Zappella et al. [Zappella 2013] et Lea et al. [Lea 2015] qui acquièrent les trajectoires et vidéos provenant du robot da Vinci afin de reconnaître les gestes exécutés. Bien que l'utilisation de données multimodales impose certaines contraintes introduites précédemment dans ce chapitre, les travaux de l'université de John Hopkins prouvent que la fusion d'information améliore significativement la reconnaissance des gestes [Zappella 2013, Lea 2015].

Référentiels d'acquisition

L'acquisition de données dépend également du référentiel choisi. Ainsi, les mesures de trajectoires ou de forces peuvent s'effectuer à plusieurs emplacements. Dans la littérature abordant la reconnaissance des gestes chirurgicaux, 84% des travaux utilisent l'acqui-

tion d'informations provenant au moins de l'instrument ou de l'outil chirurgical. Deux exemples représentatifs sont ceux des récents travaux de Rosen et al. [Rosen 2006] utilisant la mesure de trajectoires et de forces provenant de l'instrument, ainsi que les travaux de Holden et al. [Holden 2014] mesurant la position ainsi que l'orientation sous forme de quaternion des mouvements de l'aiguille.

Cependant, l'acquisition de données peut également s'effectuer au niveau de l'interface de commande, principalement dans le cadre de l'utilisation de robots de téléopération. Lin et al. [Lin 2005, Lin 2006] ont été les premiers à étudier cette thématique en utilisant les données du robot da Vinci. De manière complémentaire, la vidéo ciblant le référentiel de la tâche à réaliser peut s'utiliser en complément des données précédentes. Ainsi Tao et al. [Tao 2013], Zappella et al. [Zappella 2013] et Lea et al. [Lea 2015] se sont appuyés sur les référentiels des instruments, des interfaces mais également de la tâche à exécuter.

Néanmoins, l'acquisition de trajectoires peut s'effectuer dans un autre référentiel. Ainsi, Ahmadi et al. [Ahmadi 2009] ont acquis des données provenant d'accéléromètres placés à la taille et aux poignets du chirurgien. De manière similaire, Nomm et al. [Nomm 2008] ont placé des capteurs sur le corps, sur le bras et sur le poignet du chirurgien afin d'en reconnaître les gestes. Enfin, seuls Saggio et al. [Saggio 2011] ont introduit l'utilisation d'un gant équipé de capteurs permettant la mesure des trajectoires des mains ainsi que la mesure de pressions exercés par les doigts. Nous constatons alors que l'acquisition de données avec plusieurs référentiels permet d'étendre la perception de l'environnement afin de favoriser la compréhension et la reconnaissance des gestes chirurgicaux.

Spécifications des données

Le dernier descripteur permettant de décrire l'acquisition est la spécification des données. Ici, trois critères sont utilisés : la fréquence d'acquisition, le nombre de signaux acquis et la résolution des systèmes de capture d'images. Il est nécessaire de préciser que le nombre de signaux correspond au nombre de variables acquises durant l'expérience, provenant d'un seul instrument chirurgical, d'une seule main ou d'un seul membre.

Par ce tableau, nous remarquons que l'acquisition des données est réalisée à des fréquences comprises entre 5Hz et 100Hz. Cependant, d'après Loram et al. [Loram 2006], les mouvements humains oscillent entre 0Hz et 10Hz, avec un maximum de 4.5Hz pour les mouvements de la main [Xiong 2006]. Ainsi, une acquisition à 30Hz serait suffisante car elle permet de respecter le théorème d'échantillonnage de Nyquist-Shannon, tout en capturant l'ensemble des fréquences fondamentales nécessaires à la description de mouvements humains.

Concernant le nombre de signaux, cette valeur est disparate et dépend du système d'acquisition considéré, ainsi que des informations disponibles. Cependant, nous constatons que là où certains travaux se contentent d'utiliser 7 variables [Holden 2014], d'autres en utilisent 39 afin d'effectuer leur reconnaissance [Tao 2012].

Enfin, lorsque le système en est équipé, l'enregistrement vidéo est acquis à des résolutions

de 320x240 [Béjar Haro 2012, Tao 2013, Zappella 2013] et de 640x480 [Lea 2015], suffisantes pour la détection d'informations mais également de tailles raisonnables afin de limiter les temps de calcul.

2.2.2 Processus de reconnaissance

Le processus de reconnaissance se base sur les notions introduites précédemment dans ce chapitre. Afin de présenter l'ensemble des méthodes utilisées, nous avons décomposé notre analyse en quatre descripteurs que sont le prétraitement des données, l'extraction d'attributs, la méthodologie employée pour l'apprentissage et les spécifications du processus de reconnaissance. L'ensemble de ces informations est présenté dans le tableau 2.2.

2.2.2.1 Prétraitement des données

Suite à l'acquisition de données, il est parfois nécessaire d'appliquer des techniques de traitement de signal et d'image afin d'améliorer la qualité des données et d'augmenter leur capacité de discrimination pour l'étape de reconnaissance. Afin de décrire le prétraitement des données, nous avons considéré quatre critères : la présence d'une technique de filtrage, l'utilisation d'un fenêtrage temporel des données, l'utilisation d'une méthode de normalisation et enfin la possibilité de réduire la quantité d'informations en sortie en employant une technique de réduction de dimension.

Filtrage

L'utilisation de techniques de filtrage permet de réduire la quantité de bruit contenu dans le signal. En analysant les travaux du domaine, seules deux publications font référence à l'emploi d'un filtre. Pour Holden et al. [Holden 2014], l'utilisation d'un filtre à moyenne glissante permet de réduire la quantité de bruit blanc contenu dans les signaux acquis. Pour Van Den Berg et al. [Van Den Berg 2010], l'utilisation d'un filtre de Kalman basé sur une représentation d'état du système considéré permet de lisser les trajectoires, grâce à la double fonction prédiction/correction de la méthode, dans le cadre de la planification de trajectoire pour l'exécution de gestes.

Néanmoins, nous supposons que la plupart des travaux utilisent des techniques de filtrage afin d'améliorer le conditionnement des données et d'augmenter le rapport signal sur bruit. Cependant aucune information n'est fournie dans les autres publications.

Fenêtrage temporel

Le fenêtrage temporel est une technique permettant d'associer plusieurs échantillons, temporellement proches, dans le but de construire un unique échantillon reprenant les caractéristiques de chacun. Dans la littérature, seuls les travaux de Lin et al. [Lin 2005, Lin 2006], Reiley et al. [Reiley 2010], Varadarajan et al. [Varadarajan 2009] et Tao et al.

Réduction de dimension

Généralement, l'emploi d'un nombre important de signaux implique une quantité importante d'informations à traiter. De plus, toutes ces informations ne sont pas nécessairement représentatives du geste chirurgical. L'intérêt de la réduction de dimension est de trouver un espace, généralement de dimension inférieure, permettant de représenter l'ensemble des données afin d'augmenter leur degré de discrimination. L'une des techniques employées est l'Analyse Discriminante Linéaire. Cette méthode permet la classification, mais également l'identification de caractéristiques représentant les différents groupes, sur la base de variables explicatives [Lin 2005, Lin 2006, Reiley 2008].

2.2.2.2 Extraction d'attributs

Comme présenté précédemment dans ce chapitre, l'extraction d'attributs permet de représenter un signal ou un ensemble de signaux dans un espace de dimensions fixes, en extrayant certaines caractéristiques. L'objectif est d'augmenter le degré de discrimination des données afin d'améliorer les performances de classification. Dans la littérature concernant la reconnaissance des gestes chirurgicaux, plusieurs types de transformations ont été employés. Bien que certains attributs soient spécifiques aux travaux développés, d'autres sont plus couramment utilisés. C'est notamment le cas des caractéristiques spatio-temporels. Ces attributs permettent de caractériser une image ou une vidéo avec un fort degré de discrimination, en prenant en compte l'aspect temporel [Béjar Haro 2012, Tao 2013, Zappella 2013, Lea 2015]. Dans les travaux de Ahmidi et al. [Ahmidi 2013], l'utilisation combinée d'une transformation dans le repère de Frénet et d'un encodage descriptif pour les courbes permet de représenter les variations de trajectoires en un vecteur spécifique. Grâce à ce vecteur, la reconstruction du mouvement est également possible. D'autres travaux comme ceux de Dosis et al. [Dosis 2005] présentent une extraction d'attributs peu commune en se basant sur la description des rotations de l'avant-bras. Ces caractéristiques, plus spécifiques, permettent néanmoins d'obtenir de bons résultats dans leur contexte d'étude.

2.2.2.3 Méthodologie d'apprentissage

La méthodologie d'apprentissage est composée de deux critères qui sont les algorithmes utilisés pour la reconnaissance des gestes ainsi que le critère de connaissance a priori.

Algorithmes de reconnaissance et de reconstruction

L'état de l'art des travaux présente un ensemble d'algorithmes pour la classification qui peuvent être composés de différentes étapes distinctes (segmentation et classification), se retrouvant parfois fusionnées pour former un algorithme dit hybride. Parmi les travaux de la littérature, 72% des travaux emploient des méthodes statistiques basées sur

les réseaux bayésiens et leurs variantes temporelles, avec parmi eux, 56% qui utilisent des Modèles de Markov Cachés dans lequel le système modélisé est supposé être un processus markovien de paramètres inconnus. Dans ce cas, les états ne sont pas accessibles mais sont représentés par des observations. Ce procédé permet de représenter l'ensemble des états et leurs probabilités de transitions en prenant en compte l'aspect temporel de l'exécution des gestes. Dans ce cadre nous pouvons citer les travaux de Rosen et al. [Rosen 2001, Rosen 2002] ainsi que ceux de Padoy et al. [Padoy 2011]. Il est à noter que ce type d'algorithme est également régulièrement utilisé pour la reconnaissance d'activités et phases chirurgicales. En effet, la création du modèle mathématique est directement extraite du processus chirurgical et les paramètres sont ensuite calculés grâce à l'utilisation d'algorithmes tiers.

Outre les variantes des réseaux bayésiens, certaines approches utilisent de manière intéressante les machines à vecteurs de Support [Béjar Haro 2012, Zappella 2013] afin de résoudre le problème de discrimination. Le principal avantage de cette technique réside dans la possibilité d'utiliser un ensemble de noyaux pour la projection des données afin de trouver l'espace le plus séparateurs pour résoudre le problème de classification. D'une manière plus simplifiée, l'Analyse Discriminante Linéaire utilisé par Lin et al. [Lin 2005, Lin 2006], Reiley et al. [Reiley 2008] ainsi que Varadarajan et al. [Varadarajan 2009] permet également de résoudre le problème de discrimination. Enfin, un algorithme très employé pour le regroupement non-supervisé de données est l'algorithme des K-Moyennes. Cet algorithme peut aussi bien s'utiliser pour segmenter [Holden 2014] que pour classer les échantillons [Zappella 2013].

Enfin, certains travaux abordent la reconnaissance mais également la reconstruction du geste chirurgical à des fins d'évaluation, mais également d'assistance. Pour ce faire, Reiley et al. [Reiley 2010], Van Den Berg et al. [Van Den Berg 2010] ainsi que Padoy et al. [Padoy 2011] utilisent l'algorithme de Déformation Temporelle Dynamique permettant de calculer les variations spatio-temporelles existantes entre deux séries temporelles. Pour les autres algorithmes plus rarement employés, nous renvoyons le lecteur au tableau 2.2 et aux travaux correspondants.

Connaissances a priori

Le critère de connaissances a priori fournit une information sur les modes d'apprentissage utilisés pour le processus de reconnaissance. Dans le cas des travaux se basant sur des Modèles de Markov Cachés, la majorité d'entre eux nécessitent de connaître les couples d'entrée-sortie pour réaliser un apprentissage supervisé. De plus, il est nécessaire de définir la structure du réseau en s'appuyant sur des connaissances a priori concernant la tâche à réaliser. Dans ce cas, nous considérons que l'implémentation requiert une complète connaissance du contexte d'exécution des gestes. C'est notamment le cas dans les travaux de Dosis et al. [Dosis 2005], Padoy et al. [Padoy 2011], Saggio et al. [Saggio 2011] ainsi que Lea et al. [Lea 2015].

À l'inverse, certains processus de reconnaissance des gestes chirurgicaux ne nécessitent pas d'avoir toutes les informations pour la mise en place de l'apprentissage. C'est notamment le cas avec les travaux de Holden et al. [Holden 2014] qui adoptent une approche mélangeant apprentissage supervisé et regroupement non-supervisé, également utilisée par Rosen et al. [Rosen 2001, Rosen 2002, Rosen 2006].

2.2.2.4 Spécification du processus

La spécification permet de quantifier la complexité du processus. Elle se base sur deux critères que sont le nombre d'étape du processus et la difficulté du paramétrage pour la mise en place de la méthodologie.

Nombre d'étapes

De par la variabilité du nombre d'étapes employées afin de construire le processus, la discussion de ce critère est complexe et dépend de l'approche adoptée par les auteurs. Néanmoins, dans la plupart des travaux le processus général de segmentation présenté est retrouvé, avec toutefois quelques ambiguïtés. Ainsi, certains travaux emploient des méthodes hybrides de segmentation et classification, ne permettant pas de distinguer séparément chaque composante. Dans le cas des chaînes de Markov Cachés, la segmentation est généralement associée à la classification. Dans le cas où le nombre d'étapes est faible [Nomm 2008, Tao 2012], cette distinction peut s'avérer complexe. À l'inverse, lorsque le nombre d'étapes augmente, il est plus simple de distinguer chaque composante du processus [Ahmidi 2013, Holden 2014].

Difficulté du paramétrage

La difficulté du paramétrage est un critère qualitatif que nous avons défini afin d'appréhender la complexité des développements réalisés dans les travaux de la littérature. Il permet de fournir une indication sur la complexité du processus développé mais également la difficulté du paramétrage. Ce critère garde un lien étroit avec le nombre d'étapes incluses dans le processus. Ainsi, plus le nombre d'étapes est important, plus le paramétrage et la mise en place de la méthodologie peuvent s'avérer difficiles. On peut voir que dans certains travaux le paramétrage et la mise en place restent très simples [Hundtofte 2002, Novi 2012], alors que dans d'autres, le processus complet de reconnaissance est également agrémenté d'un processus de reconstruction qui complexifie l'implémentation [Reiley 2010, Padoy 2011].

2.2.3 Processus de validation

Afin de valider la qualité des travaux présentés dans cet état de l'art, nous souhaitons analyser les processus de validation utilisés. Pour ce faire, nous définissons deux descripteurs que sont les conditions expérimentales et la stratégie de validation employée pour

2.2. ETAT DE L'ART DES TRAVAUX POUR LA RECONNAISSANCE DES GESTES CHIRURGICAUX

	Condition expérimentale											Stratégie de validation																											
	Type de tâche							Taille du vocabulaire	Niveau de granularité		Nombre de candidats	Niveau d'expertise		Nombre de sessions	Type		Méthode référence	Nature référence	Modèle		Métrique																		
	Suture	Néod	Point	Incision	Dissection	Transfert	Insertion	Déplacement	Choix/système	Fondoplicature	Vertébroplastie	Tâche	Surgente	Devence	Expert	Intermédiaire	Novice	Vérification	Validation	Évaluation	Segmentation expert	Segmentation en ligne	Vidéo	Trajectoire	Entrain/Test	Validation croisée	Reconnaissance	Recouvrement	Confusion	Vraisemblance	Diversité	Pureté	Déviaton	Temps					
[Ahmadi2009]								x		x	2 13 14	x		1 1 1	x	x	1 10 5	x	x	x			x			x													
[Ahmadi2013]	x	x									10 7	x		8 18	x	x	39 110	x	x	x	x		x			x	x												
[BejarHaro2012]	x	x									15 15 15	x		8 8 8	x	x	39 26 36	x	x	x	x		x			x	x												
[Bodenstedt2012]												x		10				x			x																		
[Dosis2005]	x	x									8 8 8	x	x	12 12 12	x	x	59 84 57	x	x	x	x				x		x												
[Holden2014]											5 5			22 12	x	x	88 12				x				x		x	x											
[Hundthofe2002]											6 5	x		5 5			50 15	x	x				x		x		x	x											
[Kragic2003]											8							x																					
[Lea2015]	x	x									15 15 15	x		8 8 8	x	x	39 28 36				x					x	x												
[Lin2005], [Lin2006]	x										8	x		2	x	x	27	x			x					x	x												
[Nomm2008]											5	x					110	x			x						x	x											
[Novi2012]	x										10	x	x	3	x		3	x			x				x	x													
[Padoy2011]	x										6 12	x		2 2				x			x				x												x	x	
[Reiley2008]	x										11	x		7	x		15	x			x					x	x												
[Reiley2010]	x	x									16 16 16	x		11 11 11	x	x		x			x																	x	
[Rosen2001], [Rosen2002]											14 14	x		10,8 10	x	x	10,8 10	x			x																		x
[Rosen2006]		x									15	x		30	x	x	15	x			x																		
[Saggio2011]	x	x	x											15 15 15 15 15	x	x																							
[Tao2012]	x	x									11 11 11	x		8 8 8	x	x	39 26 36	x	x	x						x	x												
[Tao2013]	x	x									15 15 15	x		8 8 8	x	x	39 26 36	x	x	x						x	x												
[VanDenBerg2010]		x									3 0	x		3 3			3 3	x			x																	x	x
[Varadarajan2009]	x										11	x	x	8	x	x	30				x	x				x	x												
[Zappella2013]	x	x									15 15 15	x		8 8 8	x	x	39 26 36	x	x	x						x	x												

Tableau 2.3: Résumé des processus de validation employés la littérature pour qualifier et quantifier la reconnaissance des gestes chirurgicaux.

valider les résultats. Le tableau 2.3 fournit un résumé des différents processus de validation employés.

2.2.3.1 Conditions expérimentales

Les conditions expérimentales regroupent toutes les informations permettant de définir l'application des travaux incluant les types de tâches, la taille du vocabulaire de gestes et le niveau de granularité utilisés dans les études, le nombre de candidats ainsi que leurs niveaux d'expertise et le nombre de sessions acquises pour construire la base de données.

Types de tâches

Le type de tâche envisagée pour la validation des travaux dépend fortement du disposi-

tif utilisé lors des expériences. Nous constatons deux grandes familles de données : des données acquises durant des tâches d'entraînement, représentant 88% des travaux, et des données acquises durant des interventions sur animaux ou sur patients (12%).

Dans le cadre d'acquisition durant des sessions d'entraînement, nous remarquons que les tâches les plus récurrentes sont la suture, la réalisation d'un nœud ainsi que le transfert d'objet, et plus particulièrement d'aiguille. Effectivement, deux de ces trois tâches font partie intégrante du programme de formation adopté par la SAGES [Derossis 1998].

Pour les études employant des données acquises en environnement in-vivo [Rosen 2001, Rosen 2002, Rosen 2006, Ahmidi 2013], ce sont les procédures de cholécystectomie par voie laparoscopique, fundoplicature et vertébroplastie qui ont été retenues.

Taille du vocabulaire

Dans le cadre de la reconnaissance des gestes chirurgicaux, la taille du vocabulaire utilisé pour la description des actions a un impact direct sur la complexité du processus de validation. Ainsi, reconnaître un geste parmi un dictionnaire de gestes contenant uniquement deux surgemes n'est que peu significatif. Cependant, certains travaux présentent des dictionnaires conséquents avec 15 voire 16 surgemes, permettant ainsi d'inclure plus de variabilité et de diversité dans l'exécution [Reiley 2010, Béjar Haro 2012, Tao 2013, Zappella 2013, Lea 2015].

Niveaux de granularité

Bien que les travaux présents dans cet état de l'art traitent tous au minimum la reconnaissance de surgemes, certains s'intéressent également à la reconnaissance de niveaux de granularité inférieurs (le dexeme), mais également supérieurs (la tâche). C'est notamment le cas pour Novi et al. [Novi 2012] qui s'attachent à la reconnaissance de tâches, surgemes et dexemes. Pour Varadarajan et al. [Varadarajan 2009] et Dosis et al. [Dosis 2005], la reconnaissance est réalisée au niveau des surgemes et également des dexemes. Enfin, seuls 8% des travaux réalisent une reconnaissance multi-granularité.

Nombre de candidats, niveaux d'expertise et nombre de sessions

L'apprentissage nécessite d'acquérir une quantité importante de données contenant une variabilité non nulle, afin de généraliser au mieux l'ensemble de la population représentée. Pour ce faire, l'acquisition de données doit faire intervenir un maximum de candidats, avec des niveaux de compétences différents. Dans les travaux de la littérature, 89% d'entre eux ont précisé ces informations dans leurs publications et ont réalisé des acquisitions avec différents intervenants médicaux (généralement des chirurgiens) de niveaux de compétences variés. Ces niveaux de compétences sont caractérisés par Reiley et al. [Reiley 2009a], où un chirurgien est considéré comme expert après 100 heures de pratique sur un robot da Vinci, là où un chirurgien intermédiaire (résident) doit avoir une pratique comprise entre 0 et 100 heures. Le novice est alors décrit comme un candidat n'ayant au-

cune compétence médicale ni aucune expérience dans le domaine de la robotique. De plus, la grande majorité de ces travaux contiennent un nombre important de sessions, généralement au-delà de 25 sessions par tâche d'entraînement, nécessaire à la généralisation des performances de reconnaissance et à son évaluation sur une population variée [Dosis 2005, Béjar Haro 2012, Tao 2012, Tao 2013, Zappella 2013, Lea 2015].

2.2.3.2 Stratégie de validation

La stratégie de validation décrit le procédé employé pour qualifier et quantifier les performances de reconnaissance. Nous avons défini cinq critères pour caractériser les stratégies adoptées : le type d'analyse effectuée, la méthode de référence sélectionnée, la nature des références, le modèle de validation employé et les métriques d'évaluation utilisées pour quantifier les performances.

Types d'analyses

Jannin et al. [Jannin 2006] distinguent trois types d'analyses : La vérification, la validation et l'évaluation. La vérification consiste à analyser si un système est construit en accord avec son cahier des charges. La validation consiste à analyser un système et quantifier que ce dernier respecte ses fonctions principales. Enfin, l'évaluation consiste à analyser un système et quantifier son acceptation auprès de l'utilisateur final, durant la réalisation de ses fonctions principales. Dans cet état de l'art, seuls Ahmadi et al. [Ahmadi 2009], Lea et al. [Lea 2015] ainsi que Varadarajan et al. [Varadarajan 2009] évaluent les performances de leurs approches dans le contexte d'application présenté. Dans le cas d'Ahmadi et al. [Ahmadi 2009], l'évaluation est réalisée lors d'une intervention réelle, alors que dans les deux autres cas les auteurs ont évalué et quantifié les performances de reconnaissance dans un cadre d'entraînement, qui est l'environnement ciblé pour leur application. D'un autre côté, 75% des travaux ont effectué une vérification de leur processus de reconnaissance dans des conditions d'entraînement en se basant sur des métriques d'évaluation variées. Enfin, 17% ont effectué une simple vérification en décrivant de manière qualitative les performances de leur méthode. Enfin, seulement 12% des travaux ont réalisé une double étude de leur méthode. C'est notamment le cas de Holden et al. [Holden 2014] et de Van Den Berg et al. [Van Den Berg 2010] où deux études ont été réalisées dans chacune des publications, la première servant de vérification du procédé et la deuxième permettant de quantifier dans un environnement simulé, proche de la réalité, les performances de leur méthode.

Méthodes de référence

Afin de quantifier les performances d'une méthode, il est généralement nécessaire de comparer les résultats obtenus avec une référence, souvent appelée vérité terrain. Dans les travaux de la littérature, 96% d'entre eux effectuent une comparaison des résultats avec une segmentation des gestes réalisée par un ou plusieurs experts. Par cette approche, il est fait

l'hypothèse que ces experts possèdent les compétences et les connaissances nécessaires à la réalisation de la tâche. Cependant, la réalisation d'une segmentation manuelle a posteriori entraîne inévitablement des incertitudes et un biais dans la construction de cette vérité terrain. Ainsi, uniquement un travail de la littérature effectue une comparaison avec une segmentation en ligne réalisée manuellement. Hundtofte et al. [Hundtofte 2002] proposent alors aux candidats de réaliser les gestes et d'indiquer, durant l'exécution, les périodes de transitions d'un geste à l'autre. Dans ce cas, bien que ce type de système puisse inclure un délai de réaction, la volonté de l'utilisateur coïncide avec la segmentation des gestes qui en est faite, permettant de réduire au maximum l'incertitude et le biais durant la construction de la vérité terrain.

Natures des références

La nature des références fait appel à la notion de nature de données définies précédemment, mais utilisées cette fois-ci pour construire la référence nécessaire à l'évaluation. De ce fait, nous distinguons deux natures. D'une part la vidéo, qui est employée seule dans 82% des travaux construisant une référence pour l'analyse des performances. D'autre part, 13% des travaux segmentent manuellement les trajectoires afin d'annoter les différents surgemes. Enfin, seule une publication (5%) décrit l'utilisation d'une segmentation double provenant de la vidéo et des trajectoires [Dosis 2005].

Modèles de validation

Le modèle de validation détermine le processus utilisé pour quantifier la validité externe de la méthode. Le premier modèle employé dans les travaux de reconnaissance des gestes chirurgicaux est la séparation des données en deux bases distinctes : la base d'entraînement et la base de test. Le modèle est entraîné sur une base, et testé sur une autre dont les données n'ont jamais été observées auparavant. Cette approche n'est que peu utilisée dans les travaux de la littérature [Dosis 2005, Varadarajan 2009]. Le second modèle majoritairement employé est la validation croisée. Cette technique consiste à reprendre le concept de la première méthode, à savoir de séparer les données en deux sous-ensembles. Le premier sous-ensemble est destiné à l'entraînement du modèle et le second sert à quantifier les performances de reconnaissance. Cependant, l'approche de validation croisée réalise ensuite une rotation entre les données d'entraînement et les données d'évaluation, cela afin de tester l'ensemble des combinaisons possibles d'entraînement et d'évaluation. La validation croisée peut se faire en testant une session ou bien sur un utilisateur spécifique. Cette approche favorise la confiance que l'on peut avoir dans les résultats de reconnaissance.

Métriques de validation

La métrique de validation est une fonction permettant de mesurer la distance entre les résultats obtenus en sortie du processus de reconnaissance et la vérité terrain construite

préalablement. La plus utilisée dans la littérature, à hauteur de 67%, est le taux de reconnaissance, calculé à partir du ratio entre le nombre d'échantillons bien détectés sur le nombre d'échantillons global. Néanmoins, utilisée seule, cette métrique n'est pas complètement représentative des performances. L'emploi de l'écart-type en complément permet alors de mieux quantifier les performances [Holden 2014].

D'autres travaux utilisent la mesure de recouvrement, autorisant le calcul de la similarité de leur détection avec la vérité terrain correspondante [Lin 2005, Lin 2006, Nomm 2008, Varadarajan 2009, Tao 2013, Holden 2014]. De manière différente, les matrices de confusion permettent d'évaluer les problèmes de classification et d'utiliser les métriques de rappel et précision pour quantifier le degré de confusion entre les gestes. Cependant, le rappel et la précision ne sont pas utilisés dans les travaux de la littérature.

Enfin, la mesure de déviation, aussi bien en pixel qu'en millimètre, permet de fournir des informations complémentaires sur les performances de la méthode. D'autres mesures isolées sont également employées par Ahmadi et al. [Ahmadi 2009] ainsi que Reiley et al. [Reiley 2010].

2.2.4 Analyse et discussion

L'état de l'art des travaux abordant la thématique de reconnaissance des gestes chirurgicaux soulève de nombreux points de discussion. Ces points portent aussi bien sur les conditions d'étude qui ont été considérées que sur la méthodologie mise en place pour résoudre la problématique, faisant appel à une stratégie de validation afin de qualifier et quantifier les performances des méthodes proposées.

2.2.4.1 Conditions d'étude

La chirurgie laparoscopique est en constant développement, accéléré notamment par l'utilisation de robots chirurgicaux au sein du bloc opératoire. Bien que ces robots puissent être de différentes natures, les systèmes de téléopération restent tout de même prépondérants. Leur intérêt réside dans le confort et la précision qu'ils amènent à l'utilisation. Dans le cas du robot da Vinci, bien que son apport pour le patient soit encore discutable, il fournit néanmoins des avantages avérés pour le chirurgiens dans le cadre d'interventions laparoscopiques. De plus, la mise en place du kit de développement proposé par Kazanzides et al. [Kazanzides 2014] ouvre les capacités de ce robot. Cela simplifie alors l'acquisition de données, aussi bien dans un contexte clinique que dans un environnement d'entraînement.

Cependant, l'utilisation du da Vinci dans le cadre de la recherche est très limitée car l'accès au robot et le coût d'achat freinent drastiquement les possibilités. Des alternatives ont récemment émergé permettant d'accéder, à moindre coût, aux informations provenant des capteurs embarqués dans des robots aux architectures similaires. C'est notamment le cas du robot de téléopération Raven, développé en partenariat par les universités de Washing-

ton et Santa-Cruz [Hannaford 2013], conçu spécifiquement pour la téléopération dans le cadre de la chirurgie laparoscopique robotisée. De cette manière, l'acquisition des données est facilitée, centralisée et évite le développement d'instruments spécifiques comme ceux présentés dans la littérature [Rosen 2001, Rosen 2002], ou bien l'utilisation de systèmes d'acquisition encombrants pour le chirurgien et pouvant perturber la pratique chirurgicale [Ahmadi 2009]. Il est cependant à noter que le Raven a été initialement conçu pour la recherche en robotique chirurgicale. Son utilisation sur patient avec les versions existantes est donc impossible, ne laissant la place qu'à une utilisation dans le cadre de l'entraînement en chirurgie robotisée.

Un deuxième point à soulever est la nature des données à acquérir et du ou des référentiel(s) à considérer. D'après Zappella et al. [Zappella 2013] ainsi que Lea et al. [Lea 2015], qui avaient accès aux informations de trajectoire des instruments et interfaces ainsi qu'à la vidéo endoscopique du da Vinci, la fusion de ces deux modalités fournit d'excellents résultats pour la capture et la reconnaissance des gestes chirurgicaux. La perception de l'environnement est quasiment complète (il manquerait l'acquisition audio) et les informations sur les différents surgemes réalisés arrivent à être extraites, interprétées et reconnues. Ainsi l'acquisition de données multi-référentiels semblent fournir des résultats intéressants, permettant d'augmenter la perception de la scène et d'améliorer la compréhension de l'environnement ainsi que des événements. Toutefois, l'emploi de différentes modalités nécessite une synchronisation précise des différents signaux. Dans le cas où les modalités sont acquises par un robot, les délais d'acquisition sont réduits et des techniques de synchronisation sont mises en place afin de rendre cette stratégie d'acquisition viable. Il est à noter que le cadre de ce type de travaux reste privilégié et difficilement accessible par les autres équipes de recherche.

2.2.4.2 Processus de reconnaissance

L'attrait clinique et scientifique pour l'assistance aux gestes chirurgicaux a permis d'explorer un nombre conséquent de méthodes destinées à la fouille de données et à l'extraction de connaissances, afin de segmenter et classer les gestes chirurgicaux. Cependant, les étapes précédant la reconnaissance ne sont pas à négliger.

Un point également intéressant à soulever est le faible intérêt d'utilisation d'outils pour le prétraitement des données, et particulièrement le filtrage. La qualité des données utilisées en entrée d'un processus de segmentation suivi de classification est déterminant pour obtenir de bonnes performances. Lors de l'acquisition de données cinématiques ainsi que d'efforts, bien que les systèmes robotisés soient équipés de capteurs généralement répétables, dont la résolution permet une mesure précise des informations, l'emploi d'une chaîne d'acquisition de mesures inclut un bruit systématique. De ce fait, l'étape de filtrage semble inévitable afin d'augmenter le rapport signal sur bruit. Néanmoins, très peu de travaux aborde ce point, pourtant clé en mécatronique.

De manière similaire, très peu de publications stipulent l'emploi de techniques pour la

normalisation des signaux. Cela reste surprenant car la variabilité intrapersonnelle est rarement nulle, d'autant plus que vient s'ajouter la variabilité interpersonnelle dans le cas d'études impliquant plusieurs opérateurs avec des niveaux de compétence variés. Bien que l'extraction d'attributs permette de réduire cet effet, la normalisation reste le moyen le plus direct pour gérer le problème. Ainsi, Lin et al. [Lin 2005, Lin 2006] exploitent une méthode très intéressante pour transformer chaque signal acquis dans un référentiel commun, par utilisation de la moyenne et de la variance. L'avantage majeur de cette méthode est que les données obtenues sont indépendantes de l'unité ou de l'échelle choisie. De plus, cela n'a aucune incidence sur les profils de variation inclus à l'intérieur des signaux, et les corrélations existantes entre les signaux avant filtrage demeurent inchangées.

Un autre point d'analyse concerne la réduction de dimension. Typiquement, pour Lin et al. [Lin 2005, Lin 2006] et Varadarajan et al. [Varadarajan 2009], le nombre de variables cinématiques acquises est important et une redondance dans les données reste envisageable. De ce fait, l'utilisation de méthodes pour la réduction de dimension semble intéressante. De cette manière, les données sont projetées dans des espaces de dimensions différentes, afin de sélectionner les dimensions représentant au mieux l'organisation et la répartition des données. Cependant, les notions physiques des données (position, vitesse, accélération, etc.) sont alors perdues, remettant en question l'intérêt de l'acquisition d'un nombre conséquent de signaux. Une solution peut être d'utiliser ces données cinématiques et de trouver une représentation permettant de réduire la dépendance spatiale afin d'assurer, quel que soit le type de mouvements réalisés, l'obtention de données spatialement similaires (l'aspect temporel reste alors dépendant de l'opérateur). Ainsi, l'étape de transformation n'a plus lieu d'être et permet de conserver le sens physique des signaux, augmentant au passage la vitesse de calcul pour une possible application en temps-réel.

Abordant l'aspect temps-réel, la plupart des méthodes de reconnaissance employées dans la littérature permettent ce type d'exécution, mais nécessitent la création et l'entraînement d'un modèle au préalable. Dans cet objectif, Padoy et al. [Padoy 2011] ainsi que Holden et al. [Holden 2014] laissent envisager des possibilités d'applications intéressantes de par l'utilisation en temps-réel de processus de reconnaissance pour respectivement automatiser certains gestes chirurgicaux réalisés avec un robot, et également réaliser la détection des gestes pour la compréhension du processus médical, la formation et l'évaluation des compétences techniques. Néanmoins, ces variantes des réseaux bayésiens dynamiques restent problématiques car il faut définir en amont la structure du réseau (bien que certains algorithmes permettent d'inférer une structure à partir d'un ensemble de données) et prendre en compte l'évolution temporelle lors de l'exécution d'un geste. Ainsi, un modèle mal conçu ou paramétré ne pourra en aucun cas détecter l'enchaînement de certains gestes puisque les probabilités liant lesdits gestes seront nulles. Des représentations par sac de mots semblent alors plus intéressantes d'un point de vue reconnaissance. L'ensemble des gestes que l'on souhaite reconnaître est appris et la séquence de détection est indépendante du temps et du geste considéré, améliorant les performances de reconnais-

sance du processus.

2.2.4.3 Processus de validation

Les conditions expérimentales et les stratégies d'analyse employées dans la littérature pour la reconnaissance des gestes chirurgicaux sont variées. Cependant, bien que des pistes similaires soient fournies pour structurer le protocole d'analyse dans le cadre du traitement d'images médicales [Jannin 2006], il n'existe actuellement aucune méthode de référence pour évaluer correctement la qualité de reconnaissance des gestes chirurgicaux, évitant tout biais ou incertitudes lors de la comparaison avec une référence. De ce fait, les travaux de la communauté s'orientent de plus en plus vers une approche différente pour l'évaluation du niveau de compétences des chirurgiens afin de valider leur méthode de détection. Pour ce faire, ils comparent leur reconnaissance des gestes, sans validation spécifique, avec le niveau de compétences du chirurgien. Ce dernier est quantifié au travers des méthodes telles qu'OSATS ou GRS présentées dans le chapitre précédent. Leurs évaluations deviennent objectives mais contournent le problème de reconnaissance du geste car leurs détections ne sont pas spécifiquement confrontées avec une vérité terrain contenant les surgemes réalisés. Ce type d'approche a donc été exclue de notre analyse de l'état de l'art. Notons également que les métriques actuellement utilisées permettent de quantifier les performances globales, mais ne fournissent que peu d'indications sur la reconnaissance spécifique de chaque geste. Ainsi, l'utilisation de métriques telles que le rappel et la précision pourrait apporter des informations supplémentaires sur les performances de reconnaissance.

Parmi les travaux que nous avons analysés, la grande majorité d'entre eux réalisent uniquement une validation de leur méthode durant l'entraînement chirurgical, alors que l'objectif vise majoritairement une utilisation en clinique. Néanmoins, bien que l'acquisition de données d'entraînement reste un obstacle aux développements scientifiques, l'acquisition de données en conditions peropératoires requiert des ressources et des développements qui ne sont que peu accessibles, mais tendent à se développer avec l'intégration de nouvelles technologies au sein du bloc opératoire. Des pistes d'évolution ont été récemment proposées pour palier le manque de données. En effet, l'équipe de Hager de l'université de John Hopkins a mis gratuitement à disposition sa base de données d'entraînement acquise avec le robot da Vinci [Gao 2014]. Cette base de données a, par ailleurs, déjà été employée dans certains travaux du domaine [Béjar Haro 2012, Tao 2012, Tao 2013, Zappella 2013, Lea 2015]. Elle contient les acquisitions de trajectoires et vidéos provenant du robot da Vinci lors de la réalisation de tâches d'entraînement en chirurgie laparoscopique telle que la suture, la réalisation d'un nœud et le transfert d'aiguille d'un instrument à l'autre, de manière répétée. Ces données, acquises au travers de différents candidats avec des niveaux d'expertise variés, sont annotées et segmentées pour faciliter le travail des utilisateurs lors de l'apprentissage de leur modèle. Bien que cela soit une avancée importante

dans le domaine, ces données ne servent aujourd'hui qu'à valider la méthodologie de reconnaissance. En effet, à l'heure actuelle il n'existe aucune base de données regroupant des signaux vidéos et trajectoires, construite à partir d'acquisitions réalisées au cours d'interventions sur patient.

Enfin, le dernier point soulevé dans cette discussion concerne l'approche adoptée pour la reconnaissance et les niveaux de granularité étudiés. En mettant de côté l'approche multi-granularité de Novi et al. [Novi 2012], qui repose sur de nombreuses hypothèses et une reconnaissance basée sur la comparaison de seuils définis préalablement à la main, uniquement deux travaux de la littérature s'attèlent à la détection de dexemes [Dosis 2005, Varadarajan 2009]. En analysant la revue de Lalys et al. [Lalys 2014], les approches traitant de la reconnaissance d'étapes chirurgicales abordent cette problématique en utilisant une approche ascendante, en commençant par détecter les différentes activités. Cette reconnaissance d'activités leur permet ensuite de réaliser la reconnaissance d'étapes chirurgicales. Ce type d'approche pour la reconnaissance des gestes n'a été que peu exploré et pourrait présenter certains intérêts pour la détection.

2.3 Conclusion et problématiques de la thèse

Alors que le chapitre 1 aborde le contexte clinique des travaux et présente le besoin de systèmes et méthodes d'entraînement adaptés à la chirurgie robotisée, ce chapitre présente l'état de l'art de la reconnaissance des gestes médico-chirurgicaux. Pour ce faire, une description de la modélisation des processus chirurgicaux permet d'introduire une méthodologie générale, employée pour reconnaître plus spécifiquement les gestes. L'analyse des travaux de la littérature est structurée en trois catégories que sont les conditions d'études, les processus de reconnaissance et les processus de validation utilisés. Chaque catégorie est composée d'un ensemble de critères permettant de décrire et de classer les travaux existants, mettant ainsi en avant les stratégies adoptées pour la détection des gestes médico-chirurgicaux.

Les travaux présentés dans cet état de l'art nous ont permis de soulever plusieurs limitations concernant les méthodes actuellement développées pour la reconnaissance des gestes médico-chirurgicaux. Tout d'abord, nous remarquons que peu d'entre elles se basent sur une approche ascendante pour la reconnaissance des gestes, comme cela peut se faire avec les activités et étapes chirurgicales. La prise en compte des dexemes dans le processus de reconnaissance pourrait cependant apporter un gain d'information et augmenter les performances de reconnaissance. De plus, les algorithmes employés pour la reconnaissance nécessitent généralement l'apprentissage de modèles statistiques structurés et temporels, dépendant de la tâche. La question d'une nouvelle approche ascendante, multi-granularité et indépendante du temps peut alors se poser. Nous souhaitons également prêter une attention particulière à l'acquisition, au prétraitement des données, ainsi

qu'à la validation et l'évaluation de notre approche. Dans le cas de l'acquisition, une partie non négligeable des travaux de la littérature acquièrent des données qui peuvent s'avérer redondantes, voire possiblement non représentatives des gestes réalisés. Nous désirons donc travailler sur la mise en place d'une acquisition permettant d'extraire des signaux pertinents par rapport à notre application d'entraînement. Pour le prétraitement des données, nous souhaitons améliorer la qualité des signaux que nous acquérons en utilisant des techniques de filtrage. Nous désirons également intégrer une étape de normalisation afin de réduire la variabilité des données, permettant une plus grande généralité dans notre approche. Enfin, nous souhaitons mettre en place une analyse détaillée des performances de notre approche en utilisant certaines métriques globales introduites dans la littérature, couplées à des métriques locales permettant de quantifier les performances de chaque geste au regard de notre référence.

Une partie des travaux de ce manuscrit va donc se focaliser sur la reconnaissance des gestes chirurgicaux, en proposant une méthodologie scientifique inexplorée jusqu'à présent. Le chapitre 3 présente une approche ascendante pour la reconnaissance des surges, partant de la détection des dexemes. En se basant sur une nouvelle méthode de segmentation non-supervisée, et couplée à des algorithmes d'apprentissage, cette approche vient en complément des travaux de la littérature et s'attaque à certaines limitations présentées précédemment.

Le chapitre 1 nous a également permis de mettre en avant certaines limitations concernant les dispositifs d'entraînement actuellement disponibles, notamment de par leur coût économique. Un exemple de ce type de dispositif est le da Vinci Skills Simulator qui requiert l'utilisation d'une console da Vinci ainsi que du simulateur en complément. De plus, un robot da Vinci n'est que peu accessible pour la formation en hôpital car ce dernier est régulièrement utilisé en intervention, laissant peu de temps libre aux internes pour s'entraîner. Il est donc nécessaire de s'équiper de dispositifs, généralement de type simulateur, permettant la formation. De nouvelles alternatives, moins onéreuses et plus accessibles, ont vu le jour tel que le Xron proposé par la société BBZ Medical Technologies. Ce simulateur peut s'utiliser sur un ordinateur standard, ne nécessitant que l'achat d'une interface de commande permettant de reproduire les mouvements réalisables avec un robot da Vinci. De cette manière, chaque chirurgien et interne peut se former durant ses créneaux libres, avec son propre dispositif d'entraînement portable, accessible à moindre coût.

