Medical video compression using bandelet based on lifting scheme and SPIHT coding: In search of high visual quality

Mohammed Beladgham, Yassine Habchi, Mohamed Ben Aissa, Abdelmalik Taleb-Ahmed

Informatics in Medicine Unlocked 17 (2019) 100244

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: http://www.elsevier.com/locate/imu

Medical video compression using bandelet based on lifting scheme and SPIHT coding: In search of high visual quality

Mohammed Beladgham [a,*], Yassine Habchi [b], Mohamed Ben aissa [c], Abdelmalik Taleb-Ahmed [d]

[a] LTIT Laboratory, Electrical Engineering Department, Tahri Mohammed University, B.P 417 Bechar, 08000, Bechar, Algeria
[b] Institute of Science and Technology, University Center Salhi Ahmed, Naama, Algeria
[c] Electrical Engineering Department, University Center Belhadj Bouchaib, B.P 284, Ain Temouchent, 46000, Algeria
[d] IEMN DOAE UMR, CNRS 8520, University of Valenciennes and Hainaut Combresis, Valenciennes, France

ARTICLE INFO

Keywords:
Medical imaging
Bandelet transform
Cohen-daubechies-feauveau 9/7 (CDF9/7) filter
Lifting scheme
SPIHT encoder
Video coding

ABSTRACT

This paper investigates a coding method for medical video compression. The described technique has advantages of providing a higher visual quality of video coding, and efficiently reducing spatial redundancy. Firstly the bandelet transform, also called a non-separable transform, is studied in order to allow an efficient detection of the different complex geometries found in video. After describing the lifting operation in the bandelet transform, we reduced the artifactual effect caused by the quad-tree decomposition step in order to enhance the visual quality of the medical video sequence. Finally, the efficiency of the proposed method on medical video is tested and evaluated by means of a set of objective measurement parameters.

1. Introduction

Compression methods in the literature show that lossy coding may present a significantly undesirable effect as artifact, in the form of important changes in data content. In such cases, it is important to find a type of non-separable transform. In previous research, several transformation algorithms have been proposed for data compression: the wavelet transform [1], the curvelet transform [2], the contourlet transform [3], the wedgelet transform [4], the ridgelet transform [5], and the discrete cosine transform [6]. All of these have shown that data can be compressed simply by making changes in resolution, orientation, etc. The poor directionality of the cited transforms prevents their use in many diagnosis applications, however. To overcome these limitations, high-dimensional signals are necessary to achieve a considerable development in the theory of multi-scale geometric analysis (MGA) where several MGA transformations have appeared, in particular DBT (Discrete bandelet transform) [7,8], which is well suited to capture geometric regularity. Therefore, many researchers have studied and proved the efficiency of bandelet transforms for many applications [9–11]. The bandelet transform is accurate and practical because the search operation is not applied directly to the function, but is done to the set of blocks obtained after wavelet decomposition. The success of the bandelet is generally limited by the clarity degree in the video. The purpose of this article is to extend the bandelet transform for high-order geometry, which reduces the order of computational complexity and improves video clarity. Hence, the bandelet transform will be guided to the designated complex geometry that corresponds to the regularity of the complex geometric structures of video, but with another implementation based on a lifting scheme [12]. The advantage of the lifting scheme is its ability to develop transforms without going through the Fourier domain, by calculating simultaneously the low and high frequency sub-bands; also this structure of lifting is invertible. The lifting structure is used to improve the bandelet transform computation. The sub-band coders are used to code efficiently the generated coefficients. The Embedded Zerotree Wavelet (EZW) is firstly described by Shapiro's [13] as the first sub-band coding algorithm. Next, the EZW encoder is improved by the works of Said and Pearlman, who proposed an efficient encoder known as Partitioning Set in the Tree Hierarchical Encoder (SPIHT) [14]. In this work, motion estimation techniques have been used to reduce temporal redundancy [15]. Estimation and judgment of the recovered image quality are given by the PSNR, MSSIM, VIF and other evaluation parameters. The remainder of the paper is organized as follows: Section 2 describes the bandelet transform, Section 3 elucidates the concept of the lifting scheme, Section 4 concerns the biorthogonal wavelet CDF9/7, Section 5 presents the SPIHT algorithm, and section 6 is devoted to the proposed algorithm to illustrate its efficiency. Some

