

Evaluation des seuils de compression avec pertes pour les images médicales DICOM de type scanner et IRM.

J.Lambert¹, Y.Gaudeau^{2,4}, F. Lefevre³ et J.M Moureaux⁴

¹TELECOM Nancy, Université de Lorraine, 193 Avenue Paul Muller, Villers-lès-Nancy, 54602, France

²Université de Strasbourg, 30 Rue du Maire André Traband, Haguenau, 67500, France

³CHRU Nancy, Hôpital Central, 29 Avenue Maréchal De Lattre de Tassigny, Nancy, 54000, France

⁴CRAN, UMR 7039, Université de Lorraine, Faculté de Médecine, Bat.D – BP 184 Vandoeuvre-lès-Nancy, 54505, France

¹ julien.lambert@univ-lorraine.fr

² yann.gaudeau@unistra.fr

³ f.lefevre@chu-nancy.fr

⁴ jean-marie.moureaux@univ-lorraine.fr

Résumé – Le stockage des données médicales est aujourd’hui un enjeu particulièrement important au sein des services de santé. L’augmentation des résolutions des images et l’obligation légale de conserver au minimum 20 ans chaque donnée entraîne une saturation des espaces de stockage et de nouveaux besoins. Dans ce contexte la compression avec pertes peut proposer une solution durable. Nous présentons une étude sur la mesure de la qualité d’images médicales compressées de type scanner et IRM. L’évaluation de la qualité s’appuie sur une adaptation des spécifications de l’ITU-R [1], l’objectif étant de déterminer un seuil de compression permettant d’avoir une qualité perceptuelle irréprochable pour la bonne utilisation de ces données par les praticiens. Nous mettrons en regard deux types de compressions différentes, une compression avec pertes JPEG 2000 et une méthode basée sur une quantification algébrique avec zone morte MEDI-STORAGE. Les notations subjectives du panel de professionnels de santé sont ensuite mises en lien avec la plupart des meilleures métriques objectives actuelles. On démontrera que les images médicales sont tolérantes à la compression avec pertes et que celle-ci n’altère pas leur bonne utilisation.

Abstract – Nowadays, the storage of medical data is an important issue within health services. Increasing resolutions of images and legal obligation to keep a minimum of 20 years each data results in a saturation of storage spaces and new needs. Lossy compression can provide a long-term solution. The study presented here is dedicated to the quality assessment of compressed medical images of CT and MRI type. The quality assessment is based on an adaptation of the specifications of ITU-R [1], with the aim to determine a compression threshold for which the technical quality is irreproachable and allows the proper use of these data by practitioners. We will look two different types of compression, lossy JPEG 2000 and a method based on an algebraic quantization with deadzone called MEDI-STORAGE. The subjective ratings of health professionals panel are then linked to the most current objective metrics.

1 Introduction

L’enregistrement des images médicales au sein des hôpitaux s’effectue au format DICOM qui permet de standardiser les données obtenues par différents appareils médicaux et ainsi faciliter les échanges et la lecture. La norme DICOM n’exclue pas la compression d’images mais celle-ci est encore peu utilisée et est effectuée à l’aide d’algorithmes de compression sans pertes qui garantissent l’intégrité des données mais qui ne permettent pas de gains significatifs en terme de stockage.

En parallèle le développement des techniques d’imagerie permet pour les scanners nouvelle génération une meilleure résolution en épaisseur de coupe, ce qui

implique un nombre d’images plus élevés et donc un poids conséquent. A titre d’exemple, le CHRU (Centre Hospitalier Régional Universitaire) de Nancy produit environ 250 000 examens par an, soit un volume de 50 To brut. L’archive du site nancéen représente depuis 2003 environ 3 millions d’examens et 1 milliard d’objets soit 250 To de données qui sont compressées sans pertes actuellement dans le PACS (Picture Archiving and Communication System).

Partant de ce constat, la compression avec pertes « maîtrisées » peut être une solution à la surcharge des moyens de stockage [2]. Il nous faut cependant nous assurer que les pertes n’affectent pas la qualité des images pour l’usage régulier des praticiens. Hier encore inenvisageable, l’idée d’une compression

avec pertes semble aujourd'hui de mieux en mieux acceptée par les médecins, comme en témoigne par exemple, la Canadian Association of Radiologists (CAR) [3] qui estime que les techniques de compression avec pertes peuvent être utilisées à des taux raisonnables, sous la direction d'un praticien qualifié, sans aucune réduction significative de la qualité de l'image pour le diagnostic clinique.