Dans le cadre de la formation pour la chirurgie robotisée, nous proposons dans le chapitre 4 l'évaluation d'une nouvelle interface homme-machine manuelle et sans contact pour la commande de robots chirurgicaux, également adaptable aux logiciels de simulation. Pour ce faire, nous analysons les performances réalisées avec ce dispositif et les comparons avec un système de référence, déjà employé en environnement clinique.

Deuxième partie

Méthodes pour la reconnaissance et la réalisation de gestes destinées à l'entraînement en chirurgie robotisée

Segmentation automatique et reconnaissance des gestes chirurgicaux

Préambule

Ce chapitre aborde notre première contribution pour la reconnaissance des gestes chirurgicaux. Nous présentons dans un premier temps une nouvelle approche pour la segmentation non-supervisée des trajectoires des outils chirurgicaux. Dans un second temps, nous utilisons cette segmentation pour réaliser l'apprentissage et la reconnaissance des gestes chirurgicaux. Nous évaluons notre approche au travers de deux études permettant de qualifier et quantifier les performances de notre méthode.

Sommaire

3.1	Introduction et description de la méthode	60
3.1.1	Segmentation non-supervisée	61
3.1.2	Classification et reconnaissance	63
3.2	Approche non-supervisée pour la segmentation de trajectoires	63
3.2.1	Prétraitement des données	63
3.2.1.1	Variables cinématiques	64
3.2.1.2	Filtrage	65
3.2.1.3	Normalisation	67
3.2.2	Sélection de délimiteurs temporels	68
3.2.2.1	Notion de persistance	69
3.2.2.2	Utilisation en pratique	70
3.2.3	Analyse par similarité spatio-temporelle	70
3.2.3.1	Métriques de similarité spatio-temporelle	71
3.2.3.2	Choix de la dimension du problème	74

3.2.4	Sélection automatique	76
3.2.5	Vérification de la méthode de segmentation	77
3.2.5.1	Protocole expérimental	77
3.2.5.2	Données	78
3.2.5.3	Résultats	78
3.3	Classification et reconnaissance des surgemes	79
3.3.1	Extraction d'attributs caractéristiques	81
3.3.1.1	Transformation des signaux	82
3.3.1.2	Approximation polynômiale	82
3.3.2	Algorithmes de classification	83
3.3.2.1	K-Plus Proches Voisins	84
3.3.2.2	Machines à Vecteurs de Support	85
3.3.3	Association de dexemes	86
3.3.4	Validation de la méthode de reconnaissance	86
3.3.4.1	Type de tâche et vocabulaire	87
3.3.4.2	Stratégie de validation	87
3.3.4.3	Données	90
3.3.4.4	Métriques d'évaluation	90
3.3.4.5	Résultats	91
3.3.5	Analyse et discussion	99
3.3.5.1	Segmentation non-supervisée	99
3.3.5.2	Classification et reconnaissance	101
3.4	Conclusion	104

3.1 Introduction et description de la méthode

À partir de l'état des travaux sur la reconnaissance des gestes médico-chirurgicaux présenté dans le chapitre précédent, nous avons souligné les limitations des méthodes actuelles. Ces dernières portent notamment sur la détection des gestes à un seul niveau de granularité, l'utilisation de modèles statistiques structurés temporels dépendant de la tâche, ainsi que l'emploi de techniques d'acquisition, de prétraitement des données et de validation généralement perfectibles et incomplètes. Dans ce chapitre, nous présentons une nouvelle approche permettant de résoudre ces limitations. Cette contribution s'appuie sur la mise en place d'un processus complet pour la reconnaissance des gestes chirurgicaux à partir des informations de trajectoires des instruments chirurgicaux. Pour ce faire, nous avons tout d'abord réalisé des expérimentations afin d'acquérir un ensemble de trajectoires pour créer notre base de données. Nous avons employé la plateforme robotisée disponible au LIRMM et demandé à plusieurs sujets d'effectuer une tâche d'entraînement

robotisé de transfert d'objets telle que proposée dans la littérature. À partir des données cinématiques acquises, nous avons mis en place une méthode composée de deux phases distinctes : une phase de segmentation non-supervisée et une phase de reconnaissance pour la détection des gestes chirurgicaux. Afin de clarifier la compréhension de ces travaux et de faciliter leur lecture, la méthodologie mise en place est détaillée au travers de la figure 3.1.

3.1.1 Segmentation non-supervisée

Dans ce chapitre nous présentons une nouvelle méthode pour la segmentation de trajectoires provenant d'outils chirurgicaux. L'objectif de cette approche est de capturer les différents dexemes réalisés par l'utilisateur au travers des signaux mesurés et calculés durant l'étape d'acquisition. L'intérêt de cette approche réside dans le fait qu'elle n'utilise aucune information a priori pour segmenter les trajectoires des instruments et se base uniquement sur les informations contenues dans les signaux d'entrée. Pour accomplir cette tâche, nous avons développé une méthode de segmentation non-supervisée reposant sur quatre composantes, visibles sur la figure 3.1.

Tout d'abord, les signaux d'entrée sont prétraités (étape A de la figure 3.1). Pour cela, nous effectuons une extraction d'information à partir des données brutes afin de construire de nouveaux signaux invariants dans l'espace cartésien. Puis l'ensemble de ces signaux bruts et invariants sont filtrés et normalisés pour la suite du processus afin de réduire leur variabilité.

Une fois le prétraitement effectué, nous détectons des points caractéristiques dans ces signaux afin de définir les différents dexemes réalisés par l'utilisateur (étape B de la figure 3.1). Pour cela nous nous appuyons sur des notions récemment publiées de traitement de signal permettant d'extraire un ensemble de points persistants contenus dans des signaux. Néanmoins, ces points caractéristiques sont pertinents à l'intérieur du signal dans lequel ils ont été extraits mais peuvent être insignifiants dans les autres signaux d'entrée.

De ce fait, nous établissons ensuite une mesure de similarité entre l'ensemble des signaux en se basant sur les points caractéristiques précédemment capturés (étape C de la figure 3.1). Cette mesure de similarité permet de quantifier le degré de significativité de chaque point caractéristique par rapport aux déformations spatio-temporelles existantes dans les signaux.

Enfin, nous sélectionnons les points caractéristiques les plus pertinents au regard de leur évaluation spatio-temporelle grâce à une technique de sélection spécifique capturant uniquement les scores les plus importants (étape D de la figure 3.1). Ces points caractéristiques nous permettent de définir les limiteurs temporels des différents dexemes. Ces dexemes sont ensuite transférés à la phase de reconnaissance.

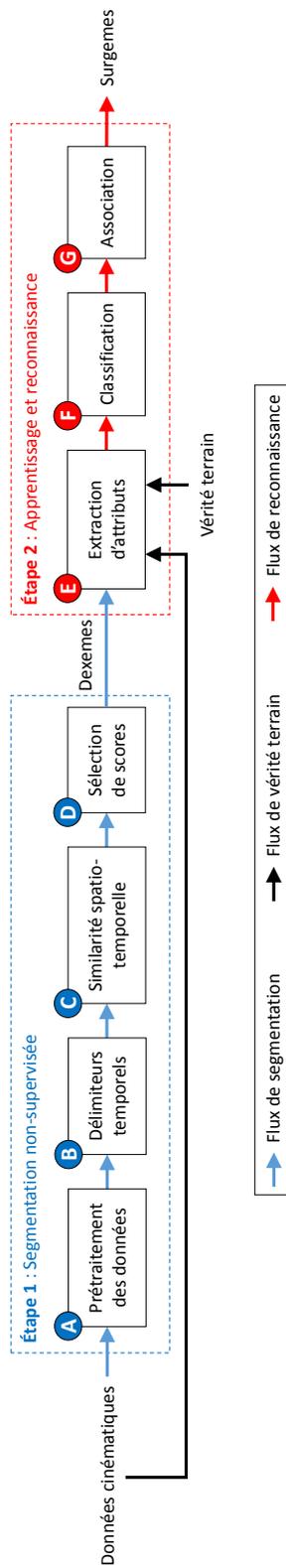


FIGURE 3.1: Schéma du processus complet de segmentation et reconnaissance des gestes chirurgicaux.

3.1.2 Classification et reconnaissance

A partir de l'ensemble des dexemes, nous cherchons à reconnaître les surgemes réalisés par l'opérateur. D'après notre hypothèse de décomposition, nous supposons qu'un surgeme peut être constitué d'un ou plusieurs dexemes. Ainsi, la détection et reconnaissance des différents dexemes nous permettent de reconstituer et reconnaître le surgeme correspondant. Pour ce faire, nous proposons une approche en trois étapes comme indiqué sur la figure 3.1.

La première étape consiste à caractériser les dexemes extraits précédemment (étape E de la figure 3.1). Nous employons une méthode permettant de caractériser chacun des signaux contenus à l'intérieur de ce dexeme, de manière à extraire un vecteur de taille fixe.

Ce vecteur est ensuite utilisé pour la classification (étape F de la figure 3.1). Cette classification permet de définir l'appartenance de ce dexeme à un surgeme existant grâce aux capacités de discrimination des algorithmes. Pour cette application, nous étudions plusieurs algorithmes existants afin d'évaluer leurs performances de reconnaissance.

Enfin, une séquence de dexemes annotés est récupérée en sortie de classification. Cependant, afin de réduire les bruits et erreurs de classification, nous réalisons un filtrage sur la séquence afin d'obtenir une détection homogène des différents dexemes (étape G de la figure 3.1). La séquence en sortie représente alors l'ensemble des dexemes annotés permettant de reconstituer les surgemes réalisés par l'opérateur.

3.2 Approche non-supervisée pour la segmentation de trajectoires

L'approche de segmentation non-supervisée proposée dans ces travaux s'appuie sur l'hypothèse qu'un surgeme est construit sur la base d'un ou plusieurs dexemes consécutifs. Dans ces travaux, nous nous efforçons de limiter la quantité de connaissances a priori nécessaire pour réaliser la segmentation des gestes.

La suite de ce chapitre suppose que les expérimentations et l'acquisition des données ont été réalisées avec la plateforme présentée dans le chapitre 4, durant une tâche d'entraînement que nous présentons dans ce chapitre. Nous disposons donc d'un ensemble de données cinématiques acquises avec un robot de téléopération destiné à la recherche ainsi qu'à l'entraînement en chirurgie robotisée.

3.2.1 Prétraitement des données

Afin de segmenter automatiquement les trajectoires, nous réalisons dans un premier temps un prétraitement des données. Ce prétraitement passe par l'utilisation de trois fonctions. La première permet d'extraire, à partir des données cinématiques, un ensemble de signaux afin d'obtenir une représentation des mouvements invariante en rotation et trans-

lation dans l'espace cartésien. Une fois cette représentation calculée, nous appliquons un filtre afin d'atténuer le bruit contenu à l'intérieur de chaque signal. Enfin, la dernière fonction permet de normaliser l'ensemble des signaux afin de réduire les variabilités intra- et interpersonnelles existantes.

3.2.1.1 Variables cinématiques

Dans cette partie, nous distinguons deux types de signaux : les signaux d'acquisition bruts provenant directement de la plateforme robotisée, et les signaux calculés en post-traitement permettant une représentation complémentaire des trajectoires des instruments chirurgicaux.

Signaux d'acquisition.

Lors de la réalisation des tâches d'entraînement, le robot acquiert à chaque instant un ensemble d'informations cartésiennes provenant des instruments chirurgicaux. Ces informations sont représentées sous formes de matrice de transformation, de la forme :

$$T = \begin{bmatrix} r_{11} & r_{12} & r_{13} & x \\ r_{21} & r_{22} & r_{23} & y \\ r_{31} & r_{32} & r_{33} & z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.1)$$

où r_{ij} correspond à l'élément i, j de la matrice de rotation et x, y, z correspondent aux positions cartésiennes de l'instrument. De plus, lors de l'acquisition, nous capturons l'angle d'ouverture et fermeture de la pince, noté α . Ces informations sont acquises pour les deux instruments chirurgicaux employés durant l'entraînement. Enfin, pour des informations de sécurité, nous relevons également l'état de la pédale de commande qui n'est cependant pas utilisé dans la suite de ces travaux.

De ces informations, nous souhaitons construire une représentation cartésienne compacte pour la suite du traitement. Cifuentes et al. [Cifuentes 2013] ont démontré que l'utilisation des quaternions pour représenter la rotation était intéressant pour l'analyse du geste lors d'exercices de rééducation. En se basant également sur l'ouvrage de Khalil et Dombre [Khalil 1999], nous transformons la matrice de rotation en quaternion en utilisant la formule suivante :

$$q = \begin{bmatrix} qx \\ qy \\ qz \\ qw \end{bmatrix} = \begin{bmatrix} k_x \sin(\theta/2) \\ k_y \sin(\theta/2) \\ k_z \sin(\theta/2) \\ \cos(\theta/2) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \text{sign}(r_{32} - r_{23}) \sqrt{r_{11} - r_{22} - r_{33} + 1} \\ \frac{1}{2} \text{sign}(r_{13} - r_{31}) \sqrt{-r_{11} + r_{22} - r_{33} + 1} \\ \frac{1}{2} \text{sign}(r_{21} - r_{12}) \sqrt{-r_{11} - r_{22} + r_{33} + 1} \\ \frac{1}{2} \sqrt{r_{11} + r_{22} + r_{33} + 1} \end{bmatrix} \quad (3.2)$$

où k_x, k_y et k_z sont les composantes unitaires du vecteur de rotation et θ l'angle de rotation autour de ce vecteur. Cette représentation plus compacte a également pour avantage d'éviter la présence de singularité dans la représentation, intégrant néanmoins de la redondance dans l'information.

Calcul d'invariants 3D.

À partir des informations cartésiennes acquises avec le robot, nous souhaitons calculer, en complément des signaux de position, des signaux représentant l'évolution temporelle des positions indépendamment de la position ou de l'orientation choisie. Nous appuyons notre approche sur les travaux de Calabi et al. [Calabi 1998], où une version revue et corrigée a été proposée par Boutin [Boutin 2000]. L'objectif est de construire le repère de Frénet pour chaque échantillon afin d'extraire les informations de courbure et de torsion des trajectoires, ainsi que leurs dérivées premières. Considérant une trajectoire libre définie par $\Gamma(t)$, avec $t \in [1, L]$ où L est la longueur de la trajectoire, une signature euclidienne S invariante en rotation et translation peut être définie par : la courbure (κ), la torsion, (τ) et leurs dérivées premières respectives (κ_s and τ_s) par rapport à l'abscisse curviligne s , sous la forme :

$$S = \{\kappa(t), \kappa_s(t), \tau(t), \tau_s(t) \mid t \in [1, L]\} \quad (3.3)$$

où

$$\kappa(t) = \frac{\|\vec{\Gamma}'(t) \wedge \vec{\Gamma}''(t)\|}{\|\vec{\Gamma}'(t)\|^3} \quad (3.4)$$

$$\tau(t) = \frac{(\vec{\Gamma}'(t) \wedge \vec{\Gamma}''(t)) \cdot \vec{\Gamma}'''(t)}{\|\vec{\Gamma}'(t) \wedge \vec{\Gamma}''(t)\|^2} \quad (3.5)$$

$$\kappa_s(t) = \frac{d\kappa(t)}{ds} = \frac{d\kappa(t)}{dt} \cdot \frac{dt}{ds} = \frac{d\kappa(t)}{dt} \cdot \frac{1}{\|\vec{\Gamma}'(t)\|} \quad (3.6)$$

$$\tau_s(t) = \frac{d\tau(t)}{ds} = \frac{d\tau(t)}{dt} \cdot \frac{dt}{ds} = \frac{d\tau(t)}{dt} \cdot \frac{1}{\|\vec{\Gamma}'(t)\|} \quad (3.7)$$

Dans le cadre de ces travaux, nous avons choisi d'employer des estimateurs numériques de chaque composante en se reposant sur des approximations locales développées par Wu et al. [Wu 2009]. De cette manière, la signature euclidienne S est construite en évitant les dérivées partielles et temporelles, gourmandes en temps de calcul. Les 12 signaux retenus pour représenter les mouvements de chaque instrument en entrée de notre méthode sont compactés dans la trame D :

$$D^i = \{x^i, y^i, z^i, qx^i, qy^i, qz^i, qw^i, \kappa^i, \tau^i, \kappa_s^i, \tau_s^i, \alpha^i \mid i \in [1, 2]\} \quad (3.8)$$

3.2.1.2 Filtrage

Une fois les signaux acquis et calculés, nous souhaitons filtrer chacun d'eux afin de réduire les bruits d'acquisition et de mesure. Nous considérons et comparons deux types de filtre : un filtre de Kalman et un filtre passe-bas à réponse impulsionnelle finie.

Filtre de Kalman.

Le filtre de Kalman employé dans ces travaux se base sur la représentation d'état discrète d'un système linéaire de la forme :

$$X_k = AX_{k-1} + W_{k-1} \quad (3.9)$$

où A représente la matrice d'état du système, X_k le vecteur d'état à l'instant k et W_k le bruit de procédé supposé blanc et suivant une distribution de probabilité normale de covariance Q . Dans notre cas, la matrice A du système a été construite à partir de l'équation linéaire du mouvement :

$$x_k = x_{k-1} + \delta t v_{k-1} \quad (3.10)$$

où x_k est la position, v_k la vitesse et δt la période d'échantillonnage du système. D'autre part, le vecteur Z de mesure est également défini tel que :

$$Z_k = HX_k + V_k \quad (3.11)$$

où H est la matrice d'observation et V_k le bruit de mesure supposé blanc et suivant une distribution de probabilité normale de covariance R . De plus, W_k et V_k sont supposés indépendants. Dans cette configuration, le filtre de Kalman joue le rôle d'estimateur récursif. En utilisant la représentation d'état du système observé, il estime l'état courant grâce à l'état précédent et la mesure courante. Le filtre possède deux phases distinctes de calcul : une phase de prédiction et une phase de correction. D'une part, la phase de prédiction utilise l'état estimé de l'instant précédent pour calculer une estimation de l'état courant de la forme :

$$\hat{X}_k^- = A\hat{X}_{k-1}^- \quad (3.12)$$

$$P_k^- = AP_{k-1}A^T + Q \quad (3.13)$$

D'autre part, une fois l'estimation réalisée, l'objectif est de corriger le modèle afin d'optimiser la prédiction en se basant sur la mesure courante :

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \quad (3.14)$$

$$\hat{X}_k = \hat{X}_k^- + K_k (Z_k - H\hat{X}_k^-) \quad (3.15)$$

$$P_k = (I_d - K_k H) P_k^- \quad (3.16)$$

où K_k est appelé gain de Kalman.

Pour notre système, nous avons estimé les matrices de covariance en fonction des données acquises à partir de la plateforme robotisée. Nous avons ensuite implémenté les équations pour filtrer chacun des signaux.

Filtre Passe-bas.

Par définition, le filtre passe-bas est adapté au filtrage des mouvements humains. En effet, ces mouvements ont des fréquences oscillant entre 0Hz et 10Hz [Loram 2006], avec un maximum de 4,5Hz pour les mouvements de la main [Xiong 2006]. Du fait que notre système d'acquisition échantillonne à une fréquence de 100Hz, ce filtre convient à nos besoins et permet de garder les fréquences fondamentales des mouvements humains tout en rejetant les perturbations. Cependant, bien que le filtre soit programmé pour fournir une réponse impulsionnelle finie grâce à l'utilisation de coefficients réels pour la convolution, il introduit un déphasage linéaire. Même si cela n'est pas problématique à l'intérieur du signal, des pertes d'informations auront lieu aux extrémités. Nous optons donc pour l'utilisation d'un filtre de type forward-backward permettant de filtrer le signal dans un sens, puis dans l'autre, afin d'obtenir un déphasage nul en sortie. Les paramètres du filtre ont été définis de la même manière que dans les travaux de Yang et al. [Yang 2013], à savoir une fréquence de coupure de 1,5Hz, un gain en bande passante unitaire et une forte atténuation après 10Hz.

Choix du filtrage.

Nous appliquons ces deux types de filtre sur les signaux acquis lors de la réalisation d'une tâche d'entraînement de transfert d'objet utilisant la plateforme robotisée. Pour simplifier la lisibilité, la figure 3.2 ne présente que le filtrage d'un signal représentant l'évolution de la position cartésienne de l'instrument chirurgical à travers le temps. En regardant le graphique, nous constatons que le filtre passe-bas permet une réduction importante des bruits, tout en évitant les dépassements dus aux erreurs de prédiction du filtre de Kalman. Bien que ces erreurs soient minimales, le filtre passe-bas correspond à nos besoins pour le filtrage des données cinématiques. De plus, il n'entraîne aucun déphasage sur le signal en sortie.

3.2.1.3 Normalisation

La dernière fonction dans notre processus de prétraitement des données est la normalisation des signaux. Cette normalisation est réalisée par un centrage réduit afin de diminuer les variabilités existantes entre les trajectoires provenant d'un même sujet, ainsi que de sujets différents. Considérant R_i un signal brut de longueur L , la normalisation appliquée est définie par :

$$N_i(t) = \frac{1}{\sigma_i^2} (R_i(t) - \mu_i) \quad t \in [1, L] \quad (3.17)$$

où μ_i et σ_i^2 sont respectivement la moyenne et la variance du signal d'origine, et N_i correspond au signal normalisé en sortie. La figure 3.3 présente la normalisation d'un signal de position cartésienne préfiltré. Le centrage et la réduction permettent alors d'exprimer l'ensemble des signaux dans un repère commun. Nous constatons que le profil et les va-

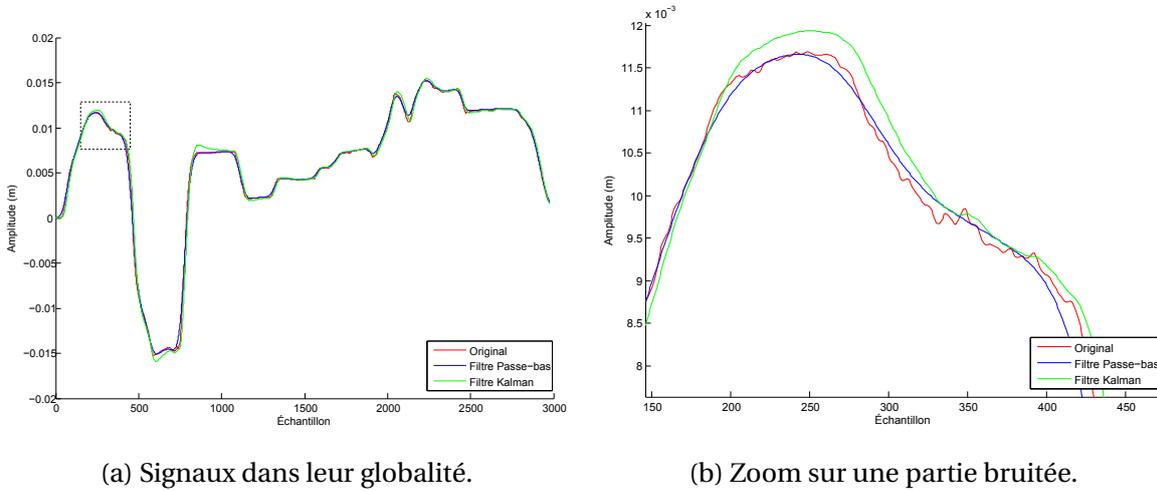


FIGURE 3.2: Filtrage de la position cartésienne de l'instrument chirurgical selon l'axe X du repère du robot.

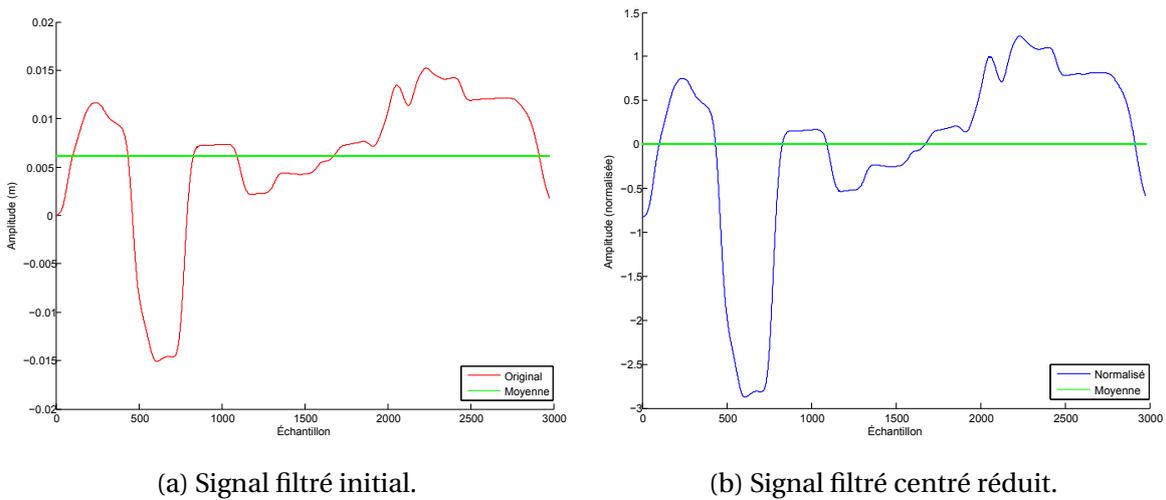


FIGURE 3.3: Normalisation du signal pour obtenir une représentation centrée réduite.

riations du signal sont conservés, impliquant que les corrélations existantes entre signaux sont également conservées.

3.2.2 Sélection de délimiteurs temporels

Dans cette étape, nous cherchons à sélectionner un ensemble de points caractéristiques représentant le temps initial et le temps final de chacun des dexemes réalisés par

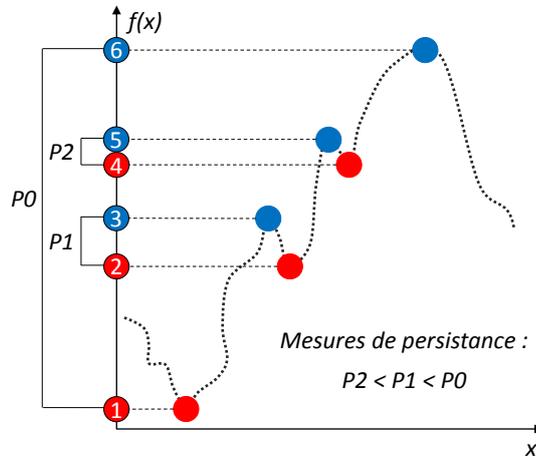


FIGURE 3.4: Illustration du principe de persistance topologique en 1D.

l'utilisateur. Grâce à ces points, nous pouvons construire un vecteur de délimiteurs temporels contenant l'ensemble des temps caractéristiques définissant lesdits dexemes.

3.2.2.1 Notion de persistance

Schulz et al. [Schulz 2010] ainsi que Lin et al. [Lin 2014] ont développé des méthodes en se basant sur les variations des vecteurs vitesse et accélération. Ainsi, les délimiteurs temporels d'un segment de trajectoire sont définis par une vitesse ou une accélération nulle, qui indique un changement d'intention de la part de l'utilisateur. Pour notre méthode, nous avons considéré une approche similaire en nous appuyant sur la détection de maximum et minimum locaux ainsi que globaux dans les trajectoires, ce qui équivaut à un changement de signe du vecteur vitesse correspondant. Pour cette approche, nous nous basons sur les travaux de Edelsbrunner et al. [Edelsbrunner 2002] introduisant la notion de persistance topologique. Ils décrivent une métrique permettant une simplification topologique de signal en se basant sur la suppression d'attributs grâce au calcul de leur durée de vie après filtrages successifs. Nous utilisons la figure 3.4 pour décrire le procédé développé.

L'idée est de créer des paires d'extrema, qui sont contenus dans le signal. Il est important de noter que les paires créées ne comportent pas des points caractéristiques obligatoirement adjacents (figure 3.4). Ces paires sont ensuite évaluées à travers la métrique de persistance. Des filtrages successifs sont appliqués au signal afin de supprimer le bruit. Si la paire d'extrema est rapidement détruite, alors la persistance temporelle associée à ces points est faible. À l'inverse, si ces points restent présents leur persistance augmente suivant le nombre de filtrage appliqués. À la fin, les paires sont triées dans un tableau suivant le score de persistance obtenu.

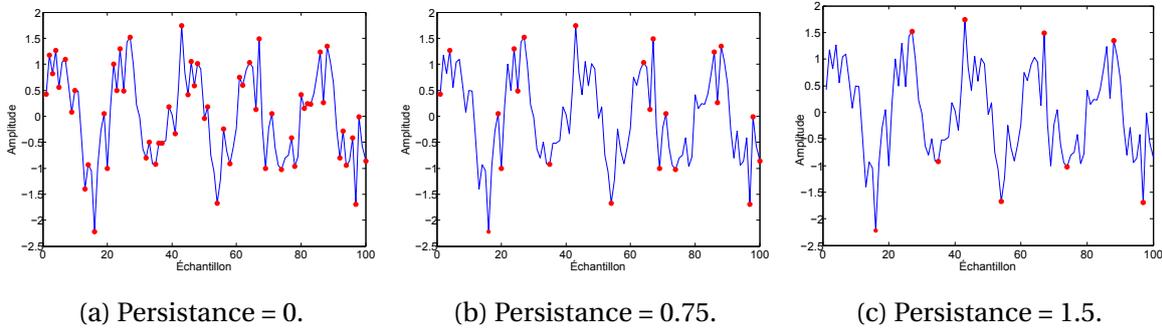


FIGURE 3.5: Évolution de la sélection des points caractéristiques en utilisant le seuil de persistance.

3.2.2.2 Utilisation en pratique

Pour l'implémentation, nous nous basons sur les travaux de Weinkauff et al. [Weinkauff 2010], avec un algorithme optimisé pour une utilisation en deux dimensions. Cependant, la partie concernant la reconstruction du signal n'est pas utilisée dans le cadre nos travaux car nous ne souhaitons pas déformer les signaux acquis. Cette implémentation nous permet de sélectionner, au travers d'un seuil de persistance, un ensemble de points caractéristiques d'un signal comme présenté sur la figure 3.5. Dans cet exemple, une sinusoïde convoluée d'un bruit blanc contient un ensemble d'extrema. Nous souhaitons retrouver les points caractéristiques permettant de décrire au mieux ce signal. Pour cela, nous modifions le seuil de persistance afin de capturer les points les plus pertinents contenus dans ce signal.

À partir de cet algorithme, nous définissons expérimentalement un seuil pour chaque signal, qui restera identique durant la durée des expérimentations. Nous extrayons ensuite l'ensemble des points caractéristiques de chaque signal comme présenté sur la figure 3.6, définissant ainsi les délimiteurs temporels. Ces points sont ensuite stockés et utilisés dans la suite de notre méthode de segmentation.

3.2.3 Analyse par similarité spatio-temporelle

Une fois l'ensemble des délimiteurs temporels défini, nous souhaitons évaluer leur pertinence vis-à-vis du critère de déformations spatio-temporelles. Autrement dit, le point sélectionné est caractéristique dans le signal, mais les segments définis par ces délimiteurs sont-ils identiques à l'échelle spatiale et également temporelle ? Pour réaliser cette étape, nous nous appuyons sur l'utilisation de métriques de similarité, que nous appellerons également métriques de dissimilarité du fait de leur utilisation opposée.

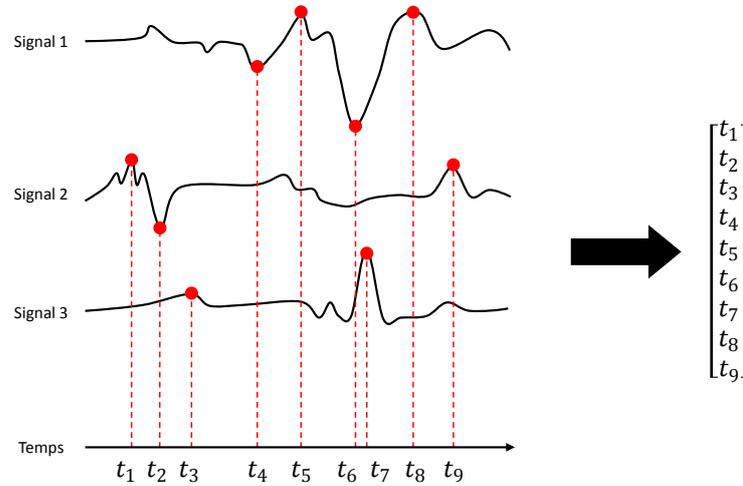


FIGURE 3.6: Extraction des points caractéristiques des signaux pour obtenir un vecteur de délimiteurs temporels.

3.2.3.1 Métriques de similarité spatio-temporelle

Les métriques de similarité spatio-temporelle sont particulièrement employées pour l'analyse, la comparaison et l'évaluation de séries temporelles. Dans ce cadre, les travaux de Wang et al. [Wang 2013] fournissent une excellente description de l'ensemble des mesures de similarité, ainsi qu'une classification suivant leurs propriétés. Dans nos travaux, nous nous concentrons sur l'utilisation et la comparaison de trois métriques de similarité pour les séries temporelles que sont la distance de Hausdorff, la distance de Fréchet et la Déformation Temporelle Dynamique ¹.

Distance de Hausdorff.

Le choix de la métrique de Hausdorff repose sur l'analyse de travaux de la littérature. Tout d'abord, les résultats obtenus par Khalid et al. [Khalid 2005b] montrent que cette métrique fournit d'excellents résultats lorsqu'elle est utilisée pour comparer deux séries temporelles. Chen et al. [Chen 2011] utilisent également cette métrique afin d'effectuer le regroupement non-supervisé de séries temporelles similaires, avec des résultats encourageants. Nous choisissons donc d'utiliser la distance de Hausdorff comme l'une des métriques possibles afin d'évaluer les segments de trajectoires. En considérant deux séries temporelles $A = \{a_1, a_2, \dots, a_M\}$ et $B = \{b_1, b_2, \dots, b_N\}$, la distance de Hausdorff est définie par :

$$h(A, B) = \max_{a \in A} (\min_{b \in B} \|a - b\|) \quad (3.18)$$

1. "Dynamic Time Warping (DTW) en anglais.

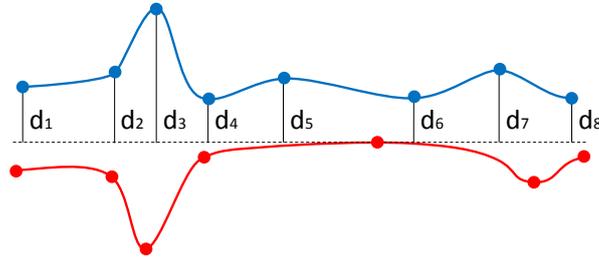


FIGURE 3.7: Illustration de la distance de Hausdorff entre deux séries temporelles.

Cependant, toute distance considérée comme métrique doit respecter la propriété de symétrie. Pour restaurer la symétrie, nous utilisons :

$$h(A, B) = \max(h(A, B), h(B, A)) \quad (3.19)$$

Schématiquement, la distance de Hausdorff correspond à la distance maximale de l'ensemble des distances minimales qui existent entre deux séries temporelles, comme présenté sur la figure 3.7. Dans cet exemple, nous pouvons distinguer 8 mesures spécifiques. Ces distances peuvent également être appelées distances de Hausdorff relatives. Néanmoins, la distance de Hausdorff correspond à la plus grande de ces distances relatives, qui dans le cas présent est la distance notée d_3 . Toutefois, nous constatons qu'en réalisant le même calcul avec la seconde courbe (en rouge), la distance de Hausdorff serait sensiblement différente, d'où l'intérêt d'instaurer une symétrie pour le calcul de la distance globale.

Distance de Fréchet.

La seconde métrique considérée dans ces travaux est la distance de Fréchet. Cette distance est initialement utilisée pour comparer deux formes continues telles que des courbes ou surfaces. Du fait que la distance de Fréchet prenne en compte la continuité des courbes, cette dernière est généralement plus satisfaisante que la distance de Hausdorff car elle permet de quantifier localement et globalement la similarité de deux séries temporelles. Une variante de cette métrique est la distance de Fréchet discrète employée pour comparer deux séries temporelles discrètes. Considérant deux séquences discrètes $P = \{p_1, p_2, \dots, p_M\}$ et $Q = \{q_1, q_2, \dots, q_N\}$, le couplage C des composantes de chaque séquence P et Q est un ensemble de composantes $C = \{c_1, c_2, \dots, c_K\}$ avec $C_r = \{p_i, q_j\}$, $k \in [1, K]$ et $i \in [1, M]$ et $j \in [1, N]$ respectant :

- Les conditions limites : $C_1 = (p_1, q_1)$ et $C_K = (p_M, q_N)$,
- La récursivité : $C_r = (p_i, q_j)$ et $C_{r+1} \in \{(p_{i+1}, q_j), (p_i, q_{j+1}), (p_{i+1}, q_{j+1})\}$.

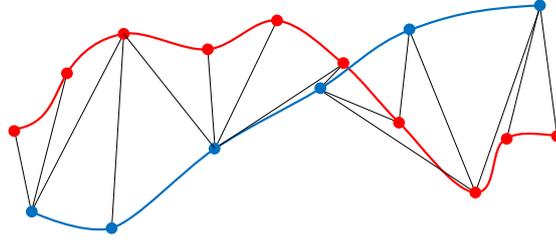


FIGURE 3.8: Illustration de la distance de Fréchet discrète entre deux séries temporelles.

La distance de Fréchet discrète est alors définie par :

$$F(P, Q) = \min_{\text{couplage } C} \max_{(p_i, q_j) \in C} |p_i - q_j| \quad (3.20)$$

Un exemple de cette distance est illustré par la figure 3.8. La distance de Fréchet est la distance minimale de la somme de couplages (appelées également distances de cordes), permettant de parcourir les deux séquences suivant leurs abscisses curvilignes respectifs. Cette métrique permet une mesure prenant en compte l'aspect global mais également local des courbes.

Déformation Temporelle Dynamique.

La dernière métrique considérée dans ces travaux pour analyser les segments consécutifs de trajectoires est la métrique de Déformation Temporelle Dynamique (DTD) [Sakoe 1978]. Dans la littérature, un grand nombre de travaux exploitent cet algorithme pour comparer des trajectoires [Sankararaman 2013], mais également pour réaliser l'alignement de séquences dans le cadre de la détection de mouvements humains [Zhou 2012] ainsi dans celui de l'analyse de gestes liés à la rééducation [Cifuentes 2013]. Les deux avantages principaux de cet algorithme sont la capacité de prise en compte des déformations spatiales et temporelles ainsi que l'alignement de signaux à des fins de classification et d'évaluation. Considérant deux séquences $A = \{a_1, a_2, \dots, a_M\}$ et $B = \{b_1, b_2, \dots, b_N\}$, l'alignement de ces séquences est calculé récursivement par :

$$D(A_i, B_j) = \delta(a_i, b_j) + \min \begin{cases} D(A_{i-1}, B_{j-1}) \\ D(A_i, B_{j-1}) \\ D(A_{i-1}, B_j) \end{cases} \quad (3.21)$$

où $\delta(a_i, b_j)$ est la norme de la distance euclidienne entre les éléments a_i et b_j , avec $i \in [1, M]$ et $j \in [1, N]$. A noter que dans ces travaux nous utilisons la distance euclidienne, plus cohérente pour la comparaison de signaux cartésiens. Cependant, d'autres distances telles que la distance de Manhattan ou la covariance peuvent également être employées. La similarité complète des deux séries temporelles est donnée par $D(A_{|A|}, B_{|B|}) = D(A_M, B_N)$.

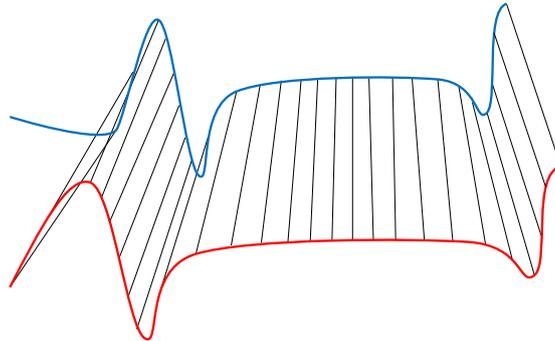


FIGURE 3.9: Illustration de l'utilisation de l'algorithme de Déformation Temporelle Dynamique entre deux séries temporelles.

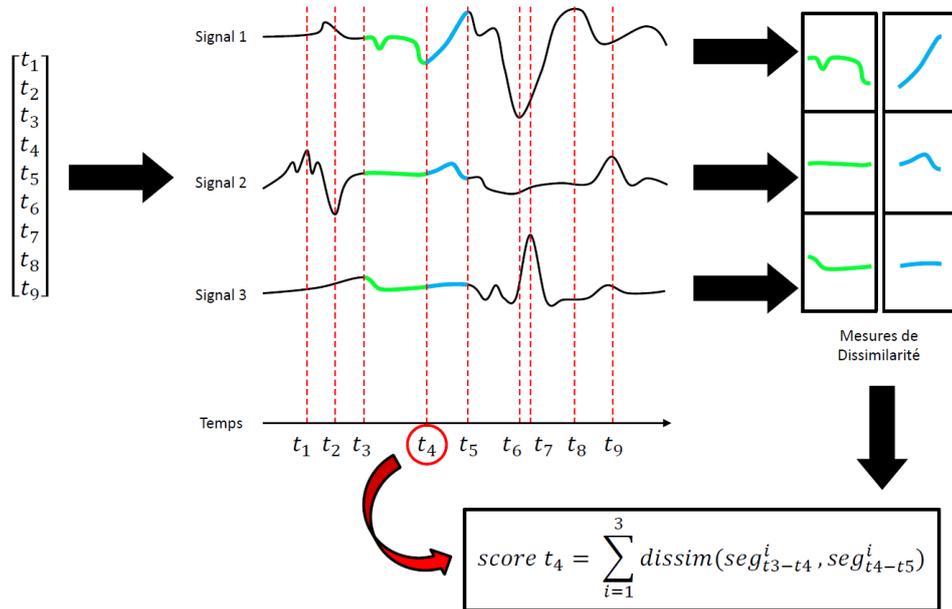
Schématiquement, l'algorithme de Déformation Temporelle Dynamique permet de rechercher le chemin le plus court permettant d'aligner les éléments de chacune des séquences. La figure 3.9 illustre le principe. Nous constatons ici que l'alignement entre les deux courbes dépend de la localisation spatiale et temporelle des points contenus dans chacune des séquences. L'alignement qui en résulte peut être, suivant les cas, une déformation non-linéaire par rapport au temps, à l'espace ou bien les deux à la fois.

