Improving 3D Medical Image Compression Efficiency Using Spatiotemporal Coherence

Matina Ch. Zerva, Michalis Vrigkas, Lisimachos P. Kondi and Christophoros Nikou
Department of Computer Science and Engineering, University of Ioannina, Ioannina, Greece
{***szerva, mvrigkas, lkon, cnikou***}***@cse.uoi.gr***

Abstract

Advanced methodologies for transmitting compressed images, within acceptable ranges of transmission rate and loss of information, make it possible to transmit a medical image through a communication channel. Most prior works on 3D medical image compression consider volumetric images as a whole but fail to account for the spatial and temporal coherence of adjacent slices. In this paper, we set out to develop a 3D medical image compression method that extends the 3D wavelet difference reduction algorithm by computing the similarity of the pixels in adjacent slices and progressively compress only the similar slices. The proposed method achieves high-efficiency performance on publicly available datasets of MRI scans by achieving compression down to one bit per voxel with PSNR and SSIM up to 52.3 *dB and* 0.7578*, respectively.*

Introduction

Medical image compression has become a prevalent tool with great impact on the diagnosis of diseases in clinical practice [1]. The problem of compression, and therefore transmitting an image in real-time, given the bandwidth of the communication channel, is of great importance, especially in a low-speed connection environment. This problem is not easy to be solved, because medical images typically contain a huge amount of important diagnostic information and therefore distortion is not allowed [2]. Applications of image compression for transmission purposes are limited by real-time constraints. On the other hand, image compression applications for storage purposes are less stringent. This is because most algorithms are not executed in real-time.

Current practice in the medical image compression field is to reduce the size of medical image files by reversible (lossless) [3] compression, which offers up to 3 times size reduction, or irreversible (lossy) [4] compression. Irreversible compression allows for a much larger (between 8 and 25 times) size reduction without significant loss of visual quality of the material.

There are many standards used for compressing medical images. One of the most popular image compression standards with important medical applications is JPEG [5]. The important feature of JPEG is that it enables compression at various levels, thus enabling the user to choose the quality of the compressed image so that information losses are not visible to physicians. JPEG 2000 [6] is the successor of JPEG standard that provides compression with no or very little information loss, so the image quality does not deteriorate but approximates the image quality without compression. Compared to the JPEG standard, the JPEG 2000 standard provides a typical compression gain of 20% on average, depending on the image features. In low bitrate applications, previous studies have shown that JPEG 2000 is superior to H.264 intra-coding [7].

There is a wide variety of image transform-based coding techniques, which amongst others include discrete cosine transform (DCT) [8] and discrete wavelet transform (DWT) [9]. For compressing volumetric medical datasets, it appears that 3D wavelet-based encoders outperform DCT-based solutions while providing the required functions such as quality scaling and resolution, random access and region coding [10]. Narmatha *et al.* [11] proposed a two-stream method for encoding and decoding medical images by dividing and merging different regions of the wavelet subbands. Amri *et al.* [12] combined into a single processing pipeline image reduction and expansion techniques of different lossless compression standards such as JPEG-LS and TIFF formats to compress medical images.

In recent years, much effort has been paid for volumetric medical image compression [13, 14, 15]. Three-dimensional medical images can be viewed as time sequences or volume tomographic slices of an object. Bruylants *et al.* [16] employed the wavelet transformation to allow support for volumetric image datasets. Ravichandran et al. [17] have also used the wavelet transform to compress 3D medical images. Based on the fact that most of the medical images are being captured in hospitals and medical organizations using 2D and 3D monitoring techniques, their simulation results have shown that 3D medical images have high-frequency patterns and therefore the waveform technique allows for achieving higher PSNR values even at the highest compression ratio than 2D medical images.

In this paper, our objectives are: to (i) take into account the spatial and temporal coherence of adjacent slices in a volumetric medical image; and (ii) improve upon previous 3D-based compression methods in terms of PSNR and SSIM [19]. To accomplish these tasks, we propose an extension of the 3D wavelet difference reduction (3D-WDR) method [20] that employs the mean co-located pixel difference (MCPD) to estimate an optimal number of slices to be efficiently compressed in one volumetric object. Given the requirements for the best possible reconstructed images, the proposed method meet these objectives.