Sur la base de ce constat, nous avons mené une étude en partenariat avec le CHRU de Nancy sur des images de type scanner et IRM. La première partie de cette étude avec le CHRU de Nancy est consacrée à des coupes 2D choisies pour essayer de constituer un panel d'images représentatives. Les aspects concernant les piles d'images proprement dits (3D) seront étudiés dans une étude à venir. Suite à l'analyse statistique des résultats des tests subjectifs, nous définissons un seuil de compression maximal jusqu'auquel la qualité des images ne sera pas remise en cause. L'objectif principal de cet article est donc de déterminer les seuils de compression limite pour chaque modalité et chaque région anatomique.

2 Matériel et méthodes pour le test

2.1 Méthodes de compression

Les images médicales de type scanner ou IRM se présentent sous forme de piles de coupes 2D. En exploitant les corrélations qui existent entre les images il est possible par des méthodes de compression dites 3D de compresser de façon plus efficace les données. Dans cette étude préliminaire, nous avons décidé de nous focaliser sur les méthodes de compression 2D. Elles présentent l'avantage d'avoir des temps de compression/décompression plus courts et pour certains standards (JPEG, JPEG 2000) d'être inclus dans la norme DICOM. Nous comparons ici un standard international reconnu JPEG 2000 [4] avec une méthode proposée par deux des auteurs et présentant des résultats intéressants sur les images médicales MEDI-STORAGE (ou QVAZM 2D) [5]. MEDI-STORAGE est une méthode basée sur une transformée en ondelettes avec une étape de quantification s'appuyant sur une Quantification Vectorielle Algébrique avec Zone Morte (QVAZM). Le point clef de cette méthode est la conception d'une zone morte multi-dimensionnelle pendant l'étape de quantification qui permet de prendre en compte les corrélations entre les pixels voisins. La zone morte vectorielle a pour but de seuiller les vecteurs de faible norme afin de quantifier plus finement les vecteurs de coefficients significatifs. Cet algorithme dans sa version 3D (intitulé QVAZM 3D [5]) sur des images scanner et IRM est compétitive avec les meilleurs codeurs 3D en terme de compromis débit-distorsion. Cette méthode permet de mieux préserver les structures fines et produit une image reconstruite d'une meilleure qualité générale. On s'intéresse ici aux performances de la version 2D pour la compression avec pertes des images médicales.

2.2 Conditions d'observation

Pour l'évaluation de la qualité subjective des images, l'ITU impose des normes très strictes sur les conditions d'observation. Ainsi les professionnels de la santé effectuent les tests dans un environnement normalisé et strictement identique pour chacun, le living lab PROMETEE (référéncé au forum des living lab en santé et autonomie). Pour formaliser les conditions d'observation, nous utilisons la norme ITU-BT.500-13 [1]. Les observateurs durant le test sont placés sur un bureau, face à un écran sur lequel sont affichées les images médicales via le logiciel de visualisation d'images DICOM OSIRIX [6]. Chaque participant dispose également d'un iPad permettant d'enregistrer les notations associées à la qualité de chaque image. La session d'essais subjectifs dure 30 minutes environ pendant lesquelles les observateurs ont notés 13 données médicales de type scanner et IRM comme celles visibles figure 1 compressées à 6 qualités pour les 2 méthodes de compression testées. Au début de la séance, on procède à 2 notations fictives pour stabiliser le jugement des observateurs dont on ne tiendra pas compte dans l'analyse des résultats. Enfin, contrairement à l'étude [3], dans laquelle les notations étaient réalisées librement par les médecins sur leurs lieux de travail (dans un temps laissé libre avec des interruptions autorisées), nous nous plaçons ici dans l'environnement normalisé PROMETEE. De plus, les conditions sont strictement identiques pour chacun des panélistes qu'il s'agisse du moniteur ou des paramètres de fenêtrage du logiciel Osirix et les observateurs procèdent aux notations en une seule session.