3.2.3.2 Choix de la dimension du problème

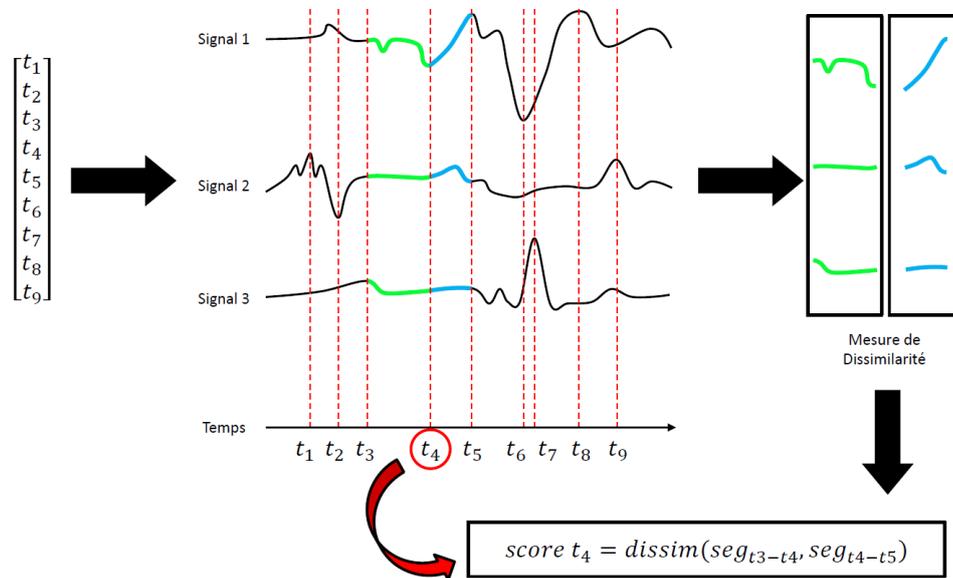
Pour quantifier la dissimilarité, nous désirons minimiser la similarité. Pour cela nous employons les trois métriques présentées précédemment et allons étudier leur performance. Cependant, les performances dépendent également de la dimension utilisée pour réaliser l'analyse. Ainsi, deux cas sont envisagés : une analyse unidimensionnelle et une analyse multidimensionnelle des signaux. Ces deux cas sont représentés par la figure 3.10.

Dans le premier cas, l'ensemble des délimiteurs temporels est utilisé en entrée afin d'évaluer leur pertinence au regard de chacun des signaux, indépendamment les uns des autres (figure 3.10a). Considérant le cas d'un délimiteur temporel, nous mesurons la dissimilarité entre le segment précédent et le segment suivant pour chaque signal. Le score du délimiteur est alors la somme des scores obtenus par chacun des signaux. Plus le score est élevé, plus le délimiteur temporel indique de fortes variations.

Dans le cas d'une comparaison multidimensionnelle, le calcul est plus complexe. On considère d'abord un signal de plusieurs dimensions (la figure 3.10b présente un exemple de trois dimensions) ainsi qu'un délimiteur temporel particulier. Le calcul de sa dissimilarité est alors réalisé en comparant l'ensemble des segments précédents avec l'ensemble des segments suivants, de manière dépendante et directe. Le délimiteur obtient donc un score unique quantifiant la similarité des segments considérés.



(a) Calcul de dissimilarité en se basant sur un ensemble de signaux unidimensionnels.



(b) Calcul de dissimilarité en se basant sur un signal multidimensionnel.

FIGURE 3.10: Illustration du calcul de dissimilarité suivant le choix de la dimension du problème.

Dans ces travaux, l'analyse de la dimension est importante car cette dernière indique le point de vue de l'analyse, qui sera local dans le cas unidimensionnel et global dans le cas multidimensionnel. En étudiant les travaux de la littérature, Ten Holt et al. [Ten Holt 2007] ainsi que de Shokoohi-Yekta et al. [Shokoohi-Yekta 2015] présentent une approche multidimensionnelle plus performante. Toutefois, pour obtenir de tels résultats, les auteurs indiquent que les signaux doivent être exprimés dans un référentiel commun. Dans notre cas, la normalisation centrée réduite permet de respecter cette contrainte, nous incitant donc à employer une analyse multidimensionnelle des signaux pour le calcul de dissimilarité.

3.2.4 Sélection automatique

La dernière étape de notre méthode de segmentation non-supervisée consiste à sélectionner les délimiteurs en se basant sur leur score de dissimilarité respectif. Afin d'éviter tout a priori sur le nombre de délimiteurs en sortie, sur la taille minimale ou maximale des segments, ainsi que sur la dispersion temporelle des points, nous désirons nous focaliser uniquement sur les scores de dissimilarité obtenus précédemment afin de fournir une segmentation indépendante de l'application considérée.

Dans notre approche, nous employons une technique de sélection basée uniquement sur les scores de dissimilarité. L'objectif est de sélectionner les délimiteurs avec les scores les plus importants, indiquant des variations importantes entre les signaux avant et après ces délimiteurs. Pour ce faire, nous utilisons la technique de Suppression de Non-Maxima² employée dans le filtre de Canny. Cependant, cet algorithme possède plusieurs variantes. Pour notre méthode, nous utilisons l'approche par paire présentée dans les travaux de Neubeck et al. [Neubeck 2006]. L'algorithme consiste à définir une fenêtre glissante, dont la taille est fixée par l'utilisateur, puis à réaliser une analyse des scores à l'intérieur de cette fenêtre. Si le score le plus important est contenu à l'intérieur de la fenêtre, ce score est gardé. Dans le cas contraire, si le score est à l'extrémité de la fenêtre, alors cette dernière se déplace pour réaliser une nouvelle analyse des scores voisins. L'intérêt de cette sélection réside dans une approche semi-locale pour la comparaison.

Un exemple de sélection automatique des scores de dissimilarité est présenté sur la figure 3.11. D'un côté les scores de dissimilarité des délimiteurs temporels sont récupérés en sortie de l'étape précédente (figure 3.11a). De ces scores, nous désirons garder les plus importants et rejeter les plus faibles. Nous utilisons ici l'algorithme de Suppression de Non-Maxima avec une fenêtre de taille 3. L'algorithme effectue alors la sélection et renvoie les scores tel que présenté sur la figure 3.11b. Nous récupérerons les délimiteurs temporels représentés par ces maxima, nous permettant ainsi de délimiter les différents dexemes au sein des signaux d'entrée.

2. "Non-Maximum Suppression (NMS) en anglais.

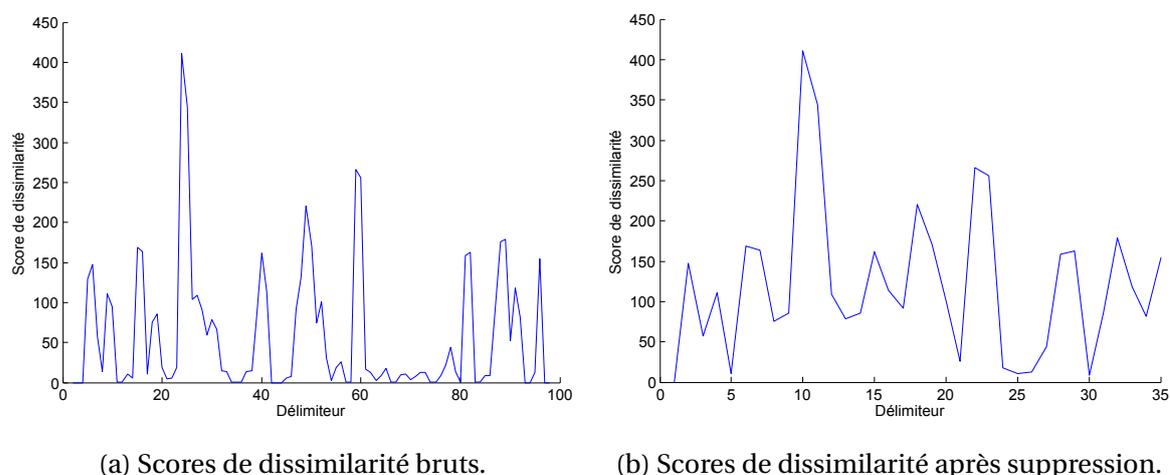


FIGURE 3.11: Sélection automatique des délimiteurs temporels en utilisant l'algorithme de Suppression de Non-Maxima.

3.2.5 Vérification de la méthode de segmentation

Nous souhaitons analyser les capacités de segmentation de l'approche non-supervisée présentée dans ce chapitre. Pour cela, nous proposons une méthode de vérification afin d'évaluer les résultats obtenus sur plusieurs trajectoires. Nous décrivons le protocole expérimental ainsi que les données recueillies pour réaliser la segmentation.

3.2.5.1 Protocole expérimental

Cette étude porte sur une vérification visuelle des résultats obtenus avec la méthode de segmentation supervisée proposée dans ce chapitre. Afin d'acquérir des données pour cette étude, nous avons dans un premier temps demandé à un sujet de réaliser une trajectoire aléatoire en utilisant la plateforme robotisée. Cette trajectoire n'a aucun objectif précis et le sujet a été libre de ses mouvements. De cette manière, nous analysons les performances de la méthode de segmentation non-supervisée sur une trajectoire générée aléatoirement. Ce jeu de données est appelé *TrajectoireAleatoire*.

Dans un second temps, nous avons demandé à trois experts de dessiner, dans le vide, la lettre "R" en utilisant la même plateforme robotisée de téléopération. Ces experts sont composés de deux urologues utilisant le robot da Vinci en intervention, ainsi qu'un expert en système de téléopération. Le dessin de cette lettre a pour objectif d'évaluer deux caractéristiques. La première est la capacité à distinguer les différentes parties de la lettre (courbes, lignes droites). La seconde est la prise en compte de variabilité dans les données, amenant à qualifier la robustesse de notre approche. Nous emploierons le terme *TrajectoiresLettre* pour décrire ce jeu de données dans la suite de l'étude.

3.2.5.2 Données

Comme décrit précédemment, nous disposons de deux jeux de données : *TrajectoireAleatoire* et *TrajectoiresLettre*. Le premier est composé d'une seule trajectoire et le second de trois trajectoires provenant de trois opérateurs différents. L'ensemble de ces trajectoires ont été réalisées en utilisant uniquement un instrument, laissé au choix de l'utilisateur. Pour la segmentation, l'algorithme se base sur les données acquises avec la plateforme robotisée. Néanmoins, dans cette étude, l'étape de prétraitement des données n'utilise qu'une partie des signaux d'acquisition présentés. En effet, l'exécution de ces trajectoires est très largement dépendante de la position de l'instrument mais en aucun cas de son orientation. De plus, les tâches à réaliser ne requièrent pas l'utilisation de la pince. Nous retirons donc l'emploi des quaternions et également l'angle d'ouverture et fermeture de l'instrument pour cette étude. La trame D utilisée pour la segmentation est alors définie par :

$$D^i = \{x^i, y^i, z^i, \kappa^i, \tau^i, \kappa_s^i, \tau_s^i \mid i = 1\} \quad (3.22)$$

où les signaux d'acquisition sont décrits dans le paragraphe 3.2.1 de ce chapitre.

3.2.5.3 Résultats

Nous présentons dans cette partie les résultats de segmentation obtenus sur les deux jeux de données en utilisant notre approche non-supervisée. Bien que cette étude ne soit qu'une vérification, elle permet néanmoins d'appréhender l'approche proposée et de visualiser les performances de segmentation sur différentes données intégrant de la variabilité. Il est également à noter que par souci de visibilité uniquement, seuls les résultats utilisant la métrique de dissimilarité Déformation Temporelle Dynamique seront présentés. En effet, en comparant les différentes métriques pour ces jeux de données, les différences de segmentation en utilisant la métrique de Hausdorff et celle de Fréchet sont respectivement de 1,2% et 2,7% par rapport à la longueur de la trajectoire, ce qui n'est pas significatif pour cette étude de vérification.

Jeu de données *TrajectoireAleatoire*.

Nous présentons deux figures afin d'analyser les résultats de segmentation pour ce jeu de données. Dans un premier temps, la figure 3.12a présente la segmentation de trajectoire obtenue à l'aide des étapes A, B et C de la figure 3.1, ne réalisant pas la sélection des délimiteurs en se basant sur les scores de dissimilarité obtenus.

Cette figure présente la trajectoire réalisée ainsi que les différents segments délimités par les points. Bien que les résultats soient intéressants en termes de cohérence visuelle de la segmentation (en prenant en compte l'aspect global de la courbe), certains délimiteurs semblent néanmoins plus pertinents que d'autres notamment lors de la réalisation de courbes avec une courbure importante. Dans ce cas, les segments retenus semblent similaires, indiquant que ces délimiteurs ne sont pas appropriés pour la segmentation.

Dans un second temps, en appliquant la technique de sélection automatique sur les scores de dissimilarité, nous obtenons une segmentation présentée sur la figure 3.12b. Dans ce cas, nous remarquons que le nombre de délimiteurs a été réduit. De plus, ces derniers permettent de segmenter de manière plus naturelle la trajectoire dans le sens où les courbes ne sont pas composées de plusieurs segments similaires. La segmentation est alors plus cohérente d'un point de vue visuel.

Jeu de données *TrajectoiresLettre*.

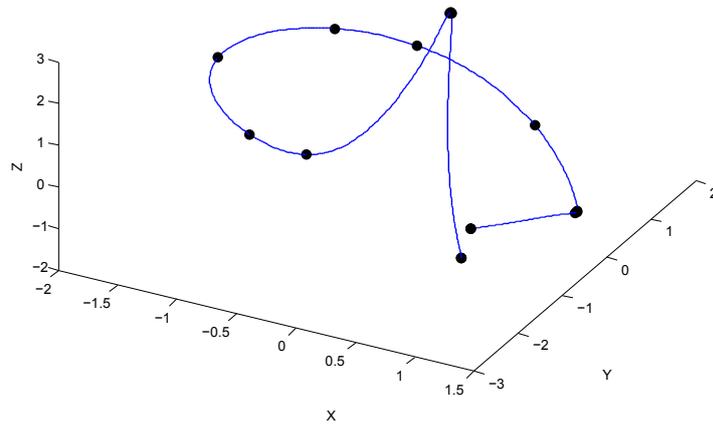
Ce jeu de données nous permet d'analyser la prise en compte de variabilité dans la segmentation. Chaque participant a dessiné la lettre "R" et nous observons les résultats de segmentation en utilisant notre approche non-supervisée. Dans cette analyse, nous utilisons la totalité du processus de segmentation, à savoir les étapes A, B, C et D de la figure 3.1. Les résultats de segmentation sont visibles sur la figure 3.13, où les trajectoires sont présentées en deux dimensions pour plus de visibilité.

En analysant les figures, nous remarquons tout d'abord que les trajectoires réalisées pour dessiner la lettre "R" comportent une variabilité interpersonnelle évidente. Néanmoins, en comparant les positions des différents délimiteurs détectés dans chacune des trajectoires, nous constatons que leurs emplacements sont similaires. Cela nous fournit alors une indication quant à la robustesse de la segmentation. Typiquement, les demi-cercles réalisés n'ont pas la même forme mais sont segmentés de façon similaire. De la même manière, la barre oblique terminant la lettre est reconnue comme un seul segment. Enfin, dans le cas de la courbe verticale permettant de débiter la lettre, nous constatons que pour l'opérateur N°3, cette dernière ne forme qu'un segment alors que pour les deux autres, deux segments sont détectés. Toutefois, nous notons que pour les opérateurs N°1 et 2, ces ondulations générées durant le tracé ne sont pas négligeables, expliquant l'apparition de ces points caractéristiques.

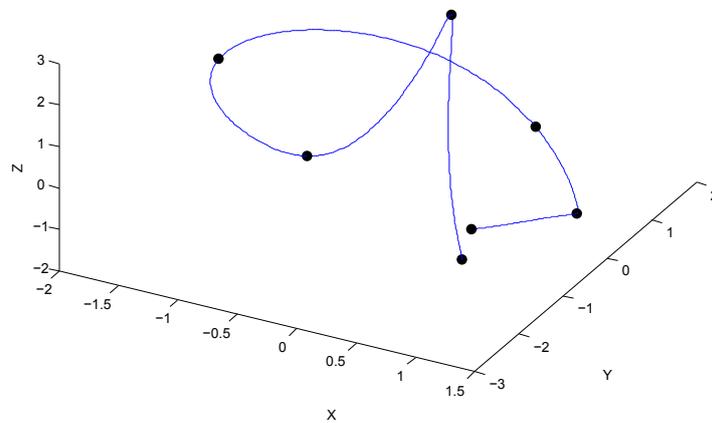
De manière générale, nous voyons par cette étude de vérification que l'approche développée dans ces travaux pour la segmentation non-supervisée de trajectoires fournit des résultats encourageants. Dans des cas simples, l'algorithme décompose les trajectoires de manière cohérente, proche de ce que pourrait faire un humain dans la réalité. Dans la suite de ce chapitre, nous utiliserons les segments définis par l'algorithme pour réaliser l'apprentissage et la reconnaissance des gestes chirurgicaux.

3.3 Classification et reconnaissance des surgemes

À partir de l'algorithme de segmentation non-supervisée introduit précédemment, nous désirons maintenant détecter les différents dexemes afin de reconnaître les surgemes réalisés. Pour cela, nous présentons une approche en trois étapes (figure 3.1) avec une pre-



(a) Segmentation non-supervisée avant la sélection de points pertinents.



(b) Segmentation non-supervisée après la sélection de points pertinents.

FIGURE 3.12: Analyse de la segmentation sur une trajectoire aléatoire. L'algorithme de sélection des scores permet une détection pertinente des points caractéristiques (en noirs), qui définissent les différentes dexemes.

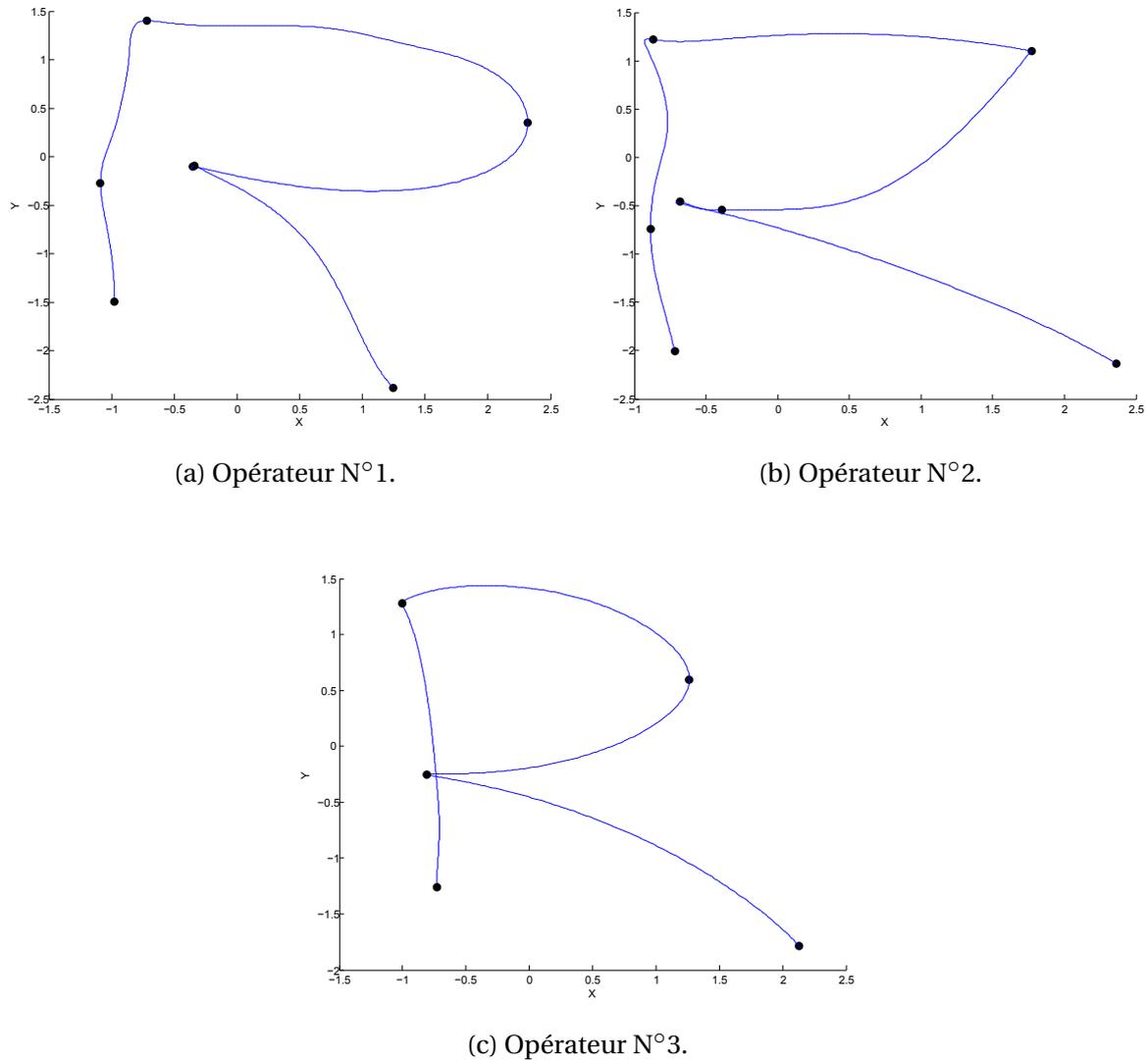


FIGURE 3.13: Trajectoires acquises durant la réalisation de la lettre "R" en utilisant la plateforme robotisée de téléopération.

mière étape d'extraction d'attributs, une seconde étape de classification et une troisième étape d'association des dexemes pour obtenir les surgemes.

3.3.1 Extraction d'attributs caractéristiques

L'extraction d'attributs est une étape importante dans le processus de reconnaissance. En effet, il est important que les données utilisées portent l'information nécessaire avec un degré de discrimination significatif. De plus, l'étape de classification ne s'applique que

dans une dimension unique, obligeant alors à extraire une signature de taille fixe quelque soit la taille des données en entrée. Ainsi, nous devons extraire des caractéristiques des dexemes dans un espace de taille fixe, indépendamment de la taille de ces derniers. Nous réalisons cette étape en deux temps. Tout d'abord, nous transformons l'ensemble des dexemes (et donc des signaux contenus à l'intérieur) pour pouvoir ensuite les exprimer dans un référentiel commun. Ceci afin d'en extraire, dans un second temps, des caractéristiques représentatives de chacun.

3.3.1.1 Transformation des signaux

Chaque dexeme est constitué d'un ensemble de signaux. Cependant, afin de réaliser une extraction homogène, nous devons exprimer ces signaux dans un même référentiel avec une même origine. Pour cela, nous avons réalisé une transformation sur chaque signal contenu dans chacun des dexemes afin de le translater dans son repère d'origine, ayant ainsi pour origine l'abscisse zéro. De ce fait, nous pouvons comparer des signaux de même nature (les positions par exemple) en évaluant leurs déformations par rapport à une même origine.

3.3.1.2 Approximation polynômiale

Les signaux représentant les dexemes sont exprimés dans un référentiel commun, avec la même origine. Nous souhaitons maintenant extraire des caractéristiques représentatives de chaque signal afin de matérialiser l'information qu'il porte, décrivant ainsi une partie du dexeme associé. L'idée est donc d'employer une modélisation mathématique pour représenter ces signaux. En analysant la littérature, il existe une multitude d'attributs permettant de caractériser un signal. Cependant, en se concentrant sur les travaux abordant la caractérisation de signaux de trajectoires, nous avons relevé des modélisations plus pertinentes et performantes que d'autres.

Khalid et al. [Khalid 2005a] ainsi que Naftel et al. [Naftel 2005] effectuent une analyse de plusieurs types d'extraction sur des trajectoires. Ils comparent la modélisation par une transformée de Fourier discrète à celle par un polynôme de Chebyshev et par une approximation polynômiale aux moindres carrés. Ils effectuent ensuite une reconnaissance (utilisant des regroupements non-supervisés) pour évaluer la qualité de l'extraction. Leurs résultats montrent que l'approximation polynômiale fournit de meilleurs résultats [Naftel 2005]. Dans nos travaux, nous utiliserons donc l'approximation polynômiale par moindres carrés.

Considérons $S(x)$ une séquence de données de taille $x \in [1, n]$. Cette dernière peut être approximée par un polynôme de la forme $y = a_0 + a_1x + \dots + a_mx^m$, où $m < n$ en utilisant la méthode des moindres carrés. Pour expliciter l'approche multidimensionnelle utilisée dans ces travaux, nous nous basons sur la figure 3.14. Chaque dexeme est représenté par un ensemble de signaux. En cherchant à caractériser le dexeme, nous cherchons à caracté-

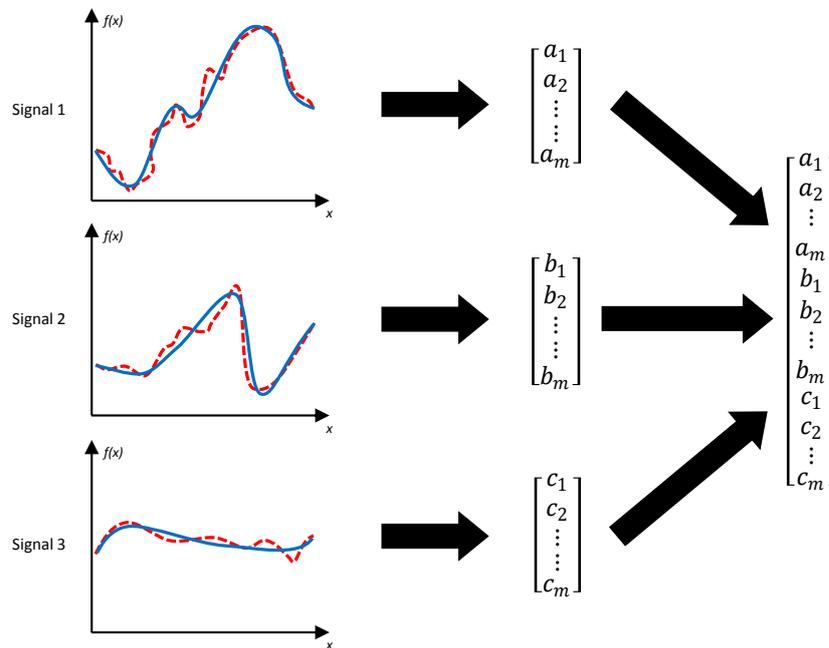


FIGURE 3.14: Schéma de construction du vecteur caractéristique représentant un dexeme à partir de données multidimensionnelles.

riser l'ensemble des signaux. Nous utilisons l'approche des moindres carrés sur chacun des signaux appartenant aux dexemes afin d'en extraire un vecteur d'attributs correspondant aux coefficients des polynômes (figure 3.14). Enfin, pour construire le vecteur d'attributs représentant le dexeme complet, nous concaténons l'ensemble des coefficients de chaque signal (figure 3.14). Ainsi, nous obtenons une signature descriptive du dexeme, représentant l'ensemble des signaux qu'il contient.

Dans la suite de ces travaux, nous évaluons l'impact de l'ordre d'approximation sur les performances de reconnaissance des surgemes.

3.3.2 Algorithmes de classification

Le processus de reconnaissance repose majoritairement sur l'étape de classification. Dans ce manuscrit, nous avons déjà argumenté notre choix d'utiliser le moins de connaissance a priori pour réaliser la reconnaissance des surgemes. Ainsi, nous n'utiliserons pas de réseaux bayésiens (dynamiques ou non) ainsi que de Modèles de Markov Cachés comme cela a été entrepris dans les travaux de la littérature. Ces algorithmes se basent sur la création d'un modèle statistique temporel et nécessitent donc soit de définir une structure du réseau, soit d'utiliser des algorithmes tiers pour réaliser l'apprentissage du modèle en se basant sur un ensemble de données et leurs annotations. Dans notre cas, nous employons

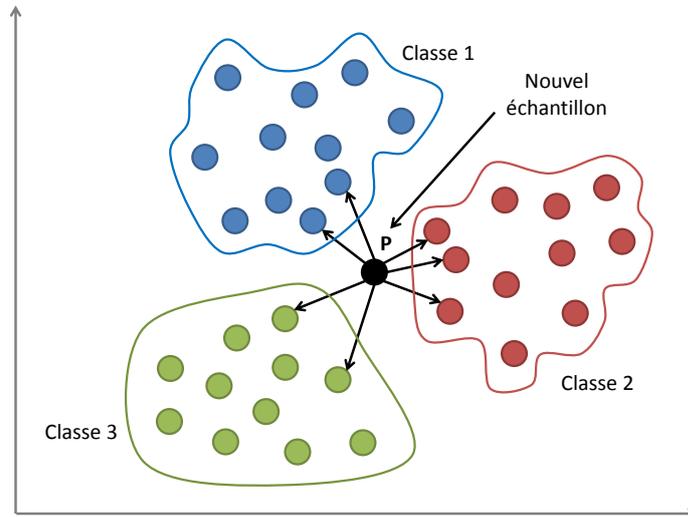


FIGURE 3.15: Schéma de fonctionnement de l'algorithme des K-Plus Proches Voisins.

des méthodes standard de classification qui ont fait leurs preuves dans de précédents travaux [Zappella 2013]. Nous étudions l'emploi de Machines à Vecteurs de Support ainsi que l'utilisation de l'algorithme des K-Plus Proches Voisins.

3.3.2.1 K-Plus Proches Voisins

L'algorithme des K-Plus Proches Voisins (K-PPV) est une technique très utilisée pour réaliser l'attribution d'étiquettes. Nous illustrons le principe au travers de la figure 3.15. Dans notre exemple, les échantillons sont exprimés dans l'espace des attributs. Chaque échantillon représente donc un vecteur de taille fixe. Nous considérons que, dans cet espace, trois classes existent, chacune d'elles étant représentée par une couleur. L'objectif est alors de classer le nouvel échantillon P. Nous paramétrons tout d'abord le nombre de voisins k que nous désirons étudier (dans notre exemple $k = 7$). Puis nous sélectionnons, en utilisant une métrique de distance, les k voisins les plus proches de l'échantillon P. Enfin, nous analysons leur étiquette et effectuons un vote majoritaire. Dans l'exemple proposé, nous voyons que deux voisins appartiennent à la classe 1, deux voisins appartiennent à la classe 3 et trois voisins à la classe 2. L'échantillon P sera donc classé comme un élément de la classe 2, correspondant alors à un surgeme spécifique.

Cette approche, très simple, admet néanmoins des variantes. Tout d'abord le nombre de voisins influe directement sur la classification. Ainsi sélectionner un paramètre d'un voisin peut être risqué dans le sens où si ce voisin est une donnée non représentative (aberrante) de la classe, les risques de mauvaise classification du nouvel échantillon sont élevés. De plus, l'attribution de l'étiquette s'effectuant par un vote majoritaire, il est préférable de

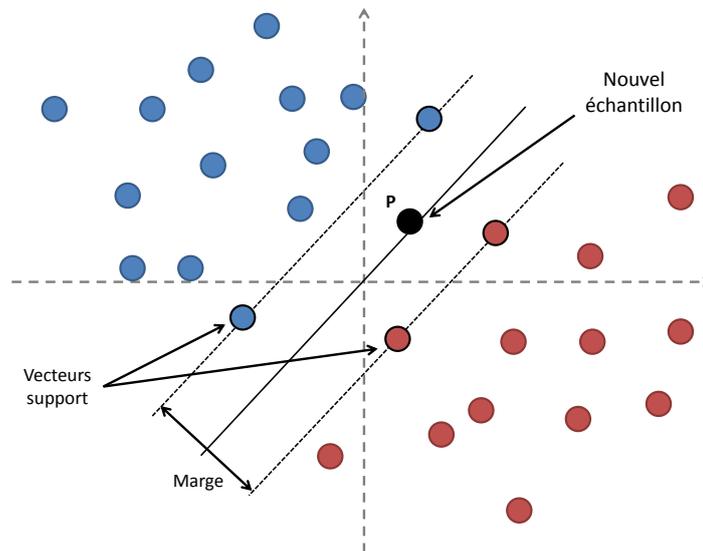


FIGURE 3.16: Schéma de fonctionnement de l'algorithme des Machines à Vecteurs de Support.

sélectionner un nombre de voisins impair. D'autre part, la métrique considérée pour sélectionner les plus proches voisins peut également influencer les performances. Il existe différentes métriques dans la littérature, dont notamment la distance Euclidienne, la distance de Manhattan ainsi que la distance de Mahalanobis. Enfin, il est à noter que pour réaliser cette classification avec ce type d'algorithme, il est nécessaire d'avoir au préalable une base de données avec la totalité des échantillons annotés.

Dans la suite de nos travaux, nous évaluons l'impact du paramètre k du nombre de voisins sur les performances de reconnaissance. De plus, nous utiliserons la distance Euclidienne comme métrique de comparaison car cette dernière semble plus cohérente pour l'analyse de trajectoires réalisées dans l'espace cartésien.

3.3.2.2 Machines à Vecteurs de Support

Le second algorithme que nous employons pour réaliser la classification est l'algorithme des Machines à Vecteurs de Support (MVS) [Bennett 2000], et plus spécifiquement de type I appelé plus communément C-MVS. Son principe est illustré par la figure 3.16. Dans notre exemple, nous considérons un cas de classification binaire simple avec deux classes : une classe bleue et une classe rouge. Cet algorithme d'optimisation cherche à séparer ces deux classes en utilisant la notion d'hyperplan. Dans le cas en deux dimensions, cet hyperplan est alors une droite séparatrice. De manière générale, cet hyperplan est construit en se basant sur des vecteurs supports, c'est-à-dire des échantillons représentant les limites de chacune des classes. Dans notre exemple, la solution du problème est

alors d'optimiser la position de cette droite afin de maximiser la marge, représentant le degré de discrimination des deux classes (figure 3.16). L'algorithme détermine une fonction linéaire de la forme $f(x) = w^T x + b$, où w est le plan séparateur, x les exemples d'apprentissage and b un offset. Cependant, certains problèmes ne sont pas linéairement séparables. Dans ce cas, une astuce est d'employer un noyau de projection afin de trouver un nouvel hyperplan dans un espace ou sous-espace différent. En notant ce noyau $\phi_k(\cdot)$, la classification est réalisée en résolvant le problème :

$$\begin{aligned} \underset{w,b,z}{\text{minimiser}} \quad & \frac{1}{2} w^T w + C \sum_{i=1}^N z_i \\ \text{sous contrainte} \quad & y_i (w^T \phi_k(x_i) - b) + z_i \geq 1 \\ & z_i \geq 0, \quad i = 1, \dots, N \end{aligned} \tag{3.23}$$

où z_i est un paramètre de relâche non-négatif qui pénalise la mauvaise classification de l'échantillon i , et C est une matrice permettant la pondération des erreurs de classification. Dans nos travaux, nous utilisons l'astuce du noyau avec des fonctions de bases radiales. Ces fonctions permettent de classer deux échantillons u et v qui ne sont pas linéairement séparables en utilisant la transformation définie par :

$$K(u, v) = \exp(\gamma \|u - v\|^2) \tag{3.24}$$

où $\gamma = -\frac{1}{2\sigma^2}$ permet la déformation des frontières (hyperplans). Nous utiliserons $\gamma = \frac{1}{\text{Nb class}}$, fournissant le meilleur ratio entre performance et complexité pour la modélisation des classes.

Dans ces travaux, nous évaluons également l'impact du paramètre C de pondération des erreurs de classification sur les performances de reconnaissance.

3.3.3 Association de dexemes

La dernière étape pour la reconnaissance consiste à réaliser l'association des différents dexemes pour reconstituer les surgemes. Cependant, afin de prévenir d'éventuelles erreurs de classification, nous appliquons un filtrage sur les étiquettes des dexemes. Prenons l'exemple d'un surgeme composé d'une séquence de dexemes. Si l'un d'eux relève d'un autre surgeme, alors ce dexeme est sujet à une erreur de classification. Pour pallier ce problème, nous appliquons un filtrage en nous basant sur un fenêtrage de trois voisins. De cette manière, les étiquettes aberrantes sont modifiées afin d'obtenir une séquence homogène, permettant de reconstituer le surgeme dans sa totalité.

3.3.4 Validation de la méthode de reconnaissance

Dans ce paragraphe, nous considérons que les données ont été acquises préalablement durant des sessions d'entraînement robotisé. Pour valider notre approche, nous avons réa-

lisé deux études distinctes. La première étude concerne la validation et la quantification des performances de la phase de classification uniquement. La seconde étude permet de valider l'approche complète de segmentation non-supervisée et reconnaissance des surgemes.

Dans les paragraphes suivants, nous présentons le type de tâche d'entraînement, le vocabulaire des surgemes que nous utilisons pour l'apprentissage, les stratégies de validation des deux études, les données et référentiels choisis, les métriques d'évaluation pour quantifier les performances et enfin nous présentons les résultats.

3.3.4.1 Type de tâche et vocabulaire

Pour décrire la construction du vocabulaire ainsi que la référence d'évaluation, nous allons préalablement définir la tâche d'entraînement réalisée à l'aide de la plateforme robotisée. Cette tâche s'inspire directement des protocoles d'évaluation utilisés par la SAGES et décrit par Derossis et al. [Derossis 1998]. Cette tâche d'entraînement implique le transfert d'objets (ici des pions en silicone) à différents emplacements (figure 3.17). Pour cela, les pions sont préalablement empilés au même endroit (point de départ en bas sur la figure figure 3.17)). L'opérateur doit ensuite réaliser la tâche en respectant les consignes suivantes :

1. Attraper le premier pion avec l'instrument gauche et l'insérer dans la cible 1 (figure 3.17) ;
2. Attraper le second pion avec l'instrument droit et l'insérer dans la cible 2 ;
3. Attraper le dernier pion avec l'instrument gauche ou droit (au choix), se déplacer au centre du plateau, transférer le pion au second instrument et insérer ce dernier dans la cible 3 ;

Durant la réalisation de cette tâche, nous effectuons l'acquisition des données cinématiques avec la plateforme robotisée et enregistrons la vidéo, capturée avec la caméra endoscopique. Les vidéos sont ensuite découpées et synchronisées manuellement avec les données cinématiques. À partir des enregistrements, nous annotons manuellement les vidéos permettant de transférer directement ces annotations sur les données cinématiques, construisant ainsi notre référence appelée vérité terrain. Pour cette référence, nous avons établi un vocabulaire de 12 surgemes présentés dans le tableau 3.1, et représentant les différents gestes à réaliser pour achever la tâche d'entraînement :

3.3.4.2 Stratégie de validation

La stratégie de validation se décompose en deux parties : la description des études et la description des paramètres pris en compte pour optimiser les performances de reconnaissance.

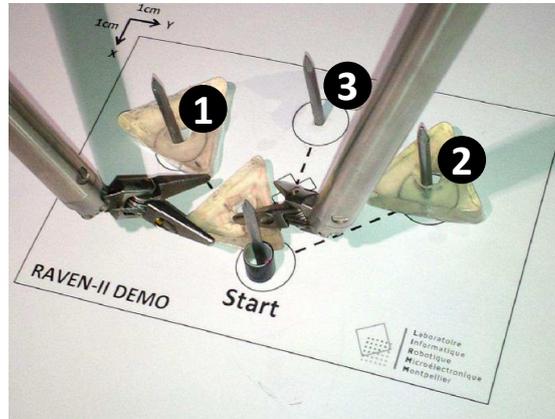


FIGURE 3.17: Illustration de la tâche d'entraînement chirurgical.

N°	Définition	N°	Définition
1	Attendre	7	Se positionner sur la cible
2	Atteindre le pion	8	Insérer le pion dans la cible
3	Se positionner pour attraper le pion	9	Relâcher le pion
4	Attraper le pion	10	Se déplacer pour attendre
5	Extraire le pion	11	Se déplacer vers le centre
6	Se déplacer vers la cible	12	Se déplacer pour terminer

Tableau 3.1: Vocabulaire de surgemes utilisé pour décrire la tâche de transfert d'objet.

Études de validation.

Notre étape de validation comporte deux études distinctes (figure 3.1). La première étude, que nous nommerons *ValidationClassification* vise à étudier et quantifier les performances de classification uniquement, en analysant la reconnaissance des surgemes et en les comparant à la vérité terrain préalablement construite. La seconde étude, que nous nommerons *ValidationReconnaissance* vise à étudier et quantifier les performances de notre processus complet de segmentation et classification, en analysant la reconnaissance des surgemes et en les comparant à la vérité terrain.

Afin d'acquérir des données pour réaliser nos validations, nous avons demandé à trois experts de réaliser la tâche d'entraînement présentée précédemment en utilisant la plateforme robotisée de téléopération. Ces experts sont composés de deux urologues utilisant le robot da Vinci en intervention, ainsi qu'un expert en système de téléopération. En amont de l'acquisition, nous avons fourni à chaque sujet des explications sur le fonctionnement du système ainsi que les consignes pour réaliser la tâche d'entraînement. Nous les avons ensuite laissés s'entraîner pendant plusieurs minutes afin de progresser dans les courbes

	Moyenne [s]	Écart-type
Opérateur N°1	44,40	3,33
Opérateur N°2	39,81	4,61
Opérateur N°3	31,13	1,44

Tableau 3.2: Statistiques sur les durées des trajectoires pour la tâche d'entraînement.

d'apprentissage (apprentissage du système et apprentissage de la tâche). Nous avons acquis un jeu de données composé de trois trajectoires par sujet, soit un ensemble de neuf trajectoires avec les annotations manuelles correspondantes. La moyenne et l'écart-type de la durée des tâches réalisées par chaque opérateur sont données dans le tableau 3.2. Nous pouvons y voir que les écarts-types sont faibles, notamment pour l'opérateur N°3. Cela indique alors que la variabilité intraopérateur n'est pas importante. Cependant, nous remarquons qu'en moyenne plus de 13 secondes séparent l'opérateur le plus rapide de l'opérateur le plus lent sur la même tâche d'entraînement, indiquant une variabilité interopérateur significativement importante.

À partir de ces données, nous avons donc mené deux études, employant un apprentissage avec différents classifieurs, et évalué les performances au travers d'un modèle de validation croisée sur tous les opérateurs. Ce modèle de validation utilise huit trajectoires pour l'apprentissage et la dernière pour réaliser les tests. Les données d'apprentissage et de test sont ensuite mélangées et nous effectuons une rotation des données afin de tester notre approche de reconnaissance sur l'ensemble des données. Dans la suite de ce chapitre, nous fournissons les résultats globaux pour chacune des expériences.

Paramètres d'optimisation.

Dans ce paragraphe, nous distinguons deux types de paramètres : les paramètres pour la segmentation et les paramètres pour la classification.

L'étude *ValidationReconnaissance* fait appel à la totalité du processus de reconnaissance décrit par la figure 3.1. Cela signifie que, dans cette étude, nous ferons varier la métrique de dissimilarité spatio-temporelle afin d'évaluer son impact sur les performances de reconnaissance. Pour rappel, les métriques de dissimilarité sont : la distance de Hausdorff, la distance de Fréchet et Déformation Temporelle Dynamique.

La phase d'apprentissage possède également plusieurs paramètres variables. Nous souhaitons, de la même manière, quantifier leur influence sur les performances de reconnaissance mais également déterminer les paramètres optimaux à employer pour notre approche. Dans les deux études, *ValidationClassification* et *ValidationReconnaissance*, nous considérons les paramètres de classification suivants :

- Degré d'approximation polynômiale, où $m \in \{1, 3, 5, 7, 9\}$;
- Nombre de voisins k pour le classifieur K-PPV, où $k \in \{1, 3, 5, 7\}$;

- Pondération des pénalités de classification C pour le classifieur C-MVS, où C est échantillonné de manière uniforme dans l'espace \log_{10} de 10^{-5} à 10^{10} .

