IOP Publishing

Imaging Modalities for Biological and Preclinical Research: A Compendium, Volume 2

Part II–IV: *In vivo* preclinical imaging: Correlated multimodality imaging and outlook

Andreas Walter, Julia G Mannheim and Carmel J Caruana

Chapter III.4.f

Data compression algorithms for biomedical images

Sérgio Faria, Francisco Cunha, João Parracho, Lucas Thomaz and Pedro Assunção

Computational image analysis and automation assume an increasing importance in today's biomedical workflows. The massive adoption of imaging protocols, combined with the advent of cutting-edge modalities, has been pushing data throughput at bioimaging facilities to unprecedented levels. Based on this, a natural requirement for exploiting the full potential of multimodal frameworks consists of simple and fast access to the wide range of acquired pictorial data. For this purpose, the design of more efficient data structures, as well as the adoption of good information management practices applied to bioimage data, constitute very relevant initiatives. One such intervention is data compression and this text provides an overview and description of the main image compression methods, as well as an exposition of the potential of certain image formats according to specific application requirements, while ensuring the preservation of the data integrity.

1 Introduction

Image compression refers to the set of image processing techniques devised to convert the acquired image data into more efficient data representations. By contributing for storage savings, data compression enhances the value of the hardware allocated to image archiving, as well as simplifying bandwidth requirements for transmission purposes. The question of the need for image compression in clinical and preclinical settings becomes especially relevant in a time that many bioimage protocols easily overload the installed storage infrastructure of imaging facilities, and the alternatives provided by cloud-based services typically involve considerable long-term costs for institutions. Given such limitations, image compression may constitute a very relevant tool in the bioimaging ecosystem.

doi:10.1088/978-0-7503-3747-2ch29

In the following subsections, the most common encoders and respective fundamentals will be discussed, as many of them are currently implemented in a wide range of imaging systems that simply encapsulate their logic through proprietary file formats.

2 Challenges in a multimodal setting

In the scope of the review of image compression algorithms that will be discussed, it is important to note that those are designed to operate on a single-modality basis, i.e. they are not prepared, by definition, to jointly encode multimodal sequences.

Nonetheless, at this time, image domain transfer and image translation constitute hot research topics that, in the short term, can deliver important foundations for the development of coding tools specifically designed for the joint compression of paired multimodal sequences. By allowing one to virtually synthesize multimodal data [1–5], it is expected that these data-driven approaches contribute to generalize the predictive coding that most compression algorithms implement, so that a supplementary dimension, which refers to the multimodality, can be handled and thus reduced.

3 Principles and main approaches

In short, compression systems involve a paired group of independent algorithms aiming to reduce the amount of some input data (coding) and to reconstruct the compressed information (decoding). Generally, the compressed information is converted into a bitstream (sequence of bits) or a file, according to a specified semantic, prior to being stored or transmitted. There are two distinct compression modalities—lossless and lossy—which differ in terms of their reversibility, i.e. the ability to fully recover the original data after compression. If not well-constrained, the information loss resulting from the (lossy) compression process may lead to the corruption of analysis outcomes (automated or not), thus making lossless compression the preferred option for bioimaging applications. It should be mentioned, however, that some applications may allow relaxing some distortion-related requirements, thus enabling the adoption of visually lossless compression schemes, which provide increased storage savings, by introducing little, non-perceptible loss of visual information.

Image compression techniques exploit the redundancy existing between adjacent pixels in the spatial, temporal, or frequency domain. These techniques may be applied to individual pixels or groups of pixels, thus allowing one to discard all the information allocated to represent redundant data. The effectiveness of image compression algorithms depends not only on the used techniques, but also on the amount of data redundancy, and therefore on the image content. However, given that distinct compression algorithms differently model such redundancy, for a given image, variable compression is achieved by distinct compression methods. Furthermore, it must be noted that the compression efficiency offered by lossless methods is far below that offered by lossy algorithms given the requirement for reversibility of the former. An intuitive metric for measuring the compression efficiency of a given compression algorithm is the ratio of the original image size to the compressed image size—compression ratio (CR). In CR units, lossless coding performance typically ranges between 1.5 and 3, whereas lossy compression algorithms may reach values starting from 10 to 100's, depending on the tolerable distortion of the reconstructions, with the visually lossless methods comprehended in the lower tens.

