On-board MRI image compression using video encoder for MR-guided radiotherapy

Jiawen Shang#, Peng Huang#, Ke Zhang#, Jianrong Dai, Hui Yan

Department of Radiation Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

Contributions: (I) Conception and design: J Shang, P Huang, K Zhang; (II) Administrative support: H Yan, J Dai; (III) Provision of study materials or patients: P Huang, K Zhang; (IV) Collection and assembly of data: P Huang; (V) Data analysis and interpretation: J Shang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

#These authors contributed equally to this work and should be considered as co-first authors.

Correspondence to: Hui Yan, PhD. Department of Radiation Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Panjiayuannanli 17, Chaoyang district, Beijing 100021, China. Email: hui.yan@cicams.ac.com.

Background: Magnetic resonance imaging (MRI) is currently used for online target monitoring and plan adaptation in modern image-guided radiotherapy. However, storing a large amount of data accumulated during patient treatment becomes an issue. In this study, the feasibility to compress MRI images accumulated in MR-guided radiotherapy using video encoders was investigated.

Methods: Two sorting algorithms were employed to reorder the slices in multiple MRI sets for the input sequence of video encoder. Three cropping algorithms were used to auto-segment regions of interest for separate data storage. Four video encoders, motion-JPEG (M-JPEG), MPEG-4 (MP4), Advanced Video Coding (AVC or H.264) and High Efficiency Video Coding (HEVC or H.265) were investigated. The compression performance of video encoders was evaluated by compression ratio and time, while the restoration accuracy of video encoders was evaluated by mean square error (MSE), peak signal-to-noise ratio (PSNR), and video quality matrix (VQM). The performances of all combinations of video encoders, sorting methods, and cropping algorithms were investigated and their effects were statistically analyzed.

Results: The compression ratios of MP4, H.264 and H.265 with both sorting methods were improved by 26% and 5%, 42% and 27%, 72% and 43%, respectively, comparing to those of M-JPEG. The slice-prioritized sorting method showed a higher compression ratio than that of the location-prioritized sorting method for MP4 (P=0.00000), H.264 (P=0.00012) and H.265 (P=0.00000), respectively. The compression ratios of H.265 were improved significantly with the applications of morphology algorithm (P=0.01890 and P=0.00530), flood-fill algorithm (P=0.00510 and P=0.00020) and level-set algorithm (P=0.02800 and P=0.00830) for both sorting methods. Among the four video encoders, H.265 showed the best compression ratio and restoration accuracy.

Conclusions: The compression ratio and restoration accuracy of video encoders using inter-frame coding (MP4, H.264 and H.265) were higher than that of video encoders using intra-frame coding (M-JPEG). It is feasible to implement video encoders using inter-frame coding for high-performance MRI data storage in MR-guided radiotherapy.

Keywords: Magnetic resonance imaging (MRI); magnetic resonance-guided radiotherapy (MR-guided radiotherapy); video encoder; auto-segment; data compression

Introduction

Magnetic resonance imaging (MRI) is a non-invasive imaging technology used for disease detection, diagnosis, and treatment monitoring. Differing from computed tomography (CT), it does not use the ionizing radiation of X-rays and has excellent resolution of the brain and soft tissues (1). MRI has shown its clinical value in radiation therapy due to its higher soft tissue contrast that allows better localization of tumor without additional radiation exposure to patient. Pre-treatment and on-board MRI allows for online verification and adjustment of treatment plan, which consequently improves the accuracy and safety of treatment in radiotherapy (2-4).

On-board imaging can provide a real-time view of how a patient's lesion is located while patient lies on the treatment couch for clinical decision-making (5-7). However, this means that medical institutions need large storage space to store these data. Similar to cone-beam computed tomography (CBCT) used in daily treatment verification of radiotherapy, on-board MRI accumulate a large amount of data in a short period of time. For MR-guided radiotherapy in our institute, at least three MRI scans are acquired before treatment, in the middle of treatment, and after treatment while patient lies on the treatment couch. If plan adjustment and verification are required, more MRI scans may be needed to double check the accuracy of the adaptive plans (8). If a treatment consisting of 5–6 fractions (average 3–5 MRI scans per fraction), the total number of MRI scans is 20–25 and the total storage space is about 3 GB. Therefore, there is a demand for high-volume MRI data storage.

Compression algorithms can significantly reduce the storage space occupied by large files, while preserving the original features of these files as much as possible. For static image, there are hundreds of compression algorithms which can be categorized into two categories: lossless and lossy. Among them, discrete cosine transform (DCT) and wavelet transform (WT) based compression algorithms are the most popular ones, which were both adopted by Digital Imaging and Communications in Medicine (DICOM) in 2001 (9,10). Lossless compression is also known as reversible compression of which the content of image can be restored completely after decompression. Lossless encoders include Huffman coding and Context-based Adaptive Binary Arithmetic Coding (CABAC) (10,11). Lossy compression is irreversible compression of which the contents of image cannot be restored completely after decompression. Lossy encoders include JPEG and MPEG-4 (MP4) (12).

