

Evaluation subjective de la qualité de vidéos encodées MPEG2 dans un contexte de télé-robotique chirurgicale

Nedia NOURI¹, Denis ABRAHAM¹, Jean-Marie MOUREAUX², Michel DUFAUT¹,

Jacques HUBERT³, Manuela PEREZ³

¹Centre de Recherche en Automatique de Nancy, Nancy-Université, CNRS

2, avenue de la forêt de Haye, 54516 Vandœuvre-lès-Nancy, France

nedia.nouri@cran.uhp-nancy.fr, denis.abraham@cran.uhp-nancy.fr, michel.dufaut@ensem.inpl-nancy.fr

²Centre de Recherche en Automatique de Nancy, Nancy-Université, CNRS

Faculté des Sciences, BP 70239, 54506 Vandœuvre-lès-Nancy Cedex, France

jean-marie.moureaux@cran.uhp-nancy.fr

³INSERM, U947, CHU Nancy Brabois

Rue du Morvan, 54511 Vandœuvre-lès-Nancy, France

j.hubert@chu-nancy.fr, m.perez@chu-nancy.fr

Résumé— La transmission de vidéos médicales sur de grandes distances est un sujet en plein développement, du fait de la « révolution numérique » que connaît le monde médical actuellement. La contrepartie de cette numérisation réside dans des volumes de données considérables nécessitant des capacités de transmission elles-mêmes très importantes, rendant la compression avec perte en amont inévitable. Celle-ci constitue un défi majeur dans un contexte aussi sensible que le contexte médical, celui de l'impact des pertes sur la qualité des données et leur exploitation. Le travail présenté ici concerne l'évaluation subjective, par un panel de chirurgiens, de séquences vidéo issues d'une application de télé-robotique chirurgicale et compressées avec pertes au format MPEG2. L'évaluation s'appuie sur les recommandations de l'ITU-R¹ à travers le protocole DSCQS², l'objectif étant de déterminer un seuil de compression permettant d'avoir une qualité perceptuelle de la vidéo chirurgicale compressée irréprochable, pour le bon déroulement de l'opération et pour assurer la sécurité des patients. Nous montrons qu'il existe un seuil de tolérance à la compression avec pertes de type MPEG2 autour de 3 Mbits/s pour les vidéos utilisées dans cette étude, ce qui équivaut à un taux de compression d'environ 90 :1 du flux vidéo initialement à 270 Mbits/s !

Mots-clés—Compression vidéo, télé-chirurgie, imagerie médicale, évaluation subjective, qualité

I. INTRODUCTION

L'évolution des techniques chirurgicales, par l'utilisation de robots de chirurgie, permet aujourd'hui des interventions mini-invasives avec une précision très importante. Le second avantage de cette évolution est qu'elle ouvre des perspectives d'interventions chirurgicales à distance, comme l'a montré la célèbre expérimentation « Opération Lindbergh » en 2001 [1]. La généralisation de ce type d'expérience n'est cependant pas

encore acquise du fait des ressources réseau qu'elle nécessite, en particulier en termes de bande passante. Actuellement, un chirurgien peut toutefois opérer son patient en étant à quelques mètres du robot, et parfois dans une autre pièce. En effet, sa console de travail lui permet de visualiser à travers une caméra endoscopique le champ d'opération et est reliée au robot par un réseau local dédié permettant une transmission de volumes importants de données, sans latence ni dégradation visuelle pouvant gêner le geste du chirurgien. Dans un contexte d'opération à grande distance, de télé-enseignement ou encore de télé-conseil (Telementoring), la transmission à travers les réseaux de communication des volumineux flux vidéo nécessite des capacités très importantes en termes de bande passante, rendant la compression avec perte de ces flux incontournable. A cet effet, il existe aujourd'hui de nombreux encodeurs vidéo, notamment les standards de la famille MPEG. Ces derniers sont déjà utilisés dans des applications grand public telles que la télévision numérique. Les algorithmes de compression vidéo tiennent compte des caractéristiques psycho-visuelles de l'œil humain pour supprimer des informations qui ne peuvent être perçues ou qui sont redondantes, en agissant à la fois sur les corrélations entre les pixels d'une image et entre ceux d'images successives de la vidéo.

Cependant, lorsque le taux de compression augmente, les encodeurs introduisent dans la vidéo compressée des artefacts pouvant affecter sa qualité visuelle. Dans le contexte de l'imagerie médicale et en particulier celui de la télé-chirurgie, cette altération de la qualité peut être rédhibitoire. C'est la raison pour laquelle le milieu médical utilise aujourd'hui la compression sans perte plutôt que la compression avec perte qui permet pourtant d'atteindre des taux de compression bien plus importants. Différents travaux émergent cependant aujourd'hui montrant qu'il existe une certaine tolérance aux pertes en imagerie médicale, y compris dans le domaine très sensible de la radiologie [2] [3]. L'ACR (American College of Radiology) recommande d'ailleurs à présent l'utilisation de techniques de compression avec perte sous la responsabilité d'un praticien qualifié.

