

Analyse de vidéos de microscopes chirurgicaux pour la reconnaissance automatique d'étapes en combinant SVM et HMM

Surgical microscope video analysis for automatic recognition of phases by combining SVM and HMM

Florent Lalys^{1,2,3}, Laurent Riffaud^{1,2,3,4}, Xavier Morandi^{1,2,3,4}, Pierre Jannin^{1,2,3}

¹ INSERM, U746, Faculty of Medicine CS 34317, F-35043 Rennes, France

² INRIA, VisAGeS Unit/Project, F-35042 Rennes, France

³ University of Rennes I, CNRS, UMR 6074, IRISA, F-35042 Rennes, France

⁴ Department of Neurosurgery, Pontchaillou University Hospital, F-35043 Rennes, France

Résumé

Dans le vaste domaine de la chirurgie guidée par l'image, l'extraction automatique d'informations émanant d'une salle d'opération a récemment été source de nombreuses études. Plus particulièrement, la reconnaissance automatique des étapes, ou plus globalement des événements, d'une procédure chirurgicale pourrait permettre d'apporter une aide intra ou post-opératoire non-négligeable aux chirurgiens. Les applications visées par ces approches sont la création automatique de rapports post-opératoires, l'enseignement, l'évaluation, l'optimisation de la gestion des salles d'opérations ou l'aide au geste chirurgical. Dans cet article, nous proposons une approche qui tend vers la reconnaissance automatique des étapes d'une chirurgie via les vidéos issues des microscopes chirurgicaux, ce qui n'a encore jamais été expérimenté. Notre méthode est basée sur la combinaison d'une classification supervisée permettant d'extraire des attributs propres à chaque chirurgie, et d'une modélisation de séries temporelles par chaîne de Markov cachée. Des validations croisées ont été menées sur deux types distincts de chirurgie : une chirurgie du cerveau et une chirurgie ophtalmologique. Des taux de reconnaissance respectifs de 90% et 93% ont ainsi été trouvés, ouvrant la possibilité à de nombreuses applications dans le domaine de la chirurgie guidée par l'image, comme l'indexation de vidéos chirurgicales.

Mots Clef

Chirurgie guidée par l'image, analyse de vidéos, microscopes chirurgicaux

Abstract

In image-guided surgery, the automatic extraction of information from the Operating Room (OR) has recently gained much interest. In particular, the automatic recognition of surgical phases, or more generally surgical events, allows bringing an additional help to surgeons, for intra or post-operative use. Clinical applications include the creation of post-operative reports, teaching, learning, surgical assessment, optimisation of OR management and of surgeries. In this article, we present a novel approach that focused on the automatic recognition of phases by microscope image analysis, which has never been done before. We used a hybrid method that combines supervised classification to extract binary visual cues and a discrete Hidden Markov Model to take into account the temporal aspect. Our framework was tested on two various datasets, including one specific type of neurosurgical intervention and one type of ophthalmological surgery. Cross-validation studies were carried out to find recognition rates of 90% for the first dataset and 93% for the second one. This will enable the system to be used in clinical applications such as post-operative surgical video indexation.

Keywords

Image-guided surgery, video analyse, surgical microscope

1 Introduction

1.1 Contexte scientifique

Devant l'apparition de nombreux outils et nouvelles technologies dans les salles d'opérations, le besoin de nouveaux systèmes assistés par ordinateur se fait actuellement ressentir. En phase pré-opératoire, ces systèmes fournissent un accès aux images multimodales, aux informations propres aux patients et aux différentes étapes de la simulation chirurgicale. Pendant la chirurgie, ils fournissent une interface de visualisation intégrant ces différentes données dans un contexte opératoire, mais aussi une aide active ou passive au geste chirurgical. Ils permettent donc d'améliorer l'ergonomie de la salle d'opération, d'augmenter la sécurité médicale, d'optimiser le temps opératoire ou encore d'aider les prises de décision. Ces systèmes assistés par ordinateur doivent se baser sur des modèles capables de suivre le déroulé de l'intervention pour adapter l'assistance en conséquence (Cleary et al., 2005). En d'autres termes, un des challenges de la chirurgie guidée par l'image est d'assister la procédure à travers la compréhension des événements, au sens large, de la salle d'opération. La reconnaissance en temps-réel de ces événements peut permettre d'apporter une aide non-négligeable aux experts. Les applications visées par ces approches sont la création automatique de rapports post-opératoires, l'enseignement, l'évaluation, l'optimisation de la gestion des salles d'opérations ou l'aide au geste chirurgical.