Our main contribution is the design of a volumetric medical image compression method that can be easily reproduced, is suitable for use in a variety of medical images such as MRI and CT scans, and achieves state-of-the-art compression results with high compression ratio and small information loss within an acceptable range. This performance originates from computing the spatiotemporal difference between adjacent slices in an image volume and compress as a single volume only those slices that exhibit the largest similarity on the pixel values. This allows us

to progressively transmit a medical image through a communication channel and also allows for a gradual improvement of image quality. This work aspires to serve as a bar in the 3D medical image compression domain that future works may improve upon. We performed extensive evaluations in publicly available datasets and achieved high compression ratios in all of them while maintaining a high visual-quality, which ensures that compression of medical images that are used for diagnosis, is of critical importance as diagnostic data are preserved.

Methods

The wavelet difference reduction (WDR) algorithm [20] follows the basic concepts of the set partitioning in hierarchical trees (SPIHT) algorithm [21] by incorporating extra features that aggregate the coefficients to an area of interest. By reducing the difference between the wavelet coefficients, it recognizes the important wavelet coefficients and improves their accuracy to achieve high compression ratios. During WDR encoding, the compressed output produced consists of the most important coefficients along with the series of bits that briefly describe the exact position of the significant values. It offers good perceptual quality and high compression rate while preserving the edges of the images. It is suitable for compressing medical images at a low bit rate per pixel.

In this paper, the structure of the data tree used by SPIHT is avoided, and the principles of integrating and partitioning the encoded bit-level without loss are preserved. Also, the WDR method implements run-length coding (RLC) that allows the encoder to achieve faster transmission of image details over networks.

Algorithm 1: WDR algorithm.

Input : Original uncompressed image.

1 Calculate the DWT of the original image.

2 **while** *(Predetermined number of bits is not reached)* **do**

3 Sort the wavelet transform coefficients from the larger scale to the finer scale.

4 Set an initial threshold: $T_n = 2^N$, with $n = 1$ and $N = \log_2(\max_{(i,j)} |\gamma(i,j)|)$, where $\gamma(i,j)$ are the wavelet coefficients in the set of non-significant coefficients and N is the total number of bit planes.

5 Sorting pass: Find the positions of the significant coefficients concerning the threshold, and keep the coefficients that satisfy the condition: $|\gamma(i,j)| \geq T_n$.

6 Improvement process: Get the improvement rates of all significant coefficients, except those found in the sorting pass step of the current iteration.

7 Update threshold: $n = n+1$; $T_{n-1} = T_n$; $T_n = \frac{T_n}{2}$.

8 **end**

Output : Compressed image.

The term "reduction of difference" refers to how the WDR algorithm encodes the positions of significant coefficient values of the wavelet transformation by using a difference-coding method. These positions are not directly encoded, but instead, the distances between the important coefficients are encoded. Thus, the WDR method encodes the path between two important coefficients. The importance of the WDR method lies in the fact that it increases the data transmission speed because the method employs the basic concepts of the run-length coding. The general model of the WDR method is shown in the following distinct steps of the Algorithm 1.

Algorithm 2: Proposed 3D-WDR-MCPD algorithm.

Input : Original uncompressed image.

1 Compute the average *MCPD* values for each slice using eq. (1) .

2 Predefine a threshold $T_{MCPD} = 0.5$.

3 Select slices with the highest spatiotemporal coherence: $MCPD \leq T_{MCPD}$ and construct the 3D volume.

4 Employ Algorithm 1 to compress the optimal 3D volume.

Output : Compressed image.

In embedded wavelet-based coding, the significance map forms a binary image; consequently, techniques that have been employed for the coding of bi-level images apply to significance-map coding. For example, run-length coding has a long history of such binary-coding use. The wavelet difference reduction algorithm combines run-length coding of the significance map with an efficient lossless representation of run-length symbols to produce an embedded image coder.

WDR was originally developed as a 2D encoder but is straightforwardly extended to 3D [22]. Also, WDR can be extended to shape adaptive by "skipping" over flat regions and not coding any significant information for them or including them in the run-length. This 3D extension deploys the run-length scanning as a 3D raster scan of each subband of the 3D discrete wavelet transform, which is easily accomplished in either dyadic or packet DWT decompositions.