* Corresponding author.
E-mail address: beladgham.tlm@gmail.com (M. Beladgham).

https://doi.org/10.1016/j.imu.2019.100244
Received 1 June 2019; Received in revised form 21 August 2019; Accepted 9 September 2019
Available online 7 November 2019

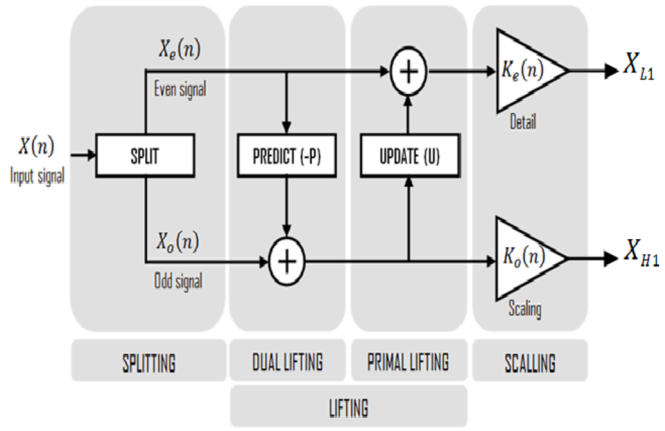

Fig. 1. Schematic diagram of the lifting-based 1D-WT.

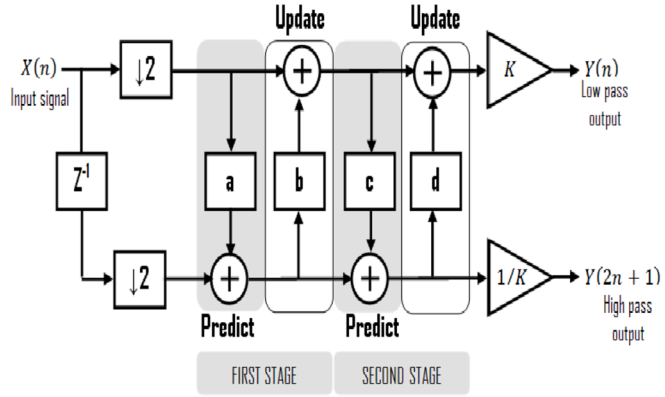

Fig. 2. Split, predict and update steps of forward CDF9/7 wavelet using lifting scheme.

simulation results are given in Section 8. Finally, Section 9 concludes the work.

2. The bandelet transform

The first generation bandelet transform was introduced by Pennec and Mallat [7]. It combines decomposition into wavelets. The estimation of the geometry is done by studying image contours. A rectification of curves is conducted, and consists in deforming the curve to make it horizontal. Finally, wavelet decomposition is applied. In practice, it is first necessary to segment each image into regions, but this transform type produces blocking artifacts. The second-generation bandelets were then introduced in Ref. [8]. The idea is again to build an orthogonal base adapted to an image. However, this time the search for geometry is not done on a geometric function directly but in the different sub-bands obtained after wavelet decomposition. In order to improve the geometry detection step, inter-scale relations can be considered via the construction of a tree connecting the coefficients, at each level of detail, corresponding to the same spatial location.

3. The lifting scheme

The wavelet transform (WT), in general, produces floating point coefficients. Although these coefficients are used to reconstruct an original image perfectly in theory, the use of finite precision arithmetic and quantization results in a lossy scheme.

Recently, reversible integer WT's have been introduced [16]. In Ref. [17], and in order to reduce computational complexity, Calderbank et al. have shown that the biorthogonal WT can be implemented in the lifting scheme, as illustrated in Fig. 1. Note that only the decomposition part of 1D-WT is depicted in Fig. 1, for the reconstruction part is the reverse version of the one in Fig. 1.