1.2 État de l'art

Typiquement, une procédure chirurgicale d'un même type est reproductible, et cette hypothèse rend possible une modélisation temporelle des procédures, dont le but est de collecter des données et de dériver des modèles issus de ces données (Jannin et al., 2007). Dans la littérature, les informations pertinentes servant de base à ces modélisations sont acquises selon deux stratégies différentes : soit par des opérateurs humains positionnés dans la salle d'opération, soit par des capteurs de façon automatique. Les approches basées opérateur humain ont la capacité de couvrir des niveaux de granularité supérieurs aux approches basées capteurs en incluant des informations sémantiques, mais ces méthodes présentent la limite de ne pas être automatisées, et donc de nécessiter beaucoup de temps et de ressources humaines. C'est pourquoi les approches basées capteurs sont de plus en plus étudiées pour automatiser ce processus d'acquisition de données. Au sein de ces approches, on retrouve les études utilisant des simulateurs ou environnements virtuels complets (Darzi et al., 2002 ; Lin et al., 2006), pour étudier la gestuelle des chirurgiens et créer des modèles de reconnaissance de gestes. On retrouve également les systèmes de reconnaissance basés sur des

capteurs installés sur les instruments (Ahmadi et al., 2007 ; Padoy et al., 2007). Des informations binaires de présence d'instruments sont analysées et des modèles graphiques probabilistes sont utilisés pour reconnaître des informations de haut niveau dans la salle d'opération.

D'autres systèmes ont aussi été testés, comme la mise en place d'un outil de suivi du regard du chirurgien (James et al., 2007), ou d'un système de tracking 3D de la position de chaque membre du staff (Nara et al., 2009). Le principal défaut de tous ces systèmes est qu'ils ne sont pas installés d'office dans les salles d'opération, et que la mise en place de nombreux outils pourrait à long terme gêner le déroulé de l'intervention. Pour pallier à ce problème, les nouvelles études se focalisent sur des sources d'information déjà installées dans la salle d'opération, telles que les signes vitaux des patients (Xiao et al., 2005) ou les vidéos des endoscopes (Speidel et al., 2008 ; Klank et al., 2008 ; Lo et al., 2003). Dans cette dernière catégorie, des outils de vision par ordinateur et de réalité augmentée sont utilisés pour extraire des informations pertinentes au chirurgien. Les images vidéos des différentes caméras installées dans une salle d'opération (endoscope, microscope) se révèlent donc être une source riche en information pouvant éventuellement remplacer les approches se basant sur des enregistrements humains.

1.3 Notre projet

Dans cet article, nous présentons une approche pour la reconnaissance automatique d'information haut-niveau d'une chirurgie via les vidéos issues des microscopes chirurgicaux, ce qui n'a encore jamais été testé. Nous proposons de combiner des méthodes de classification dites supervisées avec des modèles graphiques probabilistes pour reconnaître automatiquement les principales étapes d'une chirurgie. Nous validons ensuite notre méthode sur deux procédures différentes : un jeu de données de vidéos de neurochirurgie, et un jeu de données de vidéos de chirurgie ophtalmologique.

2 Matériels et méthodes

Le processus complet de reconnaissance est présenté sur la Fig. 1. La première étape de classification supervisée est utile pour extraire des attributs spécifiques à chaque chirurgie. Pour cela une extraction suivie d'une sélection de caractéristiques images (texture, forme et couleur) sont employées. Couplés à une base de données d'images préalablement créée, différents classifieurs sont ensuite testés pour reconnaître les attributs pertinents de la chirurgie. Puis ces informations sont introduites dans une Chaîne de Markov Cachée (CMC) qui va permettre de modéliser l'aspect temporel de la chirurgie. En sortie, l'algorithme de Viterbi nous permet d'obtenir des séquences correspondant aux différentes étapes

chirurgicales. Les paragraphes suivants présentent en détail toutes les étapes de ce processus de reconnaissance automatique.