Differing from static images, video consists of a series of images obtained in consecutive time frames. Video compression exploits redundancy of information in and between frames, i.e., intra-frame and inter-frame correlation. Accordingly, video encoders can be divided into two categories: intra-frame coding and inter-frame coding. Intra-frame coding processes each frame independently which exploits redundant information within a frame (13,14). The inter-frame coding processes multiple consecutive frames which exploits redundant information between them (15). Advanced Video Coding (AVC) and High Efficiency Video Coding (HEVC) are the most popular video encoders with inter-frame coding. AVC is Motion-JPEG-4 (MPEG-4) Part 10 or H.264. Its successor HEVC is also called H.265, which provides even higher compression ratio while retaining the equivalent video quality (16).

The feasibility to apply video encoders for compressing CBCT and 4DCT in radiotherapy was investigated in our previous studies (17-19). CBCT is routinely used for daily verification of patient position during treatment. It can be compressed by video encoders with higher compression ratio (18). To further exploit the correlation relationship among multiple image sets, the sorting methods to reorder images in the input sequence of video encoder were employed (19,20). The result showed that the higher compression ratio can be achieved by optimizing image order in the input sequence of video encoder.

In this study, the new video encoder (HEVC/H.265) was introduced and compared with the other three popular video encoders. Three cropping algorithms were used to auto-segment regions of interest for the separate data storage. The content of this paper was arranged as follows. In method, MRI data acquisition in the process of MR-guided radiotherapy was introduced. The two sorting methods and three cropping algorithms were explained in detail. In results, the compression performances of the four video encoders were compared. The effects of sorting methods and cropping algorithms on the compression
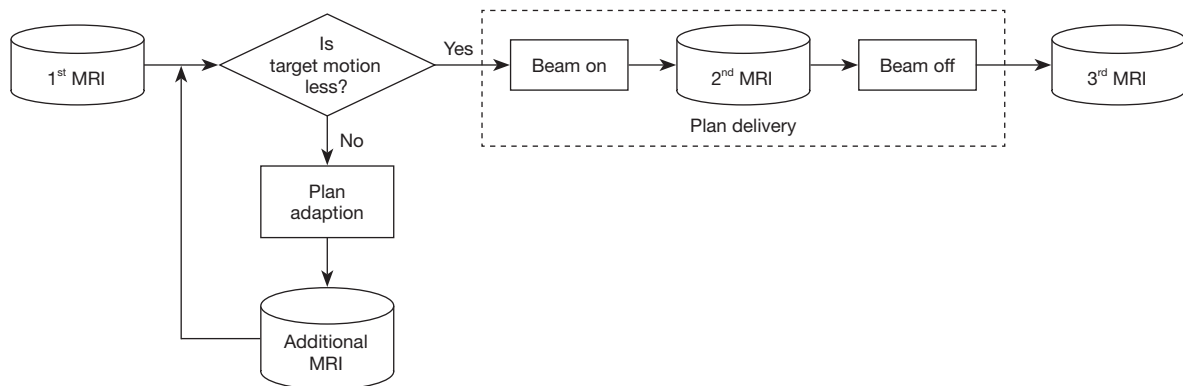
Figure 1 The scheme of MRI-guided radiotherapy on Elekta Unity. MRI, magnetic resonance imaging.

Table 1 The number of MRI scans per fraction for five patients

Patient ID	Total fractions	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
1	6	3	4	5	4	3	4
2	5	5	4	3	3	3	0
3	5	6	5	5	5	4	0
4	5	6	4	4	4	4	0
5	5	3	5	3	6	5	0

MRI, magnetic resonance imaging.

performance and restoration accuracy were evaluated and analyzed. Finally, the merits and disadvantages of the video compression method were discussed.

Methods

Data acquisition

The magnetic resonance linear accelerator (MR LINAC), Unity (Elekta Solutions AB, Stockholm, Sweden), was installed in our institute in 2018. It is a slip ring gantry-mounted 7 Mega-voltage (MV) linear accelerator combined with a 1.5 T MR scanner (21,22). The scheme of MR-guided radiotherapy on Elekta Unity is illustrated in *Figure 1*. The patient was first positioned on treatment couch and then the 1st MRI scan was performed to verify the target location. 3 mm margin was usually applied to target volume to account for the offset caused by patient positioning and movement (8). If target offset was less than 3 mm, the current plan was used and the 2nd MRI scan was performed during beam delivery to monitor the target shift. After delivery of treatment plan, the 3rd MRI scan was performed

to verify the target location again. If target offset was more than 3 mm, the new plan, also known as the adaptive plan, was generated by adjusting the current plan. Before delivery of the new plan, another MRI scan was performed and compared to the reference image to verify the target location. If the target location was correct, then the new plan was delivered to patient. Otherwise, the new plan would be adjusted again using the latest MRI. This plan adaptation process continued until target offset was within tolerance.