The lifting-based 1D-WT consists in splitting, lifting, and scaling modules, and the 1D-WT is treated as prediction error decomposition. In Ref. [18], the lifting scheme is used to improve quality of medical image compression.

In Fig. 1, $X(n)$ is the original signal and X_{L1} X_{H1} are the output signals obtained through the three steps of lifting-based 1DWT:

Splitting: In this part, the signal $X(n)$ is divided in $X_e(n) = X(2n)$ and $X_o(n) = X(2n+1)$ that denote all even and odd samples of $X(n)$, respectively.

Lifting: An estimation is used through prediction operation P to estimate $X_o(n)$ from $X_e(n)$, and results in $d(n)$, which represents the detailed part of X. Next, it is updated by applying the update operation U and the resulting signal is combined with $X_e(n)$ to $s(n)$ estimate, which represents the smooth part of $X(n)$.

Scaling: A normalization factor is applied to $d(n)$ and $s(n)$, respectively. In order to produce the wavelet sub-band X_{L1} and X_{H1}, the even part of $s(n)$ is multiplied by a normalization factor K_e , and the odd part of $d(n)$ is multiplied by K_0, respectively.

The lifting scheme of the biorthogonal transform 9/7 is made up of four steps: two operators of predictions and two update operators, as shown it Fig. 2.

The lifting scheme is one of the most efficient structures. It is used in many applications, such as in a robust and secure color image watermarking [19], audio watermarking [20], Fault diagnosis [21], and image fusion [22].

4. Biorthogonal wavelet CDF9/7

The biorthogonal wavelet Cohen-Daubechies-Feauveau 9/7 (CDF9/7) filter (also commonly known as Daubechies 9/7 filter) is part of the family of symmetric biorthogonal wavelet CDF characterized by their simplicity, biorthogonality, symmetry, and compact support. The low pass filters associated with wavelet 9/7 have $p = 9$ coefficients in the analysis, $p = 7$ coefficients to synthesize, as described in Table 1. They have $N = 4$ null moments in analysis, and $\tilde{N} = 4$ in synthesis. The wavelets 9/7 have a great number of null moments for a relatively short support. They are more symmetrical and very close to orthogonality. This is an important feature in coding, which ensures that the reconstruction error is very close to the quantization error in terms of the mean squared error. Antonini and Barlaud were the first [23] to show the superiority of the biorthogonal wavelet transform 9/7 for the decorrelation of natural images. It has been widely used in image coding [14,24,25] and is used by the JPEG-2000 codec [26].

5. The SPIHT algorithm

SPIHT [14] is one of the most advanced schemes available nowadays. It is widely used to enhance the quality of image and video. The SPIHT is a refined version of EZW used to exploit the inherent similarities across the sub-bands in a wavelet decomposition of an image. It codes the most important wavelet transform coefficients first, and transmits the bits. It delivers an excellent compression performance when applied to 1D, 2D, and 3D signals. Basically, the main advantages of SPIHT are: (i) provides a good image quality, (ii) very convenient for storage and progressive transmissions of significance coefficient, and (iii) used either for lossless or lossy compression.

The SPIHT sorts information in three ordered lists: LIP for the list of insignificant pixels, LIS for the list of insignificant sets, and LSP for the list of significant pixels. At the initialization step, the pixels of frame are added to LIP, and those with descendants are added to LIS. The LSP is an empty list. Coding starts with LIP; if the number of pixels in frame is insignificant, it remains in LIP; otherwise, pixels are moved to LSP. Similarly,

Table 1
The analysis and synthesis filter coefficients.

The analysis filter coefficients.			The synthesis filter coefficients	
i	Low-pass filter	High-pass filter	Low-pass filter	High-pass filter
0	0.6029490182363579	+1.115087052457000	+1.115087052457000	0.6029490182363579
	0.6029490182363579			
±1	+0.266864118442875	+0.591271763114250	– 0.591271763114250	– 0.266864118442875
±2	– 0.078223266528990	– 0.057543526228500	– 0.057543526228500	– 0.078223266528990
±3	– 0.016864118442875	– 0.091271763114250	+0.091271763114250	+0.016864118442875
±4	+0.026748757410810			+0.026748757410810

insignificant LIS pixels remain in LIS, while significant pixels are partitioned into significant type A and type B sets and four pixels; the type B set is added to the end of LIS whereas the rest are examined for significance.