3.3.4.3 Données

Comme décrit précédemment, nous disposons d'un même jeu de données pour les deux études de validation : *ValidationClassification* et *ValidationReconnaissance*. Ce jeu de données est constitué de neuf trajectoires, trois par opérateur. À partir des acquisitions cinématiques, nous calculons la trame complète permettant d'appliquer l'approche de segmentation proposée. La trame D est définie par :

$$D^i = \{x^i, y^i, z^i, qx^i, qy^i, qz^i, qw^i, \kappa^i, \tau^i, \kappa_s^i, \tau_s^i, \alpha^i \mid i \in [1, 2]\} \quad (3.25)$$

où les différents signaux d'acquisition sont décrits dans le paragraphe 3.2.1 de ce chapitre.

3.3.4.4 Métriques d'évaluation

Nous souhaitons quantifier les performances de nos deux études de validation. Pour cela, nous utilisons plusieurs métriques d'évaluation dont la plupart sont communes aux deux études.

Nous employons la métrique de reconnaissance pour les études *ValidationClassification* et *ValidationReconnaissance*. La reconnaissance est la capacité de prédire correctement une hypothèse validée. Elle représente la performance globale sur un problème de classification binaire ou multiclasse, et est définie par :

$$\text{Reconnaissance} = \frac{\text{Dexemes correctement attribués}}{\text{Dexemes total}} \quad (3.26)$$

Nous utilisons aussi la métrique de rappel permettant de représenter les performances locales de reconnaissance (par classe). Cette métrique est utilisée dans chacune des études de validation. Le rappel, appelé également sensibilité en statistique, est une métrique locale permettant de mesurer la capacité à fournir un résultat positif lorsqu'une hypothèse est vérifiée. Le rappel est défini par :

$$\text{Rappel}_i = \frac{\text{Dexemes correctement attribués à la classe } i}{\text{Dexemes appartenant à la classe } i} \quad (3.27)$$

En complément du rappel, nous calculons la précision, également utilisée dans chacune de nos études de validation. La précision est une métrique locale permettant de mesurer la capacité de prédiction d'une hypothèse, qu'elle soit vraie ou fausse. La précision est définie par :

$$\text{Précision}_i = \frac{\text{Dexemes correctement attribués à la classe } i}{\text{Dexemes attribués à la classe } i} \quad (3.28)$$

De plus, afin de prendre en compte les deux dernières métriques dans les études *ValidationClassification* et *ValidationReconnaissance*, nous calculons le score F^1 . Cette métrique correspond à la moyenne harmonique du rappel et de la précision en utilisant des pondérations identiques (ici équivalente à 1). Le score F^1 est défini par :

$$\text{Score}_{F_i^1} = \frac{2 * (\text{Précision}_i * \text{Rappel}_i)}{(\text{Précision}_i + \text{Rappel}_i)} \quad (3.29)$$

Enfin, nous utilisons une dernière métrique, appelée recouvrement, et qui ne sera employée que dans l'étude *ValidationReconnaissance*. Cette métrique permet d'évaluer la correspondance des annotations entre notre approche de reconnaissance et la vérité terrain, en prenant en compte l'aspect temporel de la tâche. Le score de recouvrement est défini par :

$$\text{Recouvrement}(t_i, g_i) = \frac{|\cap(t_i, g_i)|}{\text{longueur}(g_i)} \quad (3.30)$$

où \cap dénote le recouvrement entre la séquence de surgemes t_i et sa vérité terrain correspondante g_i . Le tout est normalisé par la longueur de la séquence. De cette manière, un recouvrement de 1 indique que la totalité des gestes ont été reconnus correctement durant la tâche d'entraînement, sans décalage temporel.

Il est à noter que l'ensemble des résultats associés à ces métriques sera exprimé en pourcentage (%).

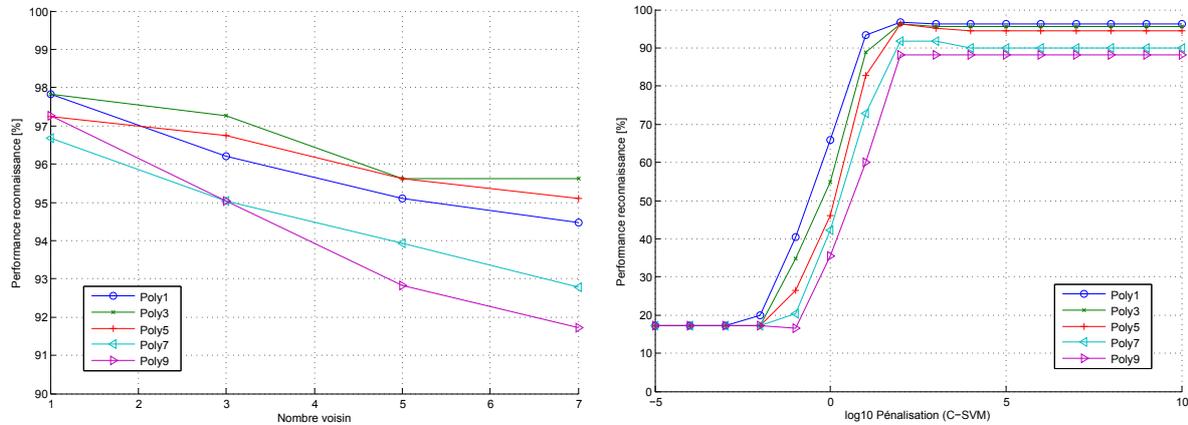
3.3.4.5 Résultats

Les différentes études ont été effectuées post-entraînement. L'analyse des trajectoires comprenant segmentation et classification a été implémentée sous MATLAB R2012b. Les calculs ont été réalisés avec un processeur Intel Core i7-3770 @ 3,40GHz, permettant d'effectuer la segmentation et la reconnaissance des surgemes provenant d'une trajectoire en moins de 5 secondes. Par souci de lisibilité, nous ne montrerons que la reconnaissance des surgemes provenant d'une des deux mains utilisables, la seconde réalisant une séquence similaire afin d'accomplir la tâche d'entraînement.

Étude *ValidationClassification*.

Cette étude utilise la vérité terrain comme entrée pour réaliser la classification. Nous apprenons donc à classer les surgemes et évaluons les performances de l'apprentissage à travers différentes métriques d'évaluation.

La figure 3.18a montre les performances de reconnaissance en utilisant le classifieur des K-Plus proches Voisins (K-PPV), faisant varier le nombre de voisin ainsi que le degré d'approximation utilisé pour l'extraction d'attributs. Nous constatons que les résultats sont élevés et qu'en utilisant peu de voisins, typiquement un, les performances atteignent 97,4% en moyenne, sans prendre en compte le degré d'approximation polynômiale. Toutefois,



(a) Performance de classification avec K-PPV. (b) Performance de classification avec C-MVS.

FIGURE 3.18: Graphique des performances de reconnaissance avec les deux classifieurs étudiés.

lorsque le nombre de voisins augmente, le taux de reconnaissance diminue. Bien qu'un nombre faible de voisins permette une meilleure classification, la réponse du classifieur se veut plus robuste lorsque le nombre de voisins est élevé car la comparaison se fait avec un plus grand nombre d'échantillons. Cependant, cela montre également que les classes de surges ne sont pas si éloignées les unes des autres dans l'espace de représentation. Employer un voisin semble donc plus performant, mais dans la pratique ce paramètre augmentera la sensibilité aux erreurs de classification. Dans notre cas, utiliser 3 ou 5 voisins semble cohérent par rapport à la taille de la base de données. En analysant le degré d'approximation employé pour réaliser l'extraction d'attributs caractéristiques, nous constatons également que les performances sont variables, favorisant alors les polynômes de degré faible, en l'occurrence *Poly1* et *Poly3*.

Dans le cas de la figure 3.18b, nous constatons que, lorsque les erreurs de classification sont peu pénalisées, les performances sont très faibles, de l'ordre de 18%, quel que soit le degré d'approximation employé. Néanmoins, lorsque nous cherchons à pénaliser les mauvaises classifications en augmentant la valeur de C , nous voyons que les performances augmentent mettant en avant l'importance des échantillons situés à l'intersection des classes. Le paramètre C affecte les frontières des classes en les déformant afin d'éviter les erreurs de classification. Ainsi, lorsque $C \in [10^3, 10^{10}]$, indiquant une forte pénalisation des erreurs, les performances de reconnaissance sont similaires avec l'emploi d'un classifieur K-PPV. De la même manière qu'avec le précédent classifieur, les performances augmentent également lorsque le degré d'approximation polynomiale est réduit. De ce fait, en utilisant un polynôme de degré 1, les performances de reconnaissance atteignent 96,2%.

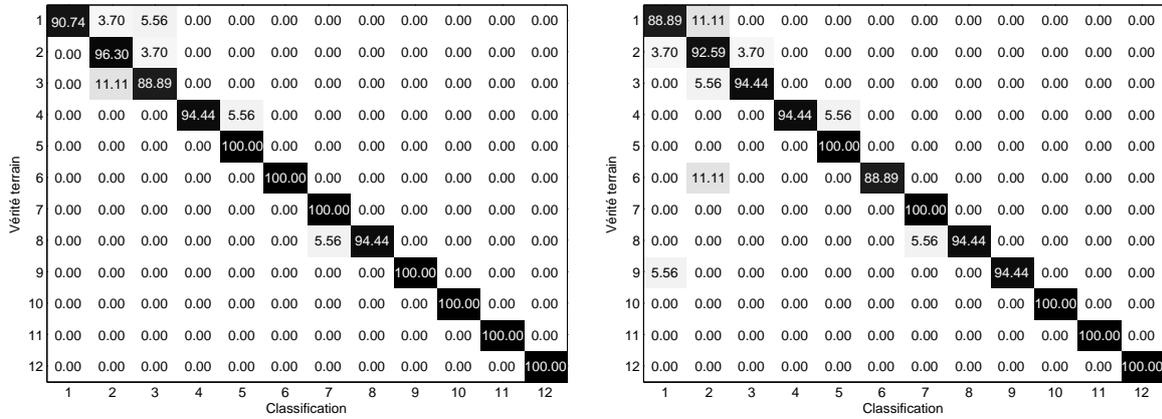
Nous remarquons donc que l'impact du degré d'approximation polynomiale influe de manière significative sur les performances des classifieurs. Là où un polynôme d'ordre 7 ou

Surgeme	Méthode	Classifieur	Précision	Rappel	F1
			[Moyenne ± Ecart-type %]	[Moyenne ± Ecart-type %]	[%]
Attendre	Vérité terrain + Classification	5-PPV	100,00 ± 0,00	90,74 ± 14,10	95,15
		MVS	95,00 ± 10,00	88,89 ± 23,57	91,84
Atteindre le pion	Vérité terrain + Classification	5-PPV	88,89 ± 16,67	96,30 ± 11,11	92,44
		MVS	85,93 ± 22,47	92,59 ± 14,70	89,13
Se positionner pour attraper le pion	Vérité terrain + Classification	5-PPV	89,81 ± 15,47	88,89 ± 22,05	89,35
		MVS	96,30 ± 11,11	94,44 ± 16,67	95,36
Attraper le pion	Vérité terrain + Classification	5-PPV	100,00 ± 0,00	94,44 ± 16,67	97,14
		MVS	100,00 ± 0,00	94,44 ± 16,67	97,14
Extraire le pion	Vérité terrain + Classification	5-PPV	96,30 ± 11,11	100,00 ± 0,00	98,11
		MVS	96,30 ± 11,11	100,00 ± 0,00	98,11
Se déplacer vers la cible	Vérité terrain + Classification	5-PPV	100,00 ± 0,00	100,00 ± 0,00	100,00
		MVS	88,89 ± 33,33	88,89 ± 33,33	88,89
Se positionner sur la cible	Vérité terrain + Classification	5-PPV	96,30 ± 11,11	100,00 ± 0,00	98,11
		MVS	96,30 ± 11,11	100,00 ± 0,00	98,11
Insérer le pion dans la cible	Vérité terrain + Classification	5-PPV	100,00 ± 0,00	94,44 ± 16,67	97,14
		MVS	100,00 ± 0,00	94,44 ± 16,67	97,14
Relâcher le pion	Vérité terrain + Classification	5-PPV	100,00 ± 0,00	100,00 ± 0,00	100,00
		MVS	100,00 ± 0,00	94,44 ± 16,67	97,14
Se déplacer pour attendre	Vérité terrain + Classification	5-PPV	100,00 ± 0,00	100,00 ± 0,00	100,00
		MVS	100,00 ± 0,00	100,00 ± 0,00	100,00
Se déplacer vers le centre	Vérité terrain + Classification	5-PPV	100,00 ± 0,00	100,00 ± 0,00	100,00
		MVS	100,00 ± 0,00	100,00 ± 0,00	100,00
Se déplacer pour terminer	Vérité terrain + Classification	5-PPV	100,00 ± 0,00	100,00 ± 0,00	100,00
		MVS	100,00 ± 0,00	100,00 ± 0,00	100,00
Performance moyenne	Vérité terrain + Classification	5-PPV	97,61 ± 4,53	97,07 ± 6,72	97,29
		MVS	96,56 ± 8,26	95,68 ± 11,52	96,07

Tableau 3.3: Rappel, précision et score F^1 pour chaque surgeme en utilisant la vérité terrain.

9 permet de modéliser avec plus de précision les signaux, un polynôme d'ordre 1 ou 3 généralisera cette modélisation. L'impact se fait ressentir puisque ce sont les polynômes d'ordre inférieur qui permettent de généraliser les signaux, autorisant des scores de reconnaissance plus élevés. Pour la suite de l'étude, nous souhaitons réaliser un compromis entre performance et généralisation. Nous utiliserons donc un polynôme de degré 5 pour modéliser les signaux dans l'étape d'extraction d'attributs. De plus, afin d'éviter une étude exhaustive, nous avons également choisi le paramètre optimal à employer pour chaque classifieur en se basant sur les résultats des figures 3.18a et 3.18b. Ainsi, le K-PPV se fera avec 5 voisins et le C-MVS utilisera une pénalisation de $C = 10^3$.

L'analyse détaillée des performances de reconnaissance de chaque surgeme est présentée dans le tableau 3.3, ainsi que les matrices de confusion correspondantes dans la figure 3.19. Les résultats présentés sont moyennés sur l'ensemble des sessions utilisées dans le modèle de validation croisée. Nous fournissons moyenne et écart-type pour chaque valeur de rappel et précision extraite à partir des matrices de confusion des figures 3.19a et 3.19b. En analysant le tableau, nous remarquons que les scores de rappel et précision sont élevés, certains atteignant 100% quel que soit le classifieur utilisé. Logiquement, les scores F^1 sont également élevés montrant la qualité d'apprentissage et de reconnaissance de l'approche développée. Néanmoins, nous remarquons que dans certains cas, l'écart-type est élevé, indiquant de fortes variations. Dans ce cas, nous supposons que cela provient des données et qu'une ou plusieurs trajectoires ne permettent pas d'obtenir des scores aussi élevés, remettant en cause la qualité desdites données. Les performances moyennes sont légère-



(a) Matrice de confusion avec les K-PPV. (b) Matrice de confusion avec les C-MVS.

FIGURE 3.19: Matrices de confusion représentant les répartitions de classification par rapport aux surgemes de référence en se basant sur la vérité terrain.

ment meilleures avec le classifieur K-PPV utilisant 5 voisins, avec lequel nous atteignons des moyennes de 97,61%, 97,07% et 97,29% en rappel, précision et score F^1 respectivement. Les matrices de confusion des figures 3.19a et 3.19b montrent les scores de reconnaissance pour chaque classe de surgeme. Nous constatons que la plupart des surgemes sont très bien reconnus (les valeurs sur la diagonale indiquant les taux de reconnaissance). Néanmoins, le surgeme *Attendre* entraîne quelques vrais négatifs, c'est-à-dire des confusions dans la prédiction avec d'autres surgemes similaires et proches, de l'ordre de 8% à 11% suivant le classifieur utilisé. Bien que ces valeurs soient minimales, elles indiquent néanmoins que nous devons prêter attention à ces surgemes dans la suite de notre étude. Au travers de ces premiers résultats, nous constatons que notre système est largement capable d'apprendre et de reconnaître les surgemes lorsque nous utilisons la vérité terrain comme référence d'entrée. Nous souhaitons donc évaluer de la même manière les performances de reconnaissance en utilisant en amont la méthode de segmentation non-supervisée proposée, afin de l'appliquer cette fois-ci aux dexemes.

Étude ValidationReconnaissance.

Cette étude utilise la totalité du processus décrit dans ce chapitre. Partant des données, nous segmentons les trajectoires pour en extraire les dexemes. Nous classifions ensuite ces derniers afin de reconnaître les surgemes réalisés. De la même manière que l'étude précédente, nous évaluons les performances de reconnaissance globale et nous analysons également les résultats obtenus pour chaque surgeme.

D'après les courbes des figures 3.20a, 3.20b et 3.20c, nous constatons que la métrique de segmentation employée pour sélectionner les meilleurs délimiteurs temporels influe sur les performances de reconnaissance. Avec la métrique de Hausdorff (figures3.20a), les per-

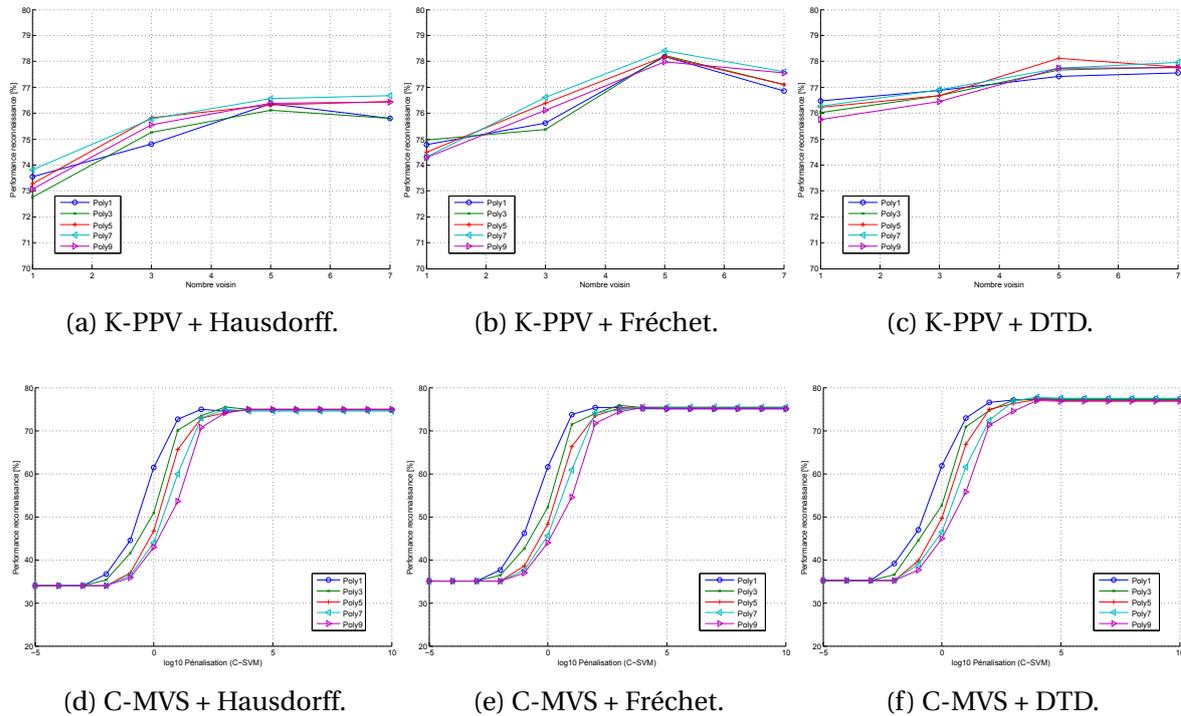


FIGURE 3.20: Graphique des performances de reconnaissance avec les deux classifieurs étudiés ainsi que les différentes métriques de segmentation utilisées.

performances de reconnaissance atteignent 76,7% pour un nombre de 7 voisins et un degré d'approximation polynômiale de 7. En utilisant la métrique de Fréchet, nous atteignons un score de 78,4% avec 5 voisins et un degré d'approximation de 7. Enfin, en utilisant la métrique de Déformation Temporelle Dynamique, les scores de reconnaissance atteignent 78,2% avec un paramètre de 5 voisins pour le K-PPV et une approximation polynômiale de degré 5. En analysant la tendance générale des courbes, nous voyons également que suivant le paramètre et la métrique de similarité choisie, l'influence des paramètres est discutable. Dans le cas de la figure 3.20c, les variations sont faibles quels que soient les paramètres d'extraction d'attributs et de classification. Cela indique donc une meilleure robustesse dans l'apprentissage, favorisant l'utilisation de cette configuration. De manière générale, ce sont cette fois-ci les approximations d'ordres 5 et 7 qui fournissent les meilleurs résultats, en se basant sur un nombre de 5 voisins pour la classification.

De l'autre côté, les figures 3.20d, 3.20e et 3.20f ne sont pas très différentes les unes des autres. En effet, la tendance générale est similaire quelle que soit la configuration utilisée, favorisant néanmoins la métrique de Déformation Temporelle Dynamique qui permet d'atteindre un maximum de 77,5% lorsque $C \in [10^3, 10^{10}]$. De plus, dans cette plage de pénalisation, nous remarquons que le classifieur C-MVS atteint des limites et que les différents paramètres n'influent que très peu sur la reconnaissance.

Surgeme	Méthode	Classifieur	Précision		Rappel		F1 [%]
			[Moyenne ± Ecart-type %]		[Moyenne ± Ecart-type %]		
Attendre	Seg. Auto + Classification	5-PPV	85,01 ± 6,25		93,77 ± 4,23		89,17
		MVS	83,72 ± 5,47		94,25 ± 6,64		88,67
Atteindre le pion	Seg. Auto + Classification	5-PPV	72,75 ± 18,73		81,39 ± 21,03		76,83
		MVS	75,93 ± 14,10		77,22 ± 24,89		76,57
Se positionner pour attraper le pion	Seg. Auto + Classification	5-PPV	31,48 ± 24,22		41,67 ± 39,53		35,86
		MVS	32,96 ± 19,82		45,37 ± 36,59		38,18
Attraper le pion	Seg. Auto + Classification	5-PPV	64,81 ± 18,99		65,74 ± 17,40		65,27
		MVS	56,30 ± 31,77		38,89 ± 19,98		46,00
Extraire le pion	Seg. Auto + Classification	5-PPV	89,07 ± 16,81		73,33 ± 11,37		80,44
		MVS	72,04 ± 13,28		85,74 ± 13,95		78,29
Se déplacer vers la cible	Seg. Auto + Classification	5-PPV	85,19 ± 22,74		83,33 ± 25,00		84,25
		MVS	82,59 ± 26,97		88,89 ± 22,05		85,63
Se positionner sur la cible	Seg. Auto + Classification	5-PPV	62,96 ± 41,48		57,41 ± 37,14		60,06
		MVS	61,30 ± 39,30		66,67 ± 41,46		63,87
Insérer le pion dans la cible	Seg. Auto + Classification	5-PPV	68,52 ± 24,22		82,22 ± 26,82		74,75
		MVS	62,96 ± 33,10		65,56 ± 36,09		64,23
Relâcher le pion	Seg. Auto + Classification	5-PPV	77,04 ± 28,06		54,63 ± 25,04		63,93
		MVS	74,07 ± 26,50		54,63 ± 13,89		62,88
Se déplacer pour attendre	Seg. Auto + Classification	5-PPV	100,00 ± 0,00		68,52 ± 24,22		81,32
		MVS	100,00 ± 0,00		79,63 ± 24,69		88,66
Se déplacer vers le centre	Seg. Auto + Classification	5-PPV	92,59 ± 14,70		94,44 ± 16,67		93,51
		MVS	96,30 ± 11,11		87,96 ± 19,14		91,94
Se déplacer pour terminer	Seg. Auto + Classification	5-PPV	100,00 ± 0,00		92,59 ± 14,70		96,15
		MVS	100,00 ± 0,00		87,04 ± 20,03		93,07
Performance moyenne	Seg. Auto + Classification	5-PPV	77,45 ± 18,02		74,09 ± 21,93		75,13
		MVS	74,85 ± 18,45		72,65 ± 23,28		73,17

Tableau 3.4: Rappel, précision et score F^1 pour chaque surgeme en utilisant notre approche de reconnaissance.

D'après cette analyse, nous remarquons que notre choix dans l'étude précédente se révèle cohérent. En effet, quel que soit le classifieur choisi ou son paramètre, l'approximation polynômiale d'ordre 5 permet un bon compromis entre généralisation des données et spécificité. De la même manière que dans l'étude précédente, nous utiliserons donc ce paramètre pour l'étude approfondie des surgemes, ainsi qu'un nombre de voisins de 5 pour le K-PPV et une pénalisation de $C = 10^3$ pour le C-MVS, nous permettant d'atteindre des scores de reconnaissance de 78,2% et 77,1% avec la métrique de Déformation Temporelle Dynamique.

Comme l'étude précédente, le tableau 3.4 et la figure 3.21 présentent une analyse détaillée des performances de notre approche de segmentation non-supervisée, combinée à la classification. D'après le tableau 3.4, nous remarquons que le K-PPV avec un nombre de 5 voisins offrent des performances légèrement supérieures, de l'ordre de 2% en moyenne, à celles du classifieur C-MVS avec le paramètre optimal. Les scores de rappel et précision sont néanmoins inférieurs ici d'environ 20% en moyenne du fait de l'utilisation de la technique de segmentation. Néanmoins, les scores sont très encourageants car certains surgemes tels que *Se déplacer pour attendre* et *Se déplacer pour terminer* sont appris puis reconnus avec un rappel et une précision de 100%. À l'inverse, d'autres surgemes tels que *Se positionner pour attraper le pion*, *Attraper le pion* et *Se positionner sur la cible* atteignent des scores de rappel et précision plus faibles. Ces résultats sont discutés dans le paragraphe suivant. Enfin, nous constatons que les surgemes les plus importants tels que *Extraire le*

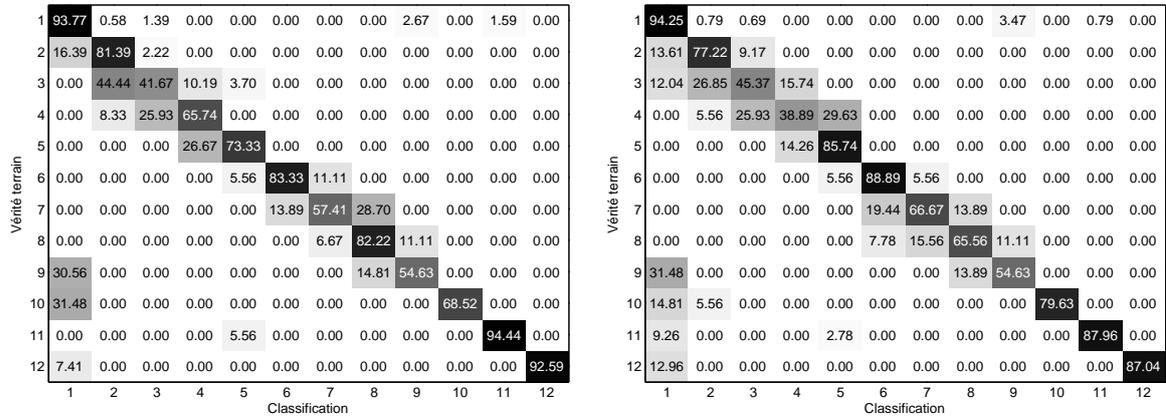
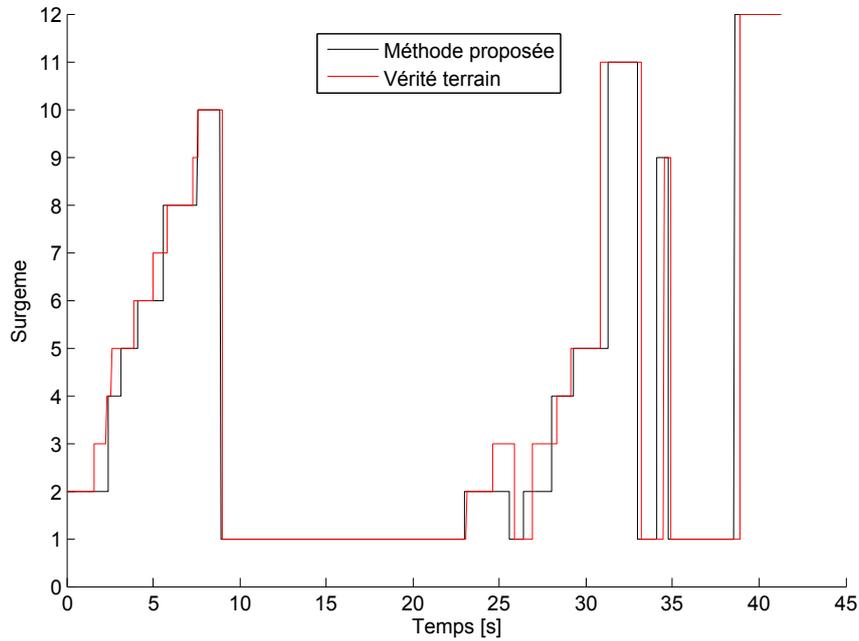
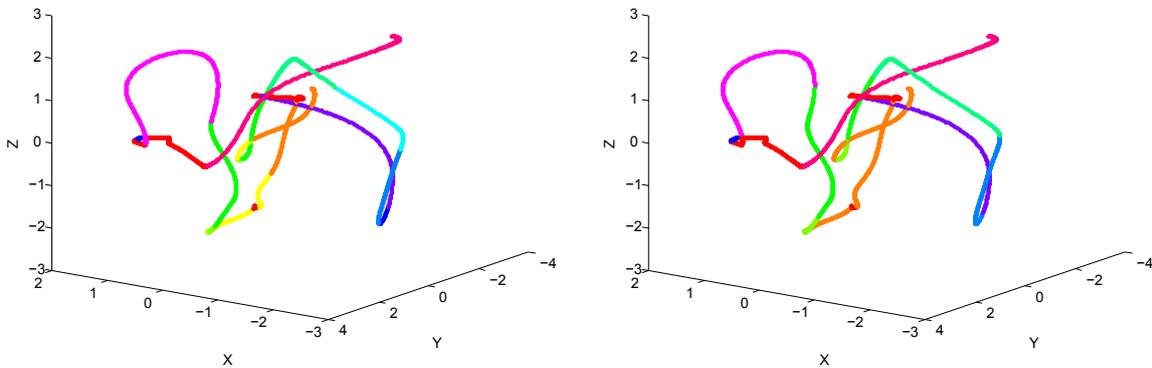


FIGURE 3.21: Matrices de confusion représentant les répartitions de classification par rapport aux surgemes de référence avec notre approche de reconnaissance.

pion, *Se déplacer vers la cible* et *Insérer dans la cible* sont reconnus avec des scores élevés. En complément, les matrices des figures 3.21a et 3.21b présentent les mêmes caractéristiques, avec des scores moins élevés que lors de l'utilisation de la vérité terrain, mais très largement acceptables. De plus, grâce à ces matrices, nous observons que le surgeme *Attendre* crée toujours des perturbations dans la reconnaissance, réduisant les scores de rappel d'autres surgemes. Ces effets sont également discutés dans le paragraphe suivant. Pour notre dernière étude, nous souhaitons également analyser la qualité de la reconnaissance en prenant en compte l'aspect temporel de la tâche. Pour cela, nous utilisons la métrique de recouvrement. La figure 3.22a permet d'analyser la chronologie des surgemes provenant de l'annotation manuelle (en rouge) avec la chronologie reconnue par notre approche de reconnaissance (en noir). Dans ce cas, la configuration utilisée repose sur le paramétrage obtenu avec la métrique DTD pour la segmentation, le classifieur K-PPV avec 5 voisins et l'approximation polynômiale d'ordre 5. Le score de recouvrement pour l'exemple présenté est de 81,1%, ce qui signifie que nous avons reconnu le bon geste au bon moment plus de 8 fois sur 10. De manière plus générale, la moyenne du recouvrement sur l'ensemble de la base de données est de 81,9% avec un écart-type de 2,4% sur la totalité des trajectoires. De plus, afin de visualiser le résultat en trois dimensions, nous utilisons la trajectoire correspondante à la chronologie de la figure 3.22a sur laquelle nous avons coloré chacun des surgemes. Cette trajectoire est visible sur la figure 3.22b. Ainsi, nous pouvons comparer cette dernière avec son homologue de la figure 3.22c, coloré grâce aux reconnaissances des surgemes en utilisant l'approche proposée dans ces travaux. Pour conclure, nous constatons que notre approche de classification fournit de très bonnes performances, et que la combinaison de la segmentation non-supervisée avec la classifi-



(a) Comparaison des chronologies de surgemes réalisés durant la tâche d'entraînement.



(b) Visualisation 3D des surgemes de référence. (c) Visualisation 3D des surgemes reconnus.

FIGURE 3.22: Comparaison spatiale et temporelle des surgemes de références et des surgemes reconnus par notre approche de reconnaissance.

cation atteint un score d'environ 75% (score F^1), très encourageant pour la poursuite de ces travaux.

3.3.5 Analyse et discussion

Dans cette partie, nous abordons certains aspects de notre méthode de reconnaissance en détaillant les points clés ainsi que les modifications envisageables. Nous discutons également des résultats présentés précédemment.

3.3.5.1 Segmentation non-supervisée

L'approche de segmentation et reconnaissance proposée dans ces travaux est innovante dans le sens où, à notre connaissance et au moment de la rédaction de ce manuscrit, aucune publication ne fait référence à une telle méthodologique pour la reconnaissance des gestes chirurgicaux. Toutefois, de récents travaux [Zhou 2015b] abordent l'aspect de segmentation en appliquant des algorithmes de classification. De cette manière, les auteurs cherchent à détecter les points caractéristiques des courbes, puis à segmenter cette dernière en apprenant à reconnaître ces points. Cette approche, complexe et particulièrement supervisée, fournit des résultats mitigés en termes de rappel et précision, montrant que l'approche développée comporte encore quelques faiblesses. De plus, aucune analyse de gestes n'a été réalisée à partir de cette méthode. L'approche proposée dans ce manuscrit diffère des travaux de la littérature en ce sens qu'elle introduit une nouvelle approche multi-granularité, indépendante de l'aspect temporel durant la réalisation de la tâche et mettant en avant les étapes nécessaires pour le prétraitement des données. Nous souhaitons dans la suite aborder certains points méthodologiques développés au travers de ce chapitre.

Prétraitement des données.

Le prétraitement des données est une étape très importante dans la phase de segmentation. Les données cinématiques et le calcul de signaux invariants en rotation et translation peuvent fournir une quantité d'informations importantes pour la reconnaissance des gestes chirurgicaux.

Lors d'analyses préliminaires, nous cherchions à évaluer l'impact des signaux sur la reconnaissance. Partant de l'intuition que seules les positions peuvent suffire pour réaliser la segmentation et la reconnaissance, nous avons réalisé les études présentées précédemment. Néanmoins, s'appuyer sur les informations de position uniquement fût un échec car les scores de reconnaissance ne dépassaient pas les 35%. Nous avons donc modifié notre structure de données en nous appuyant sur les travaux de Cifuentes et al. [Cifuentes 2013] qui utilisent les quaternions pour la représentation des orientations, réduisant le nombre de signaux à traiter. De plus, nous remarquons que la majorité des travaux développés à l'université de John Hopkins [Lin 2005, Lin 2006, Reiley 2008, Reiley 2010, Béjar Haro 2012, Tao 2012, Tao 2013, Zappella 2013, Lea 2015] utilisent une quantité d'informations cinématiques très importante (78 variables cinématiques au total), représentant notamment positions, orientations et vitesse des instruments ainsi que des interfaces de manipulation.

Afin de réaliser l'apprentissage, leurs approches emploient généralement une réduction de dimension comme présenté dans le chapitre 2, perdant alors toute information quant à la pertinence de certains signaux ainsi que des grandeurs physiques correspondantes. Dans notre cas, l'utilisation de ce type d'approche n'a apporté aucune amélioration pour la reconnaissance des gestes chirurgicaux.

Dans nos travaux de thèse, nous nous concentrons sur l'utilisation de variables significatives. Un total de 24 variables nous permet de représenter les mouvements des outils de manière complète et utilisable dans le cadre de la reconnaissance comme le prouvent les résultats présentés. Cette trame plus courte permet également d'accélérer les temps de calcul. Bien que la méthodologie ne soit actuellement applicable que hors-ligne, nous souhaiterions y remédier dans une poursuite de ces travaux.

Enfin, notre approche de filtrage, suivie d'une normalisation, comme précédemment employée par Lin et al. [Lin 2005, Lin 2006] nous permet de réduire la variabilité intra- et interpersonnelle, nous autorisant également à effectuer l'analyse des signaux de manière multidimensionnelle sans contraintes d'échelles.

Sélection des délimiteurs temporels par persistance.

Utilisant la notion de simplification topologique associée à la notion de persistance, la sélection des délimiteurs temporels est l'étape clé du processus de segmentation non-supervisée. Nous avons souhaité évaluer les performances de ce type d'approche prometteuse, permettant de fournir des informations intéressantes d'un point de vue spatio-temporel. Néanmoins, cette sélection est réalisée par l'utilisation d'un seuil de persistance qui doit être défini en amont par l'utilisateur. Nous avons entrepris plusieurs expériences afin d'estimer de manière empirique les seuils à utiliser pour chacun des signaux contenus dans la trame de traitement. Il est à noter que pour l'ensemble des expériences, les seuils de segmentation sont restés identiques.

Une amélioration de ces travaux peut également concerner l'évaluation du seuil optimal à utiliser pour chaque signal. L'utilisation d'algorithmes évolutionnistes peut être une alternative intéressante pour faire évoluer le jeu de paramètres et observer les résultats de reconnaissance [Cully 2015]. La convergence de ces algorithmes peut alors permettre de détecter le jeu de paramètres optimal à utiliser, en se basant sur une base de données plus importante.

Métriques de similarité.

Les métriques de Hausdorff et de Fréchet ont été rapidement exclues au vu de leurs taux de reconnaissance inférieurs à ce que nous pouvons atteindre en utilisant les Déformations Temporelles Dynamiques. Cette dernière métrique a d'ailleurs prouvé sa supériorité dans certains travaux de la littérature [Morris 2009, Wang 2013]. Cependant, l'algorithme de Déformation Temporelle Dynamique possède deux principaux inconvénients. Le premier est le temps de calcul qui, par l'emploi de méthodes plus évoluées et l'utilisation

de contraintes pour l'alignement des séquences temporelles pourrait être réduit. De plus, l'utilisation de cet algorithme pour l'analyse multidimensionnelle peut s'avérer complexe et demande quelques prérequis. Ten Holt et al. [Ten Holt 2007] ainsi que Shokoohi-Yekta et al. [Shokoohi-Yekta 2015] ont étudié le problème de reconnaissance de gestes humains en se basant sur des comparaisons multidimensionnelles. Leurs résultats confirment que cette approche est supérieure à une analyse unidimensionnelle seulement dans le cas où les signaux sont normalisés, comme dans notre approche. Ils affirment également que la présence de bruits dans les données peut rapidement perturber l'algorithme, faussant les résultats. Encore une fois, notre fonction de filtrage permet de minimiser ces problèmes, motivant d'autant plus l'intérêt du prétraitement des données. De futurs travaux peuvent tout de même évaluer l'influence de la dimension d'analyse en termes de performances de reconnaissance.

Étude de vérification.

Notre étude de vérification nous permet d'analyser la qualité de la segmentation d'un point de vue visuel. En effet, dans ce contexte d'étude, réaliser une comparaison avec une segmentation manuelle faite par un opérateur n'a que peu de sens, pour plusieurs raisons. La première est que nous faisons l'hypothèse que les segments de trajectoires (les *dexemes*) permettent de représenter une intention dans le geste. Nous cherchons alors des points caractéristiques permettant de définir les limites de ces intentions, nous donnant ainsi un ensemble de délimiteurs temporels de la trajectoire. Cependant, l'intention n'est pas uniquement indiquée par des points caractéristiques dans une trajectoire mais plutôt par une volonté, dirigée par l'expérience de l'opérateur, sa dextérité ainsi que ses connaissances procédurales de la tâche en cours. La seconde raison est que la segmentation manuelle est subjective et dépend de l'opérateur employé pour réaliser cette tâche. Ainsi, chaque personne effectuera une segmentation différente suivant sa perception, son apprentissage et son expertise dans le domaine. Enfin, la troisième raison est que le niveau d'échelle pour l'analyse entre notre segmentation et une référence experte est différent. Là où l'homme observe, analyse et comprend la totalité de la trajectoire, les algorithmes s'emploient à trouver des points caractéristiques ou définir des segments pertinents, ne prenant généralement en compte qu'une sous-partie de la trajectoire. Ces travaux s'appliquent uniquement à l'analyse visuelle du résultat sur les trajectoires, mais une évaluation plus poussée devra être menée afin de quantifier les performances de l'approche de segmentation proposée.

3.3.5.2 Classification et reconnaissance

Bien que la segmentation non-supervisée soit la principale contribution de ce chapitre, nous souhaitons également discuter certains points concernant l'étape de classification pour la reconnaissance.

Extraction d'attributs.

Dans ce travail, nous nous concentrons sur le calcul d'extraction d'attributs caractéristiques provenant des signaux à des fins de classification. L'approche utilisée par approximation polynômiale fournit une description complète et vraisemblablement discriminante des surges considérés dans cet entraînement. Comme nous l'avons présenté dans les résultats, l'ordre d'approximation influence fortement les performances des différents classifieurs. D'un côté, nous pouvons argumenter que l'utilisation d'un ordre élevé permet de représenter au mieux les données, avec le moins de perte d'information. Cependant, l'idée générale est de détecter des gestes humains dans une base de données impliquant de la variabilité dans l'exécution. De ce fait, nous souhaitons modéliser du mieux possible les différents gestes afin de pouvoir généraliser la reconnaissance à d'autres opérateurs ainsi que d'autres tâches d'entraînement. Dans notre cas, l'utilisation de polynômes d'ordres faibles a permis de réaliser cette généralisation, améliorant les performances de reconnaissance. Toutefois, afin de maximiser le ratio entre performance et généralisation, nous avons choisi l'utilisation d'un polynôme d'ordre 5 afin d'éviter les effets de sur-apprentissage. En sortie, nous obtenons un vecteur caractéristique de 72 coefficients par instrument (polynôme de degré 5 impliquant 6 coefficients par signal, avec 12 signaux par instrument), suffisant pour représenter l'ensemble des informations et de taille raisonnable pour pouvoir réaliser par la suite du calcul en ligne.

Néanmoins, une alternative à notre approche peut être d'étudier en détail les différents descripteurs qui existent, tels que ceux implémentés dans la librairie PCL, afin de définir un vecteur d'attribut peut être plus pertinent et performant pour notre approche.

Algorithmes de classification.