T2-weighted MRIs were collected from patients under MR-guided radiotherapy in our institute (8). Each MRI consisted of 300 slices with a pixel resolution of 480×480. The MRI slice thickness was 2 mm and in-plane spacing was 1 mm. The average size of MRI on disk was 0.2 GB. Totally, 110 image sets were acquired from five patients. These patients were diagnosed with liver cancer and treated on MR-LINAC. The slices of MRI were stored in DICOM format and their contents were extracted by DICOM parser tools in this study. The number of MRIs acquired per fraction for the five patients is listed in *Table 1*. If there was no plan adaptation, 3 MRI scans per fraction were
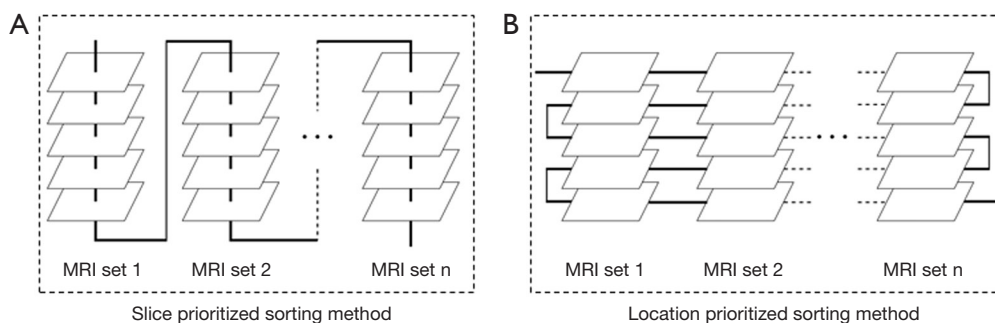
Figure 2 The illustrations of two sorting methods. (A) Slice prioritized sorting method. (B) Location prioritized sorting method. MRI, magnetic resonance imaging.

performed as shown in patients 1, 2 and 5. If plan adaptation was needed, 4–6 MRI scans per fraction were performed as shown in all five patients. This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The institutional Ethics Committee approved this study of the Cancer Hospital, Chinese Academy of Medical Sciences, and Peking Union Medical College. Informed consent was waived in this retrospective study.

Image sorting methods

As shown in *Table 1*, multiple MRIs were acquired during daily treatment of MR-guided radiotherapy. As patient lied on the treatment couch in the same position during multiple MRI scans, these scanned images were highly similar in spatially and temporarily. Therefore, their correlation relationship could be utilized to reduce the redundant information between images. Accordingly, it would be beneficial to reorder these images in the input sequence for the improved inter-frame similarity.

In this study slice-prioritized (SP) and location-prioritized (LP) sorting methods were employed (18-20). The SP method reorders the images according to their indexes in the original image sets. As shown in *Figure 2A*, for combining two image sets into one sequence, the last slice of the first image set is connected to the first slice of the second image set. The SP method is straightforward by connecting multiple image sets end to end. The LP method reorders images according to their layers along longitudinal axis of human body. As shown in *Figure 2B*, the images at the same layer of human body in multiple image sets are connected. As multiple MRIs were acquired within 1–2 hours while patient lied on the treatment couch in the same position, the same slice index in different image

sets would indicate the same layer of human body. The LP method reorder images in another way by connecting multiple image sets layer after layer.

Image cropping algorithms

As shown in *Figure 3*, the voxel outside the human body was mostly air and supporting devices such as couch and bracket. As the background is of no interest for patient positioning in MR-guided radiotherapy, it can be omitted in data storage. Cropping algorithms can identify the boundary between human body and background objects, and then fill the background with a unique value. It should be noted that cropping algorithm might wrongly identify background pixels, and would cause loss of useful clinically relevant information for diagnosis and treatment planning. Therefore, visual checking is needed.

In this study three cropping algorithms, morphology, flood-fill and level-set, were investigated (22-24). The morphology algorithm performs several consecutive open operations on the images. The noisy parts at the periphery of the images are filtered out and the body parts at the center are retained. Due to the discontinuity of the separated body area, an external rectangular box is used to limit the cropping area as shown in *Figure 3A*. The flood-fill algorithm forms regions by collecting pixels with similar properties. The algorithm first finds a seed in the region to be segmented as the starting point, and then iteratively merges the pixels in the surrounding neighborhoods that are similar to the seed pixels. As shown is *Figure 3B* certain air volume near the body surface is not removed correctly because both volumes near the body surface had similar intensities. Note that in this study seeds were automatically set at fixed points in background area far from regions
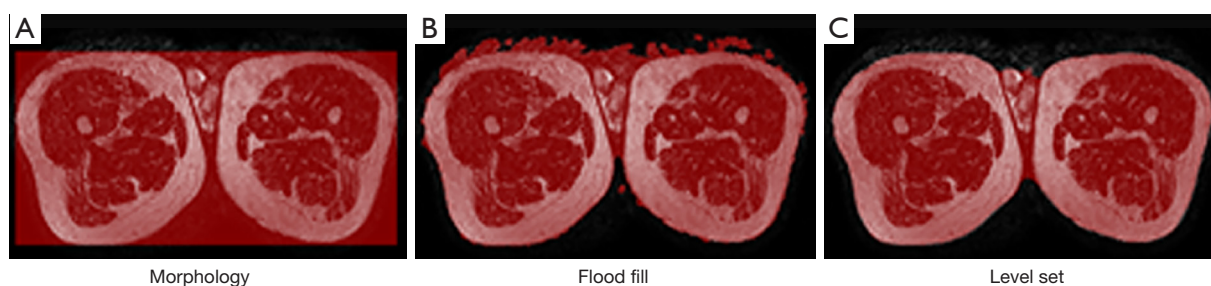