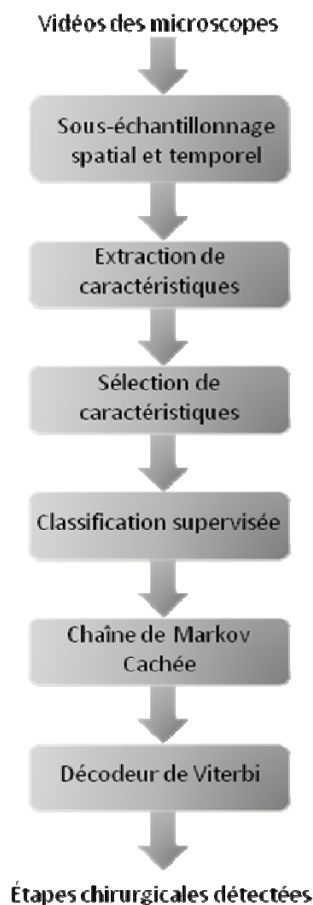


Fig. 1. Processus de reconnaissance automatique des étapes de la chirurgie

2.1 Extraction de caractéristique

Pour chaque image, une signature a été extraite, composée de caractéristiques de texture, de forme et de couleur, dans le but de chercher des similarités entre les images de même classe. La couleur a été extraite avec deux espaces complémentaires (Smeulders et al. 2000), l'espace RVB (Rouge Vert Bleu) et l'espace TSV (Teinte Saturation Valeur). Pour la texture, nous avons opté pour les matrices de co-occurrences associées aux descripteurs d'Haralick (Haralick et al., 1973). La forme a été obtenue à partir des moments spatiaux (Hu, 1962). Enfin, les coefficients de la Transformée en Cosinus Discrète (TCD) ont été calculés (Ahmed et al., 1974). Après un sous-échantillonnage spatial par un facteur de 8 (noyau Gaussien), et une normalisation statistique (centrée-

réduite), une étape de sélection de caractéristiques a été réalisée.

2.2 Sélection de caractéristiques

Les signatures d'images étant composées de 185 caractéristiques, nous avons utilisé une méthode hybride de sélection combinant une approche dit *filter* d'une approche dit *wrapper* (Duda and Hart, 1973). La première exécute la sélection en regardant les propriétés intrinsèques des données. La deuxième utilise un classifieur en sortie pour évaluer la meilleure combinaison de variables d'entrée. Ces deux types d'approches sont complémentaires et peuvent être fusionnées par l'intersection des résultats des deux algorithmes. L'information mutuelle (Hamming, 1980) et l'algorithme SVM Recursive Feature Elimination (RFE) furent choisis pour chaque approche. En parallèle, une simple Analyse en Composante Principale (ACP) fut aussi testée.

2.3 Extraction des attributs visuels

L'objectif de cette étape est d'extraire, en utilisant les caractéristiques image précédemment sélectionnées, des informations visuelles pouvant discriminer les différentes étapes de la chirurgie. Ces informations binaires sont considérées comme des attributs propres à chaque chirurgie. Il a été demandé aux chirurgiens de définir différents attributs visuels pouvant être *a priori* identifiables à travers une analyse d'images. Ces attributs peuvent être, par exemple, la présence/absence d'un objet, le zoom du microscope ou encore une couleur particulière dans la scène chirurgicale. Pour le processus de reconnaissance, plusieurs techniques classiques de classifications ont été testées : Machine à Vecteur de Support (SVM), Plus Proche Voisin (PPV), Réseaux neurones (RN), arbre de décision et Analyse Discriminante Linéaire (ADL).