The proposed method is an extension of the 3D-WDR method. Specifically, we extended the 3D-WDR method by using the mean co-located pixel difference (MCPD) to estimate the optimal number of frames that can be encoded given the best peak signal-to-noise ratio. MCPD measures the temporal difference between slices on the pixel values. The MCPD between two slices of dimensions $N \times M$ i.e., slice $\mathbf{x}$ and slice $\mathbf{y}$ for each pixel is computed as:

$$MCPD = \frac{1}{NM} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} |x(i,j) - y(i,j)|, \quad (1)$$

where $x(i,j)$ and $y(i,j)$ correspond to the pixel value at position (i,j) of slices $\mathbf{x}$ and $\mathbf{y}$, respectively.

In particular, we compute the *MCPD* of each slice i with all the slices that follow, where $i = 1, \ldots K$, and K being the total number of slices, for each MRI volume, (e.g., slice 1 with slices $2,3,4,\ldots K$, slice 2 with slices $3,4,5,\ldots K$) and construct a volumetric image consisting of only the slices that exhibit an average *MCPD* value ≤ 0.5 to keep those with similar spatial and temporal coherence. After extensive evaluation and cross-validation, the threshold of 0.5 has been found to be the optimal value for selecting the most similar slices in the spatial and temporal domain. The performance of the proposed method for new selected *MCPD* thresholds for other types of volumetric medical data should be evaluated based on cross-validation. In case of volumetric data

such as MRI and CT exams, the term "temporal coherence" may refer to the similarity between different slices. Note that the slices that are not selected to be part of any volumetric image are compressed separately using the standard WDR method. The main steps of the proposed method are summarized in Algorithm 2.

Results

Dataset: To evaluate our method, a widely-used publicly available dataset named the cancer image archive (TCIA) [18] was used. This is a collection of medical de-identified datasets related to a common disease such as lung cancer or brain tumor from 20 subjects with primary newly diagnosed glioblastoma. For each subject, two MRI exams of brain tumor are included in DICOM format containing 16-bit images with at least 20 slices per MRI exam.

Evaluation metrics: For the evaluation of the results, we computed the peak signal-to-noise ratio (PSNR) and the structural similarity index measure (SSIM) [19]. PSNR computes the peak signal-to-noise ratio between two images, in decibels (dB). This ratio is a quality measurement between the original and the compressed image. PSNR can take values up to infinity, the higher the PSNR, the better the quality of the compressed image. Since the MRI exams in the TCIA dataset contain 16-bit images, in this case, the PSNR is computed as:

$$PSNR = 10\log_{10}\left(\frac{(2^{16}-1)^2}{MSE}\right), \tag{2}$$

where $MSE = \frac{\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}(x(i,j)-\hat{x}(i,j))^2}{NM}$ with $x(i,j)$ and $\hat{x}(i,j)$ correspond to the pixel value at position (i,j) of the ground truth $\mathbf{x}$ (original uncompressed image) and the compressed image $\hat{\mathbf{x}}$ of dimensions $N \times M$, respectively. Note that, the term $2^{16}-1$ is the maximum pixel value in the input image data type. SSIM is a metric that represents a visual distortion between a reference image and the observed/compressed image. The SSIM is a function between two images $\mathbf{x}$ and $\hat{\mathbf{x}}$ and is computed between pairs of local square overlapping windows x and $\hat{x}$ of the two images, respectively:

$$SSIM(x,\hat{x}) = \frac{(2\mu_x\mu_{\hat{x}}+C_1)(2\sigma_{x\hat{x}}+C_2)}{(\mu_x^2+\mu_{\hat{x}}^2+C_1)(\sigma_x^2+\sigma_{\hat{x}}^2+C_2)}, \tag{3}$$

where μ_x and $\mu_{\hat{x}}$ denote the mean intensity of the ground truth and the compressed image, respectively, σ_x and $\sigma_{\hat{x}}$ are the standard deviations at patches x and $\hat{x}$ of the two images, $\sigma_{x\hat{x}}$ is the covariance of x and $\hat{x}$, and C_1 and C_2 are constants added to avoid instability. Values close to 1 indicate that the compressed image preserves high visual quality (i.e., identical patches). Finally, the mean SSIM index value is computed to evaluate the total image similarity.