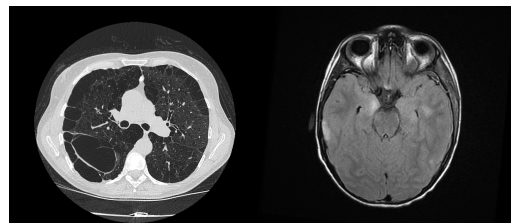


Figure 1 : exemple d'images proposées, à gauche : scanner cou-thorax, à droite : IRM de crane

2.3 Méthode et outils pour l'évaluation

Nous nous sommes là encore référés à la norme ITU-BT.500-13 [1] qui formalise également les méthodes générales d'essais ainsi que les échelles de notation. Le choix entre les différentes méthodes et les échelles dépend de l'application visée. En nous basant sur cette norme nous avons choisi d'effectuer une évaluation de la qualité à stimulus unique à l'aide d'une échelle continue sur laquelle apparait 5 repères sémantiques (Fig.2).

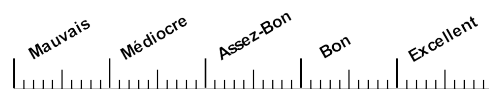


Figure 2 : échelle continue de notation de qualité

Le choix de la méthode à stimulus unique nous permet ici d'observer la réponse du panel à une donnée sans risque de comparaison avec l'original comme c'est le cas avec les méthodes à double stimulus. Le choix du simple stimulus nous permet également de diminuer la durée du test, ainsi les risques d'erreurs liées à la fatigue et de se rapprocher de l'utilisation classique de ce type de logiciel de visualisation DICOM. Il est donc présenté aux observateurs la totalité des images compressées avec JPEG 2000 et MEDI-STORAGE ainsi que les images de référence. Ces images compressées ou non sont réparties aléatoirement selon le principe des carrés gréco-latins afin que deux données issues de la même référence ne soient pas l'une après l'autre.

2.4 Méthode d'analyse des données

Comme recommandé dans la norme ITU-BT.500-13 [1], un test de cohérence est effectué pour déceler d'éventuels observateurs incohérents. Ce test a pour but de normaliser la capacité d'un observateur à répondre de façon cohérente vis-à-vis de l'ensemble du panel. S'il s'avère qu'un observateur répond systématiquement de façon éloignée par rapport au panel, il sera rejeté. Une fois ces étapes franchies nous disposons d'une base de données reprenant les notations des observateurs jugés cohérents, celle ci nous permettra de définir un score MOS (Mean Opinion Score) représentant pour chaque image la note moyenne des observateurs. Ce score d'opinion moyen est l'unité de perception subjective de qualité obtenue pour un panel d'observateurs ayant réalisé un test strictement identique. On considère que cette note MOS est le score de perception de qualité le plus fiable et il est donné par :

$$\bar{u}_{jk} = \frac{1}{N_{obs}} \sum_{i=1}^{N_{obs}} u_{ijk} \quad (1)$$

Où N_{obs} représente le nombre d'observateurs et u_{ijk} la note de l'observateur i pour le taux de compression j de l'image médicale k . On pourra par ailleurs en traçant la courbe représentant l'évolution de la qualité JPEG 2000 et MEDI-STORAGE en fonction de la note subjective retrouver pour chaque image le seuil jusqu'auquel la qualité technique de l'image sera jugée irréprochable.

3 Résultats expérimentaux

Pour cette étude il a été réuni un panel de 9 observateurs ayant de 4 à 28 années d'expérience. Le tableau 1 reprend les différentes caractéristiques de ces observateurs. La totalité du panel s'est révélée cohérente et les résultats qui suivent se baseront donc sur leurs notations subjectives.

Tab 1 : Caractéristiques des observateurs de l'étude

Obs.	Age	Spécialité	Exp.	Sexe
1	27	Radiologie	4	F
2	38	Radiologie	14	F
3	56	Radio Pédiatrie	27	F

4	32	Radiologie	7	M
5	52	Radiologie	28	M
6	39	Radiologie	14	M
7	32	Radiologie	7	M
8	52	Radiodiagnostic	25	M
9	46	Radiologie	20	M