2.4 Chaîne de Markov Cachée (CMC)

Les modèles graphiques probabilistes sont souvent utilisés pour décrire des dépendances entre des données d'observations dans des domaines tels que la biologie. Notamment, les réseaux bayésiens (RB) ont récemment prouvé leur utilité dans ces applications. Les CMC, exemples particuliers de RB, peuvent être utilisées pour modéliser des séries temporelles. Ici, nous utilisons une CMC (Rabiner, 1989) du premier ordre pour modéliser le déroulé de l'intervention. Mathématiquement, une CMC est définie par un 5-uplet (S, O, Π, A, B) , où $S = \{s_1 \dots (s_N)\}$ est un jeu fini de N états, $O = \{o_1 \dots (o_M)\}$ est un jeu de M symboles dans un

vocabulaire, $\Pi = \{\pi(i)\}$ sont les probabilités d'états initiales, $A = \{a(ij)\}$ sont les probabilités de transitions et $B = \{b_i(o(k))\}$ les probabilités de sortie. Dans notre approche, les attributs visuels détectés par la phase de classification supervisée qui précède sont utilisés comme observations pour la CMC. Ensuite, les différents paramètres du modèle sont déterminés de façon *ad hoc*, après avoir testé l'algorithme de Baum-Welch sans succès, principalement dû au manque de données d'apprentissage. Enfin, l'algorithme de Viterbi (Viterbi, 1969) trouve la séquence d'états la plus probable en sortie.

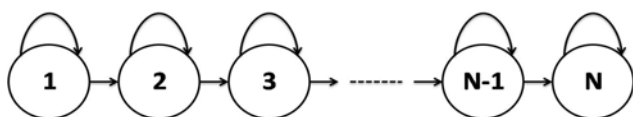


Fig. 2. CMC gauche-droite, où chaque état correspond à une étape de la chirurgie

2.5 Études de validation

Nous avons validé une à une chaque partie de notre méthode, à travers des validations croisées 10-fold. La première étude sert à sélectionner la meilleure méthode de sélection de caractéristiques ainsi que le nombre optimal de caractéristiques à garder. La deuxième étude a permis de comparer les différents classifieurs pour garder le plus adapté. Enfin, au sein d'une dernière étude, nous avons validé le processus global de reconnaissance, comprenant la détection des attributs couplée à la modélisation temporelle. Pour cela nous avons calculé le pourcentage d'étapes correctement reconnues.

2.6 Jeux de données

Nous avons évalué notre méthode sur deux jeux de données différents.

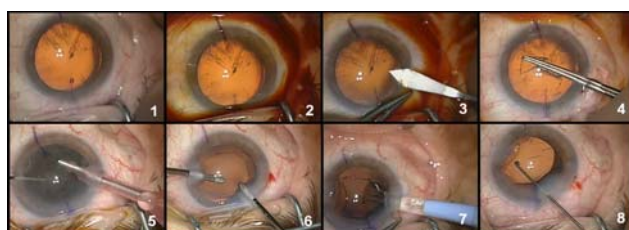
La première chirurgie étudiée fut les adénomes hypophysaires (jeu de données n°1), qui est un type particulier de neurochirurgie. Une voie d'abord transnasale est utilisée par les chirurgiens, qui atteignent ensuite l'hypophyse pour enlever la tumeur située dans cette région. Nous disposons de 16 vidéos (temps moyen de chirurgie : 50min). Le chirurgien a ainsi défini 6 étapes (Fig. 3-a) et 4 attributs visuels : le zoom du microscope, la présence des écarteurs nasaux, la présence de la cloison nasale et des compresses. La combinaison de ces quatre attributs binaires permet de discriminer les 6 étapes.

La chirurgie de la cataracte, qui est un type de chirurgie ophtalmologique, fut aussi étudiée (jeu de données n°2). Le principe est d'enlever la lentille naturelle de l'œil (la cataracte) pour la remplacer par une lentille artificielle.

Nous disposons de 18 vidéos (temps moyen de chirurgie: 15min). Huit étapes furent identifiées (Fig. 3-b) ainsi que 4 attributs visuels : la couleur de l'iris (rouge ou noir), la présence de l'antiseptique, la présence du couteau et de l'instrument pour implanter la lentille.



a)



b)

Fig. 3. Exemple d'images des microscopes pour les deux jeux de données :

- a) Chirurgie hypophysaire : 1-incision nasale, 2-installation des écarteurs nasaux, 3-exérèse de la tumeur, 4-remise en place de la cloison nasale, 5-suture, 6-installation des compresses nasales
- b) Chirurgie de la cataracte : 1-préparation, 2-injection de Bétadine, 3-incision de la cornée, 4-hydrodissection, 5-phacoemulsification, 6-aspiration corticale, 7-implantation de la lentille artificielle, 8-ajustement de la lentille

