Alternative lossless compression algorithms in X-ray cardiac images

D.R. Santos, C. M. A. Costa, A. Silva, J. L. Oliveira & A. J. R. Neves[1]
DETI / IEETA, Universidade de Aveiro, Portugal

ABSTRACT: Over the last decade, the use of digital medical imaging systems has increased greatly in healthcare institutions. Today, Picture Archiving and Communication System (PACS) is one of the most valuable tools supporting medical profession in both decision making and treatment procedures. It reduced the costs associated with the storage and management of image data and also increased both the intra and inter-institutional portability of data. One of the most important benefits of the digital medical image is that it allows the widespread sharing and remote access to medical data by outside institutions. PACS presents an opportunity to improve cooperative workgroups taking place either within or with other healthcare institutions.

Storage and transmissions costs are continuously decreasing, but, as individual digital medical studies become significantly larger, further improvements on transmission performance and on storage efficiency are critical. Image compression algorithms offer the means to reduce storage cost and to increase transmission speed.

Following previous methodologies [1], this paper provides a comparison about the application of DICOM (Digital Imaging and Communications in Medicine) lossless compression standards on Angiography cine acquisition images. A new lossless compression approach that exploits time redundancy between successive frames is also presented. Finally, the standards codecs values are compared with results obtained with the proposed method and with video lossless codecs.

1 INTRODUCTION

The compression of digital medical images is of vital importance in Picture Archiving and Communications Systems (PACS) and teleradiology applications, due to the typical huge volume of patients' data. This is especially true for grayscale images used in radiology applications (e.g. a single 240 frames angiographic cine acquisition represents approximately 60 MB of storage and a Multislice cardiac Computerized Tomography study of 2500 images per patient, may reach approximately 1.3 GB).

Digital storage of medical images can be challenging specially because of the requirement to preserve the image quality and also because the images can be very large in size and number [2, 3]. Several reasons can be pointed to stress the importance of compression [3]:

- Digital medical images databases are normally very large repositories;
- Patients' data must be stored for long periods resulting in a continuous growth of medical imaging databases;
- High resolution and many bits per pixel results in large volumes;
- The image transmission time is volume driven. The usefulness of a PACS depends greatly on appropriate transfer waiting times.

Common images compression algorithms, such as JPEG, are data-aware, i.e. they explore redundancy that is specific of this kind of two dimensional data source. In dynamic images (videos), such as angiographic cine acquisition, we should be able to take advantage of redundancy in the third dimension (time) and achieve greater compression ratios [2].

Various methods, both for lossy and lossless image compression, are proposed in the literature.

[1] This work was supported in part by the Fundação para a Ciência e Tecnologia (FCT).

Lossy compression techniques can achieve high compression ratios but they do not allow one to reconstruct exactly the original version of the input data. The use of lossy compression in medical imaging is still controversy for particular applications since it can result on decreased diagnostic accuracy and confidence [4]. The only possibility to obtain a completely reversible compression scheme is to use lossless compression methods but, in this case, the achievable compression ratios are only in the order 2:1 up to 4:1 [3].

The lossless image compressors allowed nowadays by the DICOM standard are lossless JPEG, JPEG2000 (lossless mode) and JPEG-LS (lossless mode). The last two are state of the art image compressors and achieve better performance than lossless JPEG sv 1 which is, however, the most used image compressor in DICOM images. This is due to the fact that only the most recent equipments have the capability to decode JPEG-LS and JPEG2000 image codestream.

Within the Cardiac Imaging domain, Coronary Angiography (XA) is currently a prominent X-ray image modality for coronary heart disease diagnostic and treatment [4].

Our angiographic experiment used a set of 278 DICOM digital cine angiograms, acquired at 15 fps at a matrix size of 512 X 512 pixels with an 8-bit grayscale (256 gray levels) resolution. The lengths of sequences varied from approximately 16 to 277 images, corresponding to about 1 to 18 seconds.