Figure 3 The illustration of segmentation results in axial view by three cropping algorithms. (A) Morphology. (B) Flood fill. (C) Level set.
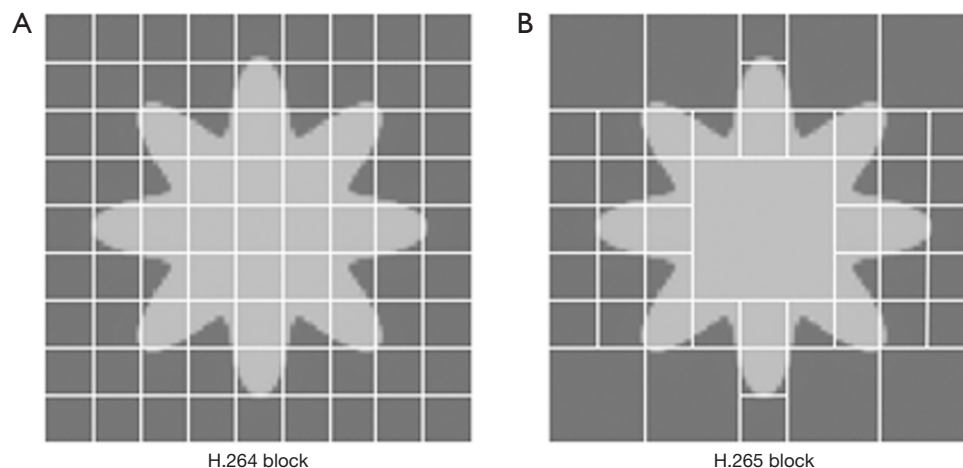

Figure 4 The block division used in video encoders H.264 and H.265. (A) H.264 block. (B) H.265 block.

of interest. The level-set method is an unsupervised segmentation method (23,24). It uses the pixel gray-scale information as energy and a gradient flow to minimize the energy function (25). Compared with the first two cropping algorithms, level-set based algorithm can avoid segmentation errors introduced by similar image intensities near the body surface. As shown in *Figure 3C* under the curve constraint specific to the level set, the body surface was identified accurately.

Video encoders

Four video encoders were investigated including: (I) M-JPEG which is a lossy DCT-based JPEG encoder using intra-frame coding; (II) MP4 which is lossy MPEG encoder using inter-frame coding; (III) AVC which is lossy encoder based upon the MPEG-4 technology using inter-frame coding and also called H.264; (IV) HEVC which is lossy encoder designed for the latest generation

of high-resolution video using inter-frame coding and also called H.265.

JPEG and MP4 are the mostly popular compression algorithms adopted by industry. Their details can be found in literatures (18-20). It should be noted that current MP4 encoder is based on H.264 standard and is going to be replaced by H.265 standard which provides better video quality at the same bit rate. Unlike H.264 algorithm that assigns all blocks the same size H.265 algorithm uses a more flexible coding structure to improve the coding efficiency. As shown in *Figure 4* it uses a recursive structure to further divide the blocks, which makes the division more dependent on the video image texture features. After division of each frame, the best matching block is found in the search range of adjacent frames so that the difference between the predicted block and the current block is minimized. The result of this search is used as a motion vector, which represents the relative offset between two images, to generate the predicted image. Thus, the video

can be compressed substantially by storing those motion vectors.

Experiments

Four video encoders were tested together with two sorting methods and three cropping methods. The MRI sets acquired on the same day was combined in a sequence by a sorting method for the input of video encoder. As shown in *Table 1*, total 26 image sequences were obtained from five patients. The performance metrics were averaged based on all 26 image sequences for a video encoder with a selected sorting method. To evaluate the performance difference among all combinations of video encoders, sorting methods, and cropping algorithms, the corrected P values for multiple comparisons were provided. Statistical analyses were performed in R (version 3.6.3).

The tests were performed on a personal computer equipped with Intel i5 CPU 2.6-GHz and 8 GB RAM. The programs for data processing were developed with Python (version 3.8.10) and ffmpeg (https://ffmpeg.org/) which is an open-source audio and video converter tool available for multiple operation platforms. For M-JPEG and MP4, the default configuration was used to allow video data to be compressed as much as possible. For H.264 and H.265, the constant rate factor (CRF) was set to 22 to maintain similar accuracy and video quality. CRF is the quality control setting for the encoders. Lower values would result in better quality at the expense of higher file sizes. The default values for H.264 and H.265 are 23 and 28, respectively.

The compression performance of video encoders was evaluated by compression ratio and compression time. Compression ratio is defined as the ratio between the sizes of image sequence and its resulted video file. Compression time is the time in processing one image and calculated by the total processing time dividing the total number of images in an input sequence.

To evaluate the average inter-frame similarity of a sequence, two similarity metrics were employed (20). The inter-frame difference (DIFF) calculates the average pixel differences for all pairs of adjacent images in a sequence (20).

$$DIFF = \frac{1}{L-1} \sum_{K=1}^{L-1} \left(\frac{\sum_{i=1}^{M} \sum_{j=1}^{N} \left| I_{ij}^k - I_{ij}^{k+1} \right|}{M \times N} \right) \quad [1]$$

Where L is the number of images in a sequence, M and N are the height and width of image, I_{ij}^k is the pixel (i,j) in

image k. The inter-frame correlation (CORR) calculates the average Pearson correlation coefficient for all pairs of adjacent images in a sequence (20).

$$CORR = \frac{1}{L-1} \sum_{K=1}^{L-1} \left(\frac{\sum_{i=1}^{M} \sum_{j=1}^{N} \left(I_{ij}^k - \overline{I^k} \right) \left(I_{ij}^{k+1} - \overline{I^{k+1}} \right)}{\sqrt{\sum_{i=1}^{M} \sum_{j=1}^{N} \left(I_{ij}^k - \overline{I^k} \right)^2} \sqrt{\sum_{i=1}^{M} \sum_{j=1}^{N} \left(I_{ij}^{k+1} - \overline{I^{k+1}} \right)^2}} \right) \quad [2]$$

$\overline{I^k}$ and $\overline{I^{k+1}}$ are the mean values of images I^k and I^{k+1}. The equation inside the bracket is the Pearson correlation coefficient between I^k and I^{k+1}.