Nos études de validation comprennent l'utilisation de deux classifieurs que sont les K-Plus Proches Voisins, ainsi que les Machines à Vecteurs de Support. Les C-MVS sont très employées dans le domaine de la classification de par leur paramétrage complet, leur rapidité de résolution d'un problème de discrimination et leur capacité de réduction des informations. Les classes ne sont ainsi plus représentées par des échantillons annotés mais par un ensemble d'hyperplans adaptatifs. De l'autre côté, l'algorithme des K-Plus Proches Voisins s'appuie sur l'ensemble des données pour la classification, rendant les calculs parfois longs lorsque l'on emploie un nombre de voisins important. Ces approches ne nécessitent aucun modèle statistique temporel de la tâche, les rendant ainsi très intéressantes dans notre application visant une méthode non-supervisée. Le K-PPV possède également des avantages indéniables que nous discutons dans la suite de ce paragraphe.

Petitjean et al. [Petitjean 2011, Petitjean 2014] montrent que l'utilisation des K-PPV associée à une analyse en utilisant l'algorithme de Déformation Temporelle Dynamique peut accélérer la classification. Dans ces travaux, les auteurs construisent grâce aux DTD une séquence moyenne des signaux, puis réalisent une comparaison avec les K-PPV basée uniquement sur cette séquence. L'intérêt est double. Tout d'abord la vitesse d'exécution est

améliorée car les comparaisons ne s'effectuent qu'entre deux échantillons : l'échantillon à tester et l'échantillon moyen, permettant de réaliser une classification par rapport à cette distance. Le deuxième avantage est sans aucun doute la construction d'une séquence moyenne en prenant en compte les caractéristiques spatio-temporelles des différentes séquences de la base de données. Une application directe de ces travaux est l'utilisation d'un tel type d'algorithme pour construire un ensemble de dexemes moyens, considérés alors comme optimaux, afin de réaliser la comparaison d'un nouvel échantillon avec ces représentations optimales. L'intérêt réside tout particulièrement dans la généralisation du geste et la vitesse de calcul, mais également dans la construction d'une trajectoire optimale à réaliser. Ainsi, ces trajectoires peuvent être employées pour fournir un retour visuel à l'opérateur. De telles applications ont déjà été introduites par Padoy et al. [Padoy 2011] où le système affiche la trajectoire qu'il réalisera automatiquement. Ce type d'approche mériterait plus d'attention dans les années à venir pour parfaire les systèmes d'entraînement en chirurgie robotisée.

Études de validation.

Nos études de validation ont été réalisées avec une base de données de neuf trajectoires provenant de trois sujets ayant des niveaux d'expertise différents. Néanmoins, cela reste peu par comparaison avec les travaux de la littérature où certains acquièrent plus de 25 répétitions par tâche [Dosis 2005, Béjar Haro 2012, Tao 2012, Tao 2013, Zappella 2013, Lea 2015]. Ainsi, cette étude n'est qu'une première étape à de futurs développements sur des bases de données plus importantes, notamment celle présentée par Gao et al. [Gao 2014].

Concernant les performances de reconnaissance des surgemes, les résultats démontrent que l'approche basée sur la segmentation non-supervisée fournit des performances inférieures à un apprentissage direct des surgemes provenant de la vérité terrain, comme c'est le cas dans la très grande majorité des travaux de la littérature. Néanmoins, les performances obtenues, de l'ordre de 75% de reconnaissance et d'un recouvrement d'environ 82%, permettent d'encourager ce type d'approche, même si actuellement une disparité existe dans les performances de reconnaissance de chaque surgeme. C'est notamment le cas avec les surgemes *Se positionner pour attraper le pion*, *Attraper le pion* et *Se positionner sur la cible* où les scores de rappel et précision sont les plus faibles. Ces erreurs sont dues à la nature du geste mais également à la méthode de segmentation.

Durant une tâche bimanuelle de téléopération, l'opérateur possède ce que nous appelons une main dominante et une main non-dominante [Ní Choisdealbha 2011]. Alors que le contrôle de la main dominante est pleinement dépendant des compétences psychomotrices de l'opérateur, la main non-dominante peut additionnellement être influée par les problèmes de coordination, de découplage des mouvements et de spatialisation [Jarrassé 2014, Dupin 2014]. Ces effets, principalement mis en avant lors d'exercices éva-

lant la latéralité bimanuelle, s'appliquent tout particulièrement dans le cadre de la télé-opération, générant ainsi des "perturbations" dans certains mouvements demandés. Dans notre exercice, lorsque la main dominante est en mouvement et que la main-dominée est supposée rester stable, cette dernière génère toutefois des trajectoires, similaire à celles de la main dominante mais à échelle réduite. Il en résulte un "bruit" dans le mouvement, non représentatif du geste à réaliser, qui affecte directement nos performances de reconnaissance. Toutefois, l'approche de segmentation non-supervisée proposée est également remise en cause. En effet, l'utilisation de l'ouverture et de la fermeture de la pince ne représente pas une caractéristique assez forte parmi les 11 autres signaux utilisés dans notre approche. Il serait donc intéressant d'augmenter son influence afin d'améliorer les capacités de détection des surgemes *Attraper le pion* et *Relâcher le pion*. De la même manière, les surgemes *Attendre* et *Se positionner pour attendre* son confondus. Cet effet est causé par la technique de segmentation qui prend en compte l'aspect spatial des données, mais néglige totalement l'aspect temporel. Ainsi, des informations complémentaires, comme la norme de la vitesse, peuvent fournir des informations discriminantes pour améliorer leurs distinctions.

Dans nos travaux, les surgemes les plus importants sont néanmoins correctement reconnus, avec des performances relativement élevées, laissant présager des applications spécifiques destinées à améliorer l'entraînement pour l'acquisition des compétences techniques nécessaires en chirurgie robotisée.

3.4 Conclusion

Dans ce chapitre, nous présentons une nouvelle approche pour la segmentation et la reconnaissance des gestes à partir des trajectoires des instruments chirurgicaux. Cette approche s'appuie dans un premier temps sur une segmentation non-supervisée, prenant en compte uniquement les variations incluses dans les signaux acquis. Nous prétraitons les données et calculons des signaux complémentaires invariants en rotation et translation permettant d'obtenir une signature complète pour le traitement et la reconnaissance des surgemes. Afin de segmenter les trajectoires, nous utilisons une méthode récente de détection de points caractéristiques que nous appliquons afin d'extraire un ensemble de délimiteurs temporels dans ces trajectoires. Ces délimiteurs sont ensuite évalués avec trois métriques caractérisant les déformations spatio-temporelles afin de garder les points permettant de décrire les segments de trajectoires les plus hétérogènes.

Dans un second temps, nous réalisons une classification des différents segments. Nous extrayons des attributs caractéristiques des signaux de chaque segment, puis nous employons des techniques d'apprentissage automatique pour la classification des segments. Enfin, nous associons l'ensemble de ces segments classés afin de reconstituer les différents gestes chirurgicaux.

Nous évaluons notre approche à travers deux études de validation permettant de quantifier

les performances de notre méthode de classification, ainsi que d'évaluer les performances complètes de la méthodologie proposée dans ce chapitre. Enfin, bien que les résultats soient très encourageants, certaines évolutions sont envisagées afin d'améliorer les performances de reconnaissance obtenues.

L'approche pour la reconnaissance des gestes chirurgicaux présentée dans ce chapitre a pour objectif d'aider à l'évaluation des compétences techniques lors de l'entraînement en chirurgie robotisée. Cependant, pour appliquer cette méthodologie, il est nécessaire d'acquérir au préalable les trajectoires des instruments chirurgicaux, à l'aide d'un système robotisé ou au minimum instrumenté. Le chapitre suivant présente notre seconde contribution scientifique qui aborde l'évaluation d'une nouvelle interface manuelle et sans contact pour la commande de robots chirurgicaux, dans le cadre de l'entraînement en chirurgie laparoscopique robotisée. Cette étude a pour objectif de mettre en avant l'utilisation de ce type d'interface, moins onéreuse et très accessible, afin de contrôler un robot ou bien un simulateur chirurgical.

Évaluation d'une nouvelle interface sans contact pour l'entraînement en chirurgie robotisée

Préambule

Ce chapitre présente notre seconde contribution pour l'entraînement en chirurgie robotisée, en évaluant une nouvelle interface de commande destinée au contrôle des robots chirurgicaux et des simulateurs. Nous réalisons dans un premier temps l'état de l'art des différents systèmes existants pour la commande de ces systèmes. Nous introduisons ensuite la plateforme robotique de téléopération mise en place pour l'entraînement aux gestes chirurgicaux. Puis, au travers d'une première étude de faisabilité, nous démontrons l'utilisation de cette interface et quantifions les performances en entraînement. Enfin, nous montrons dans une seconde étude comment ce système peut permettre d'évaluer le niveau d'expertise de l'opérateur.

Sommaire

4.1	État de l'art des interfaces homme-machine pour la commande des robots chirurgicaux	108
4.1.1	Commande par interface haptique	109
4.1.2	Commande par interface sans contact	112
4.1.3	Analyse et discussion	115
4.2	Étude d'une nouvelle interface sans contact pour l'entraînement en chirurgie robotisée	117
4.2.1	Plateforme expérimentale pour l'entraînement en chirurgie robotisée	117
4.2.1.1	Le robot Raven-II	117
4.2.1.2	Les interfaces homme-machine	119
4.2.1.3	La tâche d'entraînement chirurgical	121

4.2.2	Évaluation pour l'entraînement en chirurgie robotisée	123
4.2.2.1	Protocole expérimental	123
4.2.2.2	Données et métriques d'évaluation	123
4.2.2.3	Résultats	124
4.2.2.4	Analyse et discussion	126
4.3	Vers une analyse des trajectoires pour l'évaluation du niveau d'expertise .	128
4.3.1	Amélioration de la commande sans contact	128
4.3.1.1	Modification du modèle	128
4.3.1.2	Filtrage des tremblements	129
4.3.2	Méthodologie d'analyse des trajectoires	130
4.3.2.1	Prétraitement des données	130
4.3.2.2	Approche non-supervisée pour la classification	130
4.3.3	Vérification de l'interface et du niveau d'expertise	133
4.3.3.1	Protocole expérimental	133
4.3.3.2	Données	133
4.3.3.3	Résultats	133
4.3.3.4	Analyse et discussion	136
4.4	Conclusion	137

4.1 État de l'art des interfaces homme-machine pour la commande des robots chirurgicaux

Dans le chapitre 3, nous avons introduit le développement d'une nouvelle approche pour la segmentation et la reconnaissance des gestes chirurgicaux. Toutefois, cette méthodologie requiert au préalable l'acquisition de trajectoires des instruments chirurgicaux. Dans ce chapitre, nous introduisons l'utilisation d'une nouvelle interface de commande sans contact afin de contrôler les robots et simulateurs chirurgicaux, dans le but de générer ces trajectoires. À partir des données acquises, nous réalisons également une étude de faisabilité pour l'utilisation de cette interface dans le cadre de l'entraînement en chirurgie laparoscopique robotisée.

Depuis l'utilisation des premiers robots pour le médical [Kwoh 1988, Sackier 1994], les systèmes destinés à la chirurgie ont évolué ainsi que leurs dispositifs de commande. Ces dispositifs, appelés interfaces homme-machine, peuvent être de différentes natures selon leur utilisation et leurs caractéristiques. Dans ce chapitre, nous décrivons de manière non-exhaustive les différents types d'interfaces homme-machine existants pour les robots de téléopération en chirurgie uniquement. Nous ne traiterons pas des systèmes de réhabilitation et rééducation, ni des systèmes de comanipulation qui répondent à des besoins spécifiques et possèdent des architectures matérielles distinctes. Pour plus d'informations

sur ces dispositifs, et pour une analyse exhaustive des systèmes existants notamment haptiques, nous renvoyons le lecteur aux publications de Coles et al. [Coles 2011], Tobergte et al. [Tobergte 2011], Simorov et al. [Simorov 2012] ainsi qu'aux travaux de thèse très détaillés de Santos-Carreras [Santos-Carreras 2012b].

Après analyse de la littérature, nous avons classé les interfaces homme-machine permettant de contrôler les robots chirurgicaux en deux grandes catégories. La première représente les interfaces dites haptiques, faisant intervenir, d'une part, le canal de communication qu'est le geste et, d'autre part, nécessitant une préhension manuelle et donc un contact entre la main de l'opérateur et le dispositif. À l'inverse, la seconde catégorie regroupe l'ensemble des interfaces dites sans contact (ne nécessitant pas de contact tactile entre l'opérateur et le dispositif), faisant intervenir des canaux de communication différents telles que l'information vocale, l'information oculaire, l'information posturale ainsi que l'information gestuelle.

4.1.1 Commande par interface haptique

La majorité des dispositifs robotisés destinée à la chirurgie emploient une ou des interfaces haptiques. Ces interfaces permettent le contrôle généralement en position ou en vitesse des instruments chirurgicaux. Ce type de dispositif est majoritairement répandu par le fait qu'il autorise un retour kinesthésique provenant de l'interaction entre le corps et le dispositif en question, interaction facilitant alors le retour proprioceptif vers l'opérateur. Ces interfaces peuvent également permettre, suivant la technologie embarquée, un retour d'effort vers l'opérateur pour améliorer l'interaction entre l'homme et l'environnement. Nous présentons dans la suite de cette partie un aperçu de la littérature en décrivant les systèmes les plus couramment utilisés. Nous structurons notre analyse suivant l'architecture mécanique utilisée pour la conception de ces manipulateurs. Ces derniers peuvent être de nature série ou parallèle (figure 4.1). La différence entre ces architectures est le nombre de chaînes cinématiques employées pour relier l'opérateur à la console. Le choix d'une architecture ou d'une autre s'effectue en prenant en compte les caractéristiques techniques de fonctionnement mais également l'espace de travail disponible et souhaité pour le contexte de téléopération.

Manipulateur à structure mécanique série.

Les manipulateurs à structure mécanique série sont des systèmes robotisés composés d'un ensemble de corps mécaniques reliés par des articulations, constituant une chaîne cinématique unique comme présentée sur la figure 4.1a. Comparées aux architectures mécaniques parallèles, les architectures séries autorisent généralement un espace de travail plus important avec un faible encombrement. Toutefois, les efforts utiles appliqués sur et applicables par l'effecteur sont généralement plus faibles ainsi que leur précision. Leur principal avantage reste le prix, habituellement inférieur à celui des structures parallèles. Dans la catégorie des manipulateurs séries mécaniques, nous pouvons citer les interfaces

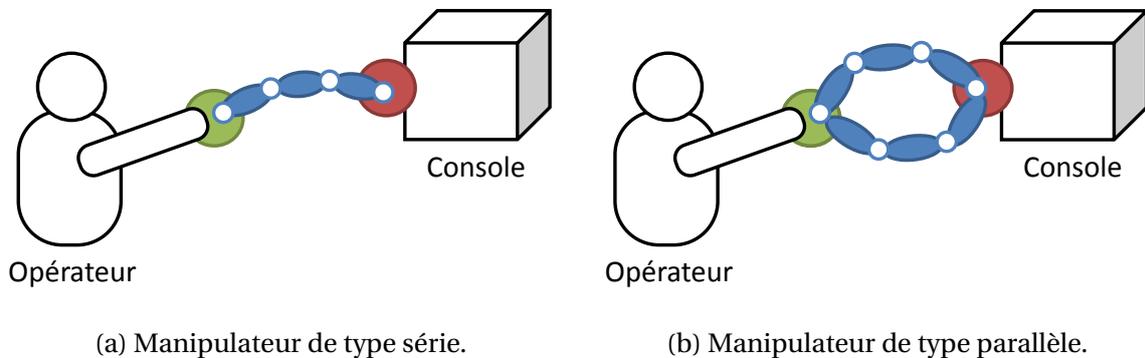


FIGURE 4.1: Les principales architectures cinématiques des manipulateurs haptiques.

du robot Zeus (figure 4.2a) et du robot da Vinci (figure 4.2b) [Guthart 2000, Freschi 2013]. Ces dispositifs ont été certifiés pour une utilisation au bloc opératoire. La principale caractéristique qui les différencie est la cinématique de l'effecteur pour la préhension du dispositif. Dans leurs versions actuelles, aucun retour d'effort n'a été développé pour ces interfaces.

La commande des robots, et plus particulièrement chirurgicaux, peut également s'effectuer à partir d'interfaces standardisées, adaptables à plusieurs plateformes robotisées. C'est principalement le cas des interfaces industrielles développées par SensAble (figure 4.3a) et Haption (figure 4.3b).

Ces interfaces possèdent 6 degrés de liberté dans l'espace opérationnel cartésien, permettant un contrôle total en rotation et translation de l'effecteur du robot. Les degrés de liberté complémentaires correspondent à la réalisation d'une action comme l'ouverture ou la fermeture de la pince du robot. En complément, certains dispositifs sont équipés pour fournir un retour d'effort dans l'espace opérationnel. Dans le cas du Phantom Omni, seuls les 3 degrés de liberté en translation sont soumis au retour d'effort, là où le Virtuose 6D fournit un retour d'effort complet sur les 6 degrés de liberté dans l'espace opérationnel.

Manipulateur à structure mécanique parallèle.

En complément des manipulateurs séries, il existe également les manipulateurs à structure mécanique parallèle. Dans ce cas, les membres de l'opérateur sont reliés à la console au travers plusieurs chaînes cinématiques comme indiqué sur la figure 4.1b. De plus en plus utilisé comme interface de téléopération, ce type de dispositifs possède plusieurs avantages. Ils permettent des mouvements plus rapides et autorisent des dynamiques plus élevées du fait de la réduction du poids, notamment grâce au déport de l'actionnement sur la base. De plus, comparée aux architectures mécaniques séries, la présence de chaînes cinématiques fermées augmente les capacités de charge du robot et garantit une meilleure précision.

Parmi les manipulateurs parallèles, nous retrouvons principalement des interfaces stan-



(a) Console du robot Zeus.
Source: Université de Pittsburgh.



(b) Console du robot da Vinci Xi.
Source: Intuitive Surgical.

FIGURE 4.2: Interfaces de commande séries de robots chirurgicaux.



(a) Dispositif Phantom Omni.
Source: SensAble.



(b) Dispositif Virtuose 6D.
Source: Haption.

FIGURE 4.3: Dispositifs mécaniques séries pour la commande de robots.

dardisées, adaptables à plusieurs plateformes robotisées. C'est notamment le cas des interfaces développées par Force Dimension, qui sont considérées comme l'une des références dans le domaine.

La principale différence entre les deux dispositifs présentés à la figure 4.4 est le nombre



(a) Dispositif Omega.7.
Source: Force Dimension.



(b) Dispositif Sigma.7.
Source: Force Dimension.

FIGURE 4.4: Dispositifs mécaniques parallèles pour la commande de robots.

de degrés de liberté dans l'espace opérationnel soumis au retour d'effort. Dans le cas de l'Omega.7, seules les translations et l'ouverture et fermeture de la pince en sont équipés. Dans le cas du Sigma.7, les 7 degrés de libertés en sont équipés. Il est à noter qu'une version modifiée de l'Omega.7, appelée Omega.3, autorise un retour d'effort sur 3 degrés de liberté dans l'espace opérationne, et est actuellement utilisé en clinique pour manipuler le Sensei Robotic Catheter System développé par Hansen Medical.

Dans le cadre de l'entraînement en chirurgie robotisée, il existe également des structures mécaniques parallèles dites à câbles. C'est notamment le cas du SPIDAR [Murayama 2004] et du dV-Trainer de chez Mimic Technologies (figure 4.5). Ces interfaces utilisent uniquement des câbles pour transmettre les mouvements de l'opérateur et permettent également un retour d'effort.

Parmi les interfaces haptiques, il existe également une minorité de systèmes ne rentrant pas dans les catégories présentées ci-dessus et faisant appel à d'autres technologies. C'est notamment le cas du Maglev 200, développé par Butterfly Haptics, qui utilise le principe de lévitation de Lorentz [Hollis 1993] et fournit 6 degrés de liberté dans l'espace opérationnel tout supprimant les frottements.

4.1.2 Commande par interface sans contact

À l'opposé des interfaces haptiques, les interfaces dites sans contact permettent la commande des robots en utilisant d'autres canaux de communication telles que l'information vocale, l'information oculaire, l'information posturale ainsi que l'information ges-



(a) Dispositif SPIDAR-G&G.

Source: Institut de Technologie de Tokyo.



(b) Plateforme d'entraînement dV-Trainer.

Source: Mimic Technologies.

FIGURE 4.5: Dispositifs à structures parallèles à câbles pour la commande de robots.

tuelle. Nous présentons dans la suite quelques exemples pour chaque canal de communication permettant la manipulation des robots chirurgicaux.

Manipulation par interface vocale.

Le contrôle des robots chirurgicaux peut se faire de manière vocale. Pour cela, le chirurgien possède un micro fixé autour du cou ou au niveau de la tête afin de définir vocalement les actions souhaitées. Initialement développé pour la commande du robot porte-endoscope AESOP [Nathan 2006], ce type de manipulation a été ensuite utilisé pour contrôler le robot LER (Light Endoscopic Robot) [Berkelman 2003]. La version industrielle du LER, nommée ViKY, est développée et produite par la société EndoControl et également certifiée pour une utilisation au bloc opératoire (figure 4.6). La commande vocale est disponible sur les deux versions du robot : la version porte-endoscope et la version pour le positionnement de l'utérus lors d'interventions gynécologiques.

À l'inverse d'une commande basée sur la gestuelle, la commande vocale ne permet pas un positionnement fin du robot. Cette contrainte est notamment due à un manque d'information lors de la communication car le vocabulaire employé est trop restreint pour réaliser des mouvements précis. Enfin, uniquement 3 degrés de liberté peuvent être facilement pilotés par la voix.

Le principal avantage de la commande vocale reste la suppression de l'interface haptique, libérant les mains du chirurgien ou de l'assistant durant l'intervention pour réaliser d'autres actions.

Manipulation par suivi de la tête ou du regard.

De manière très similaire, les mouvements de la tête peuvent servir à commander des robots. Ce type d'interface a été employé par Nishikawa et al. [Nishikawa 2001] pour contrô-



FIGURE 4.6: Robot ViKY porte endoscope contrôlé via une interface vocale.
Source: Endocontrol.

ler en temps-réel un système robotisé porte-endoscope à partir de marqueurs visuels que sont les yeux du chirurgien et d'un marqueur placé sur son calot. Dans le même objectif, Elle et al. [Elle 2002] ont équipé le chirurgien d'un capteur magnétique placé sur sa tête pour réaliser le suivi des mouvements. Cette approche a été employée pour contrôler le robot porte-endoscope EndoAssist [Gilbert 2009], comme présenté sur la figure 4.7.

Bien que ces interfaces ne requièrent aucune action manuelle de la part du chirurgien, elle peuvent entraîner de la fatigue et des troubles musculo-squelettiques (TMS) [Santos-Carreras 2012a]. De plus, ce type de commande semble moins naturel du fait que l'opérateur doit fixer le moniteur tout en déplaçant sa tête, ce qui peut entraîner perte d'attention et fatigue oculaire.

Une solution alternative est d'utiliser l'information du regard qui est plus naturelle. Despinoy et al. [Despinoy 2013] présentent une plateforme robotique pilotée grâce à une commande hybride basée sur le suivi du regard combiné à une interface haptique. Certains mouvements sont automatisés par le regard, les autres sont pilotés via l'interface manuelle. Il est à noter que dans les deux cas, pour le suivi de la tête et le suivi du regard, l'exécution de gestes fins n'est pas réalisable du fait de la faible précision dans les mouvements de la tête ainsi que du bruit important contenu dans le suivi du regard, obligeant également l'opérateur à regarder un point fixe pour éviter le mouvement du robot.

Manipulation par suivi des mains.

L'analyse de la littérature fait également apparaître de nouvelles interfaces de commande basées sur la capture de mouvements humains à partir de caméras. Plus spécifiquement dans le cadre de la téléopération, les auteurs cherchent à détecter et suivre les mains et les doigts de l'opérateur afin de reproduire leurs mouvements au travers du robot.



FIGURE 4.7: Commande du robot EndoAssist par le suivi de la tête.

Source: Prosurgics.

Du et al. [Du 2012] utilisent une caméra Kinect de Microsoft pour suivre les mouvements des mains de l'opérateur et contrôler un robot industriel. De la même manière, Dragan et al. [Dragan 2013] comparent l'utilisation d'une Kinect avec des interfaces manuelles pour contrôler les mouvements d'un bras robotisé industriel. Enfin, Hernoux et al. [Hernoux 2013] utilisent la Leap Motion pour capturer les mouvements d'un doigt de l'opérateur durant une tâche d'écriture afin de reproduire la trajectoire générée avec un robot industriel.

L'utilisation de ce type de manipulation est encore trop éloignée des applications médicales. Ainsi, Padoy et Hager de l'Université John Hopkins ont travaillé sur la mise au point d'une interface sans contact utilisant une Kinect pour contrôler les mouvements d'un bras du robot da Vinci (figure 4.8¹). Cependant, leur travail ne traite de la commande que d'un seul bras robotisé utilisé pour l'insertion d'une aiguille dans le cadre d'une suture, et nécessite quand même l'emploi des deux mains de l'opérateur, là où une interface haptique n'en requerrait qu'une. L'ouverture et la fermeture de la pince est commandée par deux états uniquement (main ouverte ou fermée, sans position intermédiaire), et tous les degrés de liberté de l'instrument chirurgical ne sont pas commandés. L'utilisation de ces travaux dans un contexte clinique n'est pas envisageable dans l'état actuel.

4.1.3 Analyse et discussion

Il existe différentes interfaces de commande dans la littérature permettant de piloter les robots, et plus particulièrement les robots chirurgicaux. Ces interfaces sont classées ici

1. À notre connaissance, ces travaux n'ont fait l'objet d'aucune publication. Cependant, une vidéo est accessible sur Internet à l'adresse suivante : https://youtu.be/YsY_A0kLh-g.

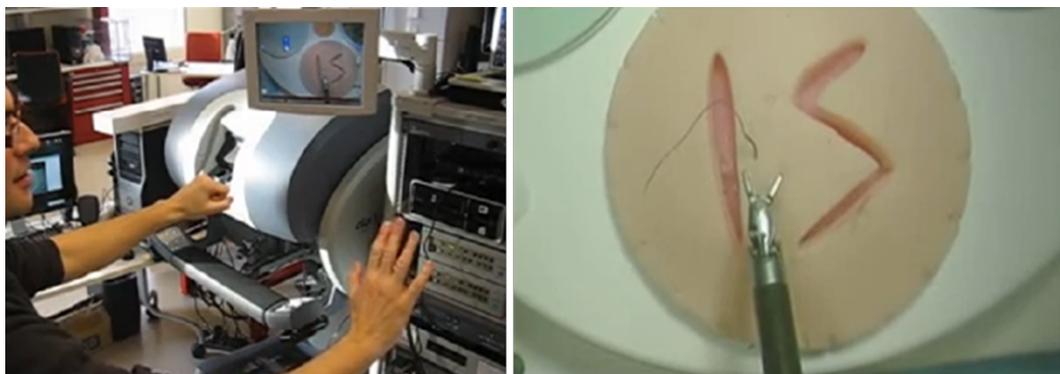


FIGURE 4.8: Interface sans contact pour la commande d'un instrument du da Vinci.
Source: Travaux de Padoy et Hager, 2011

en deux grandes catégories que sont les interfaces haptiques dites avec contact, et les interfaces sans contact ne requérant aucune interaction tactile entre la main de l'opérateur et le dispositif. Dans le cas des interfaces haptiques, le chirurgien commande l'effecteur du robot en mimant ses mouvements à travers l'interface mécanique, en contact direct avec ses mains. Ce type de manipulation permet la réalisation de gestes fins et précis ce qui est idéal pour l'utilisation en chirurgie laparoscopique robotisée. Cependant, dans le cadre d'une utilisation au bloc opératoire, la nécessité du contact entre le dispositif et la main du chirurgien ne permet plus à ce dernier d'être stérile. Une solution serait de draper l'interface dans une housse stérile, ce qui réduirait drastiquement les capacités de manipulation du clinicien. De plus, le coût de ces dispositifs est important du fait notamment des coûts de recherche et développement, des coûts du matériel embarqué et des coûts de maintenance.

Afin de résoudre ces problèmes, des développements ont été entrepris pour proposer de nouvelles solutions innovantes, permettant de manipuler les systèmes robotisés tout en supprimant l'interaction tactile entre la main de l'opérateur et le dispositif. C'est notamment le cas des systèmes commandés par la voix ainsi que par les mouvements de la tête et du regard. Néanmoins dans la pratique, ces types d'interactions ne permettent que la manipulation d'instrument (souvent la caméra endoscopique) qui nécessite moins de précision et de répétabilité qu'un geste chirurgical. Il est également à noter que ces interactions ne favorisent pas une utilisation prolongée, car elles peuvent entraîner fatigue et TMS pour le praticien.

De ce fait, certains travaux tentent de combiner les capacités de manipulation fine fournies par les interfaces haptiques avec l'emploi d'interfaces sans contact. Ces dernières, généralement basées sur des caméras (caméras stéréoscopiques ou caméras avec carte de profondeur), capturent des images permettant le suivi des membres de l'opérateur, et notamment de ses mains. Bien qu'améliorant l'espace de travail, la précision, la répétabilité

et l'ergonomie, des développements sont néanmoins nécessaires pour améliorer le suivi des mains et permettre la manipulation fine des instruments chirurgicaux à 7 degrés de liberté.

Dans cette optique, Zhou et al. [Zhou 2014, Zhou 2015a] ont présenté une nouvelle approche, chronologiquement postérieure aux travaux présentés dans ce chapitre. Ces recherches traitent du suivi des mains d'un opérateur afin de contrôler des instruments chirurgicaux robotisés. Ce type d'interaction sans contact permet une commande complète des 7 degrés de liberté de chaque instrument, tout en garantissant une hygiène et un environnement stérile lors de la manipulation. De plus, le coût moindre du ou des dispositifs employés (une ou deux caméras le cas échéant), en fait une approche intéressante pour la commande des robots chirurgicaux. Dans la suite de ce chapitre, nous présentons dans un premier temps une étude de faisabilité pour l'utilisation d'une interface manuelle sans contact associée à une plateforme robotisée destinée à l'entraînement aux gestes chirurgicaux. Une seconde étude démontre l'utilisation d'une méthodologie visant à différencier les niveaux d'expertise en se basant sur les trajectoires acquises au travers de cette interface.

4.2 Étude d'une nouvelle interface sans contact pour l'entraînement en chirurgie robotisée

Dans cette partie nous décrivons en premier lieu la plateforme expérimentale mise en place pour l'entraînement aux gestes chirurgicaux. Nous détaillons ensuite les développements apportés pour l'intégration d'une nouvelle interface manuelle sans contact afin de piloter le robot Raven-II. Nous présentons enfin une première étude de faisabilité dans laquelle nous comparons l'utilisation de cette interface sans contact avec une interface haptique de référence.

4.2.1 Plateforme expérimentale pour l'entraînement en chirurgie robotisée

La plateforme d'entraînement aux gestes chirurgicaux est composée d'un robot Raven-II et de son contrôleur, d'une console avec un ordinateur et un moniteur, ainsi que de deux interfaces homme-machine Sigma.7 (figure 4.9). Pour cette étude de faisabilité, nous comparons l'utilisation de cette interface haptique avec l'emploi d'une nouvelle interface manuelle sans contact qu'est la Leap Motion.

4.2.1.1 Le robot Raven-II

Le Raven-II est un robot de téléopération pour la chirurgie laparoscopique. Initialement développé dans le cadre d'un partenariat entre l'université de Washington et l'uni-

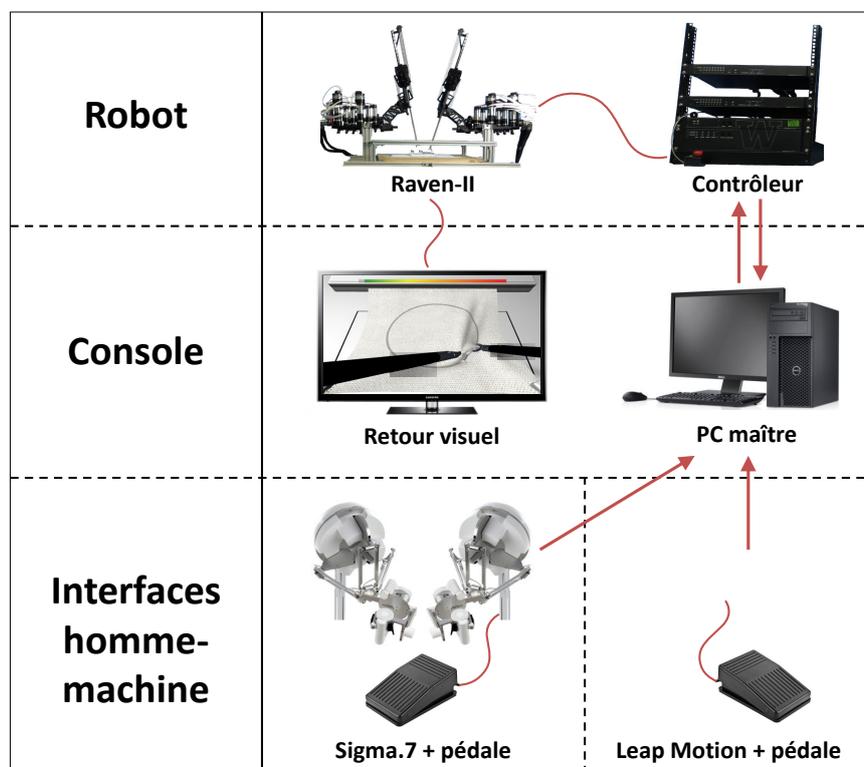


FIGURE 4.9: La plateforme robotisée pour l'entraînement aux gestes chirurgicaux.

versité de Santa Cruz, ce système est aujourd'hui produit par la société Applied Dexterity. L'objectif de ce robot est de fournir un système similaire au da Vinci, à moindre coût, autorisant la recherche et développement en robotique chirurgicale [Hannaford 2013]. Testé en conditions extrêmes de téléopération [Lum 2009], nous utilisons ici ce robot dans le cadre de l'entraînement aux gestes pour la chirurgie laparoscopique robotisée.

Le Raven-II (figure 4.10), se compose de deux bras articulés, tous deux conçus sur des architectures mécaniques à câbles actionnées par 7 moteurs. Chaque bras dispose de 7 degrés de liberté dans l'espace opérationnel permettant de reproduire, de la même manière que le robot da Vinci, tous les mouvements nécessaires à la réalisation de gestes chirurgicaux. Les principaux avantages de ce robot résident notamment dans sa conception très compacte, sa faible inertie facilitée par le déport des actionneurs grâce à une transmission par câbles, ainsi que son architecture logicielle ouverte permettant la modification du code source pour l'implémentation de commandes avancées.

Du point de vue logiciel, le robot Raven-II repose sur un système Linux 10.04 LTS temps-réel (application d'un patch préemptif) permettant d'effectuer la commande des moteurs à une fréquence de 1kHz. Dans le cadre de la téléopération, les informations cartésiennes des interfaces homme-machine sont utilisées pour réaliser la commande en position rela-

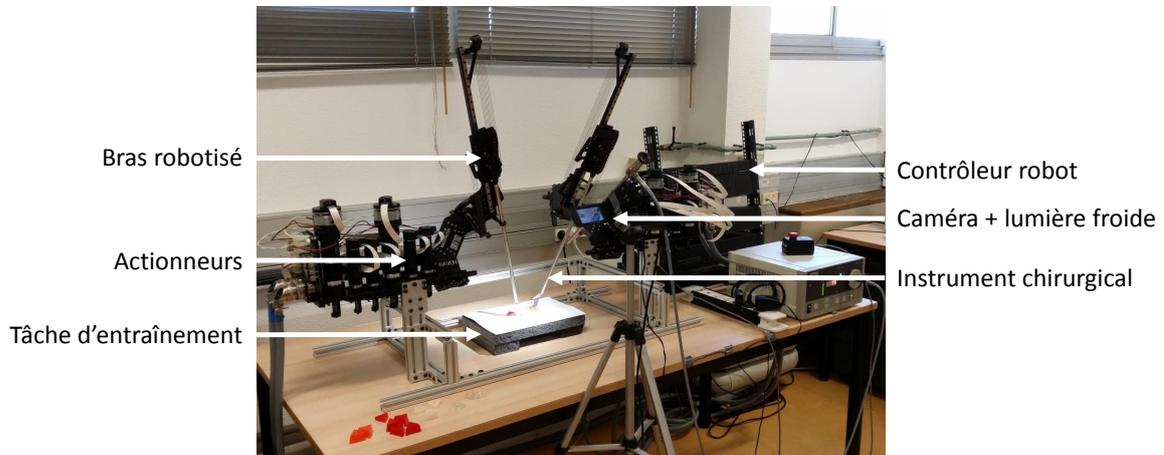


FIGURE 4.10: Le robot Raven-II utilisé pour l'entraînement aux gestes en chirurgie laparoscopique robotisée.

tive du robot.

En complément de son architecture logicielle ouverte, le Raven-II possède des instruments chirurgicaux interchangeable, qui grâce à un adaptateur autorise le montage des instruments da Vinci. Dans sa version originale, le système est livré avec deux instruments porte-aiguille que nous utiliserons dans la suite de ces travaux.

4.2.1.2 Les interfaces homme-machine

Pour commander les instruments chirurgicaux du Raven-II, nous disposons d'une console équipée de deux interfaces Sigma.7 de chez Force Dimension, de trois écrans et d'un repose bras (figure 4.11). Alors que l'écran central permet le retour visuel de l'image endoscopique (fourni ici par caméra 3D), les écrans périphériques autorisent l'affichage d'informations diverses. Cette console est également équipée d'une pédale pour activer le fonctionnement du robot, permettant également de débrayer les interfaces afin de les repositionner dans leur espace de travail. Cette fonction, appelée "débrayage", sera utilisée dans la suite de ces travaux. Dans le cadre de notre étude de faisabilité, nous avons également équipé cette console d'une Leap Motion afin de réaliser la commande manuelle sans contact du robot Raven-II.

L'interface haptique Sigma.7 dispose de 7 degrés de liberté générant avec 3 translations, 3 rotations et l'ouverture et fermeture de la pince. Idéal pour contrôler les instruments chirurgicaux du Raven-II, cette interface possède des résolutions cartésiennes élevées (0,0015mm en translation, 0,013deg en rotation et 0,006mm pour la préhension) permettant d'obtenir une excellente précision pour l'acquisition du mouvement, réalisée à 100Hz pour notre application. Les données constructeurs permettent également de représenter

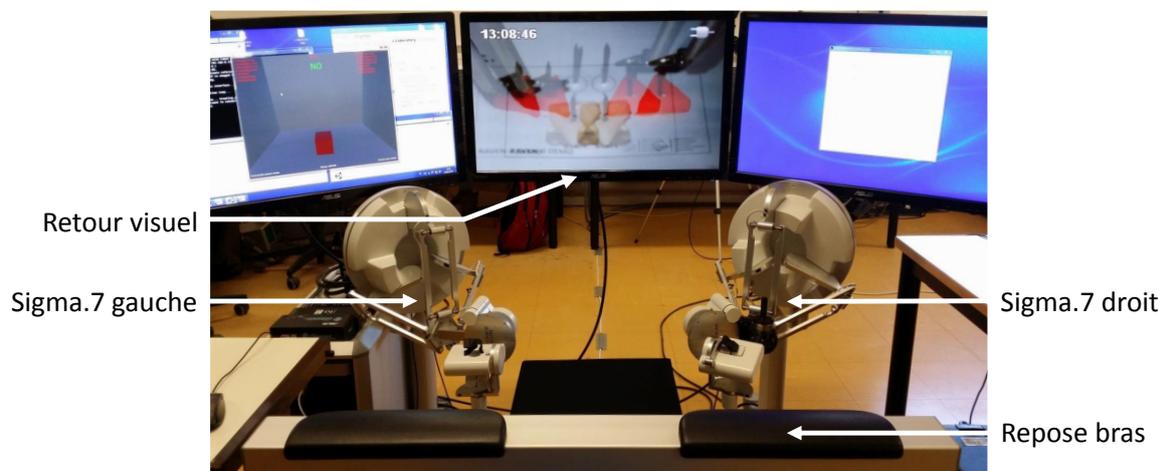


FIGURE 4.11: Console de commande avec deux Sigma.7 pour le contrôle des instruments chirurgicaux.

l'espace de travail simplifié d'une interface Sigma.7 (figure 4.12a). Il est à noter que pour notre application, les deux interfaces Sigma.7 sont agencées afin d'éviter toutes intersections entre leur espace de travail, supprimant tout risque de collision. De plus, dans la suite de ces travaux, nous n'utiliserons pas les capacités d'acquisition et de retour d'efforts autorisées par ces interfaces n'employant ainsi que les positions cartésiennes des effecteurs. La Leap Motion dispose quant à elle de deux caméras stéréoscopiques et trois LEDs infrarouges permettant la reconstruction de l'environnement en trois dimensions. À partir de ces informations et d'un algorithme interne de détection des mains, le système est capable de suivre les mouvements de l'utilisateur avec une précision inférieure au millimètre [Guna 2014], et une fréquence d'acquisition d'environ 110Hz. À partir des données constructeur, une version simplifiée de l'espace de travail de l'interface est présentée sur la figure 4.12b. Il est à noter que l'espace de travail effectif est inférieur à l'espace de travail indiqué par le constructeur du fait de la réduction du score de confiance dans la détection des mains aux extrémités. De plus, l'espace de travail est commun aux deux mains, ce qui peut générer des occlusions et des collisions durant la manipulation.

Les récents travaux de Kim et al. [Kim 2014] ont comparé les performances d'une Kinect et d'une Leap Motion afin de valider leur utilisation dans le domaine de la chirurgie téléopérée. D'après leurs analyses, aucun des dispositifs ne possède pas la précision requise pour être employé comme interface de commande. Toutefois, les conditions de manipulation présentées dans ces travaux sont différentes des nôtres. Pour notre étude, nous construisons un modèle précis à partir des informations de l'interface, nous permettant alors de contrôler précisément chaque degré de liberté de l'effecteur, incluant également l'ouver-

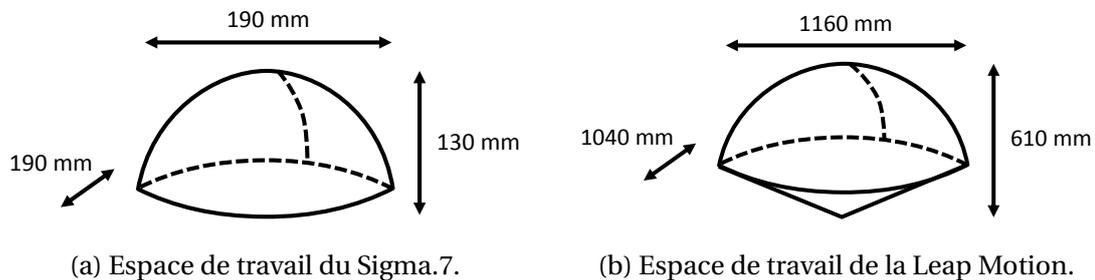


FIGURE 4.12: Représentation des espaces de travail simplifiés des interfaces homme-machine.