¹ International Telecommunication Union

² Double-Stimuli Continuous Quality Scale

Afin de définir le seuil de compression toléré par les chirurgiens, une métrique objective, telle que le très répandu rapport signal à bruit, n'est pas suffisante. En effet, ce type de mesure ne permet pas de localiser les dégradations les plus gênantes visuellement mais fournit plutôt un indicateur global qu'il est difficile de relier à la perception réelle de l'observateur. D'autres métriques objectives ont été proposées [4][5][6][7][8]. Ces métriques intègrent des propriétés du système visuel humain (SVH). En 2001, Watson et al. [4] proposent une métrique nommée Digital Video Quality (DVQ) pour la qualité des vidéos. Elle utilise une décomposition DCT, puis calcule un contraste local dans les différentes bandes de fréquences. Les coefficients sont ensuite normalisés et leur union est réalisée par une distance de *Minkowski* [8]. Une autre étude [5], propose une métrique objective de la qualité en deux étapes pour les vidéos MPEG. Dans une première étape, la distorsion de la vidéo encodée est calculée en tenant compte de la réponse du SVH. Puis, les auteurs calculent la distorsion trame par trame de la vidéo décodée par rapport à une vidéo de référence. Ce calcul inclut un filtrage spatial passe-bas, une opération de Sobel et un masquage spatial sur l'erreur brute entre images de référence et compressée. L'émulation cognitive vient en seconde étape pour simuler le traitement haut-niveau humain face aux changements de la qualité des images. Cette métrique objective donne des résultats très corrélés avec ceux donnés par la méthode de qualité subjective SSCQE. Cependant, dans des applications aussi sensibles que les applications médicales, il s'avère préférable de se référer au jugement humain par des essais subjectifs de la qualité.

Nous proposons dans cet article d'étudier le seuil de compression toléré dans une application de télé-robotique chirurgicale, à travers une évaluation subjective de la qualité de vidéos encodées MPEG2, par un panel de chirurgiens expérimentés. Cette étude, menée selon le protocole recommandé par l'ITU est basée sur le test DSCQS (Double-Stimuli Continuous Quality Scale), méthode à double stimulus utilisant une échelle de qualité continue afin d'évaluer ici la qualité des vidéos chirurgicales compressées par rapport à une référence non compressée. Dans cette étude, il est indispensable de s'approcher des conditions d'utilisation habituelles des chirurgiens afin de déterminer s'il existe un débit de compression suffisamment bas pour une transmission rapide à travers les réseaux mais assurant une qualité suffisante de la vidéo compressée pour le confort visuel du chirurgien afin qu'il opère ses patients en toute sécurité et confort.

Les principales contributions de ce travail sont les suivantes. En premier lieu, nous investiguons le domaine de la compression avec perte dans le contexte médical, domaine nécessitant encore de nombreuses études avant que ce type de compression ne soit intégré dans la pratique hospitalière, du fait des enjeux. Nous montrerons cependant qu'il existe une tolérance à la compression avec perte en télé-robotique chirurgicale : le seuil relevé dans cette étude est un taux de compression de 90:1 pour des vidéos initialement à 270 Mbits/s. En second lieu, on trouve actuellement dans la littérature très peu d'études concernant l'évaluation subjective de la qualité dans le domaine médical, du fait de la complexité de mise en œuvre et des moyens humains et matériels nécessaires. Nous proposons ici une telle étude avec un panel de 7 chirurgiens expérimentés. Enfin, ces travaux s'appuient

sur un protocole reconnu, celui de l'ITU, et à notre connaissance pas encore utilisé dans le domaine de la compression de vidéos médicales.

Cet article est organisé de la manière suivante : le contexte de ces travaux est présenté au paragraphe II. Les principes et les artefacts de la compression vidéo sont donnés au paragraphe III tandis que le paragraphe IV décrit la méthodologie d'évaluation subjective de la qualité proposée. Le paragraphe V est consacré à l'analyse des résultats de l'évaluation. Enfin, nous concluons et présentons les perspectives de ce travail au paragraphe VI.

II. CONTEXTE DE L'ETUDE

La téléchirurgie constitue un enjeu majeur pour permettre de distribuer l'expertise médicale à tous les niveaux d'un territoire ou de mutualiser des équipements lourds de chirurgie.