In order to evaluate the restoration accuracy of image decompression, mean square error (MSE), peak signal-to-noise ratio (PSNR), and video quality matrix (VQM) were employed. The MSE calculates the difference between the original image and the restored image after decompression (26).

$$MSE = \frac{1}{L} \sum_{k=1}^{L} \left(\frac{\sum_{i=1}^{M} \sum_{j=1}^{N} \left(\overline{I_{ij}^{k,de}} - \overline{I_{ij}^k} \right)^2}{M \times N} \right) \quad [3]$$

$\overline{I_{ij}^k}$ is the normalized image and $\overline{I_{ij}^{k,de}}$ is the normalized image after decompression. PSNR is the ratio between the maximum power of a signal and the power of corrupting noise that affects the fidelity of its representation as defined below (26).

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX^2}{MSE} \right) [dB] \quad [4]$$

Where MAX is the maximum possible pixel value of the image and $2^{16}-1$ for 16 bit unsigned integer in this study. Typical values for the PSNR of lossy image and video compression are between 60 and 80 dB, provided the bit depth is 16 bits. VQM is a metric to predict human perceived video quality as defined below (26).

$$VQM = \frac{1}{1 + e^{a(PSNR-b)}} \quad [5]$$

Where a =0.15 and b =19.7818 have been set experimentally. The resulting VQM is compared to fuzzy results like "excellent" (VQM <20%) or "good" (VQM <40%).

Results

Performance of video encoders

The compression performances of four video encoders are compared in *Table 2*. In general, the compression ratio of the video encoders (MP4, H.264 and H.265) with

Table 2 Comparison of compression performance of four video encoders

Encoder	Sorting	CR	T (ms)
M-JPEG	N/A	19.42±0.89	1.79±0.18
MP4	SP	24.63±1.42	2.51±0.22
	LP	20.47±0.98	2.18±0.15
H.264	SP	27.49±2.28	33.43±3.25
	LP	24.66±2.66	27.86±2.12
H.265	SP	33.43±3.25	16.42±1.02
	LP	27.86±2.12	16.08±0.95

Results were reported as mean ± standard deviation. CR, compression ratio; T, time; N/A, not available; SP, slice prioritized sorting method; LP, location prioritized sorting method.

inter-frame coding was higher than that of video encoder (M-JPEG) with intra-frame coding. Comparing with M-JPEG, the improvements of compression ratios by MP4, H.264 and H.265 with SP and LP sorting methods were 26% and 5%, 42% and 27%, 72% and 43%, respectively. Among the four video encoders, H.265 showed the highest compression ratios for both sorting methods. H.264 and MP4 showed higher compression ratios than those of M-JPEG. Comparatively, the video encoders with inter-frame coding used more time on data processing. The time spent on data processing was proportional to the magnitude of compression ratio, i.e., the higher compression ratio, the more time required on data processing. The restoration accuracy of four video encoders was shown in *Table 3*. The restoration accuracy of M-JPEG was slightly less than those of the other three video encoders (MP4, H.264 and H.265). The restoration accuracies of MP4, H.264 and H.265 were similar. H.265 had the best restoration accuracy among all four video encoders.

Effect of sorting methods

The similarity metrics (DIFF and CORR) were calculated and averaged with respect to each sorting method. The mean values of DIFF were 5.38±1.89 and 14.52±5.19 for SP and LP sorting methods, respectively. The mean values of CORR were 0.99±0.03 and 0.90±0.06 for SP and LP sorting methods, respectively. It showed that the DIFF and CORR resulted by SP sorting method were higher than those resulted by LP sorting method. This implied that SP sorting method may be beneficial for high-performance

compression tasks.

The corrected P values for compression ratios of three video encoders using two sorting methods were shown in *Figure 5A*. The SP sorting method presented higher compression ratios than those of the LP sorting method. The differences of compression ratios between SP and LP sorting methods for MP4 (P=0.00000), H.264 (P=0.00012) and H.265 (P=0.00000) were statistically significant, respectively. The differences of compression ratios among H.265, H.264, and MP4 were significant for both sorting method. H.265 has the highest value of compression ratio among these three video encoders. The corrected P values for MSE of three video encoders (H.265, H.264, and MP4) with two sorting methods were shown in *Figure 5B*. The video encoders (H.264 and H.265) using SP sorting method showed higher restoration accuracy (lower MSE, higher PSNR and lower VQM). H.265 had the lowest value of MSE among these three video encoders with inter-frame coding.