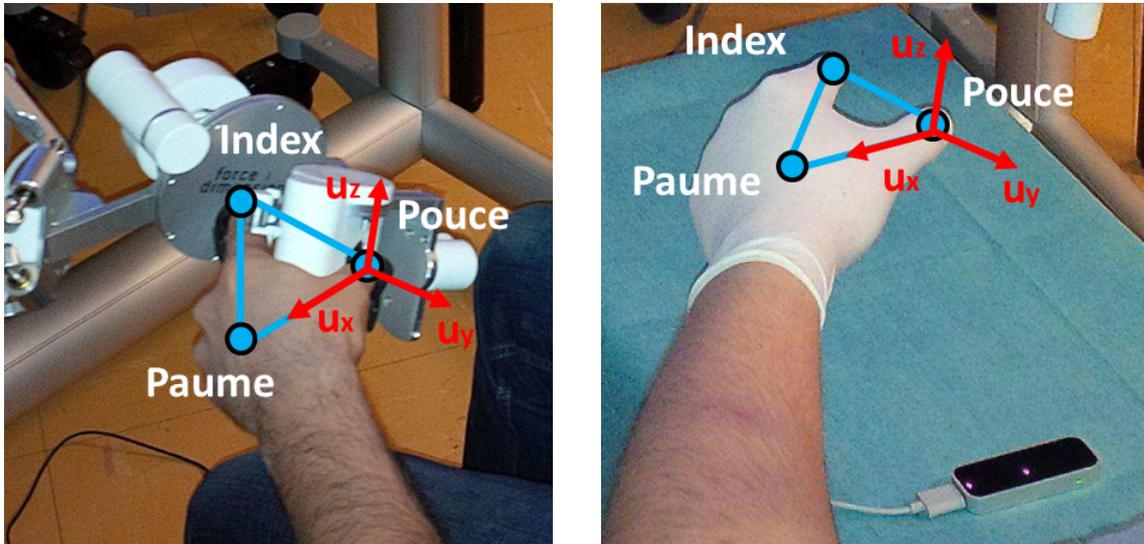
ture et la fermeture de la pince. De plus, nous ajoutons une étape de filtrage afin de réduire le bruit généré durant l'acquisition.

Nous utilisons les informations fournies par le logiciel embarqué de la Leap Motion (version 1.0.8+7665 pour cette étude) afin de reconstruire un référentiel cartésien, similaire à celui acquis avec l'interface mécanique Sigma.7 (figure 4.13a), pour la commande des instruments chirurgicaux. Grâce à la capture des informations du pouce, de l'index et de la paume des mains, nous créons pour chaque main un repère 3D comme présenté sur la figure 4.13b, autorisant 7 degrés de liberté dans l'espace opérationnel. Les 6 premiers étant les 3 translations et les 3 rotations, le septième autorise l'ouverture et la fermeture de la pince. Ce mouvement est ici défini par l'angle généré entre le pouce et l'index. Ainsi, dans cette configuration, les instruments chirurgicaux sont contrôlés par mimétisme des doigts de l'utilisateur, la pince étant représenté par le pouce et l'index.

Nous utilisons ensuite un filtre passe-bas basé sur une méthode autorégressive à moyenne mobile (ARMA), avec une atténuation de 25dB à 2Hz [Bó 2011] afin de filtrer les mouvements de la main de l'utilisateur, dont les fréquences fondamentales se situent entre 0 et 10Hz [Loram 2006]. Nécessaire lors de l'utilisation de l'interface sans contact, ce filtrage n'est néanmoins pas utilisé dans le cas des interfaces mécaniques car, l'impédance générée par le système est non nulle assurant un filtrage naturel du mouvement à basse fréquence, ce qui améliore les performances lors de la manipulation [Howard 2014].

4.2.1.3 La tâche d'entraînement chirurgical

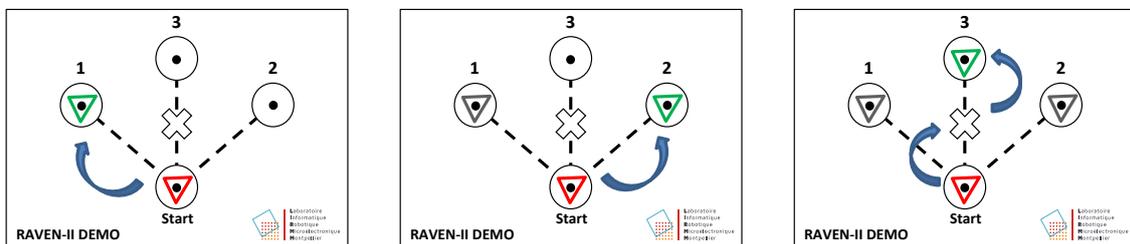
La tâche d'entraînement définie pour l'entraînement aux gestes chirurgicaux est la même que celle utilisée dans le chapitre 3. Nous allons néanmoins la redéfinir brièvement. Cette tâche s'inspire directement des protocoles d'évaluation utilisés par la SAGES [Derossis 1998]. Nous n'employons ici que l'exercice impliquant le transfert d'objets (des pions en silicone) à trois emplacements tel que présenté sur la figure 4.14. Les pions sont préalablement empilés au même endroit (point de départ en bas). L'opérateur doit ensuite réaliser la tâche en respectant les consignes suivantes :



(a) Interface Sigma.7.

(b) Interface Leap Motion.

FIGURE 4.13: Schéma d'acquisition des informations cartésiennes avec les interfaces pour la commande du robot.



(a) Transfert pion N°1.

(b) Transfert pion N°2.

(c) Transfert pion N°3.

FIGURE 4.14: Illustration de la tâche de transfert de pions.

1. Attraper le premier pion avec l'instrument gauche et l'insérer dans la cible 1 (figure 4.14a) ;
2. Attraper le second pion avec l'instrument droit et l'insérer dans la cible 2 (figure 4.14b) ;
3. Attraper le dernier pion avec l'instrument gauche ou droit (au choix), se déplacer au centre du plateau, transférer le pion au second instrument et insérer ce dernier dans la cible 3 (figure 4.14c).

4.2.2 Évaluation pour l'entraînement en chirurgie robotisée

Afin d'évaluer l'utilisation de la Leap Motion, nous entreprenons dans un premier temps une étude comparative entre cette interface et l'interface Sigma.7. Cette étude de faisabilité vise à quantifier les performances d'utilisation des deux interfaces au travers de critères quantitatifs et qualitatifs. Nous présentons ci-dessous le protocole des expériences, les données acquises, les métriques d'évaluation ainsi que les résultats.

4.2.2.1 Protocole expérimental

Pour ces expérimentations nous avons demandé à 10 chercheurs du LIRMM, possédant des compétences variées, de réaliser la tâche d'entraînement décrite précédemment en utilisant chacune des interfaces. Nous leur avons, dans un premier temps, présenté le système et fourni les instructions nécessaires à la réalisation de la tâche. Puis, l'ordre de passage des interfaces a été défini de manière aléatoire, afin d'éviter tout biais d'apprentissage dans l'analyse. À la suite de cela, chaque sujet a bénéficié de 5 minutes d'entraînement libre afin de se familiariser avec le robot, la tâche ainsi que l'interface de commande. Ce temps d'entraînement lui a permis de progresser sur sa courbe d'apprentissage et d'améliorer ses performances.

4.2.2.2 Données et métriques d'évaluation

À la suite de cet entraînement libre, nous avons acquis trois trajectoires par sujet afin d'évaluer chacune des interfaces au travers de métriques quantitatives et qualitatives. Les données, enregistrées automatiquement et pour chaque tâche, comportent la durée d'exécution, la réussite de la tâche et le nombre de débrayages. Nous définissons la durée d'exécution comme le temps compris entre le premier appui sur la pédale et le dernier relâchement pour l'arrêt du robot. La réussite de la tâche correspond au score de réalisation. Ainsi, une tâche réalisée sans faute permet d'obtenir 100%, là où une tâche réalisée avec 4 fautes ou plus entraîne un score de 0% (nous considérons que le pion ne peut tomber qu'une fois d'un instrument, la tâche obligeant à réaliser 4 préhensions). Le nombre de débrayages est enfin défini par le nombre de relâchement enchaîné par un appui sur la pédale, permettant l'arrêt puis la reprise du robot. Ces informations nous permettent de définir les critères quantitatifs associés pour évaluer les interfaces. À la fin de l'exercice, chaque sujet s'est vu évaluer également les interfaces au travers d'un questionnaire, composé de cinq critères qualitatifs que sont :

- la réactivité : capacité de réagir rapidement aux mouvements ;
- la précision : capacité d'exécution de gestes minutieux ;
- l'intuitivité : capacité de manipuler facilement, sans effort cognitif ;
- la robustesse : capacité de reproduction des gestes ;
- le confort : capacité de manipuler facilement, sans effort physique.

À partir de ces métriques, nous comparons les performances de l'interface sans contact Leap Motion avec celles du Sigma.7 pour l'entraînement en chirurgie robotisée.

4.2.2.3 Résultats

Les résultats de l'étude sont présentés en deux parties. La première aborde l'analyse des critères quantitatifs et la seconde présente les résultats obtenus avec les critères qualitatifs.

Évaluation quantitative.

Nous posons comme hypothèses de départ H_0 que pour chaque métrique considérée, les performances des interfaces sont identiques. Afin de valider ou réfuter cette hypothèse, nous utilisons le test non-paramétrique des rangs sommés de Wilcoxon. Dans le cas où l'hypothèse est réfutée, impliquant une valeur-p inférieure à 5%, nous qualifions le degré de significativité avec trois seuils de valeur-p que sont 0,001, 0,01 et 0,05 notre valeur limite.

La figure 4.15a présente les durées d'exécution réalisées avec chacune des interfaces sous forme de moyennes et écarts-types. Nous constatons dans ce cas que l'hypothèse H_0 est rejetée car la valeur-p est de 0,0172. Ainsi, nous en déduisons qu'il existe deux populations distinctes, et que le temps d'exécution moyen avec chacune des interfaces est significativement différent. Dans le cas du Sigma.7, la moyenne des temps d'exécution est de 101,66 secondes, ce qui est environ 14 secondes plus rapide que dans le cas de la Leap Motion, réalisant une moyenne de 115,76 secondes.

L'analyse du succès est représentée par la figure 4.15b. De la même manière que précédemment, nous utilisons le test de Wilcoxon pour évaluer la distinction des populations. Dans ce cas encore, les populations sont largement distinctes puisque la valeur-p est de 0,0089. Ici, c'est la Leap Motion qui permet de réaliser le moins de fautes durant la réalisation de la tâche, avec un score de 94%, là où le Sigma.7 n'atteint que 82%.

La figure 4.15c montre le nombre de débrayages réalisés en utilisant chacune des interfaces. Dans le cas Sigma.7, les sujets ont réalisé en moyenne 5,75 débrayages par entraînement, là où la Leap Motion a entraîné plus de 15,38 débrayages. De ce fait, en utilisant le même test non-paramétrique, nous constatons que la valeur-p est de 0,0017 indiquant que cet écart de 9,63 est significativement favorable à l'utilisation du Sigma.7.

Évaluation qualitative.

Pour cette évaluation, nous avons demandé à chaque sujet d'évaluer les performances de l'interface suivant cinq critères. Ces notes s'étalent sur une plage de 0 à 10, où 0 correspond à la note la plus faible et 10 la meilleure. De la même manière que pour l'évaluation quantitative, nous posons l'hypothèse que les performances des interfaces sont identiques (hypothèse H_0). Par le test de Wilcoxon, nous cherchons alors à savoir si cette hypothèse

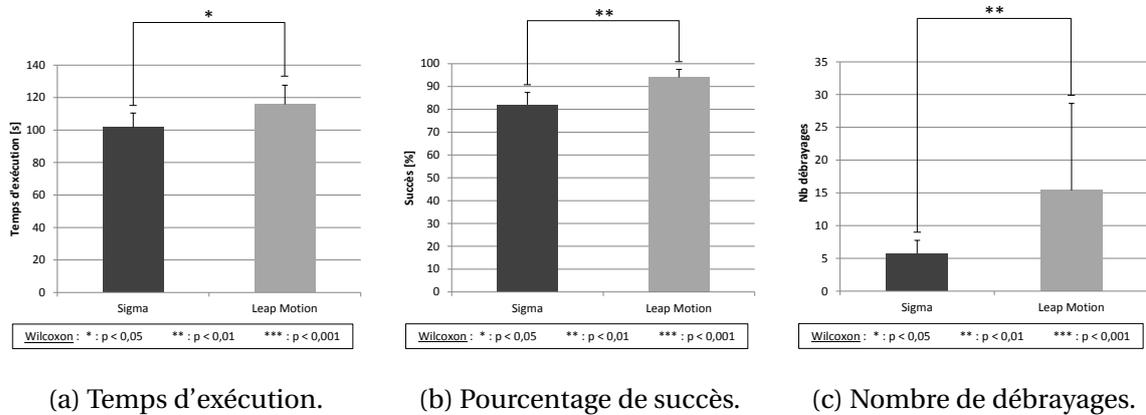


FIGURE 4.15: Comparaison des interfaces en utilisant les métriques quantitatives.

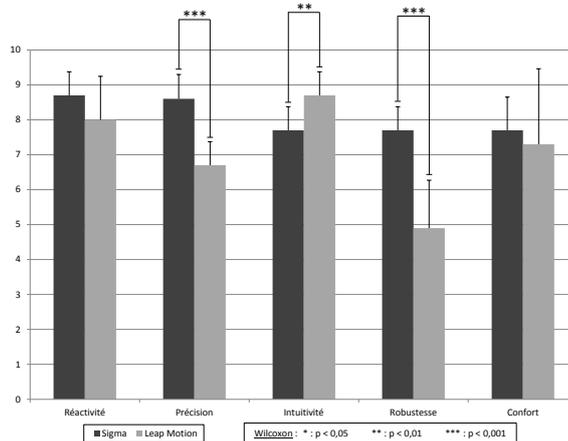


FIGURE 4.16: Comparaison des interfaces en utilisant les métriques qualitatives.

est validée ou rejetée. En analysant la figure 4.16, nous constatons que la réactivité des interfaces n'a pas été jugée significativement différente (valeur-p au-dessus de 5%). Concernant la métrique de précision, la valeur-p est de 0,0003 indiquant que les interfaces fournissent une précision significativement différente pour la réalisation des gestes. De ce fait, les interfaces Sigma.7 sont mieux adaptées à l'exécution précise des gestes. Dans le cas de l'intuitivité, la valeur-p de 0,0077 indique un écart significatif, favorisant l'utilisation de la Leap Motion. Pour la robustesse, la valeur-p est de 0,0004 indiquant alors que l'écart significatif joue largement en défaveur de la Leap Motion. Nous aborderons ce résultat dans l'analyse présentée dans le paragraphe suivant. Enfin, en analysant la métrique de confort, nous remarquons qu'aucune généralisation ne peut être faite à partir de ces expériences, avec une valeur-p supérieure à 5%.

4.2.2.4 Analyse et discussion

Dans cette partie, nous discutons des résultats et analysons les facteurs influant sur les performances. Notre analyse s'étend également l'étude de facteurs techniques et environnants non discutés jusqu'à présent.

Évaluation quantitative.

L'analyse des métriques quantitatives acquises durant les expériences montre que le Sigma.7 permet d'être plus rapide et de réaliser moins de débrayages que la Leap Motion. Cet effet est en pratique largement discutable. Tout d'abord, il est à noter que nous avons considéré comme débrayage les appuis et relâchements de la pédale ainsi que les occlusions avec la Leap Motion, nous obligeant alors à arrêter le robot afin d'assurer la sécurité du matériel. Ainsi, le nombre de débrayages calculé avec la Leap Motion est la somme des débrayages intentionnels et des débrayages générés par la manipulation avec une interface sans contact. Ce nombre de débrayages justifie également le temps d'exécution plus long avec la Leap Motion. En pratique, nous devons faire évoluer notre analyse en distinguant les débrayages intentionnels des occlusions et problèmes de suivi des mains lors du dépassement de l'espace de travail. Toutefois, bien que le nombre de débrayages réalisé avec la Leap Motion soit environ trois fois supérieur à celui réalisé avec le Sigma.7, la différence en termes de durée est seulement de 14%. Cela indique alors que même si le système est plus régulièrement arrêté, l'interface permet une exécution des gestes proche de ce que nous pouvons accomplir avec le Sigma.7. En se plaçant dans un cadre interventionnel, une différence de 14% en termes de durée d'exécution représente un allongement d'une durée de 1 heure et 30 minutes sur une chirurgie initialement réalisée en 10 heures. Bien que cela ne soit pas acceptable dans le cadre d'une intervention, 14% d'allongement sur une tâche d'entraînement durant 1 minute 30 secondes représente moins de 13 secondes d'écart ce qui est plus largement acceptable dans ce contexte. Concernant le score de réussite en faveur de la Leap Motion, nous supposons que l'attention requise pour la manipulation favorise le bon déroulement de la tâche, là où le Sigma.7 autorise plus de rapidité mais en contrepartie provoque plus d'erreurs au cours de l'exercice.

Bien que l'analyse des critères quantitatifs favorise plutôt l'utilisation de l'interface haptique Sigma.7, nous souhaitons toutefois aborder un facteur important, non discuté jusqu'à présent. En se concentrant sur le cadre d'entraînement aux gestes en chirurgie laparoscopique robotisée, la Leap Motion fournit un avantage conséquent qui est son prix, soit moins de 70 euros sur le marché grand public pour contrôler les deux instruments chirurgicaux. Une console composée de deux interfaces Sigma.7 pour une manipulation bimanuelle nécessite quant à elle un budget d'environ 100000 euros. Ce prix important, majoritairement justifié par la conception, l'utilisation de pièces mécaniques de précision et la technologie de retour d'effort embarquée, freine cependant le développement des systèmes d'entraînement pour la chirurgie robotisée, les rendant moins accessibles. Bien que des alternatives soient plus accessibles, telle que le Phantom Omni (environ 1500 eu-

ros l'unité), la Leap Motion reste plus intéressante et autorise une manipulabilité proche de ce qui est faisable avec une console Sigma.7 ou da Vinci. Le rapport prix/performance devient donc très favorable pour l'utilisation durant l'entraînement robotisé.

Évaluation qualitative.

Les résultats du questionnaire d'évaluation mettent également en avant des points que nous souhaitons analyser et discuter. Tout d'abord, l'appréciation de la précision de l'interface fait ressortir que le Sigma.7 semble plus adapté à la réalisation de tâches minutieuses. Cette remarque est justifiée par le fait que l'interface mécanique, de par l'inertie et les frottements des pièces mécaniques mises en mouvement, génère une impédance non nulle et facilite la manipulation [Howard 2014]. De plus, la résolution en translation, rotation et préhension, est largement supérieure à la résolution du suivi d'objet réalisé par la Leap Motion. Néanmoins, bien que les caractéristiques techniques soient largement en faveur de l'interface haptique, la Leap Motion réalise un score encourageant, qu'il serait intéressant de recalculer après amélioration du logiciel de suivi et de détection embarqué. Concernant la robustesse, la stabilité de la détection et du suivi des mains (dans le cas statique), bien que compensée par le filtrage mis en place, reste faible et certaines perturbations sont encore largement visibles affectant directement les résultats obtenus. De plus, la gestion des occlusions n'est pas prise en compte avec la version du micrologiciel utilisé durant cette étude. Il est donc nécessaire d'améliorer ces aspects dans le cadre de futures expériences avec la Leap Motion.

Concernant l'intuitivité, nous remarquons que ce critère penche significativement en faveur de la Leap Motion. En se basant sur les discussions avec les sujets, il en est ressorti que l'implémentation réalisée avec la Leap Motion simplifie la manipulation, rendant intuitifs les gestes à réaliser. Ici, le pouce et l'index de l'opérateur représentent directement la pince de l'instrument, facilitant ainsi le mimétisme des mains et réduisant la réflexion. Cette modélisation permet d'obtenir une manipulation très similaire à ce qui est réalisable avec un dV-Trainer ou une console da Vinci.

Enfin, la réactivité et le confort d'utilisation vont en faveur de l'interface haptique Sigma.7. Toutefois, l'étude statistique menée montre néanmoins qu'une discussion sur ces points n'est pas justifiée compte tenu des écarts non significatifs observés.

Pour conclure, cette étude met en avant les capacités de l'interface haptique par rapport à l'interface sans contact. Mis à part l'intuitivité de la commande, les interfaces Sigma.7 surpassent la Leap Motion pour la commande du robot Raven-II. Néanmoins, certains faibles écarts de performances nous incitent à penser qu'avec des améliorations, ce dispositif sans contact pourrait être utilisé au sein de systèmes d'entraînement destinés à la chirurgie laparoscopique, aussi bien sur simulateur qu'avec robot. Cette configuration permettrait alors d'améliorer l'accès à la formation en chirurgie robotisée.

Bien que cette étude de faisabilité montre certaines faiblesses de la Leap Motion, nous dé-

sirons néanmoins poursuivre le développement afin de pallier les problèmes mis en avant dans cette première étude. Pour cela, nous proposons une seconde étude se focalisant plus spécifiquement sur l'analyse des trajectoires des instruments chirurgicaux réalisées avec chacune des interfaces. L'objectif est de mettre en avant de possibles différences dans l'exécution de la tâche pour compléter notre étude de faisabilité.

4.3 Vers une analyse des trajectoires pour l'évaluation du niveau d'expertise

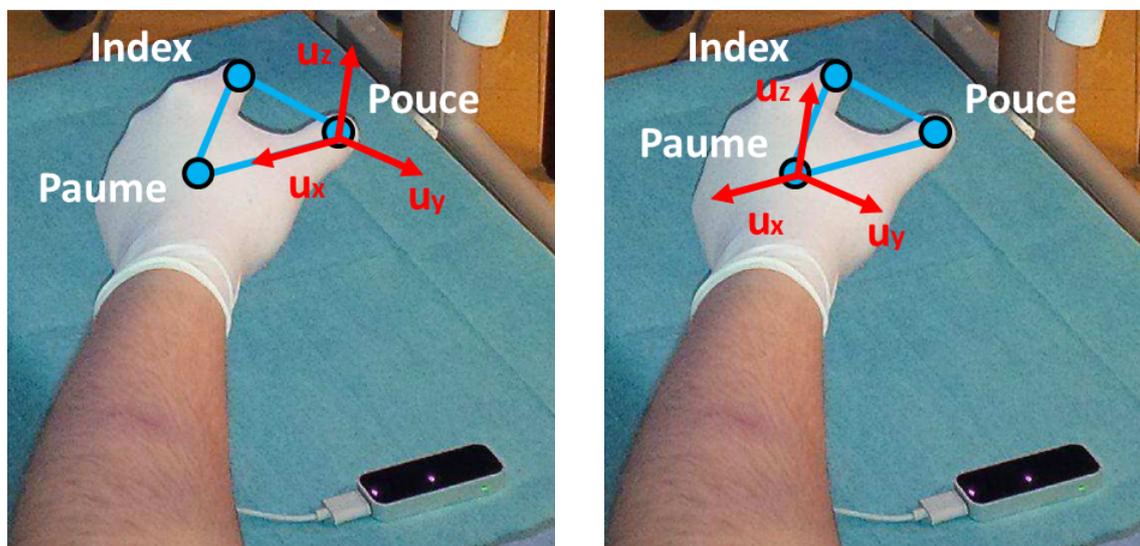
Dans cette partie, nous présentons une seconde étude avec l'interface manuelle sans contact Leap Motion. En se basant sur une approche non-supervisée, nous souhaitons analyser les trajectoires des instruments chirurgicaux réalisées avec chacune des interfaces afin de quantifier les variations dans les mouvements, et d'étudier s'il est possible de classer automatiquement les exercices suivant leur réalisation. Nous introduisons d'abord les changements opérés avec l'interface sans contact pour améliorer la commande du robot. Puis, nous présentons la méthodologie mise en place pour analyser les trajectoires afin de les classer en différents groupes. Enfin, nous décrivons l'étude visant à quantifier les variations d'exécution de la tâche, qui pourrait faire un lien direct avec le niveau d'expertise des opérateurs.

4.3.1 Amélioration de la commande sans contact

Nous réalisons quelques modifications dans le processus d'acquisition des informations avec la Leap Motion, afin d'améliorer la commande du robot de téléopération Raven-II. Pour cela, nous décrivons d'abord la modification du modèle pour la commande, puis nous introduisons une autre méthode de filtrage des mouvements afin d'améliorer les performances lors de la télémanipulation.

4.3.1.1 Modification du modèle

Lors de l'étude précédente, nous avons défini un modèle à 7 degrés de liberté permettant de piloter les mouvements de l'instrument chirurgical à partir de la Leap Motion. Cependant, au cours des expériences, nous avons remarqué des mouvements parasites, venant gêner la réalisation de certaines gestes. Le principal inconvénient de l'approche proposée réside dans le choix du pouce comme référence du système de coordonnées. Ce dernier, lors d'une phase de préhension, génère des translations involontaires, complexifiant la commande des instruments chirurgicaux. La figure 4.17 présente l'ancien et le nouveau modèle dans lequel le centre du repère a été ramené au plus proche du poignet, au niveau de la paume de la main. Cette modification du centre instantané de rotation permet de réduire au maximum les mouvements parasites involontaires et potentiellement aussi



(a) Ancien modèle basé sur le pouce.

(b) Nouveau modèle basé sur la main.

FIGURE 4.17: Schéma d'acquisition des informations cartésiennes avec l'interface sans contact Leap Motion.

d'améliorer l'intuitivité et la précision de cette interface. Afin d'évaluer ce changement, il serait néanmoins nécessaire d'entreprendre une seconde étude comme celle présentée précédemment.

4.3.1.2 Filtrage des tremblements

La fréquence d'acquisition de 110Hz utilisée lors des premiers essais étant instable, nous avons réduit à 60Hz pour réaliser un compromis entre vitesse d'acquisition (au moins deux fois plus importante que la bande passante des mouvements) et stabilité, respectant le théorème de Nyquist-Shannon.

Nous modifions également le filtrage des tremblements. Nous faisons évoluer l'emploi du filtre ARMA vers un filtre de Kalman, tel que présenté dans la méthodologie du chapitre 3. L'avantage de cette approche récursive réside dans le filtrage en temps-réel, prenant également en compte un double aspect prédiction et correction, très utile lors du filtrage des mouvements. La matrice de covariance pour les bruits de mesures a été estimée à partir de données précédemment acquises [Welch 2001]. Dans le cas de la matrice de covariances pour les bruits de procédé, nous utilisons des valeurs empiriques nous permettant le suivi de la trajectoire d'origine, tout en réduisant au maximum le bruit du procédé.

4.3.2 Méthodologie d'analyse des trajectoires

Cette partie décrit la méthodologie employée pour analyser les trajectoires du robot à partir des deux interfaces homme-machine. Une fois les trajectoires préalablement acquises, nous extrayons les positions $\{X_i, Y_i, Z_i \mid i \in 1, 2\}$ des effecteurs du robot pour l'analyse et la classification.

Notre méthodologie repose sur un processus en trois phases. La première phase consiste à prétraiter les données acquises en entrée, comme réalisé dans le chapitre 3. La seconde phase consiste à construire une matrice de distance globale entre chacune des trajectoires acquises avec les différentes interfaces. À partir de cette matrice, la troisième phase consiste à regrouper de manière non-supervisée les trajectoires similaires (dont les distances sont proches) afin de créer des classes, sans aucune connaissance a priori sur le système ou la tâche d'entraînement.

4.3.2.1 Prétraitement des données

À partir des informations de position des instruments chirurgicaux, nous utilisons la même méthode de filtrage et de normalisation que celle présentée dans le chapitre 3 : un filtre passe-bas avec une fréquence de coupure de 1,5Hz, un gain en bande passante unitaire et une forte atténuation après 10Hz [Yang 2013].

Nous normalisons ensuite les signaux afin de réduire la variabilité intra- et interpersonnelle. De ce fait nous utilisons la même approche que présentée dans le chapitre 3, à savoir une normalisation basée sur la moyenne et l'écart-type pour un centrage réduit des données.

4.3.2.2 Approche non-supervisée pour la classification

Une fois le filtrage et la normalisation appliqués, nous souhaitons regrouper les trajectoires par groupe de similitude. Pour cela, nous employons l'approche en deux étapes décrite dans les travaux de Forestier et al. [Forestier 2012, Forestier 2013], que nous appliquons aux trajectoires des instruments chirurgicaux. Cette approche consiste à construire une matrice de distances entre les différentes trajectoires en s'appuyant sur une métrique de similarité entre les signaux. La seconde étape consiste à classer les éléments de cette matrice en s'appuyant sur les distances de similarité.

Matrice de similarité.

Dans le chapitre 3, nous avons étudié trois métriques de similarité que sont Hausdorff, Fréchet et Déformation Temporelle Dynamique. Suite aux résultats, nous en avons conclu que DTD est la métrique la plus adaptée pour la mesure de similarité spatio-temporelle entre séquences. Nous utilisons donc uniquement la métrique de DTD pour l'analyse des similarités locales et globales dans les signaux de positions, reposant sur la distance Euclidienne pour la mesure entre points.

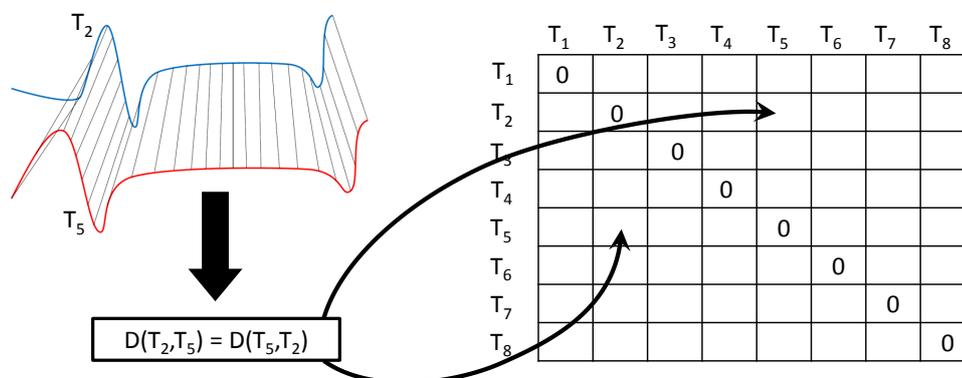


FIGURE 4.18: Illustration pour la construction de la matrice de similarité.

Nous construisons alors la matrice de similarité comme illustré par la figure 4.18. Dans un premier temps, nous calculons les distances paires à paires des trajectoires en utilisant la métrique DTD. Dans notre exemple, nous étudions les trajectoires T₂ et T₅. Puis nous remplissons les champs de la matrice de similarité correspondants M(2,5) et M(5,2), qui donnent la distance de similarité spatio-temporelle de ces deux trajectoires. A noter ici que la matrice est symétrique ce qui permet de diviser approximativement par deux le temps de construction. Le résultat final est une matrice des distances spatio-temporelles entre chacune des trajectoires.

Comme présenté dans l'analyse et discussion du chapitre 3, nous souhaitons évaluer l'impact du nombre de dimensions simultanément utilisées dans le calcul de la similarité. Nous envisageons donc deux approches. La première repose sur la comparaison indépendante des signaux de position. Dans ce cas, les signaux X, Y et Z sont comparés séparément entre les trajectoires. La distance de similarité est alors définie comme la somme des similarités de chaque signal. Nous considérons cette approche comme unidimensionnelle dans le sens où chaque signal est comparé avec son homologue d'une autre trajectoire. Le score global prend alors en compte l'ensemble des signaux. La seconde approche repose sur une analyse multidimensionnelle. Dans ce cas, nous calculons directement la distance euclidienne entre l'ensemble des signaux X, Y, Z d'une trajectoire avec ceux d'une autre. La similarité est alors estimée sans calcul intermédiaire. Dans la suite de ce chapitre, nous utilisons les abréviations 1D-DTD pour représenter l'approche unidimensionnelle et MD-DTD pour l'approche multidimensionnelle.

Classification non-supervisée.

Il existe plusieurs algorithmes permettant de réaliser du regroupement non-supervisée. L'un des plus employés est les K-Moyennes. L'objectif est de définir un nombre de groupes, puis de faire évoluer l'appartenance des points aux groupes jusqu'à convergence. L'algo-

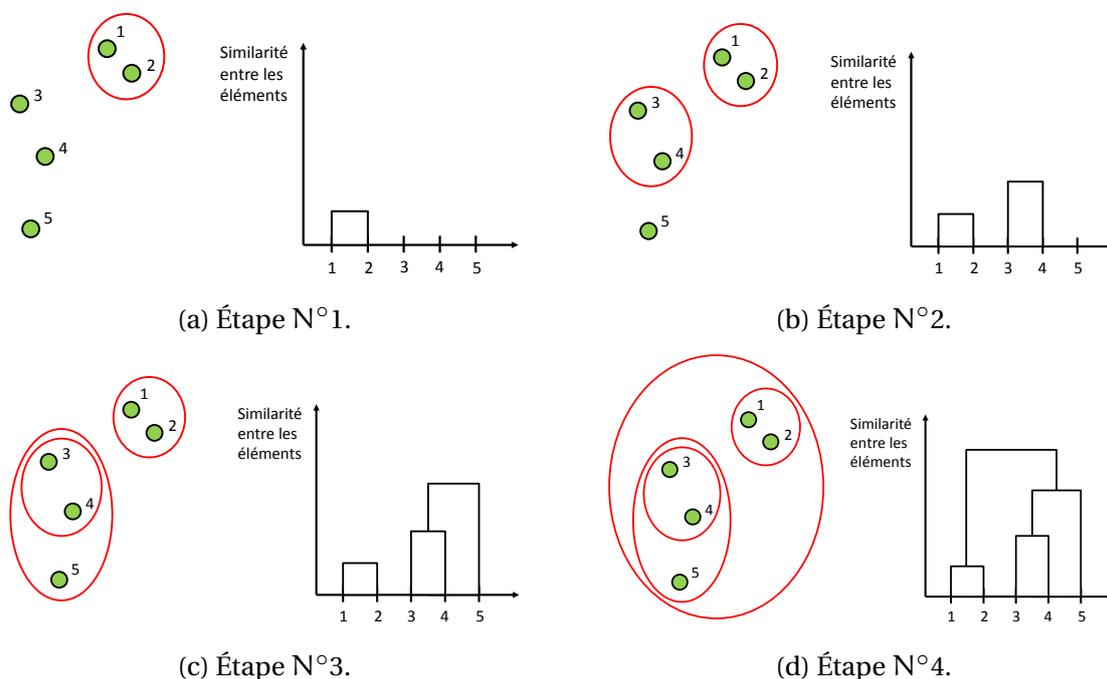


FIGURE 4.19: Schéma de fonctionnement de l'algorithme de classification ascendante hiérarchique.

l'algorithme cherche alors à maximiser la variance interclasse tout en minimisant la variance intraclasse. Bien que cet algorithme soit très simple d'emploi, une contrainte forte reste l'étape d'initialisation. D'un côté, l'algorithme peut choisir des points spécifiques dans les données pour créer le partitionnement des groupes. Dans ce cas, il devient complètement déterministe. D'un autre côté, une variante consiste à initialiser les centres des groupes aléatoirement, puis à faire évoluer ces centres afin de converger vers un partitionnement. Le souci de cette approche réside dans la sélection aléatoire qui, à chaque initialisation, pourra fournir un résultat différent sur le même jeu de données. Nous avons donc orienté notre étude vers l'utilisation d'un algorithme de regroupement plus stable, telle que la classification ascendante hiérarchique.

Le principe de la classification hiérarchique est illustré par la figure 4.19. Dans notre exemple, nous souhaitons réaliser le regroupement de 5 points. Utilisant comme entrée la matrice de similarité, nous cherchons les points les plus proches, puis les rassemblons dans un groupe. Cette étape, appelée agrégation, peut reposer sur plusieurs critères de distance. Dans notre cas, nous utilisons le critère de Ward qui cherche à maximiser l'inertie interclasse, en se basant sur la distance euclidienne. Nous itérons cette approche jusqu'à regrouper l'ensemble des points dans un unique groupe. Dans le cas présenté ici, nous employons une approche agglomérative, également appelée ascendante, où les échantillons les plus proches sont fusionnés au fur et à mesure. La classification est alors re-

présentée par le dendrogramme faisant apparaître des groupes, dont des distances spatio-temporelles sont quantifiées ici en ordonnée.

Grâce à cet algorithme, nous tentons de distinguer les trajectoires similaires permettant de créer des groupes par similarité d'exécution.

4.3.3 Vérification de l'interface et du niveau d'expertise

L'étude des compétences techniques chirurgicales est un domaine en pleine expansion. Une récente analyse de la littérature par Reiley et al. [Reiley 2011] montre que la majorité des travaux cherchent à modéliser les compétences du chirurgien aux travers de modèles mathématiques souvent complexes. Dans cette étude, nous ne faisons aucune hypothèse et analysons les résultats obtenus pour tirer des conclusions quant à l'utilisation d'une telle méthode pour l'évaluation du niveau d'expertise.

4.3.3.1 Protocole expérimental

Afin de tester les modifications apportées avec l'interface sans contact Leap Motion, nous réalisons une nouvelle étude comparative des interfaces. Dans cet objectif, nous demandons à trois experts de réaliser la tâche d'entraînement pour le transfert de pion. Ces experts sont composés de deux chirurgiens otorhinolaryngologistes utilisant le da Vinci (nommés "C3" et "C8" dans la suite de l'étude), et un expert en système de téléopération (nommé "C1"). Nous avons ensuite, de la même manière que dans l'étude précédente, demandé aux experts de réaliser la tâche d'entraînement de transfert d'objets cinq fois, en utilisant les Sigma.7 et la Leap Motion. Cette approche permet également de tester la reproductibilité de la méthode. Notre objectif est d'analyser les variations des trajectoires afin d'en déduire, le cas échéant, des différences aussi bien en termes de manipulation que du niveau de compétence.

4.3.3.2 Données

Au cours de ces sessions, nous avons acquis les positions cartésiennes $\{X, Y, Z\}$ des deux instruments chirurgicaux provenant des mouvements du robot Raven-II. À partir des données, nous construisons deux matrices de similarité, chacune utilisant une approche différente se basant sur la métrique de Déformation Temporelle Dynamique : une approche unidimensionnelle (1D-DTD) et une approche multidimensionnelle (MD-DTD).

4.3.3.3 Résultats

Nous présentons les résultats de la classification non-supervisée des trajectoires en utilisant les deux approches pour la métrique de similarité. Dans un premier temps, afin de qualifier ce que nous appelons similarité spatio-temporelle dans les trajectoires, nous présentons trois trajectoires distinctes. La figure 4.20a permet d'appréhender ce que nous

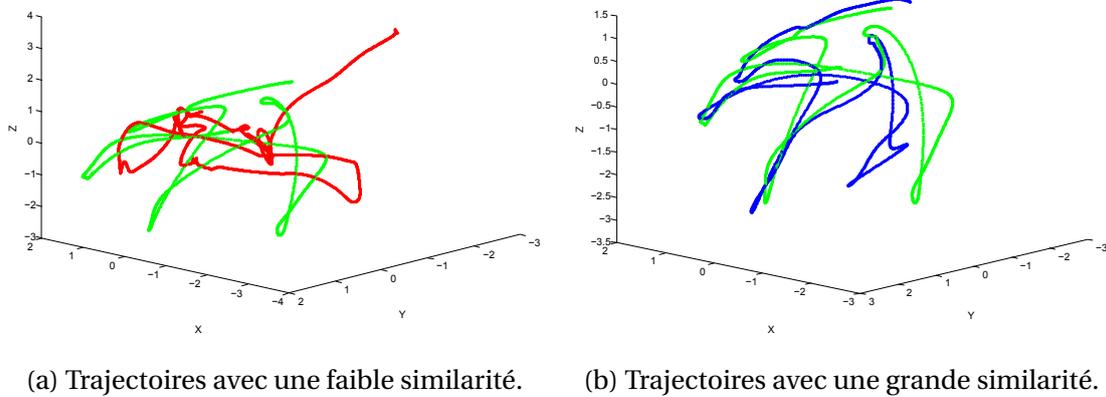


FIGURE 4.20: Analyse de la similarité entre deux trajectoires d'un même instrument, provenant de deux opérateurs différents.

considérons comme des trajectoires avec une similarité faible, où la trajectoire rouge représente l'opérateur C3 avec l'interface Leap Motion, et la trajectoire en vert représente l'opérateur C8 avec le Sigma.7. Inversement, la figure 4.20b permet d'appréhender une similarité spatio-temporelle forte, où la même trajectoire de l'opérateur C8 avec le Sigma.7 en vert est comparable avec une seconde trajectoire du même opérateur utilisant la même interface, en bleu.

En premier lieu, nous utilisons la métrique 1D-DTD pour classer les trajectoires des trois opérateurs. Sur la figure 4.21, l'axe des abscisses représente la distance spatio-temporelle. En regardant les groupes formés par la classification ascendante hiérarchique, nous constatons tout d'abord que chaque groupe se compose d'un opérateur, ce qui montre que chaque opérateur possède une manière d'exécuter la tâche qui lui est propre et qui est largement distinguable des autres opérateurs. Nous constatons également que chaque groupe est défini par une interface. Cela indique également que chaque interface possède sa propre "signature", et permet une exécution qui lui est propre. De manière globale, chaque groupe est donc caractérisé par un opérateur et une interface distincte. Plus précisément, nous remarquons que les opérateurs C1 et C8 ont réalisé des trajectoires distinctes avec le Sigma.7, mais qui ont des ressemblances qui permettent de créer une catégorie les regroupant. De la même manière, la seconde catégorie regroupe les opérateurs C1 et C8 qui ont utilisé la Leap Motion, ainsi que l'opérateur C3 utilisant le Sigma.7. Enfin, la dernière catégorie comporte uniquement les trajectoires de l'opérateur C3 avec la Leap Motion.

De manière similaire, la métrique MD-DTD construit un dendrogramme avec des groupes distincts (figure 4.22), et proches de ceux générés avec la métrique 1D-DTD, permettant de conclure que l'approche utilisée pour calculer la métrique n'influe que très peu sur la classification en sortie. Toutefois, nous préférons employer l'approche MD-DTD pour le calcul de la matrice de similarité, moins gourmande et plus efficace en temps de calcul (là

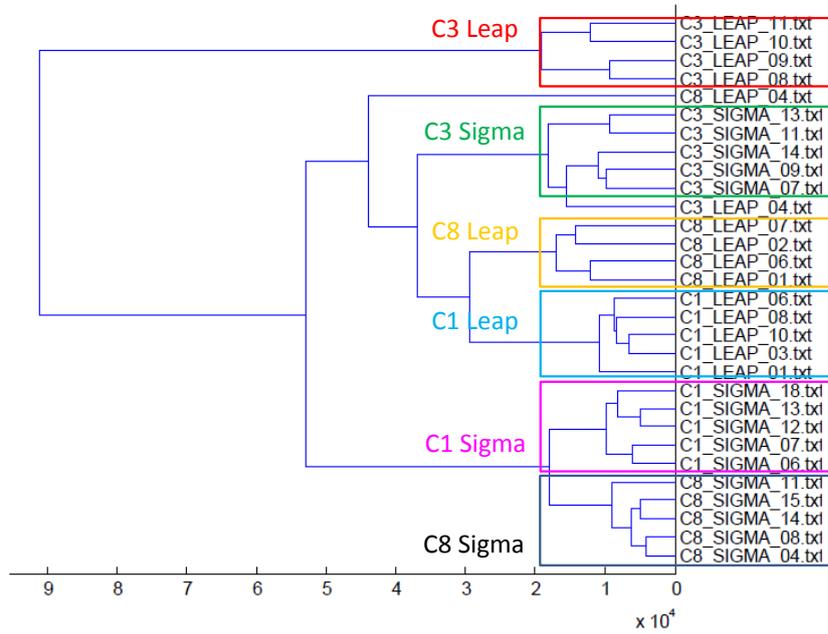


FIGURE 4.21: Regroupement des trajectoires basé sur la métrique 1D-DTD.