C'est dans ce contexte que s'inscrit le projet de télé-chirurgie RALTT³ (Robotic Assisted Laparoscopic Telementoring and Telesurgery) initié en 2005 et réunissant une équipe pluridisciplinaire du CHU de Nancy et du Centre de Recherche en Automatique de Nancy (CRAN, Nancy-Université)⁴. Ce projet vise à étudier la faisabilité de certaines techniques opératoires de chirurgie par robot piloté manuellement à distance entre deux sites hospitaliers. En effet, dans une application de chirurgie assistée par robot, le chirurgien opère son patient en visualisant le champ opératoire en 3D sur une console. Les flux vidéo visualisés proviennent d'une caméra endoscopique binoculaire, fournissant deux flux stéréoscopiques à 270 Mbits/s chacun. Dans cette étude, nous nous plaçons dans le cas de la compression avec pertes d'un seul flux vidéo. En effet, des travaux [9] ont démontré qu'il est possible de compresser indépendamment et à des taux de compression différents les deux vues d'une vidéo stéréoscopique car le système visuel humain compense naturellement cette différence de qualité en se focalisant sur la vue de meilleure qualité.

III. COMPRESSION MPEG2

A. Principes

Le standard MPEG2 [10][11] permet de transformer un signal vidéo numérisé en un train binaire destiné à être stocké ou transmis sur un réseau. MPEG 2 s'appuie principalement sur une exploitation des redondances spatiales, temporelles, subjectives et statistiques existant dans une séquence vidéo. Le train binaire obtenu est décrit selon la norme MPEG afin que l'on puisse restituer le signal par n'importe quel décodeur respectant cette même norme.

Tout codeur MPEG-2 repose sur les opérations suivantes : transformation des données, quantification, codage entropique et prédiction du mouvement.

La transformation est ici une transformée en cosinus discrète (DCT) qui permet d'exploiter les corrélations entre les pixels d'une image (redondances spatiales). C'est une transformée

³ <http://raltt.cran.uhp-nancy.fr/>

⁴ Le projet RALTT est cofinancé par la Région Lorraine et la Communauté Urbaine du Grand Nancy

réversible. La quantification est l'étape irréversible de l'algorithme. C'est elle qui produit l'essentiel des pertes de l'information.

Enfin, le codage exploite les redondances statistiques en identifiant les probabilités d'occurrence des symboles : un codage entropique à longueur variable (VLC [11]) suivi d'une détection de « zéros » (RLC [11]). Dans les deux cas la probabilité d'apparition des symboles est prise en compte afin d'éliminer les redondances (codage entropique). Le codage prédictif est dédié aux redondances temporelles [11].

Celles-ci sont exploitées à travers une prédiction des futures images en se basant sur les vecteurs de mouvement. La compression est ainsi réalisée selon trois modes qui conduisent à distinguer trois types d'images :

- une image de type I (intra) est compressée de manière indépendante des autres images, selon les principes du standard JPEG pour les images fixes [12]. Seule la redondance spatiale est réduite.
- une image de type P (prédite) est codée en utilisant une prédiction d'une image antérieure de type I ou P. Ce type d'image est moins volumineux que les images de type I.
- une image B (Bidirectionnelle) est codée par double prédiction en utilisant comme référence deux images qui la précèdent ou la suivent de type I ou P. Elles ont le taux de compression le plus élevé.

B. Artefacts typiques

Les pertes causées par la compression spatiale et temporelle des vidéos ont des conséquences sur les images compressées décodées. En effet, il existe deux types de distorsions : spatiales (qui apparaissent sur le contenu d'une image) et temporelles (qui n'apparaissent que lorsque les images défilent). Dans [13], les auteurs font un état de l'art des artefacts de compression qui peuvent apparaître aussi bien sur la composante de luminance que sur les composantes de chrominance.

➤ Distorsions spatiales

Dans un schéma de compression type MPEG, l'image est découpée en Macro-blocs eux-mêmes subdivisés en blocs, et la quantification est effectuée indépendamment sur chacun de ces blocs. Si la quantification est trop forte alors il est possible qu'une discontinuité de l'information apparaisse sur les frontières des blocs adjacents lors de leur reconstruction. Ceci crée l'effet de bloc. La gêne occasionnée par l'effet de bloc dépend non seulement du taux de quantification mais aussi du contenu du bloc et de ses voisins (plus le bloc contiendra des variations de l'information, plus la frontière sera visible).

Après la DCT, le contenu d'un bloc est représenté par une combinaison de motifs dont l'importance est donnée par les coefficients AC. Ces derniers peuvent être partiellement ou totalement éliminés après la quantification et le bloc résultant représentera alors le coefficient correspondant au motif AC (effet de motif). L'étape de quantification privilégie les coefficients basses fréquences aux coefficients hautes fréquences. Après cette étape, tous les détails sont atténués ou supprimés. Ceci a pour conséquence d'atténuer les contours ou les textures et provoque un effet de flou sur l'ensemble de l'image. Plus la quantification des blocs contenant des variations d'information sera forte, plus le flou sera important.