3 Résultats

La Fig. 4. montre que l'ACP est mieux adaptée pour un nombre de caractéristiques inférieur à 30. A partir de ce seuil, la méthode hybride de sélection donne de meilleurs résultats et atteint son maximum pour 40 caractéristiques. Ces résultats sont valables pour les deux classifieurs SVM et PPV. Lorsque l'ACP est utilisée avec un SVM, la précision diminue à partir de 40 caractéristiques, alors qu'en utilisant le même classifieur, mais l'autre approche de réduction de dimension, la précision reste pratiquement inchangée.

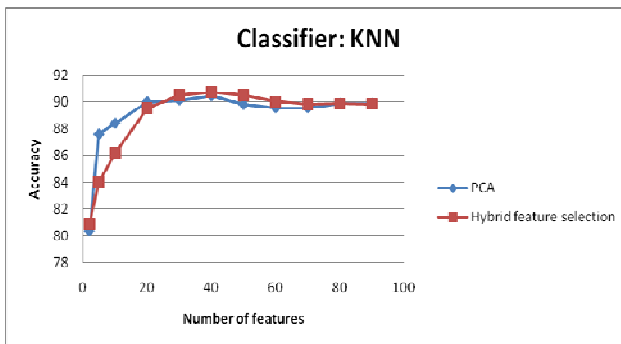
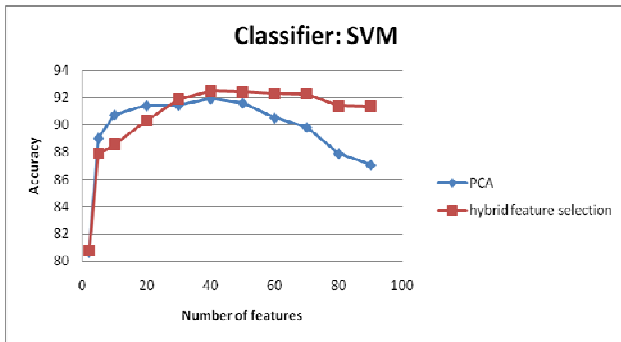


Fig. 4. Précision de la reconnaissance des attributs visuels pour 2 classifieurs, avec deux méthodes différentes de réduction de dimensionnalité (ACP et méthode hybride de sélection de caractéristiques)

Avec 40 caractéristiques, les autres classifieurs sont testés sur le Tab. 1. Les SVMs donnent les meilleurs résultats (91.5%), suivis par l'ADL et les PPV. En revanche, les arbres de décision et les réseaux de neurones donnent les moins bons résultats.

Tab. 1. Précision, sensibilité et sensibilité des différents algorithmes pour la reconnaissance des attributs visuels

Algorithmes	Précision (%)
SVM	91.5
PPV	90.2
RN	87.3
Arbre de décision	84.7
ADL	90.2

En utilisant les meilleurs paramètres (méthode hybride de sélection de caractéristiques + SVM), la précision de la classification des attributs visuels fut évaluée (Tab. 2.). Chaque détection d'attribut a une précision de reconnaissance supérieure à 88%, et nous voyons aussi que le jeu de données n°2 possède de meilleurs taux de reconnaissance que le jeu de données n°1, avec un maximum de 95.5% pour la reconnaissance du couteau.

Tab. 2. Précision de la reconnaissance des attributs visuels.