Two important techniques implemented in lossy schemes that enable such high CRs are transform-based coding and quantization. While the latter refers to the process of mapping a larger domain of image intensities into a narrower one, transform coding briefly constitutes an initial processing step aimed to convert images into more compressible representations, by projecting the image signal into different spaces that typically verify some sparseness and compactness requirements (e.g. frequency-domain). In many image encoders, the energy compactness of the transform is exploited by the quantization process to eliminate the irrelevant information, usually not perceived by the human visual system. Examples of some transform-based encoders are the JPEG and JPEG 2000 algorithms that exploit frequency-based representations obtained through the discrete cosine transform (DCT) and the discrete wavelet transform (DWT), respectively.

In lossless compression frameworks, in contrast, quantization is not allowed since it originates information loss. Additionally, transform-based coding is constrained to reversible transforms and its overall performance is limited as quantization cannot be applied to the transformed coefficients. Instead, lossless algorithms exploit the assumptions of similarity and causality in image data, in order to model image intensity statistics through predictive coding, which ultimately acts as a decorrelation technique. Predictive coding constitutes a set of techniques by which the values of specifically sampled reference pixels serve to estimate (predict) the intensity of other pixels in the neighbourhood. Since the differences between the predicted and the original values usually correspond to more sparse and compact numerical distributions, they can be represented using a fewer number of bits, thus providing storage savings. Although no universal solution is guaranteed for image compression, optimal compression for a given input with a predictive method is provided by information theory principles, thus making image compression the problem of designing data models that optimally reduce the amount of required data to encode given information.

4 Examples of application

Several standardization efforts in the field of digital imaging have been made to provide interoperability between different picture archiving and communication systems. Particularly, standardization bodies such as the International Telecommunications Union (ITU-T) and the International Standards Organization (ISO) have published several recommendations on image and video compression, with a view to foster the adoption of digital imaging in multiple domains. In the field of medical imaging, such recommendations were validated by the Digital Imaging and Communications in Medicine (DICOM) committee that brings together a set of guidelines regarding storage, visualization, and transmission of medical image and patient data.

The DICOM volume dedicated to data structures and encoding schemes provides a list of the data formats allowed to represent the compressed images and respective metadata. Typically, such formats encapsulate the definitions of some image compression standards in the form of the so-called transfer syntaxes.

Given that understanding how DICOM regulates image compression directly relates with a concise interpretation of the operating principles of such image compression standards, a brief review on the most important ones will follow.

4.1 JPEG-LS

The ISO 14495–1/ITU-T T.87 Lossless and near-lossless compression of continuous-tone still images[1], commonly referred to as JPEG-LS, is a method for coding continuous-tone, grey-scale, or colour digital still images, allowing lossless (bit-preserving) and near-lossless compression (where the error for each reconstructed sample is bounded by a pre-defined value). It is a very competitive compression algorithm given its characteristics of reduced memory footprint and fast execution. The JPEG-LS encoder implements a compression method based on context modelling of the input sample, prediction, and entropy coding. Firstly, for each input sample (e.g. pixel) a context is determined, where each context corresponds to a statistically homogeneous class. Then the current sample is predicted from a causal neighbourhood, and the prediction error is computed. In the case of near-lossless compression, the prediction error is also quantized.

The contexts modelling is based in gradient information of the input sample neighbourhood, resulting in a data model that enhances local compressibility, by allowing one to predict the probabilities of the local intensities with greater likelihood. Such an approach requires the information that characterizes the adopted contexts (some of the local image intensities and the adopted partition schemes) to be known by the decoder afterwards.

The prediction of a sample is performed by using four neighbor samples: left, top, top-left, and top-right. The effectiveness of predictive coding largely depends on how well the adopted data model predicts the local intensity distributions. Regarding this, JPEG-LS provides two operating modes—regular and run-length—which are adaptively triggered based on the variability of the local image statistics. Such versatility constitutes a very useful feature since run-length encoding provides a much faster compression on images with homogeneous background and/or little detail, such as evidenced in some examples of fluorescence microscopy images. When the context information indicates that the neighborhood of the current pixel contains details, the regular operating mode is selected, where entropy coding is carried out using Golomb codes or arithmetic coding.