Dans l'espace (YCbCr), l'information de couleur est donnée par les composantes Cb et Cr. Les informations Cb et Cr peuvent être représentées sur deux axes de couleurs différents. Lors de la réunification des trois composantes, une fausse couleur peut être représentée : la valeur correspondant à une couleur sur un axe peut être modifiée par une autre valeur sur l'autre axe après quantification, on parle de l'effet des fausses couleurs.

➤ Distorsions temporelles

La compensation de mouvement est obtenue par la combinaison de l'estimation de mouvement et du vecteur d'erreur de prédiction. Ces informations sont codées comme le contenu d'un bloc et soumises à la quantification. Une mauvaise estimation de mouvement ou une quantification trop importante du vecteur d'erreur peuvent être responsables d'un codage différent d'une même information d'une image à l'autre. Ceci peut causer une fluctuation autour des contours lorsqu'ils sont animés. Ce phénomène de battement est comparé au battement des ailes d'un moustique (effet moustique). Dans le cas d'une mauvaise estimation de mouvement, le vecteur d'erreur aura beaucoup d'information à transmettre. De plus, si ce vecteur est fortement quantifié, l'information sera non seulement transmise à la mauvaise position mais également avec un faux contenu. Ceci peut créer une disparition de l'information. La limite de la zone de recherche dans laquelle peut être estimé le vecteur mouvement peut parfois être insuffisante. Dans ce cas, le mouvement ne sera pas représenté d'une image à l'autre et la fluidité du mouvement au cours du temps sera cassée provoquant un mouvement saccadé.

IV. EVALUATION SUBJECTIVE DE LA QUALITE

Pour l'évaluation subjective de la qualité des images, l'ITU recommande les méthodes générales d'essai, les échelles et les conditions d'observation. Le choix entre les différentes méthodes de test dépend de l'application considérée. Ainsi, lorsqu'on dispose d'une séquence de référence, la méthode à double stimulus utilisant une échelle de qualité continue DSCQS [14] est la méthode la plus utilisée, en particulier pour mesurer la qualité d'un système par rapport à une référence ou encore comparer des systèmes entre eux. C'est la raison pour laquelle nous avons retenu cette méthode dans le cadre de notre étude.

Dans ce qui suit nous décrivons les conditions d'observation et les séquences de test au paragraphe A, ainsi que la méthode d'évaluation elle-même au paragraphe B.

A. Conditions d'observation et séquences de test

Les conditions d'observation sont basées sur la norme ITU pour la télévision numérique décrite dans [15]. Dans l'esprit de la norme, quelques modifications doivent être apportées afin d'adapter les recommandations de l'ITU à notre application : la chirurgie à distance assistée par robot. En effet, l'environnement du bloc opératoire n'est pas aussi maîtrisé qu'un laboratoire d'essais subjectifs tel qu'il est préconisé par l'ITU. Nous estimons néanmoins être en conformité avec l'idée directrice de la norme. Pour juger de la qualité des vidéos chirurgicales, il est nécessaire que les chirurgiens soient dans des conditions de fonctionnement habituelles. Par conséquent, le test s'est déroulé dans un bloc

opératoire et les vidéos sont visualisées dans la console de visualisation du chirurgien comme c'est le cas en routine. Par ailleurs, pour l'évaluation de la qualité de la télévision numérique, la norme définit un nombre minimum de 15 observateurs s'ils sont non experts, ce nombre pouvant être plus faible (minimum 4) lorsque les observateurs sont expérimentés. Nous avons ainsi soumis ces évaluations à 7 observateurs experts (chirurgiens).

La session d'essais subjectifs est limitée à 35 minutes pendant lesquelles les 7 observateurs ont noté 38 présentations provenant de 4 scènes typiques de chirurgie selon un enchaînement décrit dans la figure 3. Au début de la séance, on procède à 5 présentations fictives pour stabiliser les jugements des observateurs et dont on ne tiendra pas compte dans le dépouillement des résultats. De plus, certaines séquences peuvent être présentées plusieurs fois afin de conforter le jugement des observateurs.