The SPIHT use the following sets of coordinates:

$O(i,j)$: set of coordinates of all offspring of node (i,j).
$D(i,j)$: set of coordinates of all descendants of node (i,j).
H: set of coordinates of all spatial orientation tree roots (nodes in the highest pyramid level).

$$L(i,j) = D(i,j) - O(i,j)$$

A LIS entry is of type:

- A if it represents $D(i,j)$,
- B if it represents $L(i,j)$.

SPIHT uses the function $S_n(T) = \begin{cases} 1 & \max_{(i,j)\in T}\{|c(i,j)|\} \geq 2^n \\ 0 & \text{otherwise} \end{cases}$

To indicate the significance of a set of coordinates T. The steps of the algorithm are as follows:

(A) Initialization:
output $n = \left\lfloor \log 2\left(\max_{(i,j)}\left\{\left|C_{i,j}\right|\right\}\right)\right\rfloor$;
set the LSP as an empty list;
add the coordinates $(i,j) \in H$ to the list LIP, and only those with descendants also to the LIS, as type A entries.

(B) Sorting pass:
(B.1) for each entry (i,j) in the LIP do:
(B.1.1) transmit $Sn(i,j)$;
(B.1.2) if $Sn(i,j) = 1$ then move (i,j) to the LSP and transmit the sign of $C_{i,j}$;
(B.2) for each entry (i,j) in the LIS do:
(B.2.1) if the entry is of type A then
- transmit $Sn(D(i,j))$;
- if $Sn(D(i,j)) = 1$ then
  - for each $(k,l) \in O(i,j)$ do:
    - transmit $Sn(k,l)$;
    - if $Sn(k,l) = 1$ then add (k,l) to the LSP and output the sign of $C_{k,l}$;
    - if $Sn(k,l) = 0$ then add (k,l) to the end of the LIP;
  - if $L(i,j) \neq \varnothing$ then move (i,j) to the end of the LIS, as an entry of type B, and go to step (2.2.2); otherwise, remove entry (i,j) from the LIS;

(B.2.2) if the entry is of type B then
- transmit $Sn(L(i,j))$;
- if $Sn(L(i,j)) = 1$ then
  - add each $(k,l) \in O(i,j)$ to the end of the LIS as an entry of type A;
  - remove (i,j) from the LIS.

(C) Refinement pass:
for each entry (i,j) in the LSP, except those included in the last sorting pass, output the nth most significant bit of $\left|C_{i,j}\right|$;
(D) Quantization-step update:
decrement n by 1 and go to step (2).

6. The proposed algorithm

The detection of complex geometrics, which has been dealt with by few authors, is an important issue to address. In video sequences, due to contours, poor visual quality between successive frames of videos often occurs. Fig. 3 shows the block diagram of the proposed medical video compression method using bandelet transforms based on the lifting scheme. The input medical video is decomposed in approximation and details through the two-dimensional CDF9/7 wavelet [27] based on the lifting scheme. The approximation is a coarser version of the original frame, while the other sub-band details represent the horizontal, vertical, and diagonal directions, respectively. The resulting wavelet transform of each frame of video is partitioned into smaller dyadic squares by

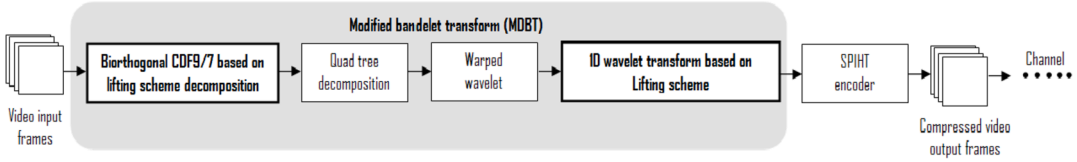

Fig. 3. The block diagram of the proposed algorithm.