[1] ISO/IEC 14495–1 ITU-T Rec. T.87, Information technology – Lossless and near-lossless compression of continuous-tone still images: baseline, 1998. (Source code available at: http://libjpeg.sourceforge.net/)

4.2 JPEG 2000

Resulting from an effort of the JPEG committee to provide an irreversible-to-reversible compression alternative with features such as resolution scalability, partial decoding (random access, number of components), and region-of-interest coding, JPEG 2000 constitutes an ISO standardized transform-based coding system[2], built-upon the adoption of DWT as a decorrelation step. Furthermore, extensions such as the Multi-Component Transform (Part 2 of the ISO/IEC 15444 standard), which allows one to generalize the operating principle of the proposed single-component compression for multichannel or multi-frame images makes this coding system compatible with a wide range of image sequences, such as diffusion-weighted MRI. Another relevant extension[3] proposed by the standard JPEG 2000 coding system involves the volumetric expansion of the DWT to fully exploit the axial redundancy that some volumetric sequences may present, thus representing an advantage when compared to the per-image operation of the usual still-image encoders. This extension also supports the lossy-to-lossless coding functionality (quality scalability), resolution scalability and region-of-interest coding. In fact, this feature makes JPEG 2000 (Part 10 of the ISO/IEC 15444) to fit the requirements for optimal compression of inherently 3D single-channel data such as CT and MRI images, although any other type of volumetric datasets may benefit from the adoption of the 3D extension of the JPEG 2000 coding system. Furthermore, given that JPEG 2000 enables quality scalability through progressive-coding, combined with random access capabilities, decisively contributes for being the preferred choice for light-weight browsing of very high-resolution images like digital versions of pathology slides, in the same way it can be adopted to provide enhanced visualization experience of high-resolution images such as high-resolution episcopic microscopy (HREM).

The JPEG 2000 coding pipeline is comprised of the following modules: multi-component transformation, tiling, wavelet-transform, quantizer, and entropy-coding. Finally, the rate allocator decides which compressed data is inserted in the bitstream. The multicomponent transformation module has reversible and irreversible modes, allowing for up to 16 384 components. In the tiling mode, each component can be divided into up to 65 535 rectangular tiles, each with a maximum size of $2^{32} \times 2^{32}$. The sample bit depth can be represented by up to 38 bits/sample and can be different for each component. An L-Level wavelet transform is applied separately to each component (of each tile), generating a set of sub bands of wavelet coefficients. The reversible mode uses a (5,3) biorthogonal filter and the irreversible mode employs the (9,7) biorthogonal filter. The wavelet coefficients can be quantized differently in each sub band, but the step is one in the reversible mode. The wavelet coefficients are grouped into rectangular non-overlapped sets called

[2] ISO/IEC 15444–1 ITU-T Rec. T.800, Information technology—JPEG2000 image coding system: Core coding system, 2004. (Source code available at https://www.openjpeg.org)

[3] ISO/IEC 15444–10 ITU-T Rec. T.809, Information technology—JPEG2000 image coding system: Extensions for Three-dimensional Data, 2011. (Source code available at https://www.openjpeg.org)

code-blocks, which are entropy encoded employing a bit-plane-based binary arithmetic coding. Because it is possible to obtain the distortion reduction and the required bitrate during each coding pass, an optimal rate allocation is achievable if the whole image is encoded. This method also allows one to encode arbitrary shaped regions of interest (ROI), with higher quality than the background, transmitting them first. In JPEG 2000, the MAXSHIFT method is employed to identify the wavelet coefficients in the ROI and up-scaling them by a user pre-defined value. As ROI coefficients and background are in different bit-planes and have different quantization indices, the rate allocator and the decoder can identify them. This feature is particularly important to encode bioimages since the region of interest and the background can be differently encoded.