First, we computed PSNR and SSIM values for the 20 consecutive slices to build the 3D model used for compression. Table 1 shows the results in terms of PSNR and SSIM for 20 slices using the 3D-WDR algorithm at bit rate one bit per voxel. As it can be seen, PSNR varies from 46.9228 to 51.3733 and SSIM varies from 0.6403 to 0.7841.

Table 2 depicts the average *MCPD* values for all slices. It can be observed that *MCPD* is less than 0.5 for the slices $2,4,6,15,18,19,20$, in that sense a threshold of 0.5 is considered as a reasonable choice to form our 3D model.

Table 1. Evaluation of standard 3D-WDR model in terms of PSNR and SSIM for the 20 slices of the MRI exam.

Slice index	PSNR	SSIM
1	48.0704	0.7484
2	50.8077	0.7263
3	49.2582	0.7670
4	48.7388	0.7124
5	51.3304	0.7490
6	48.9123	0.6838
7	48.0288	0.7515
8	49.3198	0.7133
9	49.4259	0.6935
10	46.9228	0.7815
11	49.7743	0.7841
12	48.7383	0.6992
13	51.0344	0.7106
14	47.5173	0.7577
15	49.2868	0.7049
16	48.3499	0.7481
17	51.3733	0.6467
18	49.0812	0.6800
19	49.3517	0.6917
20	48.0205	0.6403

Table 2. Estimated MCPD values for the 20 slices of the MRI exam.

Slice index	MCPD average
1	0.6009
2	0.4716
3	0.5357
4	0.4882
5	0.5173
6	0.4784
7	0.5229
8	0.5806
9	0.5325
10	2.7478
11	0.6987
12	0.6331
13	0.6603
14	0.5302
15	0.4501
16	0.6162
17	0.6703
18	0.3852
19	0.2561
20	0.0000

In Table 3, we report the results of the proposed 3D-WDR-MCPD method in terms of PSNR, SSIM, and PSNR improvement. Specifically, PSNR varies from 50.7433 to 52.2776 and SSIM varies from 0.6834 to 0.7810. Also, note that the PSNR improvement with respect to the standard 3D-WDR algorithm (Table 1) is remarkably high in terms of absolute dB difference. For instance, PSNR for slice 2, when the standard 3D-WDR

Table 3. Evaluation of the proposed method in terms of PSNR, SSIM and PSNR improvement for the seven slice indexes originated from the computed MCPD values in Table 2 given the predefined threshold.

Slice index	PSNR	SSIM	PSNR Improvement (in dB)
2	52.2776	0.7578	1.5
4	50.6824	0.7810	1.9
6	50.7814	0.7503	1.9
15	51.0053	0.7596	1.7
18	51.9189	0.7742	2.8
19	50.7433	0.7375	1.4
20	51.8223	0.6834	3.8

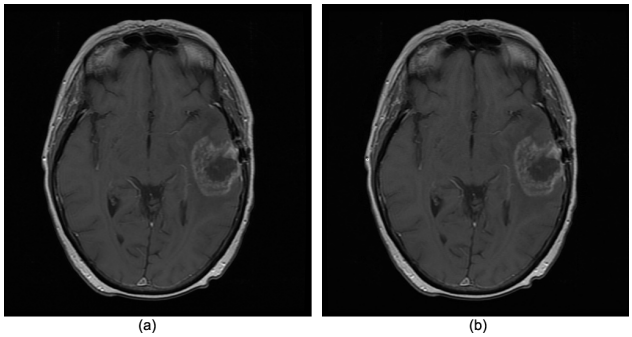

Figure 1. *Slice 2: (a) Original uncompress slice and (b) compressed slice using the 3D-WDR-MCPD method for compression.*