Pour commencer nous mesurons l'évaluation subjective de la qualité des images compressées. Nous avons donc tracé deux courbes reprenant l'évolution de la note MOS en fonction de la qualité pour JPEG2000 et également pour MEDI-STORAGE pour la totalité des images médicales. On note qu'il arrive parfois que l'image compressée soit mieux notée par le panel que la référence non compressée, cela peut s'expliquer de part des phénomènes de lissage qui tendent à atténuer le bruit présent dans l'image référence. De ces courbes nous déterminons pour chaque image et chaque méthode de compression le seuil de compression maximal permettant une qualité technique irréprochable des images. Afin de trouver ce seuil, pour chaque courbe nous avons tracé la droite ayant pour ordonnée la note donnée à l'image originale moins 10% de l'échelle de qualité. Cette variation de 10% du MOS nous permet de nous assurer que le seuil trouvé reste dans une zone où la qualité est jugée parfaite par le panel. L'intersection de cette droite avec la courbe nous donne ensuite le seuil de compression.

Les résultats de la détermination du seuil par cette méthode sur une image médicale de type scanner apparaissent sur la figure 3. Ici par exemple on obtient pour JPEG 2000 un seuil maximal de 1.005 Mbits/pixel soit un taux de compressions de 15.92 : 1, et pour MEDI-STORAGE un seuil maximal de 0.945 Mbits/pixel soit un taux de compression de 16.93 : 1.

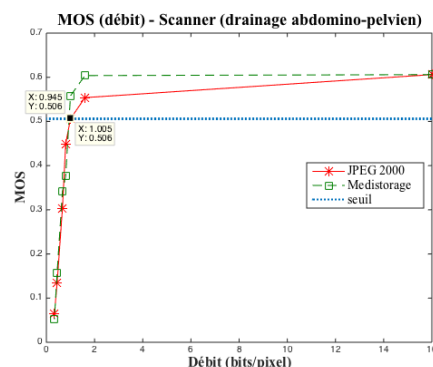


Figure 3 : Note moyenne obtenue pour un SCAN drainage abdomino-pelvien compressé avec les méthodes JPEG2000 et MEDI-STORAGE en fonction du débit en Mbits/pixel.

Les seuils de compression obtenus pour chacune des images et chacune des deux méthodes de compression sont regroupés tableau 2. On peut observer que pour les images de type scanner le taux de compression limite suivant les régions anatomiques varient de 2 à 23 avec JPEG2000 et de 4 à 25 avec MEDI-STORAGE. Pour les IRM ce taux est de 43 à 71 avec JPEG2000 et de 43 à 80 avec MEDI-STORAGE

Les images IRM acceptent donc une compression jusqu'à trois fois plus importante que les scanners et pour ces deux types d'images médicales la méthode

MEDI-STORAGE devance dans la plupart des cas la compression JPEG2000. Les taux de compression trouvés sont cohérents avec l'étude canadienne de 2009 [3].

Tab 2 : Taux de compressions limites pour JPEG2000 et MEDI-STORAGE.

Type	Image médicale	TC limite JPEG2000	TC limite MEDI-STORAGE
512x512 12 bits	cardio-thoraco-abdomino-pelvien	23.36 : 1	24.81 : 1
	crâne #1	19.88 : 1	21.33 : 1
	crâne #2	4.82 : 1	10.03 : 1
	abdomino-pelvien #1	18.18 : 1	18.71 : 1
	abdomino-pelvien #2	23.70 : 1	22.38 : 1
	drainage abdomino-pelvien	15.92 : 1	16.93 : 1
	cou-thorax #1	10.13 : 1	4.60 : 1
	cou-thorax #2	2.37 : 1	11.00 : 1
512x512 9 ou 11 bits	crâne #1 9 bits	71.11 : 1	76.19 : 1
	crâne #2 9bits	43.24 : 1	45.71 : 1
	foie #1 11bits	54.24 : 1	47.06 : 1
	foie #2 11bits	57.14 : 1	72.73 : 1
	Biliaire 9 bits	71.11 : 1	78.05 : 1

Tab 3 : Coefficients de corrélation de Pearson entre mesures objectives et subjectives pour les différentes images médicales - LCC : moyenne des coefficients de corrélation de Pearson