- a) Jeu de données n°1
b) Jeu de données n°2

	Vue Global/Zoom	Présence rétracteurs nasaux	Présence cloison nasale	Présence compresseurs
Précision (std) (%)	88.9 (2.2)	89.4 (1.1)	94.8 (1.3)	88.3 (1.6)

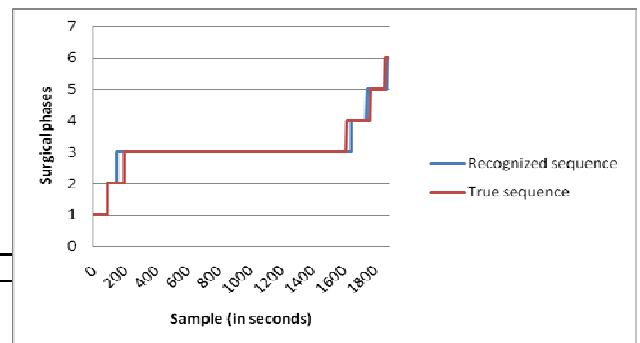
a)

	Couleur iris	Présence antiseptique	Présence couteau	Présence instrument lentille
Précision (std) (%)	89.4 (1.4)	92.3 (1.9)	95.5 (1.3)	94.8 (1.1)

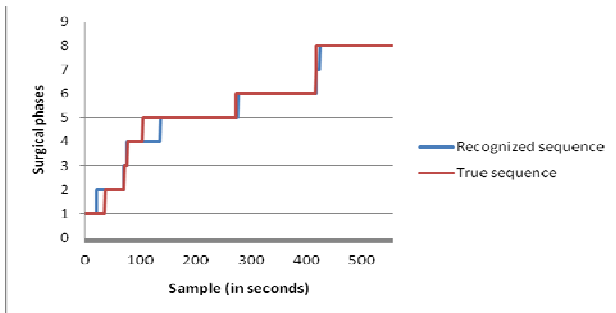
b)

Tab. 3. Pourcentage d'étapes correctement reconnu pour les deux jeux de données

	Moyenne (%)	Std (%)	Min (%)	Max (%)
Chirurgie hypophysaires	92.9	5.3	84.8	97.5
Chirurgie de la cataracte	90.2	8.4	78.1	99.9



a)



b)

Fig. 5. Séquence reconnue par notre processus de reconnaissance et séquence réelle pour une vidéo :
a) jeu de donnée n°1
b) jeu de données n°2

D'après le Tab. 3., le jeu de données n°1 (chirurgie hypophysaire) possède un meilleur taux de reconnaissance des étapes (92.9%), comparé au jeu de données n°2 (90.2%). De plus, les résultats pour le jeu de données n°2 sont assez dispersés (std=8.4%). Une séquence reconnue par le système, comparée à la séquence réelle est montrée sur la Fig. 5., pour les deux jeux de données.

4 Discussion

Notre méthode de reconnaissance automatique des étapes de la chirurgie a montré qu'il était possible d'utiliser seulement les vidéos des microscopes comme source d'information. Des taux de reconnaissances de 93% et 90% furent trouvés pour respectivement le jeu de données n°1 et le jeu de données n°2.

4.1 Données vidéos des microscopes

Comme expliqué par Bouarfa et al. (2010), les informations extraites des salles d'opérations doivent être discriminantes, facilement contrôlables, invariantes selon le chirurgien qui opère et ne pas demander beaucoup de ressources. Les données vidéos issues des microscopes réunissent toutes ces contraintes. Elles sont tout d'abord assez discriminantes pour pouvoir obtenir des taux de classification de l'ordre de 90% pour la détection des attributs image. La performance de cette classification est largement liée à la puissance de discrimination de la base de données image. Cette contrainte soulève le problème de l'adaptabilité du système pour les différents hôpitaux. Au sein d'un même service, les couleurs des instruments ou les particularités de la scène chirurgicale sont les mêmes, mais pour d'autres services hospitaliers, cela peut ne pas être le cas. Par exemple, le couleur des tissus utilisés à Rennes peut ne pas être la même que celle

d'autres services. La solution serait alors d'entraîner une base de données pour chaque service, qui prendrait en compte l'environnement local. Un autre problème pourrait se poser avec la variabilité entre les chirurgiens. Mais en supposant que l'environnement est le même pour chaque chirurgien au sein d'un même service, cette variabilité est réduite au temps de chirurgie, qui ne biaise pas du tout notre modèle. Ensuite, les données sont toujours disponibles selon le même format, ce qui facilite l'utilisation qui va en être fait derrière. Une simple réduction de dimensionnalité permet par exemple d'obtenir des temps de calcul pour la reconnaissance d'une image de 0.8s. Ce temps est compatible avec une utilisation *on-line* si on regarde le sous-échantillonnage temporel effectué sur les données (1 image/s ou 1 image/2s). Enfin, l'atout principal de ce projet est l'utilisation du microscope. Cet appareil est non seulement déjà installé dans les salles d'opération, mais en plus l'enregistrement n'a pas à être contrôlé par une personne pendant l'intervention. Ce type de données est donc une bonne solution pour la création de nouveaux systèmes de chirurgie guidée par l'image incluant une connaissance explicite et formalisée de l'activité chirurgicale en cours.