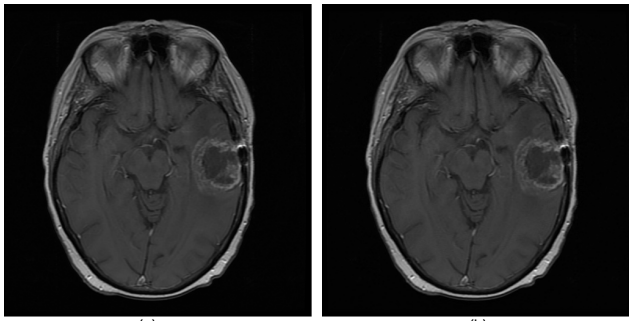

Figure 2. *Slice 4: (a) Original uncompress slice and (b) compressed slice using the 3D-WDR-MCPD method for compression.*

method was employed (Table 1), is 50.8077 and SSIM is 0.7263, while when the proposed 3D-WDR-MCPD method is used PSNR and SSIM increase to 52.2776 and 0.7578, respectively. This corresponds to a PSNR improvement of approximately 1.5 dB. The maximum PSNR improvement is achieved for slice 20 as it reaches 3.8 dB, which indicates that compressing similar slices (7 slices) in one volume per time is more efficient than directly compressing the 3D image as a whole.

Figures 1 and 2 depict qualitative results when the proposed 3D-WDR-MCPD method is employed for compressing slice two and slice four, respectively. For the depicted MRI images the highest compression rate is achieved at one bit per voxel (16 : 1) compression ratio. Also, the visual perception of the compressed image is retained and is very close to the original uncompressed image. The main reason behind this is the smooth transition of images throughout the MRI slices.

Conclusions

Medical image compression plays an important role in efficient storage, transmission, and management of these datasets. In this paper, a 3D image compression model based on discrete wavelet transform is proposed, which is a clinically acceptable option for medical image compression. In particular, we extended the standard WDR method using *MCPD* to select the optimal number of slices that exhibit the highest similarity in the spatial and temporal domain. The slices with large spatiotemporal coherence are then encoded together as one volume in terms of higher PSNR and SSIM. It is found that the perceptual quality of the medical image is remarkably high. The results indicate that the PSNR improvement over existing schemes may reach up to 3.8 dB and they may guide us through the implementation of a mobile and web platform that may be used for the compression and transmission of medical images in real-time.

Acknowledgments

This work has been co-funded by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH-CREATE-INNOVATE (project code: T1EDK-03895). The authors gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan X and Titan Xp GPUs used for this research. All statements of fact, opinion or conclusions contained herein are those of the authors and should not be construed as representing the official views or policies of the sponsors.

References

[1] R. Duszak, "Medical imaging: Is the growth boom over?" in Neiman Report, Harvey L. Neiman Health Policy Institute: Reston, VA, USA, 2012.

[2] A. J. Hussain, A. Al-Fayadh, and N. Radi, "Image compression techniques: A survey in lossless and lossy algorithms," Neurocomputing, vol. 300, pp. 4469, Jul. 2018.

[3] L. F. R. Lucas, N. M. M. Rodrigues, L. A. da S. Cruz, and S. M. M. de Faria, "Lossless Compression of Medical Images Using 3-D Predictors," IEEE Transactions on Medical Imaging, vol. 36, no. 11, pp. 22502260, Nov. 2017.

[4] A. AbuBaker, M. Eshtay, and M. AkhoZahia, "Comparison Study of Different Lossy Compression Techniques Applied on Digital Mammogram Images," Int. J. Adv. Comput. Sci. Appl. IJACSA, vol. 7, no. 12, 2016.

[5] G. K. Wallace, "The JPEG still picture compression standard," in IEEE Transactions on Consumer Electronics, vol. 38, no. 1, pp. xviii - xxxiv, Feb. 1992.

[6] M. Rabbani and R. Joshi , "An overview of the JPEG 2000 still image compression standard," Signal Processing: Image Communication, vol. 17, no. 1, pp. 3-48, January 2002.

[7] Q. Cai, L. Song, G. Li and N. Ling, "Lossy and lossless intra coding performance evaluation: HEVC, H.264/AVC, JPEG 2000 and JPEG LS," in Proc. Asia Pacific Signal and Information Processing Association Annual Summit and Conference, Hollywood, CA, pp. 1-9, December 2012.

[8] W. B. Pennebaker and J. L. Mitchell, "JPEG: Still Image Data Compression Standard," Springer Science & Business Media, 1992.

[9] R. Leung and D. Taubman, "Transform and embedded coding techniques for maximum efficiency and random accessibility in 3-D scalable compression," IEEE Trans. Image Process., vol. 14, no. 10, pp. 16321646, Oct. 2005.

[10] P. Schelkens, A. Munteanu, J. Barbarien, M. Galca, X. Giro-Nieto, and J. Cornelis, "Wavelet coding of volumetric medical datasets," IEEE Trans. Med. Imaging, vol. 22, no. 3, pp. 441458, Mar. 2003.