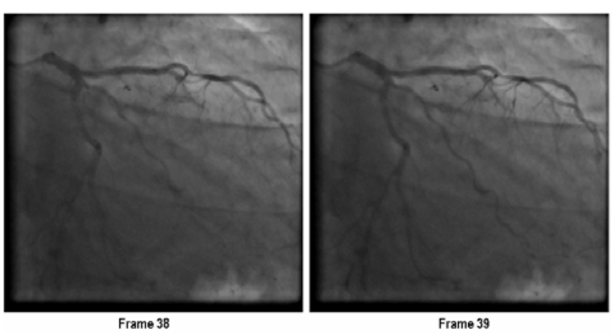

Figure 1 – Two successive XA frames.

Observing successive XA frames (Figure 1) it is possible to recognize that two consecutive images are very similar to each other and the exploitation of this redundancy would possibly lead to a better compression approach. On the other hand, the majority of data volume produced in an angiographic study is relative to dynamic captures (i.e. videos). So, the alternative compression method described in this paper is based on the hypothesis that the compression of the differences between successive images would obtain better results than the compression of the individual images. This solution is very similar to Differential Pulse Code Modulation (DPCM) used, for example, in digital coding and also on some video encoders/decoders (codec). The improvement is obtained since the differences between pixels of consecutive images are smaller than the actual pixel amplitudes.

The normal approach would be to obtain the *differential images* from the natural image acquisition order ($dif_{i,j} = im_i - im_j$ in which i and j are the images acquisition number and i = j + 1). Several compression experiments were carried out using different compression algorithms but results were not satisfactory mainly due to noise presence. Instead using the difference between image that are separated in time, we have aggregated all frames together (maintaining the normal frame order) to form a tridimensional volume (x, y, t) of 512 x 512 x N where N is the number of frames of the XA cine acquisition (Figure 2). The acquisition number (N) of each frame is equivalent to the instant (t) it was acquired so a higher acquisition number imply a later acquisition. This volume process was performed using Matlab® software. Successive images are obtained from the volume but in the ***xt*** plane - Figure 3b (in opposition to the normal temporal image sequence – ***xy*** plane - Figure 3a).

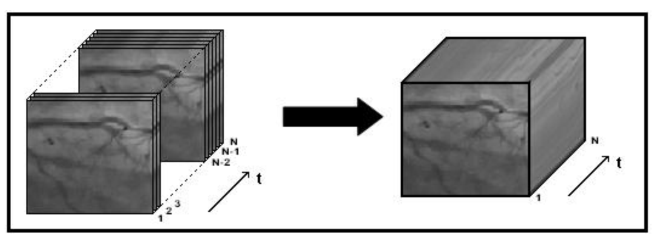

Figure 2 – Volume creation.

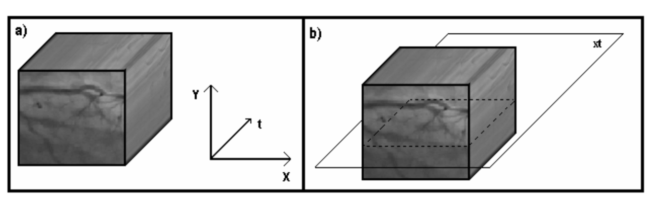

Figure 3 – xt images extraction from volume.

It can be observed, in Figure 4, the two types of images, the original frame (left) and the ***xt*** plane (right). The later is the one used on d*ifferential method*. *Differential images* are obtained by subtracting an ***xt*** plane image from the previous one. The first ***xt*** plane is the top of the volume and the last ***xt*** plane is the bottom of it.

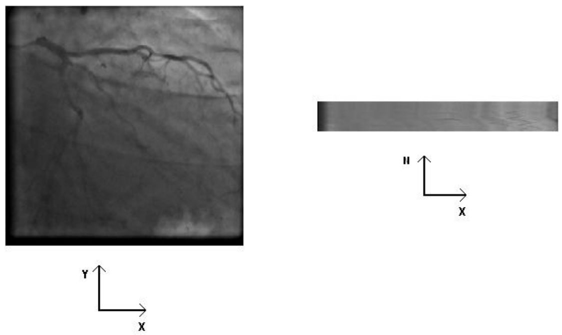

Figure 4 – Original frame (left) and xt frame (right).