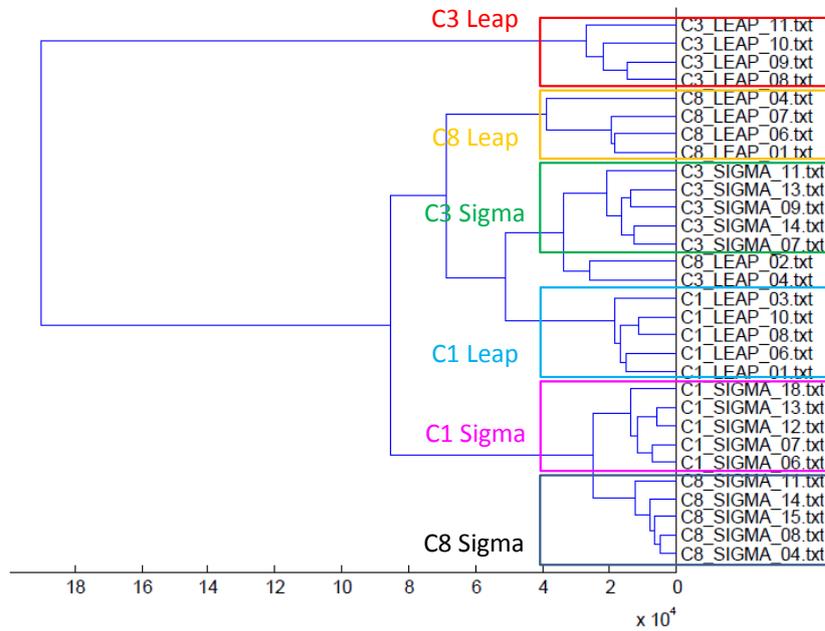


FIGURE 4.22: Regroupement des trajectoires basé sur la métrique MD-DTD.

où nous calculons une mesure de similarité entre deux trajectoire, nous en calculons trois avec 1D-DTD). De plus, nous remarquons que les groupes sont plus homogènes, plaçant les deux trajectoires "aberrantes" dans un groupe séparé.

4.3.3.4 Analyse et discussion

D'après ces résultats, nous pouvons extraire quatre informations majeures. La première est que la variabilité intraopérateur est faible, montrant que chaque opérateur est capable de reproduire l'exercice, sans inclure d'importantes variations dans ses mouvements. À l'inverse, la variabilité interopérateur est forte, permettant de construire un groupe distinct par opérateur. Ajouté à cela, nous constatons également que la variabilité intrainterface est faible, indiquant que chaque interface permet une répétabilité dans la réalisation de la tâche. Enfin, nous remarquons que la variabilité interinterface est forte, permettant d'associer chaque groupe à une interface distincte.

Nous pouvons déduire que l'association d'un opérateur et d'une interface crée une signature spécifique dans la trajectoire, distinguable à travers l'utilisation d'autres interfaces et la comparaison avec d'autres opérateurs. De plus, les trajectoires des instruments chirurgicaux provenant des interfaces Sigma.7 et de la Leap Motion sont significativement différentes, indiquant que la Leap Motion n'autorise pas une exécution similaire à ce que peut proposer l'interface haptique Sigma.7 pour un exercice d'entraînement de transfert d'objets.

Dans l'état actuel des travaux, les compétences des opérateurs n'ont pas été préalablement évaluées au travers de métriques cliniques dédiées à l'évaluation de compétences techniques telles que GOALS ou OSATS. L'association de ces données aux résultats présentés précédemment aurait pu nous permettre de conforter notre intuition, à savoir que la méthode proposée est capable de distinguer l'interface utilisée ainsi que l'opérateur, impliquant alors un niveau de compétence spécifique.

L'approche méthodologique proposée dans cette étude admet deux principaux avantages. D'une part, la méthode de classification proposée est générique, ne reposant sur aucune information a priori. Cela indique donc qu'elle est applicable à une autre plateforme, sous réserve d'acquérir les trajectoires des instruments chirurgicaux. D'autre part, d'après les premiers résultats obtenus, l'approche proposée semble montrer un chemin vers une évaluation automatique des compétences techniques de l'opérateur, n'utilisant que les trajectoires comme données d'analyse. Toutefois, il sera nécessaire d'entreprendre une nouvelle étude afin de valider le lien entre le niveau d'expertise de l'opérateur et le classement réalisé par notre approche non-supervisée.

4.4 Conclusion

Dans ce chapitre, nous proposons l'analyse et l'étude d'une nouvelle interface homme-machine sans contact qu'est la Leap Motion, pour réaliser une tâche d'entraînement en chirurgie laparoscopique robotisée. Après le développement d'une méthodologie permettant l'utilisation du dispositif pour la commande du robot, nous présentons une première étude de faisabilité permettant de quantifier et qualifier les performances de cette interface ainsi que celle d'une interface haptique de référence qu'est le Sigma.7. Suite aux résultats, nous constatons que le dispositif autorise l'exécution précise de gestes, mais admet des performances en dessous de l'interface haptique. Toutefois, bien que cette dernière surpasse la Leap Motion en termes d'utilisation, l'écart reste faible et parfois non significatif. Dans une seconde étude, nous avons analysé les trajectoires des instruments chirurgicaux lors de la manipulation avec chacune des interfaces afin d'étudier les variations induites lors de l'exécution. L'approche de classification non-supervisée proposée dans ces travaux permet de créer des groupes, décrit par un opérateur et une interface spécifique. Ces groupes semblent donc générer une signature unique, distinguable au sein d'une base de données de trajectoires. La méthodologie proposée semble montrer un chemin vers une évaluation automatique du niveau d'expertise, en se basant uniquement sur les trajectoires des instruments chirurgicaux. Toutefois, des études complémentaires sont nécessaires pour corréler les groupes obtenus avec une représentation formelle du niveau de compétence acquis au travers de techniques actuellement validées en clinique. Ce chapitre nous permet de conclure que l'utilisation de la Leap Motion dans le cadre de l'entraînement en chirurgie robotisée est faisable. Bien que les analyses présentent les faiblesses du dispositif, nous avons réussi à proposer un cadre d'entraînement utilisable par les chirurgiens. Ces travaux ouvrent donc des portes vers l'utilisation de ce type d'interface pour le déploiement de nouveaux systèmes d'entraînement, accessibles à moindre coût, et autorisant une manipulation proche de ce qui peut être obtenu avec les systèmes de référence existants.



Conclusion et perspectives des travaux

Dans ce manuscrit, nous avons présenté des travaux en lien direct et étroit avec la robotique chirurgicale et la modélisation de processus chirurgicaux. Nous avons abordé ces thématiques en présentant deux contributions majeures visant l'assistance aux gestes pour l'entraînement en chirurgie laparoscopique robotisée. Dans ce chapitre, nous rappelons brièvement chacune de ces contributions avant de présenter des perspectives d'évolutions à court terme, mais également à long terme.

Contributions principales

Ces travaux présentent deux contributions principales. La première est la mise en place d'une méthodologie pour la segmentation automatique de trajectoires des instruments chirurgicaux et la reconnaissance associée aux gestes dans le cadre de l'entraînement en chirurgie robotisée. La seconde contribution traite en premier lieu de l'étude d'une nouvelle interface homme-machine sans contact pour les systèmes de téléopération, ainsi qu'une méthodologie pour l'analyse des trajectoires réalisées dans le cadre de l'entraînement en chirurgie robotisée.

Segmentation automatique et reconnaissance des gestes chirurgicaux

Dans notre première contribution, nous nous sommes intéressés à l'analyse des gestes chirurgicaux dans le cadre de l'entraînement. Pour cela, nous avons créé un environnement d'entraînement physique similaire à ce que nous pouvons trouver sur les simulateurs en utilisant la plateforme robotisée disponible au LIRMM composée d'un robot Raven-II et des interfaces Sigma.7. Nous avons ensuite construit une tâche de transfert d'objets, ici des pions en silicone, afin de travailler et entraîner les compétences de manipulation et

de coordination au travers d'un système de téléopération. Enfin, durant la réalisation de sessions d'entraînement, nous avons acquis les trajectoires des instruments chirurgicaux afin de développer notre méthode de reconnaissance des gestes chirurgicaux.

À partir de ces trajectoires, nous développons une méthode de segmentation des gestes chirurgicaux, appelés surgemes, en un ensemble de composantes primitives appelées dexemes. Cette approche de segmentation non-supervisée repose sur quatre étapes et ne nécessite aucune connaissance a priori, ni sur la tâche d'entraînement, ni sur le système robotisé. La première étape réalise un prétraitement des données pour calculer de nouveaux signaux invariants en rotation et translation, et permet également de filtrer et normaliser les signaux pour réduire les variabilités intra- et interpersonnelles. La deuxième étape consiste à détecter des points caractéristiques dans chacun des signaux en utilisant la notion de persistance développée dans des travaux de la littérature. Cette méthode nous a permis de construire un ensemble de points que nous considérons comme des délimiteurs temporels des dexemes. La troisième étape permet de mesurer la pertinence de ces délimiteurs en utilisant des métriques d'analyse spatio-temporelle des signaux. Pour cela, nous avons comparé trois métriques de similarité que sont les distances de Hausdorff, de Fréchet et les Déformations Temporelles Dynamiques. Grâce à ces métriques, nous quantifions la similarité des signaux de trajectoire définis par les délimiteurs temporels. La dernière étape de la segmentation consiste à sélectionner les délimiteurs temporels ayant les scores les plus importants. Pour cela, nous avons employé une méthode de Suppression des Non-Maxima sur les scores de similarité afin d'obtenir en sortie un ensemble de dexemes représentant des portions de trajectoires distinctes les unes des autres. Nous avons vérifié notre approche sur des jeux de données acquis avec le système robotisé, au travers de plusieurs opérateurs. Nous en avons conclu que notre approche fournit une segmentation robuste et cohérente d'un point de vue observateur humain.

Nous avons ensuite combiné l'approche de segmentation développée avec une phase d'apprentissage et de reconnaissance, composée de trois étapes. La première étape permet d'extraire des attributs caractéristiques de l'ensemble des signaux composant un dexeme. Ces attributs sont construits en utilisant une approximation polynômiale de chaque signal, afin d'en récupérer une signature unique composée des coefficients du polynôme. Grâce à cette signature, nous pouvons apprendre les différents dexemes présents dans la trajectoire. Nous avons ensuite testé deux algorithmes de classification que sont les K-Plus Proches Voisins et les Machines à Vecteurs de Support. Grâce à ces algorithmes d'apprentissage automatique, nous sommes capables d'annoter chaque dexeme avec son appartenance à un surgeme. La dernière étape consiste à robustifier la séquence de dexemes en sortie. Pour ce faire, nous filtrons la séquence de dexeme en sortie de notre processus afin de supprimer les bruits de classification. En sortie, la séquence de dexemes générée permet de reconstituer l'ensemble des surgemes réalisés par l'opérateur.

À travers deux études de validation, nous avons quantifié notre approche pour la reconnaissance des gestes. Ainsi, nous avons mesuré les performances de notre phase d'apprentissage et reconnaissance en utilisant une référence de surgemes construite manuel-

lement. Dans un second temps, nous avons mesuré les performances de notre approche complète incluant la segmentation non-supervisée et la reconnaissance. Nous sommes capables de distinguer la séquence de surgemes réalisée durant l'entraînement avec un taux de reconnaissance de 81,9%. Nous avons constaté que grâce à cette approche, les gestes les plus importants réalisés durant la tâche d'entraînement peuvent être très bien reconnus. Toutefois, nous souhaitons proposer des évolutions à ces travaux, visant à étendre ces recherches et à améliorer la méthodologie mise au point destinée à la reconnaissance des gestes chirurgicaux à des fins d'évaluation des compétences techniques chirurgicales.

Étude d'une nouvelle interface sans contact pour la téléopération

Pour notre seconde contribution, nous nous sommes intéressés à l'utilisation d'une nouvelle interface manuelle sans contact qu'est la Leap Motion, afin de commander un robot chirurgical pour l'entraînement aux gestes. Dans ce contexte, nous avons utilisé la plateforme de téléopération du LIRMM. Cette plateforme est composée d'un robot Raven-II destiné à la recherche en robotique chirurgicale, auquel nous avons connecté l'interface sans contact pour piloter les instruments. Nous avons ensuite mis en place des méthodes fonctionnelles pour réaliser la commande du robot à partir de cette interface. Dans un premier temps, nous filtrons les signaux acquis par l'interface afin de réduire les mouvements parasites et capturer uniquement les fréquences fondamentales des mouvements humains. Nous avons ensuite défini un modèle permettant de contrôler finement les 7 degrés de liberté des deux instruments chirurgicaux du robot Raven-II.

À travers une première étude de faisabilité, nous avons quantifié les performances de ce dispositif en le comparant à une interface haptique de référence qu'est le Sigma.7. Pour ce faire, nous avons mis en place la tâche d'entraînement de transfert d'objets liée à l'entraînement aux gestes chirurgicaux. Puis, nous avons acquis plusieurs variables quantitatives durant les exercices. Nous avons également relevé des informations sur le ressenti des sujets au travers d'un questionnaire, afin d'évaluer les dispositifs via différents critères qualitatifs. Suite à cette première étude, nous en avons conclu que, bien que les performances globales de l'interface en contact soit en dessous des performances de l'interface haptique Sigma.7, les faibles écarts nous incitent à poursuivre ces travaux afin d'améliorer la commande des instruments.

Dans une seconde étude, nous avons modifié certaines caractéristiques afin d'améliorer la manipulation offerte par l'interface Leap Motion. Grâce à cela, nous avons réalisé de nouvelles expériences visant cette fois-ci à analyser les trajectoires des instruments chirurgicaux réalisées avec chacune des interfaces. À partir de nouvelles acquisitions provenant d'exercices d'entraînement de transfert d'objets, nous avons mis en place une méthodologie permettant d'évaluer les caractéristiques spatio-temporelles des trajectoires afin de les classer par groupe de similarité. Pour cela, nous construisons au préalable une matrice de distance, et appliquons ensuite une technique de classification non-supervisée sur ces données. Sans aucune information a priori sur le système, l'opérateur et la tâche à réaliser,

nous sommes capables de distinguer les différents sujets, ainsi que l'interface employée durant l'exercice. Cette approche, générique et indépendante de la plateforme d'entraînement, nous a également révélé que chaque interface de commande possède sa propre signature lors de l'exécution, nous permettant de distinguer chacune d'elle. Nous en avons conclu que la Leap Motion ne permet pas la réalisation de trajectoires similaires à celles du Sigma.7. Toutefois, les améliorations apportées à la commande basée sur la Leap Motion nous laissent à penser que cette interface, accessible à moindre coût, peut être une alternative très intéressante pour la commande de robot ou de simulateur dans le cadre de l'entraînement aux gestes en chirurgie laparoscopique robotisée.

Ces approches, complémentaires, contribuent à l'amélioration des techniques et dispositifs existants destinés à la reconnaissance des gestes pour l'apprentissage des compétences et la formation en chirurgie robotisée.

Perspectives des travaux

Les contributions développées dans ce manuscrit amènent de nombreuses perspectives, aussi bien cliniques et technologiques que scientifiques. Nous tentons d'en fournir un aperçu global afin de proposer des pistes d'améliorations directes de nos travaux, ainsi que des perspectives à plus long terme pour l'intégration et l'utilisation des méthodes dans un cadre d'entraînement aux gestes, mais également dans un cadre clinique.

Perspectives scientifiques à court terme

Les contributions présentées dans ce manuscrit sont des premières étapes pour l'évolution des méthodes et dispositifs destinés à l'entraînement en chirurgie robotisée. Nous proposons ci-après quelques perspectives scientifiques à court terme permettant l'amélioration des recherches entreprises jusqu'à présent.

Évolution de la base de données de trajectoires.

Bien que ces travaux introduisent de premiers développements pour la mise en place de techniques et dispositifs pour l'entraînement en chirurgie robotisée, nous sommes conscients que les bases de données employées ne contiennent qu'un faible nombre d'observations et de sujets. L'analyse de l'état de l'art dans le chapitre 2 a mis en avant le développement de méthodes utilisant des données acquises lors de multiples entraînements, avec plusieurs sujets possédant des niveaux d'expertise variés. L'une des évolutions directes de nos travaux peut être l'utilisation d'une base de données de trajectoires plus importante, comportant plus de gestes, d'exercices, et également de variabilité grâce à un plus grand nombre d'opérateurs. Les travaux de Gao et al. [Gao 2014] ont ouvert de nouvelles perspectives pour la communauté grâce à la mise à disposition d'un jeu de données conséquent, qui a déjà été utilisé dans une partie des travaux de la littérature traitant de

la reconnaissance des gestes chirurgicaux. A partir de cette initiative, plusieurs développements peuvent être entrepris afin de tester et d'évaluer notre approche sur un plus grand nombre de données. De plus, la création et la mise à disposition d'un système de "Benchmark" à la communauté permettrait de comparer et d'évaluer les performances des différents algorithmes développés par les équipes à travers le monde, sur une même base de données.

Utilisation de données cinématiques complémentaires.

Ayant concentré une partie de nos efforts sur le choix de signaux pertinents pour la segmentation et la reconnaissance, une possibilité d'amélioration peut être l'intégration de signaux complémentaires telles que la vitesse, l'accélération ou le jerk afin de fournir plus d'informations pour l'étape de segmentation non-supervisée des primitives de gestes. De plus, ces données pourraient également servir au perfectionnement de l'apprentissage des gestes en améliorant leur degré de discrimination, ce qui suggérerait de meilleures performances pour la reconnaissance des gestes chirurgicaux.

Extraction d'attributs caractéristiques locaux et globaux.

Dans notre approche de reconnaissance, nous avons employé l'approximation polynômiale afin d'extraire certaines caractéristiques des signaux. Alors que d'autres travaux ont tenté d'utiliser les polynômes de Tchebychev [Naftel 2005], ainsi que la transformée de Fourier [Khalid 2005a] pour obtenir une représentation pertinente des signaux dans un espace de dimension fixe, une possibilité d'amélioration peut être l'étude et l'évaluation d'un ou d'une combinaison de descripteurs pour l'extraction d'attributs. Dans cet objectif, l'étude des descripteurs locaux et globaux implémentés dans la librairie PCL² pourrait fournir des améliorations dans le processus de reconnaissance et augmenter les scores de rappel et de précision pour chaque geste considéré.

Reconnaissance des gestes en temps-réel.

Nos travaux de segmentation et reconnaissance des gestes chirurgicaux ont été appliqués dans un contexte de reconnaissance a posteriori de la tâche d'entraînement. L'un des principaux freins à une utilisation en temps-réel réside dans la technique de segmentation employée faisant appel au critère de persistance. L'approche de sélection des points caractéristiques par persistance requiert la totalité des signaux, et dans sa version actuelle, ne peut être implémentée pour une segmentation en temps-réel. Toutefois, des modifications sont possibles pour mettre de côté cette approche et utiliser une méthodologie de détection de points caractéristiques adaptée aux contraintes temps-réel. Dans ce cas, une fois la segmentation et l'apprentissage du modèle réalisés, le système pourrait être capable d'effectuer la reconnaissance des gestes chirurgicaux au cours de l'entraînement, ouvrant

2. Point Cloud Library : <http://pointclouds.org/>

alors un champ d'application intéressant pour le développement d'un dispositif d'entraînement aux gestes chirurgicaux avec retour d'information en temps-réel.

Amélioration de l'étude de faisabilité.

L'étude de faisabilité présentée dans le chapitre 4 présente les premiers développements pour l'utilisation d'une interface sans contact destinée à la commande de robots et simulateurs chirurgicaux. Toutefois, nous sommes également conscients que suite à la première étude, des modifications ont été apportées pour la commande basée sur la Leap Motion, qui nécessiteraient de nouvelle analyse pour qualifier et quantifier les améliorations en termes de manipulation. De plus, des possibilités d'améliorations du protocole expérimental pour l'évaluation des dispositifs est envisageable. Ainsi, l'utilisation de métriques plus adaptées comme celles présentées dans le livre de Samur [Samur 2012] autoriserait une évaluation complémentaire à celle présentée dans ces travaux, permettant ainsi d'analyser plus en détail les performances de la Leap Motion pour la commande de systèmes chirurgicaux robotisés.

Évolutions de l'interaction sans contact.

Dans sa version actuelle, la Leap Motion permet l'acquisition de données cartésiennes de la main grâce à la détection et au suivi des mouvements. L'un des intérêts de ce type d'interaction est que ce système s'adapte à toutes les morphologies de main, évitant ainsi de possibles complications d'adaptation aux dimensions de la main de l'opérateur. De plus, dans un cadre d'utilisation clinique, l'asepsie et l'hygiène ne sont plus des contraintes puisque toute interaction physique est supprimée avec l'interface. Toutefois, nous avons constaté que l'utilisation d'une interaction sans contact supprimant tout retour haptique vers l'opérateur pouvait perturber la manipulation, du fait d'une impédance nulle ne favorisant ni la proprioception ni la stabilisation des mouvements. Howard et al. [Howard 2014] s'attachent à évaluer l'impact de certains canaux de communication lors de la réalisation d'une tâche d'entraînement pour la chirurgie laparoscopique. Leur étude a permis de mettre en avant l'intérêt du retour visuel de la scène, ainsi que l'importance d'un retour kinesthésique vers l'opérateur (via des moteurs vibrotactiles) améliorant la proprioception lors de la manipulation. Une poursuite de nos travaux pourrait viser à combiner les retours d'informations développés dans leurs travaux afin de compenser le manque d'information haptique durant la manipulation. De plus, une étude plus poussée pourrait permettre de distinguer et quantifier l'importance de chaque canal d'interaction lors de l'interaction sans contact pour la commande d'un système de téléopération.

Perspectives d'applications à long terme

Les contributions introduites dans ce manuscrit amènent également de nombreuses perspectives à plus long terme, principalement orientées vers l'intégration et l'application

des méthodes dans un cadre clinique. Nous proposons ci-après quelques pistes de développement à partir des travaux introduits dans ce manuscrit.

Construction de bases de données expertes.

Notre approche pour la segmentation non-supervisée et la reconnaissance des gestes chirurgicaux peut s'intégrer dans le cadre de la construction de bases de données expertes, utilisables pour l'entraînement aux gestes mais également adaptables à une utilisation clinique. L'enregistrement de multiples sessions peut permettre, grâce à la détection et la reconnaissance des gestes, de construire une base de données conséquente permettant de futures évaluations. Cette base de données peut comporter plusieurs exemples du même geste, ou en combinant notre approche avec les travaux de Petitjean et al. [Petitjean 2011], créer une séquence moyenne optimale pour chaque geste. Cette séquence peut ensuite être utilisée pour la reconnaissance des gestes, permettant ainsi de construire des gestes optimaux, compressant la taille de la base de données et facilitant la recherche et la classification d'éléments [Petitjean 2014].

Apprentissage personnalisé des gestes.

Notre approche pour la segmentation non-supervisée et la reconnaissance des gestes chirurgicaux peut également s'intégrer dans le cadre d'un apprentissage personnalisé des gestes. Il serait intéressant d'utiliser ce type d'approche pour le développement d'outils pédagogiques destinés à l'apprentissage des gestes chirurgicaux. Nous donnons ci-après un exemple concret dans le cadre d'un entraînement pour la chirurgie laparoscopique robotisée. Lorsqu'un opérateur novice exécute un geste, le système pourrait détecter et reconnaître le geste en cours, mais également évaluer la qualité du geste. De ce fait, si ce geste est mal réalisé le système peut proposer de présenter le geste optimal à exécuter. Muni d'une base de données experte conséquente, le système pourrait alors superposer en réalité augmentée le geste à réaliser afin que le novice puisse visuellement l'appréhender et la reproduire. En s'appuyant sur les récents travaux de Padoy et al. [Padoy 2011] ainsi que de Pauly et al. [Pauly 2014], il serait possible de mettre en place un retour visuel via de la réalité augmentée pour améliorer et faciliter l'apprentissage de la gestuelle.

Évaluation des compétences procédurales et documentation automatique.

Notre contribution pour la reconnaissance des gestes chirurgicaux a pour principal objectif d'aider à l'évaluation des compétences techniques, et notamment de la dextérité. Toutefois, ce type de méthodologie peut également s'employer pour réaliser l'apprentissage des compétences procédurales. En modifiant notre approche de classification et en nous basant sur des algorithmes nécessitant un modèle temporel de la tâche en cours, l'approche proposée serait adaptable pour évaluer la qualité ainsi que l'enchaînement des gestes, aussi bien dans un cadre d'entraînement que dans un contexte clinique lors d'une intervention robotisée. En complément, ce type d'approche pourrait permettre la création

de compte-rendus automatisés afin de faciliter les démarches administratives pour la gestion du patient.

Évolution des systèmes d'entraînement pour la chirurgie robotisée.

Notre contribution pour l'analyse et l'évaluation d'une nouvelle interface homme-machine sans contact pour la téléopération annonce des perspectives d'application intéressantes. En effet, dans ces travaux nous utilisons un dispositif très peu coûteux qu'est la Leap Motion. Nous montrons alors par ces premiers résultats que ce type d'interface peut être utilisé dans le cadre de l'entraînement pour la chirurgie robotisée. L'intérêt principal de ce système réside dans la commande sans contact préservant l'hygiène et l'asepsie durant la manipulation. Toutefois, nous pouvons également envisager l'utilisation de ce dispositif dans le cadre de la création d'un simulateur robotisé portable à moindre coût. Sa portabilité et son coût en font une interface de choix pour ce marché, permettant ainsi aux apprentis de s'entraîner facilement, sans contrainte d'accès à un matériel tel que le da Vinci. Cette interface peut également servir à la formation sur simulateur, dans le cadre du compagnonnage. Un exemple très direct est la mise en place d'une seconde console pour l'aide à la gestuelle. Dans ce scénario, l'apprenti est aux commandes du simulateur et l'expert connecte l'interface sans contact pour lui enseigner de manière directe la gestuelle à adopter pour réaliser une tâche. Dans ce cas, le dispositif sert alors d'interface d'appoint à l'accompagnement et l'apprentissage des compétences techniques.

Nous espérons que les développements entrepris dans ce manuscrit contribueront au déploiement de nouvelles méthodes, dispositifs et applications afin d'améliorer les conditions durant l'entraînement aux gestes pour la chirurgie laparoscopique robotisée.



Bibliographie

- [Aggarwal 2007] R. Aggarwal, J. Ward, I. Balasundaram, P. Sains, T. Athanasiou et A. Darzi. *Proving the effectiveness of virtual reality simulation for training in laparoscopic surgery*. *Annals of Surgery*, vol. 246, no. 5, pages 771–779, 2007. Cité page 17.
- [Agha 2003] R. Agha et G. Muir. *Does laparoscopic surgery spell the end of the open surgeon ?* *Journal of the Royal Society of Medicine*, vol. 96, no. 11, pages 544–546, 2003. Cité page 6.
- [Ahmadi 2009] Seyed-ahmad Ahmadi, Nicolas Padoy, Kateryna Rybachuk, Sandro Michael Heining et Nassir Navab. *Motif Discovery in OR Sensor Data with Application to Surgical Workflow Analysis and Activity Detection*. *Modeling and Monitoring of Computer Assisted Interventions*, 2009. Cité pages 36, 38, 39, 41, 49, 51, et 52.
- [Ahmed 2011] K. Ahmed, D. Miskovic, A. Darzi, T. Athanasiou et G. B. Hanna. *Observational tools for assessment of procedural skills : A systematic review*. *The American Journal of Surgery*, vol. 202, no. 4, pages 469–480, 2011. Cité pages 14, 19, et 20.
- [Ahmidi 2013] Narges Ahmidi, Yixin Gao, Benjamín Béjar, S. Swaroop Vedula, Sanjeev Khudanpur, René Vidal et Gregory D. Hager. *String Motif-Based Description of Tool Motion for Detecting Skill and Gestures in Robotic Surgery*. *Medical Image Computing and Computer-Assisted Intervention*, vol. 8149, pages 26–33, 2013. Cité pages 44, 46, et 48.
- [Altieri 2015] M. S. Altieri, J. Yang, D. A. Telem, J. Zhu, C. Halbert, M. Talamini et A. D. Pryor. *Robotic approaches may offer benefit in colorectal procedures, more controversial in other areas : a review of 168,248 cases*. *Surgical Endoscopy*, vol. 7, 2015. Cité page 11.

- [Andre-Obrecht 1988] Regine Andre-Obrecht. *A New Statistical Approach for the Automatic Segmentation of Continuous Speech Signals*. IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 36, no. 1, pages 29–40, 1988. Cité page 33.
- [Barnes 1987] R. W. Barnes. *Surgical Handicraft : Teaching and Learning Surgical Skills*. The American Journal of Surgery, vol. 153, no. 5, pages 422–427, 1987. Cité page 14.
- [Béguin 1997] P. Béguin et A. Weill-Fassina. *La simulation en ergonomie : connaître, agir et interagir*. Octares édition, 1997. Cité page 14.
- [Béjar Haro 2012] Benjamín Béjar Haro, Luca Zappella et René Vidal. *Surgical Gesture Classification from Video Data*. Medical Image Computing and Computer-Assisted Intervention, vol. 7510, pages 34–41, 2012. Cité pages 42, 44, 45, 48, 49, 54, 99, et 103.
- [Benenson 2012] Rodrigo Benenson, Markus Mathias, Radu Timofte et Luc Van Gool. *Pedestrian detection at 100 frames per second*. IEEE Conference on Computer Vision and Pattern Recognition, pages 2903–2910, 2012. Cité page 33.
- [Benenson 2014] Rodrigo Benenson, Omran Mohamed, Jan Hosang et Bernt Schiele. *Ten Years of Pedestrian Detection, What Have We Learned?* European Conference on Computer Vision - CVRSUAD Workshop, 2014. Cité page 33.
- [Bennett 2000] Kristin P Bennett et Colin Campbell. *Support Vector Machines : Hype or Hallelujah ?*, 2000. Cité page 85.
- [Berkelman 2003] P. Berkelman, P. Cinquin, E. Boidard, J. Troccaz, C. Létoublon et J-M. Ayoubi. *Design, control and testing of a novel compact laparoscopic endoscope manipulator*. Journal of Systems and Control Engineering, vol. 217, no. 4, pages 329–341, 2003. Cité page 113.
- [Bholat 1999] O. S. Bholat, R. S. Haluck, W. B. Murray, P. J. Gorman et T. M. Krummel. *Tactile feedback is present during minimally invasive surgery*. Journal of the American College of Surgeons, vol. 189, no. 4, pages 349–355, 1999. Cité page 7.
- [BlueDragon 2015] BlueDragon. http://bionics.seas.ucla.edu/research/surgery_device_1.html. 2015. Cité page 18.
- [Bó 2011] Antnio Padilha Lanari Bó, Philippe Poignet et Christian Geny. *Pathological Tremor and Voluntary Motion Modeling and Online Estimation for Active Compensation*. IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 19, no. 2, pages 177–185, 2011. Cité page 121.
- [Botden 2007] S. M B I Botden, Sonja N. Buzink, Marlies P. Schijven et Jack J. Jakimowicz. *Augmented versus Virtual Reality Laparoscopic Simulation : What Is the Difference?* World Journal of Surgery, vol. 31, no. 4, pages 764–772, 2007. Cité page 16.

- [Boutin 2000] Mireille Boutin. *Numerically Invariant Signature Curves*. International Journal of Computer Vision, vol. 40, no. 3, pages 235–248, 2000. Cité page 65.
- [Breedveld 1999] P. Breedveld, H. G. Stassen, D. W. Meijer et L. P. S. Stassen. *Theoretical background and conceptual solution for depth perception and eye-hand coordination problems in laparoscopic surgery*. Minimally Invasive Therapy & Allied Technologies, vol. 8, no. 4, pages 227–234, 1999. Cité page 7.
- [Breedveld 2001] P. Breedveld et M. Wentink. *Eye-hand coordination in laparoscopy - an overview of experiments and supporting aids*. Minimally Invasive Therapy & Allied Technologies Allied Technologies, vol. 10, no. 3, pages 155–162, 2001. Cité page 7.
- [Calabi 1998] E Calabi, P Olver, C Shakiban, A Tannenbaum et S Haker. *Differential and Numerically Invariant Signature Curves Applied to Object Recognition*. International Journal of Computer Vision, vol. 26, no. 2, pages 107–135, 1998. Cité page 65.
- [Chen 2011] Jinyang Chen, Rangding Wang, Liangxu Liu et Jiatao Song. *Clustering of Trajectories Based on Hausdorff Distance*. International Conference on Electronics, Communications and Control, pages 1940–1944, 2011. Cité page 71.
- [Chiniara 2007] G. Chiniara. *Simulation médicale pour acquisition des compétences en anesthésie*. Congrès national d’anesthésie et de réanimation, pages 41–49, 2007. Cité page 16.
- [Chipman 2009] J. G. Chipman et C. C. Schmitz. *Using Objective Structured Assessment of Technical Skills to Evaluate a Basic Skills Simulation Curriculum for First-Year Surgical Residents*. Journal of the American College of Surgeons, vol. 209, no. 3, pages 364–370, 2009. Cité page 21.
- [Chmarra 2010] M. K. Chmarra, S. Klein, J. C. F. De Winter, F. W. Jansen et J. Dankelman. *Objective classification of residents based on their psychomotor laparoscopic skills*. Surgical Endoscopy and Other Interventional Techniques, vol. 24, no. 5, pages 1031–1039, 2010. Cité page 22.
- [Cifuentes 2013] Jenny Cifuentes, Minh Tu Pham, Richard Moreau, Flavio Prieto et Pierre Boulanger. *An Arc-length Warping Algorithm for Gesture Recognition Using Quaternion Representation*. IEEE International Conference on Engineering in Medicine and Biology Society, pages 6248–6251, 2013. Cité pages 64, 73, et 99.
- [Coles 2011] Timothy R. Coles, Dwight Meglan et Nigel W. John. *The role of haptics in medical training simulators : A survey of the state of the art*. IEEE Transactions on Haptics, vol. 4, no. 1, pages 51–66, 2011. Cité page 109.

- [Cooper 2013] Michol A. Cooper, Andrew Ibrahim, Heather Lyu et Martin A. Makary. *Underreporting of Robotic Surgery Complications*. Journal for Healthcare Quality, pages 1–5, 2013. Cité page 11.
- [Cotin 2002] S. Cotin, N. Stylopoulos, M. P. Ottensmeyer, P. F. Neumann, D. W. Rattner et S. Dawson. *Metrics for Laparoscopic Skills Trainers : The Weakest Link!* Medical Image Computing and Computer-Assisted Intervention, vol. 2488, pages 35–43, 2002. Cité page 22.
- [Cully 2015] Antoine Cully et Jean-baptiste Mouret. *Evolving a Behavioral Repertoire for a Walking Robot*. Evolutionary Computation, no. 2005, 2015. Cité page 100.
- [Datta 2002] V. Datta, A. Chang, S. Mackay et A. Darzi. *The relationship between motion analysis and surgical technical assessments*. American Journal of Surgery, vol. 184, no. 1, pages 70–73, 2002. Cité page 14.
- [Derossis 1998] A. M. Derossis, G. M. Fried, M. Abrahamowicz, H. H. Sigman, J. S. Barkun et J. L. Meakins. *Development of a Model for Training and Evaluation of Laparoscopic Skills*. The American Journal of Surgery, vol. 175, no. 6, pages 482–487, 1998. Cité pages 22, 31, 48, 87, et 121.
- [Despinoy 2013] Fabien Despinoy, Jonathan Leon-torres, Marie-Aude Vitrani et Benoît Herman. *Toward Remote Teleoperation with Eye and Hand : A First Experimental Study*. Joint Workshop on New Technologies for Computer/Robot Assisted Surgery, 2013. Cité page 114.
- [Dosis 2005] Aristotelis Dosis, Fernando Bello, Duncan Gillies, Shabnam Undre, Rajesh Aggarwal et Ara Darzi. *Laparoscopic task recognition using Hidden Markov Models*. Studies in health technology and informatics, vol. 111, pages 115–122, 2005. Cité pages 39, 40, 44, 45, 48, 49, 50, 55, et 103.
- [Dragan 2013] Anca D Dragan, Siddhartha S. Srinivasa et Kenton C. T. Lee. *Teleoperation with Intelligent and Customizable Interfaces*. Journal of Human-Robot Interaction, vol. 2, no. 2, pages 33–57, 2013. Cité page 115.
- [Du 2012] Guanglong Du, Ping Zhang, Jianhua Mai et Zeling Li. *Markerless Kinect-Based Hand Tracking for Robot Teleoperation*. International Journal of Advanced Robotic Systems, vol. 9, no. 36, pages 1–10, 2012. Cité page 115.
- [Dupin 2014] Lucile Dupin, Vincent Hayward et Mark Wexler. *Direct coupling of haptic signals between hands*. Proceedings of the National Academy of Sciences of the United States of America, vol. 112, no. 2, pages 619–24, 2014. Cité page 103.
- [DVSS 2015] DVSS. http://www.intuitivesurgical.com/products/skills_simulator/, 2015. Cité page 18.

- [DvTrainer 2015] DvTrainer. <http://www.mimicsimulation.com/products/dv-trainer/>. 2015. Cité page 18.
- [Edelsbrunner 2002] Herbert Edelsbrunner, David Letscher et Afra Zomorodian. *Topological Persistence and Simplification*. Discrete & Computational Geometry, vol. 28, no. 4, pages 511–533, 2002. Cité page 69.
- [Egi 2008] H. Egi, M. Okajima, M. Yoshimitsu, S. Ikeda, Y. Miyata, H. Masugami, T. Kawahara, Y. Kurita, M. Kaneko et T. Asahara. *Objective Assessment of Endoscopic Surgical Skills by Analyzing Direction-Dependent Dexterity Using the Hiroshima University Endoscopic Surgical Assessment Device (HUESAD)*. Surgery Today, vol. 38, no. 8, pages 705–710, 2008. Cité page 18.
- [Elle 2002] O. J. Elle, M. G. Gulbrandsen, E. Samset, G. ten Cate, L. Aurdal, A. Austad, T. K. Lien et E. Fosse. *Head tracking of a surgical robotic scopeholder - a user involvement test of the system*. Computer Assisted Radiology and Surgery, pages 188–193, 2002. Cité page 114.
- [Fang 2014] Hao Fang, Saurabh Gupta, Forrest Iandola, Rupesh Srivastava, Li Deng, Piotr Dollár, Jianfeng Gao, Xiaodong He, Margaret Mitchell, John C. Platt, C. Lawrence Zitnick et Geoffrey Zweig. *From Captions to Visual Concepts and Back*. IEEE Conference on Computer Vision and Pattern Recognition, pages 1–10, 2014. Cité page 33.
- [Flin 2013] R. Flin. *Non-technical skills for anaesthetists, surgeons and scrub practitioners (ANTS, NOTSS and SPLINTS)*. The Health Foundation, pages 1–9, 2013. Cité page 13.
- [FLSTrainer 2015] FLSTrainer. <http://www.fl-products.com/products/category/fls-trainers>. 2015. Cité page 18.
- [Forestier 2012] Germain Forestier, Florent Lalys, Laurent Riffaud, Brivael Trelhu et Pierre Jannin. *Classification of surgical processes using dynamic time warping*. Journal of Biomedical Informatics, vol. 45, no. 2, pages 255–264, 2012. Cité page 130.
- [Forestier 2013] Germain Forestier, Florent Lalys, Laurent Riffaud, D. Louis Collins, Jurgen Meixensberger, Shafik N. Wassef, Thomas Neumuth, Benoit Goulet et Pierre Jannin. *Multi-site study of surgical practice in neurosurgery based on surgical process models*. Journal of Biomedical Informatics, vol. 46, no. 5, pages 822–829, 2013. Cité page 130.
- [Francis 2001] N. K. Francis, G. B. Hanna et A. Cuschieri. *Reliability of the Advanced Dundee Endoscopic Psychomotor Tester for bimanual tasks*. Archives of surgery, vol. 136, no. 1, pages 40–43, 2001. Cité page 18.

- [Freschi 2013] C. Freschi, V. Ferrari, F. Melfi, M. Ferrari, F. Mosca et A. Cuschieri. *Technical review of the da Vinci surgical telemanipulator*. The International Journal of Medical Robotics and Computer Assisted Surgery, vol. 9, no. 4, pages 396–406, 2013. Cité pages 10 et 110.
- [Gao 2014] Yixin Gao, S. Swaroop Vedula, Carol E. Reiley, Narges Ahmidi, Balakrishnan Varadarajan, Henry C. Lin, Lingling Tao, Luca Zappella, Benjamín Béjar, David D. Yuh, Chi Chiung Grace Chen, René Vidal, Sanjeev Khudanpur et Gregory D. Hager. *JHU-ISI Gesture and Skill Assessment Working Set (JIGSAWS) : A Surgical Activity Dataset for Human Motion Modeling*. Modeling and Monitoring of Computer Assisted Interventions, pages 1–10, 2014. Cité pages 54, 103, et 142.
- [Gilbert 2009] J. M. Gilbert. *The EndoAssist robotic camera holder as an aid to the introduction of laparoscopic colorectal surgery*. Annals of the Royal College of Surgeons of England, vol. 91, no. 5, pages 389–393, 2009. Cité page 114.
- [Guna 2014] Jože Guna, Grega Jakus, Matevž Pogačnik, Sašo Tomažič et Jaka Sodnik. *An Analysis of the Precision and Reliability of the Leap Motion Sensor and Its Suitability for Static and Dynamic Tracking*. Sensors, vol. 14, no. 2, pages 3702–3720, 2014. Cité page 120.
- [Guthart 2000] Gary S. Guthart et Kenneth Salisbury. *The Intuitive Telesurgery System : Overview and Application*. IEEE International Conference on Robotics and Automation, vol. 1, pages 618–621, 2000. Cité pages 8 et 110.
- [Hamdorf 2000] J. M. Hamdorf et J. C. Hall. *Acquiring Surgical Skills*. British Journal of Surgery, vol. 87, no. 1, pages 28–37, 2000. Cité page 13.
- [Hannaford 2013] Blake Hannaford, Jacob Rosen, Diana W. Friedman, Hawkeye King, Philip Roan, Lei Cheng, Daniel Glozman, Ji Ma, Sina Nia Kosari et Lee White. *Raven-II : An Open Platform for Surgical Robotics Research*. IEEE Transactions on Biomedical Engineering, vol. 60, no. 4, pages 954–959, 2013. Cité pages 52 et 118.
- [HAS 2012] HAS. *La simulation en santé*. Haute Autorité de santé, no. Janvier, 2012. Cité pages 14 et 15.
- [HAS 2014] HAS. *Bilan de l'enquête concernant les robots chirurgicaux Da Vinci de la société Intuitive Surgical à destination des professionnels de santé*. Agence Nationale de Sécurité du Médicament et des Produits de Santé, no. Février, 2014. Cité pages 10 et 11.
- [HAS 2015] HAS. *Évaluation des dimensions clinique et organisationnelle de la chirurgie robot-assistée dans le cadre d'une prostatectomie totale*. Haute Autorité de santé, no. Avril, 2015. Cité page 11.