Effect of cropping algorithms

The effects of cropping methods on compression performance and restoration accuracy of H.265 were shown in *Table 4*. The compression ratios of H.265 using three cropping algorithms were improved significantly. Among the three cropping algorithms, level-set algorithm showed the highest compression ratios (35.71 and 29.81 for SP and LP sorting methods, respectively). The compression ratios of H.265 (35.36 and 29.67 for SP and LP sorting methods, respectively) using flood-fill algorithm was close to the compression ratios of H.265 (35.71 and 29.81 for SP and LP sorting methods, respectively) using level-set algorithm. The restoration accuracies of H.265 [(1.26E–05, 97.15, 9.14E–06) and (1.46E–05, 95.50, 1.01E–05) for SP and LP sorting methods, respectively] using flood-fill algorithm were higher than those of H.265 [(1.29E–05, 97.06, 9.27E–06) and (1.49E–05, 96.44, 1.02E–05) for SP and LP, respectively] using level-set algorithm.

The corrected P values for compression ratios of H.265 with three cropping algorithms were shown in *Figure 6A*. The compression ratios of H.265 between two sorting methods were significant for three cropping algorithms. The SP sorting method showed higher compression ratio of H.265 than LP sorting method. The improvements of compression ratios were statistically significant with the applications of morphology algorithm (P=0.01890 and P=0.00530), flood-fill algorithm (P=0.00510 and P=0.00020)

5214

Shang et al. Compressing on-board MRI by advanced video coding methods

Table 3 Comparison of restoration accuracy of four video encoders

Encoder	Sorting	MSE	PSNR	VQM
M-JPEG	N/A	2.19E−05±5.03E−06	94.83±0.92	1.3E−05±1.90E−06
MP4	SP	1.6E−05±5.34E−07	96.09±0.14	1.07E−05±2.31E−07
	LP	1.6E−05±5.92E−07	96.09±0.16	1.07E−05±2.57E−07
H.264	SP	1.51E−05±1.78E−06	96.38±0.51	1.03E−05±7.91E−07
	LP	1.69E−05±1.78E−06	95.88±0.45	1.11E−05±7.53E−07
H.265	SP	1.38E−05±1.58E−06	96.75±0.49	9.71E−06±7.19E−07
	LP	1.57E−05±1.97E−06	96.19±0.53	1.06E−05±8.53E−07

Results were reported as mean ± standard deviation. MSE, mean square error; PSNR, peak signal-to-noise ratio; VQM, video quality matrix; N/A, not available; SP, slice prioritized sorting method; LP, location prioritized sorting method.
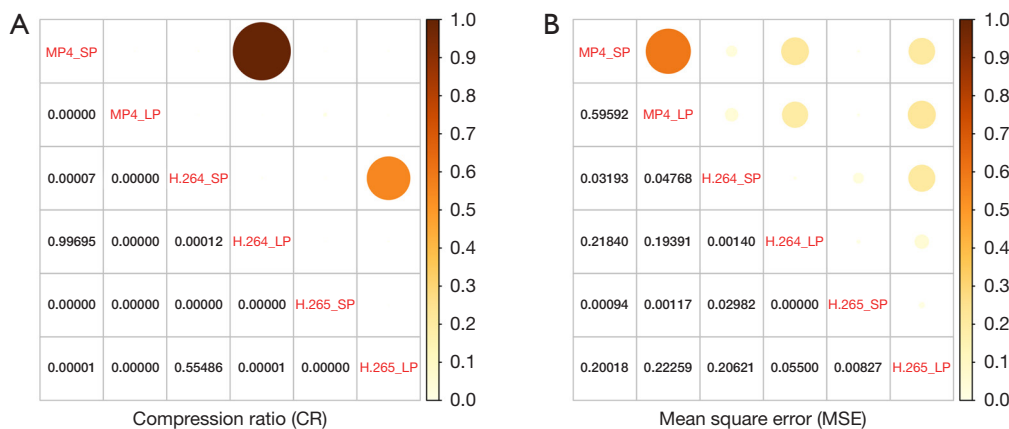

Figure 5 The corrected P value for multiple comparison in evaluating the performance differences of three video encoders (MP4, H.264, and H.265). (A) Compression ratio and (B) mean square error. A level of P<0.05 was considered statistically significant.

Table 4 The effect of three cropping algorithms on compression performance and restoration accuracy of H.265 video encoder

Cropping	Sorting	CR	T (ms)	MSE	PSNR	VQM
H.265	SP	33.43±3.25	16.42±1.02	1.38E−05±1.58E−06	96.75±0.49	9.71E−06±7.19E−07
	LP	27.86±2.12	16.08±0.95	1.57E−05±1.97E−06	96.19±0.53	1.06E−05±8.53E−07
H.265_M	SP	34.95±3.60	15.51±1.06	1.42E−05±1.50E−06	96.62±0.46	9.9E−06±6.78E−07
	LP	29.07±2.39	15.16±0.99	1.63E−05±1.87E−06	96.04±0.49	1.08E−05±8.02E−07
H.265_F	SP	35.36±3.56	15.48±0.95	1.26E−05±1.35E−06	97.15±0.47	9.14E−06±6.38E−07
	LP	29.67±2.37	14.94±1.01	1.46E−05±1.73E−06	95.50±0.50	1.01E−05±7.69E−07
H.265_L	SP	35.71±3.65	17.40±2.49	1.29E−05±1.34E−06	97.06±0.45	9.27E−06±6.28E−07
	LP	29.81±2.39	17.93±4.28	1.49E−05±1.72E−06	96.44±0.49	1.02E−05±7.62E−07