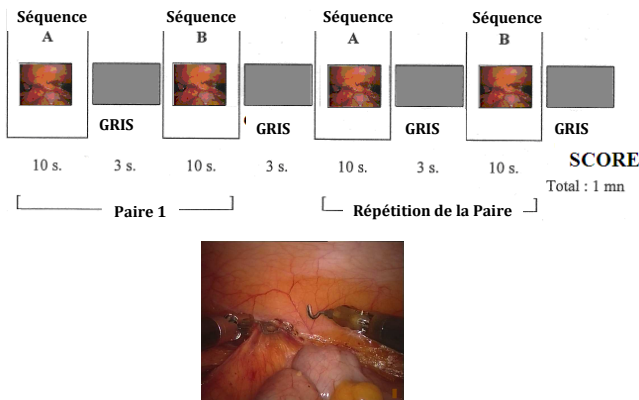


Fig. 1. Structure d'une séance d'évaluation et extrait de séquence vidéo utilisé dans l'étude

Comme le montre la figure 3, chaque stimulus (paire d'images) est présenté deux fois consécutives afin de conforter le jugement des participants. De plus, une paire (séquence A, séquence B) contient obligatoirement une séquence de référence n'ayant subi aucune dégradation et une séquence dégradée (dans notre cas du fait de la compression).

TABLE I
DEBITS ASSOCIES A CHAQUE SEQUENCE (EN Mbits/s)

Scène	Débit en Mbits/s							
1	1.02	3	4.02	5.04	6	8.04		
2	1.5	2.56	4.5	5.52	6.54	7.5		
3	1.74	2.76	3.78	4.74	5.76	6.78	7.74	
4	1.2	2.04	2.28	3.24	4.25	5.28	6.24	7.2

La position de la séquence de référence dans une paire varie de manière pseudo-aléatoire. A la fin de chaque présentation, les observateurs expriment leur jugement sur une paire d'échelles verticales continues. A ces échelles graduées de façon continue sont associés des qualificatifs du protocole DSCQS : excellent, bon, assez bon, médiocre et mauvais, comme le montre la figure 4. Il est cependant erroné d'associer aux notes un seul qualificatif car elles ne sont pas absolues mais représentent la différence de note entre l'image de référence et l'image dégradée.

Le test réalisé se compose de L présentations (L paires), chaque présentation représente un ensemble J de conditions de test (i.e. débits). De plus, chaque paire de séquences peut

être répétée R fois afin de pouvoir déceler les incohérences et de telle sorte qu'une séquence n'influe pas sur le jugement de l'observateur pour la séquence suivante. Il faut noter que seulement K inférieure à L présentations sont réellement effectives (si l'on tient compte des cinq séquences fictives au début du test et des séquences répétées). Dans cette étude, $L=38$, $J=27$, $K=28$ et $R=3$. Nous avons également 4 scènes différentes compressées à un débit variant de 1.2 Mbits/s et 8.04 Mbits/s (cf. TABLE I). Différents débits ont été choisis pour les 4 scènes de test, comme le montre la table I. En effet, le choix des valeurs du débit de compression doit prendre en compte la capacité des espaces de stockage et de la bande passante des réseaux actuels. Par ailleurs, les 4 scènes sont représentatives de la diversité des vidéos chirurgicales :

- scène 1 : représente un mouvement d'outil laparoscopique,
- scène 2 : représente un cas de coagulation du sang,
- scène 3 : représente une zone grasseuse,
- scène 4 : représente l'application d'une compresse.

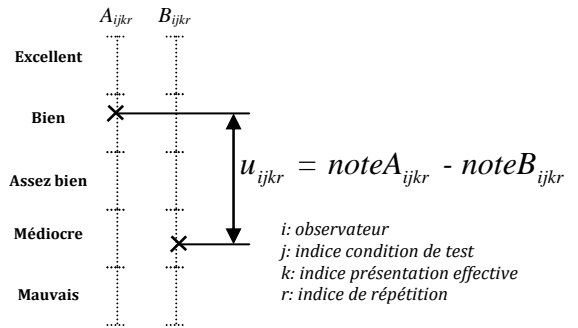


Fig. 2. Echelles d'évaluation et exemple de notation

B.Méthode d'évaluation

L'ITU préconise que les données recueillies au cours des essais subjectifs soient traitées selon les mêmes techniques statistiques recommandées. En effet, l'analyse des résultats issus des tests subjectifs dépend de la méthode utilisée. Ainsi, pour la DSCQS, la cohérence des résultats sera vérifiée en étudiant les notes données par le même observateur à la même séquence pendant la même séance. Si les notes diffèrent de 2 points ou plus (pour une échelle continue allant de 1 à 5), ces notes seront rejetées, l'observateur étant jugé « non fiable ». Après chaque séance, on calcule les valeurs moyennes et les écarts-type associés à chaque niveau de dégradation. La valeur moyenne est donnée par :

$$\bar{u}_{jkr} = \frac{1}{N_{obs}} \sum_{i=1}^{N_{obs}} u_{ijk}$$

Où N_{obs} représente le nombre d'observateurs et u_{ijk} la note de l'observateur i pour la dégradation j de la séquence k et la répétition r .