quad-tree decomposition, where the transformed video is recursively divided into four smaller squares of varying dyadic sizes. To define the singularities, a global regularity condition is imposed on the squares results of the quad-tree decomposition algorithm. The blocks should be under satisfactory Lagrangian conditions defined as:

$$\sum_j \left(\left\| f_j - f_{jR} \right\|^2 + \lambda \left(R_{jS} + R_{jG} + R_{jC} \right) \right) \tag{1}$$

where

$\left\| f_j - f_{jR} \right\|^2$: The Euclidean norm restricted to the region S_j
$\lambda = \frac{3}{28}$: The Lagrange multiplier
R_{jS}: The number of bits needed to encode the dyadic sub-square
R_{jG}: The number of bits needed to encode the geometric parameters
R_{jC}: The number of bits needed to encode the warped coefficient

The Lagrangian of the subdivided square must be small enough to build the best quad-tree and to minimize the probability that a square contains singularities. The geometric flow is applied for each bloc as a result of the quad-tree decomposition to follow the different complex geometry with precision. In addition, the successive division of the square is directly linked to the presence of the uniformed geometrical flow in the sub-square. Then, a rectification step is carried out, the deformed wavelet coefficients are calculated with an operation that consists in deforming the curve to make it horizontal if it is rather horizontal, and inversely, after a mono-dimensional wavelet transformation based on structure lifting is applied consecutively. The Lifting structure is used to improve and reduce the limits of the bandelet, and therefore allows extending the wavelet theory in a non-linear framework, and simply permits the construction of non-linear transforms. Finally, the resultant modified bandelet coefficients are encoded using the SPIHT algorithm.

The proposed algorithm is formulated as follows:

1. Decomposition of the input medical video through 2D wavelet transforms (biorthogonal CDF9/7).
2. Applying quad-tree decomposition recursively for each frame and for each sub-band as a result of biorthogonal CDF9/7 to restrict complex geometrics in dyadic square.
3. Warping the uniformed geometric flow applied in each dyadic square with a warped operator to horizontal and vertical directions.
4. Computing the warped wavelet coefficients.
5. Performing 1D wavelet transform based on lifting scheme.
6. The resulting bandelet coefficients are encoded using the SPIHT encoder.

7. Video quality parameters assessment

The objective metrics are used to assess the performance of the proposed algorithm and the measure the quality of the reconstructed video.

7.1. The peak signal-to-noise ratio (PSNR)

The PSNR represents an objective evaluation parameter for the measurement of visual quality. It is defined as follows:

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

where

- $(2^n - 1)$ is the dynamic of the signal (the maximum possible value for a pixel). In the standard case of an image where the components of a pixel are coded on $n = 8$ bits, $(2^n - 1) = 255$.
- MSE represents the mean square error between two frames, namely the original frame $f(i,j)$ and the recovered frame $f_r(i,j)$ of size $M \times N$ and is given by

$$MSE = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (f(i,j) - f_r(i,j))^2 \tag{3}$$

7.2. The mean SSIM (MSSIM)

The PSNR is an easy, fast, and very popular quality measurement metric, widely used to compare the quality of video encoding and decoding. Although a high PSNR generally means a good quality reconstruction, this is not always the case. Indeed, PSNR requires the original image for comparison but this may not be available in every case; also PSNR does not correlate well with subjective video quality measures, therefore it is not very suitable to perceived visual quality. Hence, there is an interest in considering a measure of structural similarity (SSIM) adapted to the human visual system. The SSIM index introduces three key features: luminance l, contrast c, and structure s.

$$SSIM(f, f_r) = l(f, f_r) c(f, f_r) s(f, f_r) \tag{4}$$

The luminance comparison function l is determined by the following expression:

$$l(x,y) = \frac{2M_x M_y + C_1}{M_x^2 + M_y^2 + C_1} \tag{5}$$

where

M_x and M_y are the mean intensity of the signal x and y defined by $M_x = \frac{1}{N} \sum_{i=1}^{N} x_i$ and $M_y = \frac{1}{N} \sum_{i=1}^{N} y_i$ respectively.