4.3 H.265/HEVC

The H.265/HEVC (ISO/IEC 23008–2 Standard)[4] is a video coding standard named High Efficiency Video Coding, which inherently allows one to exploit both spatial and temporal redundancies in a lossless-to-lossy architecture. It was designed as a successor of the widely used Advanced Video Coding (AVC, H.264, or MPEG-4 Part 10), to support resolutions up to 8192×4320 pixels, and offering from 25% to 50% better data compression at the same level of video quality. This performance gain is achieved due to the use of additional prediction modes and partitioning schemes. In order to cope with the additional complexity, some tools have been introduced to take advantage of the parallel processing architectures. Besides exploiting spatial redundancy through intra-frame prediction techniques, the temporal redundancy of the sequence is also addressed through motion estimation and compensation techniques, commonly referred to as inter-frame prediction. The core system employs a hierarchical content-based image partitioning scheme, blocks from 4×4 to 64×64, prediction followed by DCT-based transformation and quantization of the residuals and context-adaptive binary arithmetic coding, which refers to an efficient entropy coding technique.

To operate with the lossless profile, the HEVC main profile (HEVC) bypasses transform, quantization, and in-loop filters. When compared to the non-lossless HEVC coding mode with the smallest quantization parameter value (i.e., 0 for 8-b video and -12 for 10-b video), the lossless mode provides perfect fidelity and an average bitrate reduction of 3.2%-13.2%. In order to further improve the coding efficiency, a sample-based angular intra prediction (SAP) method is employed instead of the block-based one. This one uses adjacent neighbours for better intra prediction accuracy and performs prediction sample by sample. The experimental results reveal that the SAP provides an additional bit-rate reduction of 1.8%–11.8% on top of the HEVC lossless coding mode.

Analogously to the above-mentioned standard, some extensions have been included in the H.265/HEVC coding standard—Range Extensions (RExt) [6]—in

[4] ISO/IEC 23008–2 ITU-T Rec. H.265, Information technology—High efficiency video coding and media delivery in heterogeneous environments, 2013. (Source code available at: https://hevc.hhi.fraunhofer.de/)

order to provide alternative coding features and to make the coding framework compatible with superior image resolutions, bit-depths, as well as alternative colour encoding formats. The operation of H.265/HEVC system must be configured to match the characteristics of a given source image sequence, namely considering the spatial and temporal resolution, the number of components and the bit depth, among others. Also, the operating conditions must be taken into consideration regarding the visualization and analysis systems. While the latter may require the original raw data, the former may work with visually lossless methods. In fact, some mobile devices currently provide H.265 decoders, however the limited availability of embedded profiles make them unsuitable for computer assisted diagnosis supported by image analysis, since such applications typically require the original raw data. In this context, H.265 also provides scalable encoding versions which have been adopted by DICOM, offering a more efficient way to store and manage image data. The targeted applications include using HEVC streams to store and exchange raw images but also to produce lossy content out of the archived lossless streams. The 'HEVC/H.265 Scalable Monochrome' enables the storage of monochrome sequences of still pictures with resolution up to 4096×2160, with a bit depth varying from 8 to 16 bits, at 50 Hz/60 Hz, in a lossless or lossy format and comprising a detachable lossy and lighter version. CT scanners typically produce such content. The 'HEVC/H.265 Scalable Main 4:4:4 Image Compression' tier to encode and decode RGB still pictures and video content with a resolution up to full HD with a bit depth of 8 bits per component. This will enable the storage of RGB video files and sequences of still pictures with a resolution of 1920×1080 and 8 bits per component, at 50 Hz/60 Hz, in a lossless or lossy format and comprising a detachable lossy and lighter version. Such content can typically be produced when rendering 3D views in 2D. Another relevant aspect is related to the encoding profiles more suitable to encode sequences of images, namely HEVC/HEVC-RA (random-access), where the third dimension is temporal (video) or spatial (volumetric). In such a case, a video encoder is able to exploit the redundancy between images along the sequence. To achieve higher compression ratios, the prediction techniques are applied to groups of images causing some dependencies between them, thus limiting the instantaneous (random) access to each image individually at the decoder. This means that, depending on how fine-grained the random access is required to be, one must select the appropriate number of reference images, which are used for predictive coding of the sequence, as well as the number of images between those that can be individually decoded.

In the following, we will revisit some up-to-date alternatives to the DICOM-compliant image compression algorithms provided that superior compression performance has been reported in the literature.