Afin d'évaluer au mieux la fiabilité des résultats, on associe à chaque moyenne un intervalle de confiance. En général, il est convenu d'utiliser l'intervalle de confiance à 95% donné par :

$$\left[\bar{u}_{jkr} - \delta_{jkr}, \bar{u}_{jkr} + \delta_{jkr} \right] \text{ où } \delta_{jkr} = 1,96 \cdot \frac{\sigma_{jkr}}{\sqrt{N_{obs}}}$$

L'écart-type de chaque présentation σ_{jkr} est donné par :

$$\sigma_{jkr} = \sqrt{\frac{\sum_{i=1}^{N_{obs}} (\bar{u}_{ijkr} - u_{ijkr})^2}{(N_{obs} - 1)}}$$

Ces valeurs moyennes reposent sur une distribution dont les deux variables sont les scènes et les observateurs. Afin de vérifier si cette distribution suit une loi normale, on calcule le coefficient d'aplatissement (kurtosis) de la fonction, défini comme étant le rapport entre le moment d'ordre 4 et le carré

du moment d'ordre 2, soit : $\beta_{2_{jkr}} = \frac{m_4}{(m_2)^2}$

Avec
$$m_x = \frac{\sum_{i=1}^{N_{obs}} (u_{ijkr} - \bar{u}_{jkr})^x}{N_{obs}}$$

La norme indique que si β_2 est compris entre 2 et 4, on peut considérer la distribution comme normale.

Les résultats de chaque distribution sont alors à comparer à la valeur moyenne dans un intervalle selon l'algorithme donné ci-dessous. Chaque fois que les résultats d'un observateur se situent en dehors de cet intervalle, il faut les enregistrer dans un compteur associé à chaque observateur. Il faut donc deux intervalles pour les valeurs supérieures et inférieures. Cette procédure est récapitulée dans [9] et peut s'exprimer comme suit :

```

pour chaque observateur i
  pour j, k, r = 1 → J, K, R
    si  $2 \leq \beta_{2_{jkr}} \leq 4$  alors
      si  $u_{ijkr} \geq \bar{u}_{jkr} + 2 \cdot \sigma_{jkr}$  alors  $P_i = P_i + 1$ 
      si  $u_{ijkr} \leq \bar{u}_{jkr} - 2 \cdot \sigma_{jkr}$  alors  $Q_i = Q_i + 1$ 
    sinon
      si  $u_{ijkr} \geq \bar{u}_{jkr} + \sqrt{20} \cdot \sigma_{jkr}$  alors  $P_i = P_i + 1$ 
      si  $u_{ijkr} \leq \bar{u}_{jkr} - \sqrt{20} \cdot \sigma_{jkr}$  alors  $Q_i = Q_i + 1$ 
    fin pour
  si  $\frac{P_i + Q_i}{J \cdot K \cdot R} > 0.05$  et  $\left| \frac{P_i - Q_i}{P_i + Q_i} \right| < 0.3$  alors
    rejeter l'observateur i
  fin pour

```

avec :

J : nombre de conditions de test y compris la référence

K : nombre d'images ou séquences de test effectives

R : nombre de répétitions

L : nombre de présentations de test

Il faut noter que le test du β_2 décrit ci-dessus est réalisé systématiquement afin de vérifier la cohérence des observateurs.

V. RESULTATS EXPERIMENTAUX

Dans cette étude nous utilisons des séquences issues de flux vidéo au format PAL provenant d'une caméra tri-CCD et numérisés dans la chaîne de transmission du robot. Les signaux numériques obtenus sont échantillonnés selon la norme 4:2:2 et compressés par une carte de compression

MPEG-2 (hardware) à différents débits (cf. Table I). Ces informations sont alors transmises à la console du robot pour être visualisée par les sept observateurs. La vidéo initiale est constituée de 720 pixels sur 576 lignes codés sur 10 bits (voir figure 3), ce qui représente un débit de 270 Mbits/s.

Après la réalisation du test subjectif selon le protocole DSCQS où l'observateur juge les images deux par deux (image de référence, image dégradée) sur une échelle continue, les résultats représentent la différence entre les notes de l'image de référence et l'image dégradée. Ces résultats sont normalisés à une valeur comprise entre 0 et 100 et ne peuvent évidemment être calculés que pour les observateurs dont le jugement est cohérent. Ainsi, lors de l'exploitation de la base de données des résultats, un observateur atypique est parti d'une fausse hypothèse pour la notation qui a engendré une incohérence de ses notes. Il a donc été écarté. Pour des raisons évidentes, le test du β_2 a été réalisé sur l'ensemble des observateurs restants et n'a rejeté aucun observateur supplémentaire.