4.2 Réduction de dimensionnalité

Deux méthodes furent testées pour réduire la dimension des données images : l'ACP et une méthode hybride de sélection de caractéristiques combinant une approche *wrapper* avec une approche *filter*. Intuitivement, les méthodes *wrapper* semblent plus avantageuses, puisqu'elles utilisent les résultats des classifications pour faire la sélection. En revanche, la principale limitation reste le temps de calcul, qui augmente de façon exponentielle avec la taille des données. Pour les approches *filter*, la sélection est faite sans regarder les résultats des classifications, mais seulement sur les données d'entrée, en évaluant le pouvoir prédictif de chaque variable. La principale limite se trouve dans l'incapacité de telles méthodes de prendre en compte des combinaisons de caractéristiques, ce qui affecte la précision de la sélection. L'avantage de la combinaison est donc de bénéficier des avantages des deux méthodes. Une étude de comparaison (Fig. 4.) de cette approche avec l'ACP a clairement montré la supériorité de cette méthode.

4.3 Algorithmes de classification

Pour extraire les attributs visuels des images, les SVMs ont montré de meilleurs résultats (Fig. 5.). La précision des SVMs, associée à une faible déviation standard, a montré la robustesse de cet algorithme pour ce type d'images. Les bonnes performances, en grande partie expliquées par leurs capacités de généralisation, ne sont pas surprenantes si on regarde la récente explosion de son

utilisation. L'ADL, même si ses performances se dégradent rapidement, restent aussi une bonne méthode de classification. En revanche, les arbres de décision et l'algorithme des PPV ont montré leurs limites pour nos jeux de données. Ces derniers étaient vraisemblablement trop variables en couleur et texture et pas assez discriminants pour utiliser des outils de classification simples. Le résultat des réseaux de neurones peut paraître surprenant, compte tenu du type de données disponibles. Les algorithmes non-linéaires sont généralement adaptés aux systèmes très complexes, en revanche les algorithmes linéaires sont plus faciles et rapides à utiliser, ce qui les rend adaptés à notre problématique.

4.4 Applications cliniques

L'idée générale est d'assister la chirurgie à travers la compréhension des événements chirurgicaux dans les salles d'opérations. Le but serait d'aider à mieux gérer les salles d'opérations (Xiao et al., 2005) et la chirurgie (Jannin et al., 2007 ; Speidel et al., 2008). Par exemple, il y a une réelle demande d'outils permettant de spécifier le type d'information, par exemple, d'images médicales multimodales du patient, dont le chirurgien a besoin.

Dans sa version actuelle, le système pourrait être intégré en phase post-opératoire pour indexer automatiquement les vidéos. Les vidéos de chirurgies sont utiles pour des problématiques d'enseignement et d'apprentissage, mais les chirurgiens ne les utilisent pas souvent, compte tenu de la grosse quantité de données disponibles et de la difficulté de stockage. Cet outil pourrait donc apporter une solution à ces problèmes en fournissant des vidéos déjà indexées qui seraient plus facilement utilisables. De même, les rapports post-opératoires pourraient être pré-remplis et ainsi permettre d'économiser du temps au chirurgien.

5 Conclusion

Ce projet original de reconnaissance automatique est un premier pas vers la création de systèmes intelligents pour la chirurgie. L'idée est d'assister la procédure grâce à la compréhension des événements de la chirurgie. Il doit aussi servir à prouver qu'il n'est pas nécessaire de rajouter une multitude de capteurs supplémentaires dans une salle d'opération pour en extraire des informations pertinentes. Dans un premier temps, l'indexation post-opératoire de vidéos peut être mise en place dans un esprit d'optimisation du temps, en particulier dans la création des rapports. D'autres méthodes de vision par ordinateur (tracking, segmentation) pourront aussi être expérimentées pour apporter des informations complémentaires au processus de reconnaissance.