4.4 CALIC

Up to the emergence of JPEG-LS, the Context Adaptive Lossless Image Codec (CALIC) [7] was the best in-class lossless compression scheme given its compression efficiency. In practice, JPEG-LS shares some of the CALIC functionalities since both methods adopt context modelling. However, CALIC is built-upon a larger number of modelling contexts, which are used in prediction and in residual coding, making it a more complex system. The initial prediction employed in CALIC is based on the local gradient information that is subsequently refined when the prediction context and the associated residual bias are determined. Finally, after a process of contexts quantization and modelling, the prediction residuals are entropically coded. In addition to the encoding mechanism for 2D images, some volumetric expansions of the CALIC coding system have been proposed.

4.5 MRP

The minimum rate predictors (MRP) is a state-of-the art lossless compression technique and here we refer to its volumetric extension—3D-MRP- which was proposed as a superior compression method for volumetric medical images [8]. Comparing to JPEG-LS, JPEG 2000, HEVC, and CALIC, the authors reported compression gains ranging from 12% to 15%. The 3D-MRP algorithm uses an adaptive block-based predictive scheme that dynamically assigns linear predictors accordingly to the local content of the image blocks. Unlike other techniques that address the problem of optimal design of predictors, the 3D-MRP implements a volumetric optimization scheme aimed to iteratively design linear predictors that minimize the number of necessary bits to encode the resulting residues. Its adoption in the context of volumetric image compression allows one to reach increased compression performance, however at the cost of increased computational complexity.

4.6 JPEG-XL

The JPEG XL coding system currently constitutes a standardization (ISO/IEC 18181)[5] effort to provide compression efficiency superior to existing standard codecs, while keeping a set of desirable features such as resolution scalability and support for images with up to 4096 channels. It also has support for: 16-bit integer and 16-bit float inputs, and outputs with corresponding precision, multi-channel lossless coding and 'visually lossless' coding (as defined in AIC-2 (ISO/IEC 29170–2)), without requiring mathematically lossless coding.

The codec architecture is backward-compatible with the legacy JPEG format and provides efficient lossy-to-lossless compression of high-quality imagery. In its core, by extending the fixed size DCT blocks implemented in the legacy JPEG into variable-size DCT blocks, combined with a progressive DCT-based decomposition, JPEG XL allows for both efficient and versatile image decorrelation and decoding. Furthermore, as opposed to the legacy JPEG format, JPEG XL implements an adaptive predictor that exploits the images bi-dimensionality in such way that it optimally updates the prediction mode, thus leading to storage savings. However, the most remarkable feature of JPEG XL consists of an alternative entropy coding

[5] ISO/IEC CD 18181 Information technology—JPEG XL image coding system, (under development). (Source code available at: https://gitlab.com/wg1/jpeg-xlm)

scheme, called Asymmetric Numeral Systems, that achieves significantly faster operation without sacrificing compression efficiency. For now, the group involved in the JPEG XL standardization claims that its adoption can lead to storage savings up to approximately 60%, in common use at equivalent subjective quality [9]. This combined with the implemented group splitting principle, which allows for parallel decoding, makes JPEG XL a very convenient solution for biomedical image compression.

4.7 VVC

Versatile Video Coding (VVC) is currently under development by the Joint Video Experts Team (JVET) of ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG) to become the successor of H.265/HEVC as the next generation video coding standard [10]. The development of VVC intends to meet the requirements to ultra-high definition (4 K to 16 K) and 360-degree video computation, with a bit depth from 10 to 16 bits per component. It also has support for still picture coding, scalable video coding (temporal, spatial, and quality), as well as for lossless and subjectively lossless compression.

VVC inherits most of its coding architecture from HEVC/H.265 with the improvement of several techniques such as intra prediction with the addition of more directional modes, inter prediction by using neural networks, and focussing on affine models for motion compensation, the partition structure, in-loop filter, quantization and entropy coding. It also uses new high complexity elements such as affine motion compensation, rectangular, and triangular image partitions, as well as some prediction in decoder side. With the following improvements VVC aims to achieve a 30%–50% bit-rate reduction over H.265/HEVC Main Profile for the same perceptual quality.