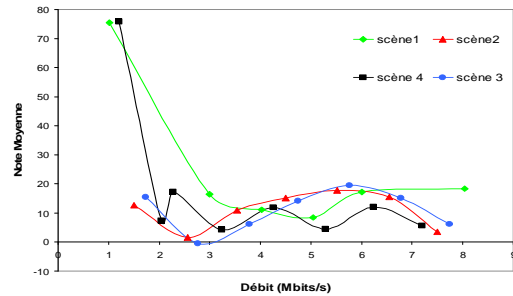


Fig. 3. Courbes note moyenne en fonction du débit pour les 4 scènes

La figure 6 représente les courbes de la moyenne des notes en fonction du débit associé à chaque séquence. Si cette note est faible, la qualité de la vidéo a peu diminué alors que si la note tend vers 100, on est face à d'importantes pertes de qualité.

La figure 6 montre que les scènes 2 et 3 ne sont pas sensibles à la variation de la qualité des vidéos. En effet, pour ces scènes, les chirurgiens observateurs n'ont pas mis en évidence une perte de qualité significative malgré les dégradations introduites par la compression pour des débits faibles (1.5 Mbits/s). On remarque également que la note moyenne pour ces deux scènes oscille entre 0 et 20 (sur une échelle de 100).

Au contraire, les courbes correspondant aux séquences 1 (Figure 7) et 4 (Figure 8) permettent d'identifier une perte de qualité significative avec la réduction du débit après compression.

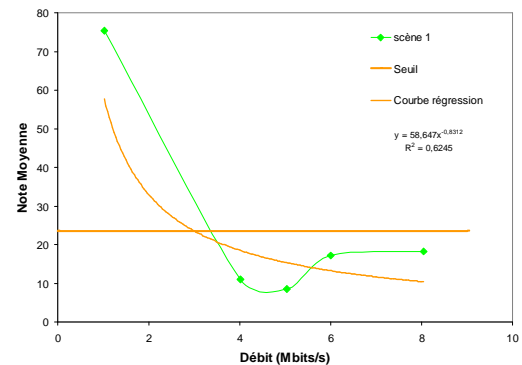


Fig.4. Note moyenne de la scène 1 en fonction du débit avec courbe de régression ($R^2=0,6246$)

La modélisation des données expérimentales par la méthode des moindres carrés pour estimer la dispersion des valeurs du nuage de points obtenu (score moyen en fonction du débit) permet de déterminer une courbe de régression ainsi que le coefficient de détermination pour chacune des scènes 1 et 4. D'une valeur comprise entre 0 et 1, le coefficient de détermination est un indicateur qui permet de mesurer l'adéquation entre le modèle et les valeurs observées. Les résultats permettent d'obtenir un coefficient de détermination $R_1=0,79$ pour la scène 1 et $R_4=0,95$ pour la scène 4. Ces valeurs montrent une homogénéité des observations et nous permettent de déterminer avec précision la valeur du débit seuil au delà duquel les observateurs ne distinguent pas de perte de qualité de la vidéo compressée.

Ce seuil correspond au point à partir duquel la détection de perte de qualité n'est plus perçue par les observateurs dans un intervalle de confiance à 95%. Ainsi, pour la scène 1, on obtient un débit seuil de 3,2Mbit/s pour un intervalle de confiance à 95%. Tandis que pour la scène 4, on obtient une valeur seuil du débit égale à 2,9 Mbits/s pour le même intervalle de confiance.

Dans le contexte de cette étude, on peut conclure qu'à partir d'un débit de 3,2 Mbits/s, aucun praticien n'observe une baisse de qualité sur les images compressées et donc aucune gêne. Etant donné que chacune des vidéos nécessitait initialement 270 Mbits/s pour être transmise, un taux de compression autour de 90 :1 peut être adopté sans risques dus à la compression des vidéos dans un contexte de chirurgie robotisée.

Ces résultats sont prometteurs, dans un contexte médical et plus précisément dans un contexte de chirurgie à distance. Ils permettent d'ouvrir des perspectives pour une compression des vidéos chirurgicales avec pertes et d'envisager des applications médicales à grandes distances : le télé-enseignement, l'aide chirurgicale à distance (Telementoring) ou encore la téléchirurgie.