The pixel values of the difference between ***xt*** plane images use a direct method $pixel\_dif_{i,j}(m,n) = pixel_i(m,n) - pixel_j(m,n)$ but since differential images have, as original ones, 8 bits resolution the difference value can not be directly assigned because of negative differences. A similar coding scheme to JPEG-LS was used – Equation 1 [5]:

$$pixel(\varepsilon) = 2|\varepsilon| - \mu(\varepsilon) \quad (1)$$

where $\mu(\varepsilon) = 1$ if $\varepsilon < 0$ or 0 otherwise

This scheme may, in some cases, seem to have the disadvantage that the *differential image* needs 2 bytes per pixel but this has never happened in all the procedures tested because their pixel values are relatively small. The reconstruction operation is straightforward since an even $pixel(\varepsilon)$ value indicates a positive difference and an odd indicates a negative difference. It is important to focus that the *differential images* consider both spatial and temporal redundancy since the original ***xt*** images have intrinsic temporal information. Two consecutive difference images of a 64 frames cine angiogram are presented on Figure 5. The *differential images* were compressed with several compressors, including state of the art image compression algorithms in order to obtain the best compression ratio possible and to compare their performance when *differential images* are provided to them. Because the *differential images* are not typical images we also decide to include non-image non commercial lossless compression algorithms in our analysis.

Figure 5 – xt plane division result for 64 frames cine angiogram (512 x 64).

Finally, the analysis was extended to video lossless codecs to compare the compression ratios with the ones obtained with still image lossless compressors in both methods mentioned above.

2 RESULTS

Several compression algorithms were tested in this work in order to compare the best image and non-image based lossless compressors. Implementations relied on publicly available software libraries. The compression effectiveness presented in Table 1 and Table 2 was obtained using the ratio between the raw pixel data size (usually 262144 Bytes in XA) and the compressed file size since it is more indicative of compression performance in the real world.

The tables present the compression results achieved by applying state of art lossless compressors in the original XA frames (Table 1) and general lossless compressors in *differential images* (Table 2). The general PPMd (Prediction by Partial Matching) compressor has been chosen as the reference non-image based lossless compression algorithm.

Table 1 – Angiographic compression results (Original images)

Number of Procedures	278
Number of Frames	23152
Total of Space Volume	6.077.106.590 Bytes (6 GB)
Compressor	Average Compression Ratio
Lossless JPEG sv1	2,67
PPMd	**2,96**
JPEG 2000	3,37
JPEG-LS	3,52
PAQ8i	3,78
BMF2.0	3,80

Table 2 – Angiographic compression results (Differential images).

Compressor	Average Compression Ratio
BMF2.0	2,74
BZ2	2,99
PPMd	**3,38**
PAQ8f	3,43

From Table 1 and Table 2 observation we can conclude several aspects:

- The JPEG-LS is the best lossless compressor between the three image compressors adopted by DICOM standard on Angiographic studies.
- However, there are two image lossless codecs (BMF2.0 and PAQ8i) that compress more than actual codecs supported by the DICOM standard. Yet, their performance decreased on *differential images* (see BMF2.0 in Table 2)
- The *differential method* with PPMd compressor achieves, in average, 27% higher compression than JPEG lossless sv1. It also has a slightly better compression ratio than JPEG2000 and loses only 4% on JPEG-LS.
- The general predictive compressor PPMd (Bold), i.e. a non-image based compressor, performs excellent results with *differential images*, approximately equal to the best DICOM standard codec. We must emphasize to the fact that JPEG-LS was specially developed for image compression.
- The PAQ8 family general compressors were among the best on both strategies. However, they were excluded as practical solution because they are heavy time and resource consuming compressors so they are used in this test only as benchmark compressors.

After several tests we have concluded that expected higher compression ratios were not possible with the *differential image* method because of noise presence in the images. Some previous experiences with other image modality, namely Multislice cardiac Computerized Tomography, strength that idea.