- [Hayter 2009] M. A. Hayter, Z. Friedman, M. D. Bould, J. G. Hanlon, R. Katznelson, B. Borges et V. N. Naik. *Validation of the Imperial College Surgical Assessment Device (ICSAD) for labour epidural placement*. Canadian Journal of Anesthesia, vol. 56, no. 6, pages 419–426, 2009. Cité page 18.
- [Hernoux 2013] F. Hernoux, R. Béarée, L. Gajny, E. Nyiri, J. Bancalin et O. Gibaru. *Leap Motion pour la capture de mouvement 3D par spline L1 - Application à la robotique*. Conférence Groupe de Travail en Modélisation Géométrique, pages 1–6, 2013. Cité page 115.
- [Holden 2014] Matthew Stephen Holden, Tamas Ungi, Derek Sargent, Robert C. McGraw, Elvis C. S. Chen, Sugantha Ganapathy, Terry M. Peters et Gabor Fichtinger. *Feasibility of Real-Time Workflow Segmentation for Tracked Needle Interventions*. IEEE Transactions on Biomedical Engineering, vol. 61, no. 6, pages 1720–1728, 2014. Cité pages 36, 39, 41, 42, 45, 46, 49, 51, et 53.
- [Hollis 1993] R. L. Hollis et S. E. Salcudean. *Lorentz Levitation Technology: A New Approach to Fine Motion Robotics, Teleoperation, Haptic Interfaces, and Vibration Isolation*. 1993. Cité page 112.
- [Horeman 2014] T. Horeman, J. Dankelman, F. W. Jansen et John J. Van Den Dobbelen. *Assessment of Laparoscopic Skills Based on Force and Motion Parameters*. IEEE Transactions on Biomedical Engineering, vol. 61, no. 3, pages 805–813, 2014. Cité page 22.
- [Howard 2014] Thomas Howard et Jérôme Szewczyk. *Visuo-haptic feedback for 1-D Guidance in laparoscopic surgery*. International Conference on Biomedical Robotics and Biomechatronics, pages 58–65, 2014. Cité pages 121, 127, et 144.
- [Hundtofte 2002] C.S. Hundtofte, G.D. Hager et A.M. Okamura. *Building a Task Language for Segmentation and Recognition of User Input to Cooperative Manipulation Systems*. Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, pages 225–230, 2002. Cité pages 36, 39, 46, et 50.
- [Jannin 2001] P. Jannin, M. Raimbault, X. Morandi, E. Seigneuret et B. Gibaud. *Design of a Neurosurgical Procedure Model for Multimodal Image-Guided Surgery*. International Congress Series, vol. 1230, pages 102–106, 2001. Cité page 27.
- [Jannin 2006] Pierre Jannin, Christophe Grova et Calvin R. Maurer. *Model for defining and reporting reference-based validation protocols in medical image processing*. International Journal of Computer Assisted Radiology and Surgery, vol. 1, no. 2, pages 63–73, 2006. Cité pages 49 et 54.

- [Jannin 2007] P. Jannin et X. Morandi. *Surgical models for computer-assisted neurosurgery*. NeuroImage, vol. 37, no. 3, pages 783–791, 2007. Cité pages 27 et 28.
- [Jannin 2008] Pierre Jannin et Werner Korb. *Assessment of image-guided interventions*. In Image-Guided Interventions : Technology and Applications, pages 531–549. 2008. Cité page 38.
- [Jarrassé 2014] Nathanaël Jarrassé, Adriano Tacilo Ribeiro, Anis Sahbani, Wael Bachta et Agnes Roby-Brami. *Analysis of hand synergies in healthy subjects during bimanual manipulation of various objects*. Journal of neuroengineering and rehabilitation, vol. 11, no. 1, page 113, 2014. Cité page 103.
- [Judkins 2009] T. N. Judkins, D. Oleynikov et N. Stergiou. *Objective evaluation of expert and novice performance during robotic surgical training tasks*. Surgical Endoscopy and Other Interventional Techniques, vol. 23, no. 3, pages 590–597, 2009. Cité page 22.
- [Kamakshi Prasad 2004] V. Kamakshi Prasad, T. Nagarajan et Hema A. Murthy. *Automatic segmentation of continuous speech using minimum phase group delay functions*. Speech Communication, vol. 42, no. 3-4, pages 429–446, 2004. Cité page 33.
- [Kazanzides 2014] Peter Kazanzides, Zihan Chen, Anton Deguet, Gregory S Fischer, Russell H Taylor et Simon P Dimaio. *An Open-Source Research Kit for the da Vinci Surgical System*. IEEE International Conference on Robotics and Automation, pages 6434–6439, 2014. Cité page 51.
- [Khalid 2005a] Shehzad Khalid et Andrew Naftel. *Classifying Spatiotemporal Object Trajectories using Unsupervised Learning of Basis Function Coefficients*. ACM International Workshop on Video Surveillance and Sensor Networks, pages 45–52, 2005. Cité pages 82 et 143.
- [Khalid 2005b] Shehzad Khalid et Andrew Naftel. *Evaluation of Matching Metrics for Trajectory-based Indexing and Retrieval of Video Clips*. IEEE Workshop on Applications of Computer Vision, pages 242–249, 2005. Cité page 71.
- [Khalil 1999] Wisama Khalil et Etienne Dombre. *Modélisation identification et commande des robots*. 1999. Cité page 64.
- [Kim 2014] Yonjae Kim, Peter C W Kim, Rebecca Selle, Azad Shademan et Axel Krieger. *Experimental Evaluation of Contact-Less Hand Tracking Systems for Tele-Operation of Surgical Tasks*. IEEE International Conference on Robotics and Automation, pages 3502–3509, 2014. Cité page 120.

- [Kragic 2003] D. Kragic et G.D. Hager. *Task Modeling and Specification for Modular Sensory Based Human-Machine Cooperative Systems*. IEEE International Conference on Intelligent Robots and Systems, pages 3192–3197, 2003. Cité pages 36 et 39.
- [Kranzfelder 2014] Michael Kranzfelder, Armin Schneider, Adam Fiolka, Sebastian Koller, Silvano Reiser, Thomas Vogel, Dirk Wilhelm et Hubertus Feussner. *Reliability of sensor-based real-time workflow recognition in laparoscopic cholecystectomy*. International Journal of Computer Assisted Radiology and Surgery, vol. 9, no. 6, pages 941–948, 2014. Cité page 27.
- [Kwoh 1988] Y. S. Kwoh, J. Hou, E. A. Jonckheere et S. Hayati. *A Robot with Improved Absolute Positioning Accuracy for CT Guided Stereotactic Brain Surgery*. IEEE Transactions on Biomedical Engineering, vol. 35, no. 2, pages 153–160, 1988. Cité pages 8 et 108.
- [Lalys 2011] Florent Lalys, Laurent Riffaud, Xavier Morandi et Pierre Jannin. *Surgical phases detection from microscope videos by combining SVM and HMM*. Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging, vol. 6533, pages 54–62, 2011. Cité page 27.
- [Lalys 2012] Florent Lalys, Lalys Riffaud, David Bouget et Pierre Jannin. *A framework for the recognition of high-level surgical tasks from video images for cataract surgeries*. IEEE Transactions on Biomedical Engineering, vol. 59, no. 4, pages 966–976, 2012. Cité page 31.
- [Lalys 2014] Florent Lalys et Pierre Jannin. *Surgical process modelling : A review*. International Journal of Computer Assisted Radiology and Surgery, vol. 9, pages 495–511, 2014. Cité pages 26, 29, 30, 31, et 55.
- [LapMentor 2015] LapMentor. <http://symbionix.com/simulators/lap-mentor/platforms/>. 2015. Cité page 18.
- [LaproTrain 2015] LaproTrain. <http://www.laprotrain.com/index.aspx>. 2015. Cité page 18.
- [LapSim 2015] LapSim. <http://www.surgical-science.com/lapsim-the-proven-training-system/>. 2015. Cité page 18.
- [LapVR 2015] LapVR. <http://caehealthcare.com/eng/interventional-simulators/lapvr>. 2015. Cité page 18.
- [Lea 2015] Colin Lea, Gregory D. Hager et René Vidal. *An Improved Model for Segmentation and Recognition of Fine-grained Activities with Application to Surgical Training Tasks*. IEEE Winter Conference on Applications of Computer Vision, pages 1–7, 2015. Cité pages 38, 40, 41, 42, 44, 45, 48, 49, 52, 54, 99, et 103.

- [Lepor 2009] H. Lepor. *Status of Radical Prostatectomy in 2009 : Is There Medical Evidence to Justify the Robotic Approach ?* Reviews in urology, vol. 11, no. 2, pages 61–70, 2009. Cité page 10.
- [Lin 2005] Henry C. Lin, Izhak Shafran, Todd E. Murphy, Allison M. Okamura, David D. Yuh et Gregory D. Hager. *Automatic Detection and Segmentation of Robot-Assisted Surgical Motions*. Medical Image Computing and Computer-Assisted Intervention, vol. 3749, pages 802–810, 2005. Cité pages 41, 42, 43, 44, 45, 51, 53, 99, et 100.
- [Lin 2006] Henry C. Lin, Izhak Shafran, David Yuh et Gregory D. Hager. *Towards automatic skill evaluation : Detection and segmentation of robot-assisted surgical motions*. Computer Aided Surgery, vol. 11, no. 5, pages 220–230, 2006. Cité pages 27, 31, 41, 42, 43, 44, 45, 51, 53, 99, et 100.
- [Lin 2014] Jonathan Feng-Shun Lin et Dana Kulić. *Online Segmentation of Human Motion for Automated Rehabilitation Exercise Analysis*. IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 22, no. 1, pages 2881–2884, 2014. Cité page 69.
- [Loram 2006] Ian D Loram, Peter J Gawthrop et Martin Lakie. *The frequency of human, manual adjustments in balancing an inverted pendulum is constrained by intrinsic physiological factors*. The Journal of physiology, vol. 577, pages 417–432, 2006. Cité pages 41, 67, et 121.
- [Lum 2009] Mitchell J. H. Lum, Diana C. W. Friedman, Ganesh Sankaranarayanan, Hawkeye King, Kenneth Fodero, Rainer Leuschke, Blake Hannaford, Jacob Rosen et Mika N. Sinanan. *The RAVEN : Design and Validation of a Telesurgery System*. The International Journal of Robotics Research, vol. 28, no. 9, pages 1183–1197, 2009. Cité page 118.
- [Lyons 2013] C. Lyons, D. Goldfarb, S. L. Jones, N. Badhiwala, B. Miles, R. Link et B. J. Dunkin. *Which skills really matter ? proving face, content, and construct validity for a commercial robotic simulator*. Surgical Endoscopy and Other Interventional Techniques, vol. 27, no. 6, pages 2020–2030, 2013. Cité page 22.
- [Mackenzie 2001] C. L. Mackenzie, J. A. Ibbotson, C. G. L. Cao et A. J. Lomax. *Hierarchical decomposition of laparoscopic surgery : a human factors approach to investigating the operating room environment.*, 2001. Cité pages 27, 28, et 29.
- [MaestroAR 2015] MaestroAR. <http://www.mimicsimulation.com/products/maestro-ar/>. 2015. Cité page 18.
- [Malpani 2014] Anand Malpani, S. Swaroop Vedula, Chi Chiung Grace Chen et Gregory D. Hager. *Pairwise Comparison-based Objective Score for Automated Skill Assessment*

- of Segments in a Surgical Task*. International Conference on Information Processing in Computer-Assisted Interventions, vol. 8498, pages 138–147, 2014. Cité page 35.
- [Martin 1997] J. A. Martin, G. Regehr, R. Reznick, H. Macrae, J. Murnaghan, C. Hutchison et M. Brown. *Objective structured assessment of technical skills (OSATS) for surgical residents*. British Journal of Surgery, vol. 84, pages 273–278, 1997. Cité pages 20 et 21.
- [Matsuda 2012] T. Matsuda, E. M. McDougall, Y. Ono, R. Hattori, S. Baba, M. Iwamura, T. Terachi, S. Naito et R. V. Clayman. *Positive Correlation Between Motion Analysis Data on the LapMentor Virtual Reality Laparoscopic Surgical Simulator and the Results from Videotape Assessment of Real Laparoscopic Surgeries*. Journal of Endourology, vol. 26, no. 11, pages 1506–1511, 2012. Cité page 17.
- [McCormick 1997] Robert McCormick. *Conceptual and Procedural Knowledge*. International Journal of Technology and Design Education, vol. 7, no. 1, pages 141–159, 1997. Cité page 13.
- [Megali 2006] Giuseppe Megali, Stefano Sinigaglia, Oliver Tonet et Paolo Dario. *Modelling and evaluation of surgical performance using hidden Markov models*. IEEE Transactions on Biomedical Engineering, vol. 53, no. 10, pages 1911–1919, 2006. Cité page 35.
- [Mitchell 2013] L. Mitchell, R. Flin, S. Yule, J. Mitchell, Kathy Coutts et George Youngson. *Development of a behavioural marker system for scrub practitioners’s non-technical skills (SPLINTS system)*. Journal of Evaluation in Clinical Practice, vol. 19, no. 2, pages 317–323, 2013. Cité page 13.
- [MLabs 2015] MLabs. <http://www.mimicsimulation.com/products/mlabs/>. 2015. Cité page 18.
- [Moorthy 2003] K. Moorthy et Y. Munz. *Objective assessment of technical skills in surgery*. British Medical Journal, vol. 327, no. 7422, pages 1032–1037, 2003. Cité page 13.
- [Moorthy 2004] K. Moorthy, Y. Munz, A. Dosis, F. Bello, A. Chang et A. Darzi. *Bimodal assessment of laparoscopic suturing skills : Construct and concurrent validity*. Surgical Endoscopy and Other Interventional Techniques, vol. 18, no. 11, pages 1608–1612, 2004. Cité page 22.
- [Morris 2009] Brendan Morris et Mohan Trivedi. *Learning Trajectory Patterns by Clustering : Experimental Studies and Comparative Evaluation*. IEEE Conference on Computer Vision and Pattern Recognition, pages 312–319, 2009. Cité page 100.

- [Murayama 2004] Jun Murayama, Laroussi Bougrila et Yi Luo. *SPIDAR G&G: A two-handed haptic interface for bimanual VR interaction*. Eurohaptics, pages 138–146, 2004. Cité page 112.
- [Naftel 2005] Andrew Naftel et Shehzad Khalid. *Classification and Prediction of Motion Trajectories using Spatiotemporal Approximations*. International Workshop on Human Activity Recognition and Modelling, pages 17–26, 2005. Cité pages 82 et 143.
- [Naik 2007] V. N. Naik, A. Perlas, D. B. Chandra, D. Y. Chung et V. W. S. Chan. *An Assessment Tool for Brachial Plexus Regional Anesthesia Performance : Establishing Construct Validity and Reliability*. Regional Anesthesia and Pain Medicine, vol. 32, no. 1, pages 41–45, 2007. Cité page 21.
- [Nathan 2006] Cherie-Ann O. Nathan, Vinaya Chakradeo, Kavita Malhotra, Horacio D’Agostino et Ravish Patwardhan. *The voice-controlled robotic assist scope holder AESOP for the endoscopic approach to the sella*. Skull base : official journal of North American Skull Base Society, vol. 16, no. 3, pages 123–131, 2006. Cité page 113.
- [Neubeck 2006] Alexander Neubeck et Luc Van Gool. *Efficient Non-Maximum Suppression*. International Conference on Pattern Recognition, vol. 3, pages 850–855, 2006. Cité page 76.
- [Neumuth 2006a] Thomas Neumuth, Nico Durstewitz, Milos Fischer, Gero Strauß, Andreas Dietz, Jürgen Meixensberger, Pierre Jannin, Kevin Cleary, Heinz U. Lemke et Oliver Burgert. *Structured recording of intraoperative surgical workflows*. The International Society for Optical Engineering, vol. 6145, 2006. Cité page 27.
- [Neumuth 2006b] Thomas Neumuth, Gero Strauß, Jürgen Meixensberger, Heinz U Lemke et Oliver Burgert. *Acquisition of Process Descriptions from Surgical Interventions*. International Conference on Database and Expert Systems Applications, vol. 4080, pages 602–611, 2006. Cité pages 27, 28, et 29.
- [Ní Choisdealbha 2011] Áine Ní Choisdealbha, Nuala Brady et Corrina Maguinness. *Differing roles for the dominant and non-dominant hands in the hand laterality task*. Experimental Brain Research, vol. 211, pages 73–85, 2011. Cité page 103.
- [Nishikawa 2001] A. Nishikawa, T. Hosoi, K. Koara, D. Negoro, A. Hikita, S. Asano, F. Miyazaki, M. Sekimoto, Y. Miyake et M. Yasui. *Real-time visual tracking of the surgeon’s face for laparoscopic surgery*. Medical Image Computing and Computer-Assisted Intervention, vol. 2208, pages 9–16, 2001. Cité page 113.
- [Nix 2010] J. Nix, A. Smith, R. Kurpad, M. E. Nielsen, E. M. Wallen et R. S. Pruthi. *Prospective Randomized Controlled Trial of Robotic versus Open Radical Cystectomy for Bladder Cancer : Perioperative and Pathologic Results*. European Urology, vol. 57, no. 2, pages 196–201, 2010. Cité page 10.

- [Nomm 2008] Sven Nomm, Eduard Petlenkov, Juri Vain, Juri Belikov, Fujio Miyawaki et Kitaro Yoshimitsu. *Recognition of the Surgeon's Motions During Endoscopic Operation by Statistics Based Algorithm and Neural Networks Based ANARX Models*. The International Federation of Automatic Control, pages 14773–14778, 2008. Cité pages 38, 39, 41, 46, et 51.
- [Novi 2012] Gianluca D. E. Novi, Ryan Bardsley, Rajesh Shah, Mark P. Ottensmeyer, John Cho Moore et Bummo Ahn. *Event-driven Surgical Gesture Segmentation and Task Recognition for Ocular Trauma Simulation*. International Conference on Intelligent Environments, pages 341–352, 2012. Cité pages 36, 39, 43, 46, 48, et 55.
- [OeSim 2015] OeSim. <http://www.eosurgical.com/collections/simulators-institutions>. 2015. Cité page 18.
- [Oropesa 2013] I. Oropesa, P. Sánchez-González, M. K. Chmarra, P. Lamata, Á. Fernández, J. A. Sánchez-Margallo, F. W. Jansen, J. Dankelman, F. M. Sánchez-Margallo et E. J. Gómez. *EVA : Laparoscopic instrument tracking based on endoscopic video analysis for psychomotor skills assessment*. Surgical Endoscopy and Other Interventional Techniques, vol. 27, no. 3, pages 1029–1039, 2013. Cité page 22.
- [Padoy 2011] Nicolas Padoy et Gregory D. Hager. *Human-Machine Collaborative Surgery Using Learned Models*. IEEE International Conference on Robotics and Automation, pages 5285–5292, 2011. Cité pages 38, 45, 46, 53, 103, et 145.
- [Padoy 2012] Nicolas Padoy, Tobias Blum, Seyed-Ahmad Ahmadi, Hubertus Feussner, Marie-Odile Berger et Nassir Navab. *Statistical modeling and recognition of surgical workflow*. Medical Image Analysis, vol. 16, no. 3, pages 632–641, 2012. Cité page 27.
- [Pandey 2006] V. A. Pandey, J. H. N. Wolfe, C. D. Liapis et D. Bergqvist. *The examination assessment of technical competence in vascular surgery*. British Journal of Surgery, vol. 93, no. 9, pages 1132–1138, 2006. Cité page 21.
- [Pauly 2014] Olivier Pauly, Benoît Diotte, Séverine Habert, Simon Weidert, Ekkehard Euler, Pascal Fallavollita et Nassir Navab. *Relevance-based Visualization to improve Surgeon Perception*. International Conference on Information Processing in Computer-Assisted Interventions, vol. 8498, pages 178–185, 2014. Cité page 145.
- [Penet 2013] Cédric Penet. *De l'indexation d'évènements dans des films - Application à la détection de violence*. PhD thesis, 2013. Cité page 33.
- [Perez 2012] Manuela Perez. *Chirurgie Robotique : de l'Apprentissage à l'Application*. PhD thesis, Université de Lorraine, 2012. Cité page 8.

- [Petitjean 2011] François Petitjean, Alain Ketterlin et Pierre Gançarski. *A global averaging method for dynamic time warping, with applications to clustering*. Pattern Recognition, vol. 44, no. 3, pages 678–693, 2011. Cité pages 102 et 145.
- [Petitjean 2014] François Petitjean, Germain Forestier, Geoffrey I Webb, Ann E Nicholson, Yanping Chen et Eamonn Keogh. *Dynamic Time Warping Averaging of Time Series allows Faster and more Accurate Classification*. IEEE International Conference on Data Mining, pages 470–479, 2014. Cité pages 102 et 145.
- [PoPTrainer 2015] PoPTrainer. <http://www.optimist.at/pop-trainer/>. 2015. Cité page 18.
- [Quelloc 2014] Gwénolé Quelloc, Mathieu Lamard, Béatrice Cochener et Guy Cazuguel. *Real-Time Segmentation and Recognition of Surgical Tasks in Cataract Surgery Videos*. IEEE Transactions on Medical Image Analysis, vol. 33, no. 12, pages 2352–2360, 2014. Cité page 31.
- [Reader 2006] T. Reader, R. Flin, K. Lauche et B. H. Cuthbertson. *Non-technical skills in the intensive care unit*. British Journal of Anaesthesia, vol. 96, no. 5, pages 551–559, 2006. Cité page 13.
- [RedDragon 2015] RedDragon. http://bionics.seas.ucla.edu/research/surgery_device_3.html. 2015. Cité page 18.
- [Reiley 2008] Carol E. Reiley, Henry C. Lin, Balakrishnan Varadarajan, Balazs Vagvolgyi, Ssanjeev Khudanpur, David D. Yuh et Gregory D. Hager. *Automatic Recognition of Surgical Motions Using Statistical Modeling for Capturing Variability*. Studies in health technology and informatics, vol. 132, no. 1, pages 396–401, 2008. Cité pages 31, 38, 44, 45, et 99.
- [Reiley 2009a] Carol E Reiley et Gregory D Hager. *Decomposition of Robotic Surgical Tasks : An Analysis of Subtasks and Their Correlation to Skill*. Modeling and Monitoring of Computer Assisted Interventions, 2009. Cité page 48.
- [Reiley 2009b] Carol E. Reiley et Gregory D. Hager. *Task versus Subtask Surgical Skill Evaluation of Robotic Minimally Invasive Surgery*. Medical Image Computing and Computer-Assisted Intervention, vol. 5761, pages 435–442, 2009. Cité pages 32 et 35.
- [Reiley 2010] Carol E. Reiley, Erion Plaku et Gregory D. Hager. *Motion generation of robotic surgical tasks : Learning from expert demonstrations*. IEEE Engineering in Medicine and Biology Society, pages 967–970, 2010. Cité pages 42, 45, 46, 48, 51, et 99.
- [Reiley 2011] C. E. Reiley, H. C. Lin, D. D. Yuh et G. D. Hager. *Review of methods for objective surgical skill evaluation*. Surgical Endoscopy, vol. 25, no. 2, pages 356–366, 2011. Cité pages 14 et 133.

- [Reznick 1993] R. K. Reznick. *Teaching and testing technical skills*. American journal of surgery, vol. 165, no. 3, pages 358–361, 1993. Cité page 13.
- [RobMentor 2015] RobMentor. <http://symbionix.com/simulators/robotix-mentor/>. 2015. Cité page 18.
- [Rosen 2001] J. Rosen, B. Hannaford, C. G. Richards et M. N. Sinanan. *Markov Modeling of Minimally Invasive Surgery Based on Tool/Tissue Interaction and Force/Torque Signatures for Evaluating Surgical Skills*. IEEE Transactions on Biomedical Engineering, vol. 48, no. 5, pages 579–591, 2001. Cité pages 36, 38, 39, 40, 45, 46, 48, et 52.
- [Rosen 2002] Jacob Rosen, Massimiliano Solazzo, Blake Hannaford et Mika Sinanan. *Task Decomposition of Laparoscopic Surgery for Objective Evaluation of Surgical Residents’s Learning Curve Using Hidden Markov Model*. Computer Aided Surgery, vol. 7, no. 1, pages 49–61, 2002. Cité pages 36, 38, 39, 40, 45, 46, 48, et 52.
- [Rosen 2006] Jacob Rosen, Jeffrey D. Brown, Lily Chang, Mika N. Sinanan et Blake Hannaford. *Generalized Approach for Modeling Minimally Invasive Surgery as a Stochastic Process Using a Discrete Markov Model*. IEEE Transactions on Biomedical Engineering, vol. 53, no. 3, pages 399–413, 2006. Cité pages 36, 38, 39, 41, 46, et 48.
- [RoSSII 2015] RoSSII. <http://www.simulatedsurgeries.com/ross2.html>. 2015. Cité page 18.
- [Sackier 1994] J.M. Sackier et Y. Wang. *Robotically assisted laparoscopic surgery : From concept to development*. Surgical Endoscopy, vol. 8, no. 8, pages 63–66, 1994. Cité pages 8 et 108.
- [Saggio 2011] G. Saggio, G. L. Santosuosso, P. Cavallo, C. A. Pinto, M. Petrella, F. Giannini, N. Di Lorenzo, A. Lazzaro, A. Corona, F. D’Auria, L. Iezzi et A. L. Gaspari. *Gesture recognition and classification for surgical skill assessment*. IEEE International Workshop on Medical Measurements and Applications, pages 662–666, 2011. Cité pages 39, 40, 41, et 45.
- [Sakoe 1978] Hiroaki Sakoe et Seibi Chiba. *Dynamic Programming Algorithm Optimization for Spoken Word Recognition*. IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 26, no. 1, pages 43–49, 1978. Cité page 73.
- [Samur 2012] Evren Samur. Performance Metrics for Haptic Interfaces. 2012. Cité page 144.
- [Sankararaman 2013] S. Sankararaman, P. K. Agarwal, T. Molhave et A. P. Boedihardjo. *Computing Similarity between a Pair of Trajectories*. Computational Geometry, pages 1–18, 2013. Cité page 73.

- [Santos-Carreras 2012a] L. Santos-Carreras, M. Hagen, R. Gassert et H. Bleuler. *Survey on Surgical Instrument Handle Design : Ergonomics and Acceptance*. *Surgical Innovation*, vol. 19, no. 1, pages 50–59, 2012. Cité pages 7 et 114.
- [Santos-Carreras 2012b] Laura Santos-Carreras. *Increasing Haptic Fidelity and Ergonomics in Teleoperated Surgery*. PhD thesis, École Polytechnique Fédérale de Lausanne, 2012. Cité page 109.
- [Schulz 2010] Sebastian Schulz et Annika Woerner. *Automatic Motion Segmentation for Human Motion Synthesis*. *International Conference on Articulated Motion and Deformable Objects*, vol. 6169, pages 182–191, 2010. Cité page 69.
- [Secin 2010] F. P. Secin, C. Savage, C. Abbou, A. Taille, J. Rassweiler, M. Hruza, F. Rozet, X. Cathelineau, F. Nassar, I. Turk, A. J. Vanni, I. S. Gill, P. Koenig, Ji. H. Kaouk, L. M. Pineiro, V. Pansadoro, P. Emiliozzi, A. Bjartell, T. Jiborn, C. Eden, A. J. Richards, R. V. Velthoven, R. Rabenalt, Li-ming Su, C. P. Pavlovich, A.W. Levinson, K.A. Touijer, A. Vickers et B. Guillonneau. *The Learning Curve for Laparoscopic Radical Prostatectomy : An International Multicenter Study*. *The Journal of Urology*, vol. 184, no. 6, pages 2291–2296, 2010. Cité page 7.
- [SEPBASIC 2015] SEPBASIC. <http://www.simsurgery.com/basic.html>. 2015. Cité page 18.
- [SEPROBOT 2015] SEPROBOT. <http://www.simsurgery.com/robot.html>. 2015. Cité page 18.
- [Shokoohi-Yekta 2015] Mohammad Shokoohi-Yekta, Bing Hu, Hongxia Jin, Jun Wang et Eamonn Keogh. *Generalizing Dynamic Time Warping to the Multi-Dimensional Case Requires an Adaptive Approach*. *Data Mining and Knowledge Discovery*, 2015. Cité pages 76 et 101.
- [Simorov 2012] Anton Simorov, R. Stephen Otte, Courtney M. Kopietz et Dmitry Oleynikov. *Review of surgical robotics user interface : what is the best way to control robotic surgery?* *Surgical Endoscopy*, vol. 26, no. 8, pages 2117–2125, 2012. Cité page 109.
- [SimPro 2015] SimPro. <http://www.simendo.eu/services/laparoscopy-training/>. 2015. Cité page 18.
- [Spaner 1997] S.J. Spaner et G. L. Warnock. *A Brief History of Endoscopy, Laparoscopy, and Laparoscopic Surgery*. *Journal of Laparoendoscopic and Advanced Surgical Techniques*, vol. 7, no. 6, pages 369–373, 1997. Cité page 6.
- [Stauder 2014] Ralf Stauder, Asli Okur, Loic Peter, Armin Schneider, Michael Kranzfelder, Hubertus Feussner et Nassir Navab. *Random Forests for Phase Detection in Surgical Workflow Analysis*. *International Conference on Information Processing in Computer-Assisted Interventions*, vol. 8498, pages 148–157, 2014. Cité page 27.

- [Steinberg 2008] P. L. Steinberg, P. A. Merguerian, W. Bihrlé, J. A. Heaney et J. D. Seigne. *A da Vinci Robot System Can Make Sense for a Mature Laparoscopic Prostatectomy Program*. Journal of the Society of Laparoendoscopic Surgeons, vol. 12, no. 1, pages 9–12, 2008. Cité page 10.
- [SurgWork 2015] SurgWork. http://www.msrg.org/medical_simulation_equipment/79.htm. 2015. Cité page 18.
- [Tang 2006] B. Tang, G. B. Hanna, F. Carter, G. D. Adamson, J. P. Martindale et A. Cuschieri. *Competence assessment of laparoscopic operative and cognitive skills: Objective Structured Clinical Examination (OSCE) or Observational Clinical Human Reliability Assessment (OCHRA)*. World Journal of Surgery, vol. 30, no. 4, pages 527–534, 2006. Cité page 20.
- [Tao 2012] Lingling Tao, Ehsan Elhamifar, Sanjeev Khudanpur, Gregory D. Hager et René Vidal. *Sparse Hidden Markov Models for Surgical Gesture Classification and Skill Evaluation*. Information Processing in Computer-Assisted Interventions, vol. 7330, pages 167–177, 2012. Cité pages 41, 46, 49, 54, 99, et 103.
- [Tao 2013] Lingling Tao, Luca Zappella, Gregory D. Hager et René Vidal. *Surgical Gesture Segmentation and Recognition*. Medical Image Computing and Computer-Assisted Intervention, vol. 8151, pages 339–346, 2013. Cité pages 40, 41, 42, 43, 44, 48, 49, 51, 54, 99, et 103.
- [Ten Holt 2007] G A Ten Holt, M J T Reinders et E A Hendriks. *Multi-Dimensional Dynamic Time Warping for Gesture Recognition*. Annual Conference of the Advanced School for Computing and Imaging, 2007. Cité pages 76 et 101.
- [Tobergte 2011] Andreas Tobergte, Patrick Helmer, Ulrich Hagn, Patrice Rouiller, Sophie Thielmann, Sébastien Grange, Alin Albu-Schäffer, François Conti et Gerd Hirzinger. *The sigma.7 haptic interface for MiroSurge : A new bi-manual surgical console*. IEEE International Conference on Intelligent Robots and Systems, pages 3023–3030, 2011. Cité page 109.
- [Tokunaga 2012] M. Tokunaga, H. Egi, M. Hattori, M. Yoshimitsu, T. Kawahara, M. Okajima et H. Ohdan. *Approaching time is important for assessment of endoscopic surgical skills*. Minimally Invasive Therapy & Allied Technologies, vol. 21, no. 3, pages 142–149, 2012. Cité page 22.
- [Trelhu 2011] Brivaël Trelhu. *Modélisation, analyse et comparaison de procédures chirurgicales, application à la neurochirurgie*. PhD thesis, 2011. Cité page 29.
- [Troccaz 2012] Jocelyne Troccaz. Medical Robotics. 2012. Cité page 8.

- [Tserie 2015] Tserie. <http://www.3-dmed.com/product/minimally-invasive-training-system-mits-series>. 2015. Cité page 18.
- [Tsuda 2009] S. Tsuda, D. Scott, Je. Doyle et D. B. Jones. *Surgical Skills Training and Simulation*. Current problems in surgery, vol. 46, no. 4, pages 271–370, 2009. Cité pages 7, 16, et 24.
- [Uemura 2014] M. Uemura, M. Tomikawa, R. Kumashiro, T. Miao, R. Souzaki, S. Ieiri, K. Ohuchida, A. T. Lefor et M. Hashizume. *Analysis of hand motion differentiates expert and novice surgeons*. Journal of Surgical Research, vol. 188, no. 1, pages 8–13, 2014. Cité pages 22, 24, 27, et 35.
- [Van Den Berg 2010] Jur Van Den Berg, Stephen Miller, Daniel Duckworth, Humphrey Hu, Andrew Wan, Xiao Yu Fu, Ken Goldberg et Pieter Abbeel. *Superhuman performance of surgical tasks by robots using iterative learning from human-guided demonstrations*. IEEE International Conference on Robotics and Automation, pages 2074–2081, 2010. Cité pages 39, 42, 45, et 49.
- [Van Der Schatte Olivier 2009] R. H. Van Der Schatte Olivier, C. D. P. Van't Hullenaar, J. P. Ruurda et I. A. M. J. Broeders. *Ergonomics, user comfort, and performance in standard and robot-assisted laparoscopic surgery*. Surgical Endoscopy and Other Interventional Techniques, vol. 23, no. 6, pages 1365–1371, 2009. Cité page 10.
- [van Velthoven 2006] F. R. van Velthoven et P. Hoffmann. *Methods for laparoscopic training using animal models*. Current Urology Reports, vol. 7, no. 2, pages 114–119, 2006. Cité page 15.
- [Varadarajan 2009] Balakrishnan Varadarajan, Carol Reiley, Henry Lin, Sanjeev Khudanpur et Gregory Hager. *Data-Derived Models for Segmentation with Application to Surgical Assessment and Training*. Medical Image Computing and Computer-Assisted Intervention, vol. 5761, no. 1, pages 426–434, 2009. Cité pages 32, 38, 42, 45, 48, 49, 50, 51, 53, et 55.
- [Vassiliou 2005] M. C. Vassiliou, L. S. Feldman, C. G. Andrew, S. Bergman, K. Leffondré, D. Stanbridge et G. M. Fried. *A global assessment tool for evaluation of intraoperative laparoscopic skills*. American Journal of Surgery, vol. 190, no. 1, pages 107–113, 2005. Cité page 20.
- [Wang 2013] Xiaoyue Wang, Abdullah Mueen, Hui Ding, Goce Trajcevski, Peter Scheuermann et Eamonn Keogh. *Experimental comparison of representation methods and distance measures for time series data*. Data Mining and Knowledge Discovery, vol. 26, no. 2, pages 275–309, 2013. Cité pages 71 et 100.

- [Weinkauff 2010] Tino Weinkauff, Yotam Gingold et Olga Sorkine. *Topology-based Smoothing of 2D Scalar Fields with C1-Continuity*. Eurographics Conference on Visualization, vol. 29, no. 3, pages 1221–1230, 2010. Cité page 70.
- [Welch 2001] Greg Welch et Gary Bishop. *An Introduction to the Kalman Filter*. In Special Interest Group in Graphics, volume 8, pages 1–81. 2001. Cité page 129.
- [Wilson 1997] M. S. Wilson, R. Stone, A. Middlebrook, R. F. McCloy et C. Sutton. *MIST VR : a virtual reality trainer for laparoscopic surgery assesses performance*. Annals of the Royal College of Surgeons of England, vol. 79, pages 403–404, 1997. Cité page 18.
- [Wolf 2013] R. Wolf. *Quantification de la qualité d'un geste chirurgical à partir de connaissances a priori*. PhD thesis, Université de Grenoble, 2013. Cité pages 19 et 23.
- [Wu 2009] Shandong D. Wu et Y. F. Li. *Flexible signature descriptions for adaptive motion trajectory representation, perception and recognition*. Pattern Recognition, vol. 42, pages 194–214, 2009. Cité page 65.
- [Xiong 2006] Yingen Xiong et Francis Quek. *Hand Motion Gesture Frequency Properties and Multimodal Discourse Analysis*. International Journal of Computer Vision, vol. 69, no. 3, pages 353–371, 2006. Cité pages 41 et 67.
- [Xperience 2015] Xperience. <http://www.mimicsimulation.com/products/xperience/>. 2015. Cité page 18.
- [Xron 2015] Xron. <http://www.bbzsrl.com/xron>. 2015. Cité page 18.
- [Yang 2013] Sungwook Yang, Trent S. Wells, Robert A. Maclachlan et Cameron N. Riviere. *Performance of a 6-Degree-of-Freedom Active Microsurgical Manipulator in Hand-held Tasks*. IEEE International Conference on Engineering in Medicine and Biology Society, pages 5670–5673, 2013. Cité pages 67 et 130.
- [Yule 2006a] S. Yule, R. Flin, S. Paterson-Brown et N. Maran. *Non-technical skills for surgeons in the operating room : A review of the literature*. Surgery, vol. 139, no. 2, pages 140–149, 2006. Cité page 13.
- [Yule 2006b] S. Yule, R. Flin, S. Paterson-Brown, N. Maran et D. Rowley. *Development of a rating system for surgeons' non-technical skills*. Medical Education, vol. 40, no. 11, pages 1098–1104, 2006. Cité page 13.
- [Zappella 2013] Luca Zappella, Benjamín Béjar, Gregory Hager et René Vidal. *Surgical gesture classification from video and kinematic data*. Medical Image Analysis, vol. 17, no. 7, pages 732–745, 2013. Cité pages 40, 41, 42, 44, 45, 48, 49, 52, 54, 84, 99, et 103.

- [Zhou 2012] Feng Zhou et Fernando De La Torre. *Generalized Time Warping for Multimodal Alignment of Human Motion*. IEEE Conference on Computer Vision and Pattern Recognition, pages 1282–1289, 2012. Cité page 73.
- [Zhou 2014] Tian Zhou, Maria E. Cabrera et Juan P. Wachs. *Touchless telerobotic surgery : A comparative study*. International Conference on Intelligent Robots and Systems, Telerobotics Workshop, 2014. Cité page 117.
- [Zhou 2015a] Tian Zhou, Maria E. Cabrera et Juan P. Wachs. *Touchless Telerobotic Surgery : Is It Possible at All?* AAI Conference on Artificial Intelligence, pages 4228–4229, 2015. Cité page 117.
- [Zhou 2015b] Yun Zhou, Ioanna Ioannou, James Bailey Sudanthi Wijewickrema, Gregor Kennedy et Stephen O’Leary. *Automated segmentation of surgical motion for performance analysis and feedback*. Medical Image Computing and Computer-Assisted Intervention, 2015. Cité page 99.

Abstract

Integration of robotic systems in the operating room changed the way that surgeries are performed. It modifies practices to improve medical benefits for the patient but also brought non-traditional aspects that can lead to serious adverse events. Recent studies from the French authorities for hygiene and medical care highlight that these events mainly come from surgeons' technical skills, which question surgical robotic training and teaching. To overcome this issue, surgical simulators help to train practitioner through different training tasks and provide feedback to the operator. However the feedback is partial and do not help the surgeon to understand gestural mistakes. Thus, we want to improve the surgical robotic training conditions. The objective of this work is twofold. First, we developed a new method for segmentation and recognition of surgical gestures during training sessions based on an unsupervised approach. From surgical tools kinematic data, we are able to achieve gesture recognition at 82%. This method is a first step to evaluate technical skills based on gestures and not only on the global execution of the task as it is done nowadays. The second objective is to provide easier access to surgical training and make it cheaper. To do so, we studied a new contactless human-machine interface to control surgical robots. In this work, the interface is plugged to a Raven-II robot dedicated to surgical robotics research. Then, we evaluated the performance of such system through multiple studies, concluding that this interface can be used to control surgical robots. In the end, one can consider to use this contactless interface for surgical training with a simulator. It can reduce the training cost and also improve the access for novice surgeons to technical skills training dedicated to surgical robotics.

Keywords: *Medical robotics, Teleoperation, Human-machine interface, Machine learning, Surgical gesture, Laparoscopy.*

Résumé

L'intégration de systèmes robotiques au sein du bloc opératoire a modifié le déroulement de certaines interventions, laissant ainsi place à des pratiques favorisant le bénéfice médical rendu au patient en dépit des aspects conventionnels. Dans ce cadre, de récentes études de la Haute Autorité de Santé ont mis en avant les effets indésirables graves intervenant au cours des procédures chirurgicales robotisées. Ces erreurs, majoritairement dues aux compétences techniques du praticien, remettent ainsi en cause la formation et les techniques d'apprentissage pour la chirurgie robotisée. Bien que l'utilisation abondante de simulateurs facilite cet apprentissage à travers différents types d'entraînement, le retour fourni à l'opérateur reste succinct et ne lui permet pas de progresser dans de bonnes conditions. De ce fait, nous souhaitons améliorer les conditions d'entraînement en chirurgie laparoscopique robotisée. Les objectifs de cette thèse sont doubles. En premier lieu, ils visent le développement d'une méthode pour la segmentation et la reconnaissance des gestes chirurgicaux durant l'entraînement en se basant sur une approche non-supervisée. Utilisant les données cinématiques des instruments chirurgicaux, nous sommes capables de reconnaître les gestes réalisés par l'opérateur à hauteur de 82%. Cette méthode est alors une première étape pour l'évaluation de compétences basée sur la gestuelle et non sur l'ensemble de la tâche d'entraînement. D'autre part, nous souhaitons rendre l'entraînement en chirurgie robotisée plus accessible et moins coûteux. De ce fait, nous avons également étudié l'utilisation d'une nouvelle interface homme-machine sans contact pour la commande des robots chirurgicaux. Dans ces travaux, cette interface a été couplée au Raven-II, un robot de téléopération pour la recherche en robotique chirurgicale. Nous avons évalué les performances du système à travers différentes études, concluant ainsi à la possibilité de téléopérer un robot chirurgical avec ce type de dispositif. Il est donc envisageable d'utiliser cette interface pour l'entraînement sur simulateur afin de réduire le coût de la formation, mais également d'améliorer l'accès et l'acquisition des compétences techniques spécifiques à la chirurgie robotisée.

Mots clefs : *Robotique médicale, Téléopération, Interface homme-machine, Apprentissage automatique, Geste chirurgical, Laparoscopie.*