5 Comparative overview

The previously described image and video coding algorithms have been proposed to fulfil different image coding requirements, using the technology available at the time. Currently, some of these algorithms are no longer the state-of-the-art for natural images, but due to their functionalities they are particularly suited to encode certain types of images, as in the area of bioimaging.

In this context, table 1 shows the ability of these compression methods to encode volumetric data or to perform coding of ROIs. Additionally, a ranking of the

Table 1. Features comparison between coding algorithms.

	JPEG LS	JPEG 2000	H.265/HEVC	CALIC	3D MRP	JPEG XL
Volumetric coding	✗	✓	✓	✓	✓	✗
ROI coding	✗	✓	✓	✗	✗	✓
Compression efficiency	6	5	2	4	1	3
Low complexity	2	3	4	5	6	1

methods is presented (1-best and 6-worse), to classify each method in terms of compression efficiency and coding complexity. As can be seen, it may happen that the most efficient algorithm presents the highest complexity. Therefore, for each application it is necessary to evaluate the operating conditions.

6 Conclusions

The massive adoption of digital imaging in research and medical practice currently constitutes a data bottleneck for institutions, which increasingly demand compression systems for efficient storage and distribution. In this context, the adoption of image compression standards (some included in the DICOM standard) has enabled superior compatibility and interoperability between picture archiving and communication systems. Nevertheless, besides achieving state-of-the-art compression efficiency, alternative compression methods must be considered for specific applications, given the richer set of features that may provide.

We focussed on the most relevant lossless coding schemes by describing their principles and highlighting core functionalities, but also on recent research aiming to improve the modest compression performance provided by lossless techniques. Concurrently, even though no mature guideline exists for regulating the use of lossy compression in clinical settings, the adoption of visually lossless techniques, capable of delivering high-quality visual contents at reduced bit rates, constitutes a natural trend, particularly suitable for supervised image inspection purposes and semi-automated analysis. Furthermore, the growing adoption of hybrid multimodal imaging systems such as CT/PET and MRI/PET, combined with the advent of deep learning, has been opening new research paths to adopt cross-modality image synthesis as a tool to encode complementary, yet correlated, datasets in a more compact format.

Combined with the wide range of preclinical imaging modalities, image-based applications such as quantification and segmentation can constitute important research platforms for quantitatively characterizing the impact of data compression on specific routine protocols. In fact, in the short term, image compression for preclinical imaging might involve data-specific coding schemes (not necessarily related to the human visual system) to optimize the resulting bitstreams against objective quality metrics determined by post-processing applications.

References and further reading

[1] Han L and Yin Z 2018 *Mach. Vision Appl.* **29** 1257–67

[2] Dong X, Wang T, Lei Y, Higgins K, Liu T, Curran W, Mao H, Nye J and Yang X 2018 *Phys. Med. Biol.* **64** 215016

[3] Armanious K, Jiang C, Abdulatif S, Küstner T, Gatidis S and Yang B 2019 *Proc. of EUSIPCO 2019—27th European Signal Proc. Conf. 2019* 09

[4] Nadarajan G and Doyle S 2020 Medical imaging 2020: digital pathology *Proc. of SPIE* 11320

[5] Ben-Cohen A, Klang E, Raskin S, Soffer S, Ben-Haim S, Konen E, Amitai M and Greenspan H 2019 *Eng. Appl. Artif. Intel.* **78** 186–94

[6] Flynn D, Marpe D, Naccari M, Nguyen T, Rosewarne C, Sharman K, Sole J and Xu J 2016 *IEEE Trans. Circuits Syst. Video Technol.* **26** 4–19
[7] Wu X and Memon N 1997 *IEEE T. Comm.* **45** 437–44
[8] Lucas L F R, Rodrigues N M M, da Silva Cruz L A and de Faria S M M 2017 *IEEE Trans. Med. Imaging* **36** 2250–60
[9] Alakuijala J *et al* 2019 Applications of digital image processing XLII *Proc. of SPIE* 11137
[10] Bross B, Chen J and Liu S 2019 *JVET-N1001: Versatile Video Coding (Draft 5), Joint Video Experts Team (JVET), 14th Meeting* (Geneva, CH, Technical Report)