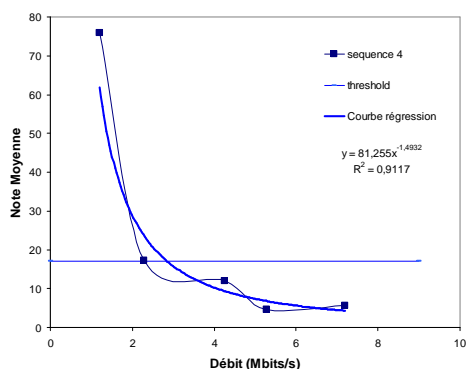


Fig.5. Note moyenne de la scène 4 en fonction du débit avec courbe de régression ($R^2=0,9117$)

VI. CONCLUSION ET PERSPECTIVES

Cette étude a permis de mettre en exergue la possibilité de compresser avec pertes des vidéos médicales en identifiant le débit seuil à partir duquel une opération chirurgicale à distance est envisageable tout au moins concernant la qualité des vidéos compressées en vue de leur transmission. Ce débit

de compression MPEG2 fixé ici à 3,2 Mbits/s a été déterminé en soumettant 7 observateurs chirurgiens à des séances d'essais subjectifs de la qualité visuelle s'inspirant des recommandations de l'ITU. La particularité de l'environnement chirurgical nous a conduits à adapter la norme à l'application. Ainsi, les tests se sont déroulés dans un bloc opératoire dans les conditions de fonctionnement habituelles avec un panel d'observateurs experts.

Si le choix du codeur MPEG 2 a été motivé à la fois par la maturité de cette norme de compression et pour ses bonnes performances en termes de débit-distorsion, l'apparition de normes plus récentes comme H.264, nous incitent à poursuivre la mise en œuvre de ce type d'étude. En effet, H.264 introduisant des artefacts différents de ceux engendrés par MPEG2, il y a fort à parier que le seuil de tolérance à la compression ne sera pas nécessairement le même.

REMERCIEMENTS

Nous remercions Christophe Burlot, Ingénieur de l'Ecole Nationale Supérieure d'Electricité et de Mécanique, ainsi que toutes les personnes ayant accepté de participer aux essais subjectifs. Nous remercions également la Région Lorraine et la Communauté Urbaine du Grand Nancy qui cofinancent le projet RALTT.

RÉFÉRENCES

- [1] Marescaux J., Leroy J., Rubino F., Vix M., Simone M. et Mutter D., Transcontinental Robot Assisted Remote Telesurgery : Feasibility and Potential Applications. *Annals of Surgery*, n° 235, pp. 487-92, 2002.
- [2] Gaudreau Y. et Moureaux JM. Lossy Compression of Volumetric Medical Images with 3D Dead-Zone Lattice Vector Quantization. *Annals of Telecommunications*, Springer, ISSN 0003-4347, vol. 64, n° 5-6, June 2009.
- [3] Schelkens P., Munteanu A., Barbarien J., Galca M. et Giro-Nieto X. Wavelet Coding of Volumetric medical Datasets. *IEEE Transactions on Medical Imaging*, vol. 22, n° 3, pp. 441-458, March 2003.
- [4] Watson A.B., Hu J. et McGowan III J.F. DVQ : A digital video quality metric based on human vision. *Journal of Electronic Imaging*, vol. 10, no. 1, pages 20-29, 2001.
- [5] Hekstra A.P., Beerends J.G., Ledermann D., De Caluwe D., Kohler S., Koenen R.H., Rihs S., Ehsam M., Schlauss D., PVQM - A perceptual video quality measure, *Signal Processing: Image Communication*, Volume 17, Issue 10, November 2002, Pages 781-798.
- [6] Winkler S., Issues in vision modeling for perceptual video quality assessment, *Signal Processing*, Volume 78, Issue 2, October 1999, pp 231-252.
- [7] Tan K., Ghanbari M., Pearson D., An objective measurement tool for MPEG video quality, *Signal Processing*, Volume 70, Issue 3, 1998, Pages 279-294.
- [8] Wang W., Sheikh H.R. et Bovik A., Objective video quality assessment, *The Handbook of Video Database: Design and applications*, pp. 1041-1078, CRC Press, 2003.
- [9] Stelmach L., Wa James T. et al. Stereo Image Quality : Effects of Mixed Spatio-Temporal Resolution. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, pp. 188-193, March 2000.
- [10] MPEG-2 ISO IEC. MPEG2 Standard. Rapport Technique, ISO IEC, 1993.
- [11] Watkinson J. *The MPEG Handbook*. Elsevier, 2004.
- [12] Penebaker WB. et Mitchell JL. *JPEG : Still Image Data Compression Standard*. Van Nostrand, 1993.
- [13] Yuen M. et Wu H.R., A survey of hybrid mc/dpcm/dct video coding distortions, *Signal Processing*, Volume 70, Issue 3, 1998, pp 247-278.
- [14] Abraham D., Ardito M., Boch L., Messina A., Stroppiana M. et Visca M. Attempts at correlation between DSCQS and objective measurements. *EBU Technical Review*, Spring 1997.
- [15] Recommandation UIT-R BT-500-10 : Méthodologie de l'évaluation subjective de la qualité des images de télévision, 2000.