[11] C. Narmatha, P. Manimegalai, and S. Manimurugan, "A Lossless Compression Scheme for Grayscale Medical Images Using a P2-Bit Short Technique," Journal of Medical Imaging and Health Informatics, vol. 7, no. 6, pp. 11961204, Oct. 2017.

[12] H. Amri, A.Khalfallah, M. Gargouri, N. NebhaniJean, J.C. Lapayre and M.S. Bouhlel, "Medical Image Compression Approach Based on Image Resizing, Digital Watermarking and Lossless Compression," Journal of Signal Processing Systems, vol. 87, no. 203, 2017.

[13] R.K. Senapati, P.M. Prasad, G. Swain and T.N. Shankar, "Volumetric medical image compression using 3D listless embedded block partitioning," Springerplus, vol. 5, no. 1, pp. 2100, 2016.

[14] X. Tang and W.A. Pearlman "Three-Dimensional Wavelet-Based Compression of Hyperspectral Images," in: Motta G., Rizzo F., Storer J.A. (eds) Hyperspectral Data Compression. Springer, Boston, MA, 2006.

[15] N. Sriram and R. Shyamsunder, "3D medical image compression using 3-D wavelet coders," Digital Signal Process. no. 21, pp. 100109, 2011.

[16] T. Bruylants, A. Munteanu, and P. Schelkens, "Wavelet based volumetric medical image compression," Signal Processing: Image Communication, vol. 31, pp. 112133, 2015.

[17] D. Ravichandran, M. G. Ahamad, and M. R. A. Dhivakar, "Performance analysis of three-dimensional medical image compression based on discrete wavelet transform," in Proc. 22nd International Conference on Virtual System Multimedia, Kuala Lumpur, Malaysia, pp. 18, 2016.

[18] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, L. Tarbox and F. Prior, "The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository," Journal of Digital Imaging, vol. 26, no. 6, pp. 1045-1057, 2013.

[19] Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing. vol. 13, no. 4, pp. 600612, 2004.

[20] J. Tian and R. O. Wells, Jr., "A Lossy Image Codec Based on Index Coding." in Proc. IEEE Computer Society Data Compression Conference. , Washington, DC, 1996.

[21] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," IEEE Transactions on Circuits Systems and Video Technology, vol. 6, no. 3, pp. 243250, Jun. 1996.

[22] J. E. Fowler, "Qccpack: An open-source software library for quantization, compression, and coding," in Proc. IEEE Computer Society Data Compression Conference. pp.554, 2000.

Author Biography

Matina Ch. Zerva is a Ph.D. student in the Department of Computer Science and Engineering of the University of Ioannina, Greece. She received her M.Sc. in Applied Mathematics and Computer Science from the department of Mathematics and her B.Sc. in Computer Science from department of Computer Science and Engineering of the same institution. Her main areas of interest are image enhancement and medical image compression.

Michalis Vrigkas received the Ph.D. degree from the Department of Computer Science and Engineering, University of Ioannina, Greece in 2016. He is currently a Postdoctoral Fellow at the University of Ioannina. In the past, he was an Adjunct Lecturer at the same Institution(2018-2019) and a Postdoctoral Fellow at the University of Houston, USA (2016-2018). His research interests include image and video processing, computer vision, machine learning, and biomedical image analysis.

Lisimachos P. Kondi received the Ph.D. degree in electrical and computer engineering from Northwestern University, Evanston, IL, USA, in 1999. He is currently Professor in the Department of Computer Science and Engineering, University of Ioannina, Greece. His research interests are in the general areas of signal and image processing and communications, including image and video compression and transmission over wireless channels and the Internet, sparse representations and compressive sensing, and super-resolution of video sequences.

Christophoros Nikou received the Ph.D. degree in image processing and computer vision from Louis Pasteur University, Strasbourg, France, in 1999. Since 2018 he has been a Professor in the Department of Computer Science and Engineering, University of Ioannina, Greece. Prof. Nikou was the General Chair of IEEE International Conference on Image Processing 2018 (ICIP'18). His research interests mainly include computer vision, pattern recognition, image processing and analysis, and their application to medical imaging.