	JPEG2000				MEDI-STORAGE			
	LCC	Rang	LCC	Rang	LCC	Rang	LCC	Rang
	SCAN		IRM		SCAN		IRM	
SSIM [7]	0.938	8	0.926	2	0.908	7	0.898	2
PSNR [7]	0.946	5	0.838	7	0.940	5	0.842	7
MSE [7]	0.924	9	0.942	1	0.892	9	0.900	1
MSSIM [7]	0.940	7	0.914	3	0.917	6	0.889	3
NQM [7]	0.952	2	0.850	4	0.894	8	0.866	5
PSNRHVS [2]	0.951	4	0.839	5	0.943	2	0.864	6
PSNRHVM [2]	0.944	6	0.838	6	0.942	4	0.874	4
VIF [7]	0.952	2	0.806	9	0.943	2	0.819	9
VIFP [7]	0.954	1	0.830	8	0.947	1	0.826	8

Les tests subjectifs étant coûteux à mettre en œuvre nous avons également en parallèle observé les résultats donnés par des métriques objectives de qualité. Nous avons observé la corrélation qui existe entre ces différentes métriques objectives et les notations subjectives obtenues par le panel, ceci dans le but de définir la ou les métriques les plus adaptées à chaque type d'images médicales (modalité, région anatomique). Pour chacune d'elles on calcule le coefficient de corrélation de Pearson LCC (Linear Correlation Coefficient) indiquant la qualité de la régression linéaire. L'ensemble de ces métriques est listé dans le tableau 3, et la plupart d'entre elles sont disponibles dans la bibliothèque Matlab Metrix Mux [7].

On remarque que ce ne sont pas les mêmes métriques qui sont corrélées aux résultats subjectifs pour les scanners et les IRMs. Ainsi pour les scanners VIF et VIFp se retrouvent en tête, ces deux métriques exploitent les statistiques de scènes naturelles et la pondération par une fonction de sensibilité aux

contrastes (CSF). Au contraire pour les IRMs les métriques les plus corrélées sont SSIM et MSSIM qui se basent sur les changements structurels de l'image ainsi que MSE qui est le calcul de l'erreur quadratique moyenne donc qui effectue une comparaison pixel par pixel de deux images.

4 Conclusions et perspectives

De cette étude nous confirmons que la compression avec pertes est souhaitable pour des images médicales (IRM, scanner) sans altérer leur qualité technique et leur bonne utilisation. Nous avons également comparé deux méthodes de compression et nous avons pu observer que sur les images de notre étude la méthode MEDI-STORAGE était plus efficace que la méthode JPEG 2000. Nos perspectives portent sur l'extension de cette étude aux méthodes 3D ainsi que sur le test de nouveaux algorithmes de compression comme HEVC.

5 Remerciements

Nous tenons à remercier l'ensemble des médecins qui ont pris part à notre étude en acceptant de réaliser des sessions d'essais subjectifs. Par ailleurs, le Living Lab PROMETEE bénéficie du soutien du projet européen Celtic Plus E3 (E-health services Everywhere and for Everybody).

Références

- [1] ITU-R. *Recommendation 500-13, Methodology for the subjective assessment of the quality of television pictures*. ITU-R Rec - BT.500 (2012).
- [2] Gaudeau Y., Lambert J., Labonne N., and Moureaux J.M. *Compressed image quality assessment: application to an interactive upper limb radiology atlas*. IEEE International Conference on Image Processing, ICIP 2014, Paris, France, Oct. 2014.
- [3] D. Koff and al. *Pan-Canadian Evaluation of Irreversible Compression Ratios (Lossy Compression) for Development of National Guidelines* Journal of Digital Imaging, vol. 22, 6, pp. 569-578, Déc. 2009.
- [4] D.S. Taubman, M.W. Marcellin. *JPEG2000: Image Compression Fundamentals, Standards, and Practice*. Kluwer Academic Publishers, Boston, 2002.
- [5] Y. Gaudeau et JM. Moureaux. *Lossy compression of volumetric medical images with 3D dead-zone lattice vector quantization*. Annals of Telecommunications, vol. 64, 5-6, pp. 359-367, Avr. 2009.
- [6] A.Rosset, L.Spadola, O.Ratib. *An Open-Source Software for Navigating in Multidimensional DICOM Images*. J Digit Imaging. 17(3), pp. 205-216 Sep. 2004.
- [7] M. Gaubatz. *Metrix Mux Visual Quality Assessment Package*, Disponible sur foulard.ece.cornell.edu