Results were reported as mean ± standard deviation. H.265_M is H.265 using morphology algorithm; H.265_F is H.265 using flood-fill algorithm; H.265_L is H.265 using level-set algorithm. CR, compression ratio; T, time; MSE, mean square error; PSNR, peak signal-to-noise ratio; VQM, video quality matrix; SP, slice prioritized sorting method; LP, location prioritized sorting method.
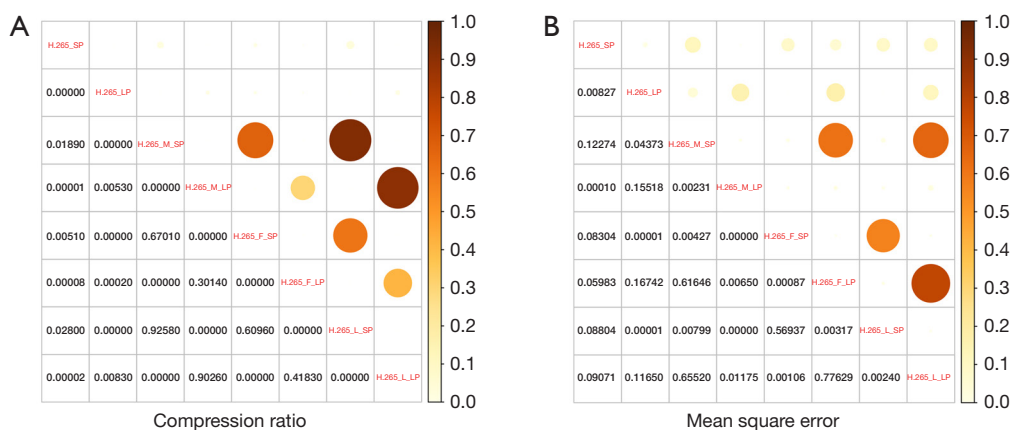
Figure 6 The corrected P value for multiple comparison in evaluating the performance differences of three cropping algorithms and two sorting methods. (A) Compression ratio and (B) mean square error. A level of P<0.05 was considered statistically significant.

and level-set algorithm (P=0.02800 and P=0.00830) for both sorting methods. The compression ratio differences among the three cropping algorithms were insignificant. Among the three cropping algorithms, the level-set algorithm had the highest value of compression ratio. The corrected P values for MSE of H.265 using the three cropping algorithms were shown in *Figure 6B*. The video encoders with SP sorting method showed higher restoration accuracies than those of video encoders with LP sorting method. Among the three cropping algorithms, the flood-fill method had the lowest value of MSE.

Discussion

While comparing video encoders with inter-frame coding to that with intra-frame coding, it was found that the compression ratio of formers was consistently higher than that of latter. This is because the redundant information presented between frames can be effectively removed by motion prediction mechanism employed in the inter-frame coding algorithms (18-20). For movie images, this is caused by the continuity of content that exists in temporal space. For medical images such as CT and MRI, this continuity is a natural property that exists in spatial space. Therefore, video encoders with inter-frame coding are highly preferred for high-performance medical image compression tasks.

The most recent video encoders provide higher compression ratio than the older ones. In this study, the compression ratios of H.264 were higher than that of MP4 by 11.6% and 24.4% for SP and LP sorting methods, respectively. Its successor, H.265, showed even higher

compression ratios than that of MP4 by 35.7% and 36.1% for SP and LP sorting methods. The compression ratios of H.265 were also higher than that of H.264 by 21.4% and 13.1% for SP and LP sorting methods, respectively. Comparing to our previous results that the compression ratios of H.264 for 4DCT/CBCT were ~200–300 (18-20), the compression ratios (~20–30) of H.264 for MRI were relatively low. This is because 4DCT and CBCT usually present more bony structures and less noise, while MRI comes with a large quantity of noise, contrast difference, and deformation of soft tissues. These factors would result in poor inter-frame correlation between slices (15).

Two sorting methods were employed in this study for improving the inter-frame correlation quality of the input sequences for video encoder. The inter-frame similarity metrics (DIFF and CORR) were evaluated on both sorting methods. In general, the SP sorting method provided higher compression ratio and restoration accuracy than those of the LP sorting method. It indicates that the similarity between two slices at consecutive locations in one image set was higher than that between two slices at the same layer in two image sets. The reason could be that there is higher amplitude of organ motion during treatment which causes poor correlation between the slices at the same layer of two image sets.

The cropping algorithms were introduced to further improve the compression ratio by removing the volume of air around the human body in the image. All three cropping algorithms showed significant improvements of compression ratios for H.265. Among them, the level-set algorithm showed the highest value of compression ratio.

However, it is worth noting that the level-set algorithm took longer time to obtain the best segmentation due to its higher computational complexity. With less processing time the flood-fill algorithm provided similar compression performance and restoration accuracy, which is an efficient alternative to level-set algorithm.

In term of the processing time, M-JPEG with intra-frame coding took least time due to the fact that its algorithm does not deal with inter-frame correlation and motion prediction. The video encoders with inter-frame coding took more time on block searching, motion vector prediction, data interpolation, etc. It should be noted that with the introduction of the new video encoder, the compression time could be increased significantly. For example, the processing time of H.264 was 2.64 times to that of MP4, while the processing time of H.265 was 6.5 times to that of MP4. As a tradeoff, the increase of processing time would lead to better compression ratios and restoration accuracy.