Finally, to realize the video experiment three non commercial and freely available lossless video codec were applied on original XA DICOM procedures. The video codec results are presented in Table 3.

Table 3 – Video lossless codec results.

Encoder/Decoder	Average Compression Ratio
Alparysoft lossless [6]	3,02
Lagarith lossless [7]	3,37
MSU lossless [8]	3,43

MSU lossless codec has a slightly better average compression ratio but the compression time is very high, disabling the usage of this codec (with maximum compression configuration) in real environment. This codec can be tuned so that the time will be similar (but still slower) than the others codecs but the compression ratio falls between Lagarith and Alparysoft codec. Lagarith codec was noticeably faster than the other two.

3 CONCLUSION

This study shows that better compression ratios can be obtained using other lossless image compressors (PAQ8 family and BMF2.0) than the ones allowed by the DICOM standard (JPEG family). The purpose of this work was to test several compressors including the ones used nowadays by Picture Archiving and Communications Systems.

This paper also presents an alternative compression strategy based on preprocessing the pixel raw data of XA DICOM images and using publicly available non-image based compressor libraries. Relevant compression ratios were obtained with this method, taking in consideration that we are working with lossless products. The results obtained suggest that lossless compression of this kind of images is reaching a limit with little improvement from previous compression schemes (compression ratios between 3 and 4) and that general purpose compression schemes are acceptable for image compression depending on the way that data is provided.

This study also suggests that image noise removal is crucial to obtain higher compression ratio but this must be accomplished without compromising the full diagnostic information contained in the image. Therefore, if noise bits removal doesn't affect image diagnostic accuracy and confidence, storage space can be saved. The method, however, will become visually lossless. Bit discarding techniques such as the one described in [9, 10] could result in noise removal without affecting the visual content and permitting a considerable increment in compression ratio

Combined strategies of lossy and lossless compression can also play a significant role in the compression of medical images and achieve much greater impact than pure lossless methods [11].

Available lossless video codecs do not improve compression ratio of XA since they do not take into account the fact that XA files are 8 bits grayscale and it is not necessary to save three color-space components. The development of a grayscale video codec is of great importance and could lead to better compression ratios. Since X-ray imaging is necessarily grayscale, the applications of this codec could be extended to almost every X-ray medical imaging modality.

4 REFERENCES

1. Costa, C.M.A., et al., *Himage PACS: A new approach to storage, integration and distribution of cardiologic images.* Proceedings of the SPIE, 2004. **5371**: p. 277-287.
2. Clunie, D.A., *Lossless compression of grayscale medical images – effectiveness of traditional and state of the art approaches.* Proceedings of the SPIE, 2000. **3980**.
3. Kivijarvi, J., et al., *A comparison of lossless compression methods for medical images.* Computerized Medical Imaging and Graphics, 1998. **22**: p. 323-339.
4. Kerensky, R.A., et al., *American College of Cardiology/European Society of Cardiology international study of angiographic data compression phase I: The effects of lossy data compression on recognition of diagnostic features in digital coronary angiography.* Journal o the American College of Cardiology, 2000. **35**(5): p. 1370-1379.
5. Weinberger, M.J., G. Seroussi, and G. Sapiro, *The LOCO-I Lossless Image Compression Algorithm: Principles and Standardization into JPEG-LS.* IEEE Transactions on Medical Imaging, 2000. **9**(8): p. 1309-1324.
6. http://www.alparysoft.com/.
7. http://lags.leetcode.net/codec.html.
8. http://www.compression.ru/video/.
9. Keh-Shih, C., et al., *Content Analysis in Clinical Digital Images.* Radiographics 1994. **14**(2): p. 397-403.
10. Chen, T.-J., et al., *A Blurring Index for Medical Images.* Journal of Digital Imaging, 2006. **19**(2): p. 118-125.
11. Ashraf, R. and M. Akbar, *Absolutely lossless compression of medical images.* Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, 2005: p. 4006-4009.