Une version préliminaire de cette étude a été présentée à la conférence MICCAI (2010). Les auteurs voudraient aussi remercier la société Carl Zeiss Meditec pour le financement du projet.

Bibliographie

- Ahmadi, S.A., Sielhorst, T., Stauder, R., Horn, M., Feussner, H., Navab, N. Recovery of surgical workflow without explicit models. Proc of MICCAI. Berlin: Springer, 420-428 (2006)
- Ahmed, N., Natarajan, T., Rao, K.R. Discrete Cosine Transform. IEEE Trans Comp. 90--93 (1974)
- Cleary, K., Chung, H.Y., Mun, S.K.: OR 2020: The operating room of the future. Laparoendoscopic and Advanced Surgical Techniques. 15(5), 495--500 (2005)
- Darzi, A., Mackay, S. Skills assessment of surgeons. Surg, 131(2)121--124 (2002)
- Hamming, R.W. Coding and Information Theory. Prentice-Hall Inc (1980)
- Haralick, R.M., Shanmugam, K., Dinstein, I. Textural features for image classification. IEEE Trans. on Systems, Man, and Cybernetics. 3(6), 61--621 (1973)
- Duda, R.O. and Hart, P.E. Pattern classification and scene analysis. Guyon, John Wiley & Sons (1973)
- Hu, M.K. Visual pattern recognition by moment invariants. IRE Trans on Information Theory. 8(2), 179--187 (1962)
- Jannin, P., Morandi, X. Surgical models for computer-assisted neurosurgery. Neuroimage, 37(3), 783--91 (2007)
- James, A., Vieira, D., Lo, B.P.L., Darzi, A., Yang, G-Z.: Eye-gaze driven surgical workflow segmentation. In: Proc. of Medical Image Computing and Computer-Assisted Intervention (MICCAI), 110--117 (2007)
- Klank, U., Padoy, N., Feussner, H., Navab, N. Automatic feature generation in endoscopic images. Int J Comput Assist Radiol Surg. 3(3-4), 331-339 (2008)
- Lin, H.C., Shafran, I., Yuh, D., Hager, G.D. Towards automatic skill evaluation: Detection and segmentation of robot-assisted surgical motions. Computer Aided Surgery, 11(5), 220--230 (2006)
- Lo, B., Darzi, A., Yang, G. Episode Classification for the Analysis of Tissue-Instrument Interaction with Multiple Visual Cues. International Conference on Medical Image Computing and Computer-Assisted Intervention (2003)
- Nara, A., Izumi, K., Iseki, H., Suzuki, T., Nambu, K., Sakurai, Y. Surgical workflow analysis based on staff's trajectory patterns. M2CAI workshop, MICCAI, London (2009)
- Padoy, N., Horn, M., Feussner, H., Berger, M., Navab, N. Recovery of surgical workflow: a model-based approach. Int j Comput Assist Radiol Surg, 2(1), 481-482 (2007)

- Rabiner, LR. A tutorial on Hidden Markov Models and selected applications in speech recognition, Proc of IEEE, 77(2) (1989)
- Smeulders, AW., Worrin, M., Santini, S., Gupta, A., Jain, R. Content-based image retrieval at the end of the early years. IEEE Trans on pattern analysis and machine learning intelligence. 22(12), 1349--1380 (2000)
- Speidel, S., Sudra, G., Senemaud, J., Drentschew, M., Müller-stich, BP., Gun, C., Dillmann, R. Situation modelling and situation recognition for a context-aware augmented reality system. Progression in biomedical optics and imaging. 9(1), 35 (2008)
- Viterbi, A. Errors bounds for convolutional codes. IEEE TIT, 13(2), 260--269 (1967)
- Xiao, Y., Hu, P., Hu, H., Ho, D., Dexter, F., Mackenzie, CF., Seagull, FJ. An algorithm for processing vital sign monitoring data to remotely identify operating room occupancy in real-time. Anesth Analg, 101(3), 823-832 (2005)