Conclusions

The video encoders with inter-frame coding provide an effective way to store high volume of MRI data accumulated in MR-guided radiotherapy. It can significantly reduce the storage space on disk while retaining the vast majority of image content. For multiple daily MRI image sets, SP sorting method provided a higher compression ratio than that of LP sorting method. Among the four video encoders, H.265 demonstrated the best performance in both compression and restoration accuracy, which would be an ideal choice for high-performance MRI data storage.

Acknowledgments

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims. amegroups.com/article/view/10.21037/qims-22-1378/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The institutional Ethics Committee approved this study of the Cancer Hospital, Chinese Academy of Medical Sciences, and Peking Union Medical College. Informed consent was waived in this retrospective study.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Yousaf T, Dervenoulas G, Politis M. Advances in MRI Methodology. Int Rev Neurobiol 2018;141:31-76.
2. Pommier P, Gassa F, Lafay F, Claude L. Image guided radiotherapy with the Cone Beam CT kV (Elekta): experience of the Léon Bérard centre. Cancer Radiother 2009;13:384-90.
3. Bergom C, Prior P, Kainz K, Morrow NV, Ahunbay EE, Walker A, Allen Li X, White J. A phase I/II study piloting accelerated partial breast irradiation using CT-guided intensity modulated radiation therapy in the prone position. Radiother Oncol 2013;108:215-9.
4. Witt JS, Rosenberg SA, Bassetti MF. MRI-guided adaptive radiotherapy for liver tumours: visualising the future. Lancet Oncol 2020;21:e74-82.
5. Chen C, Qin C, Qiu H, Tarroni G, Duan J, Bai W, Rueckert D. Deep Learning for Cardiac Image Segmentation: A Review. Front Cardiovasc Med 2020;7:25.
6. Seo H, Badiei Khuzani M, Vasudevan V, Huang C, Ren H, Xiao R, Jia X, Xing L. Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications. Med Phys 2020;47:e148-67.
7. Sotiras A, Davatzikos C, Paragios N. Deformable medical image registration: a survey. IEEE Trans Med Imaging 2013;32:1153-90.
8. Spindeldreier CK, Klüter S, Hoegen P, Buchele C, Rippke C, Tonndorf-Martini E, Debus J, Hörner-Rieber

J. MR-guided radiotherapy of moving targets. Radiologe 2021;61:39-48.

9. Kajiwara K. JPEG compression for PACS. Comput Methods Programs Biomed 1992;37:343-51.

10. Xu R, Pattanaik SN, Hughes CE. High-dynamic-range still-image encoding in JPEG 2000. IEEE Comput Graph Appl 2005;25:57-64.

11. Noreña T, Romero E. Medical image compression: a review. Biomedica 2013;33:137-51.

12. Thayammal S, Selvathi D. A Review on Transform Based Image Compression Techniques. International Journal of Engineering Research & Technology 2013;2:1589-96.

13. Poynton C. 45 – JPEG and motion-JPEG (M-JPEG) compression. In: Digital Video and HD (Second Edition). Boston: Morgan Kaufmann; 2012:491-504.

14. Bjontegaard G, Lillevold KO, Danielsen R. A comparison of different coding formats for digital coding of video using MPEG-2. IEEE Trans Image Process 1996;5:1271-6.

15. Nosratinia A, Mohsenian N, Orchard MT, Liu B. Interframe coding of magnetic resonance images. IEEE Trans Med Imaging 1996;15:639-47.

16. Lan C, Shi G, Wu F. Compress compound images in H.264/MPGE-4 AVC by exploiting spatial correlation. IEEE Trans Image Process 2010;19:946-57.

17. Malhotra M, Singh AV, Matam R. Comparative Performance Issues with H.264 vs H.265. 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon). Faridabad: IEEE; 2019.

18. Yan H, Li Y, Dai J. Evaluation of video compression methods for cone-beam computerized tomography. J Appl Clin Med Phys 2019;20:114-21.

19. Yan H, Li Y, Dai J. Exploring correlation information for image compression of four-dimensional computed tomography. Quant Imaging Med Surg 2019;9:1270-7.

20. Yan H, Li Y, Dai J. Four-Dimensional Cone-Beam Computed Tomography Image Compression Using Video Encoder for Radiotherapy. J Digit Imaging 2020;33:1292-300.

21. Zhang K, Tian Y, Li M, Men K, Dai J. Performance of a multileaf collimator system for a 1.5T MR-linac. Med Phys 2021;48:546-55.

22. Xia W, Liu Z, Yan L, Han F, Hu Z, Tian Y, Cui W, Ren W, Guo C, Miao J, Dai J. A longitudinal evaluation of improvements in treatment plan quality for lung cancer with volumetric modulated arc therapy. J Appl Clin Med Phys 2020;21:33-43.

23. Li X, Luo S, Li J. Liver Segmentation from CT Image Using Fuzzy Clustering and Level Set. Journal of Signal and Information Processing 2013;4:36-42.

24. Xu C, Prince JL. Snakes, shapes, and gradient vector flow. IEEE Trans Image Process 1998;7:359-69.

25. Li C, Xu C, Gui C, Fox MD. Distance regularized level set evolution and its application to image segmentation. IEEE Trans Image Process 2010;19:3243-54.

26. Frankewitsch T, Söhnlein S, Müller M, Prokosch HU. Computed Quality Assessment of MPEG4-compressed DICOM Video Data. Stud Health Technol Inform 2005;116:447-52.