$C_i = K_i^2 D^2$, $i = 1, 2$, and K_i is a constant, such as $K_i << 1$ and D is the dynamic range of the pixel values ($D = 255$ corresponds to a grey-scale digital image when the number of bits/pixel is 8).

The contrast comparison function c takes the following form:

$$c(x,y) = \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \tag{6}$$

where

$$\sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^{N} \left(x_i^2 - (M_x)^2 \right) \right)^{1/2}$$

The structure comparison function s is defined as follows:

$$s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3} = \frac{cov(x,y) + C_3}{\sigma_x \sigma_y + C_3} \tag{7}$$

$$C_3 = \frac{C_2}{2} and\ cov\left(x, y\right) = M_{xy} - M_x M_y$$

with

$$M_{xy} = \frac{1}{N-1}\sum_{i=1}^{N} x_i y_i$$

Hence, the explicit expression of the structural similarity (SSIM) index is:

$$SSIM(x,y) = \frac{\left(2M_x M_y + C_1\right)\left(2\sigma_{xy} + C_2\right)}{\left(M_x^2 + M_y^2 + C_1\right)\left(\sigma_x^2 \sigma_y^2 + C_2\right)} \tag{8}$$

where

$$\sigma_{xy} = \left(\frac{1}{N-1}\sum_{i=1}^{N}\left(\left(x_i y_i\right)^2 - \left(M_{xy}\right)^2\right)\right)^{1/2}$$

Generally, over the whole video coding, a mean value of SSIM is required as Mean SSIM (MSSIM):

$$MSSIM(f, f_r) = \frac{1}{L}\sum_{i=1}^{L} SSIM(f_i, f_{ri}) \tag{9}$$

where f_i and f_{ri} are the contents of frames (original and recovered respectively) at the ith local window (or sub-image), and L is the total of local windows number in frame. The MSSIM values exhibit greater consistency with the visual quality.

7.3. The visual information fidelity (VIF)

In 2006, Sheikh and Bovik proposed a new paradigm for video quality assessment; visual information fidelity (VIF). This criterion quantifies the Shannon information that is shared between the original and recovered images according to the contained information in the original image itself. It uses natural scene statistics modelling in conjunction with an image-degradation model and a human visual system (HVS) model.

Visual Information Fidelity uses the Gaussian scale mixture model (GSM) in the wavelet domain. To obtain VIF, one performs a scale-space-orientation wavelet decomposition using the steerable pyramid, and models each sub-band in the source as $C = SU$, where S is a random field of scalars and U is a Gaussian vector.

The distortion model is $D = GC + v$, where G is a scalar gain field and v is an additive Gaussian noise.

VIF then assumes that the distorted and source images pass through the human visual system and the HVS uncertainty is modelled as visual noise N and $N^{'}$ for the source and distorted image, respectively.

The model is then:

$$\text{Reference signal} \quad E = C + N \tag{10}$$

$$\text{Test signal} \quad F = D + N \tag{11}$$

where E and F denote the visual signal at the output of the HVS model from the reference and the test videos respectively, from which the brain extracts cognitive information.

The VIF measure takes values between 0 and 1, where 1 means perfect quality and is given by:

$$VIF = \frac{\sum_j I\left(C^j;\ F^j / s^j\right)}{\sum_j I\left(C^j;\ E^j / s^j\right)} \tag{12}$$

where, $I(X; Y/z)$ is the conditional mutual information between X and Y, given z; C denotes the random field from a channel in the original image, s^j is a realization of S^j for a particular image and the index j runs through all the sub-bands in the decomposed image.

7.4. The edge measurement (EDGE)

This type of quality measurement can be obtained from,

$$Edge = \frac{1}{M \times N}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(Q\left(i,j\right) - \widehat{Q}\left(i,j\right)\right)^2 \tag{13}$$

where $Q(i,j)$ and $\widehat{Q}(i,j)$ are the gradients of the original and compressed image.

7.5. The weighted peak signal to noise ratio (WPSNR)

This parameter is based on the fact that the human eye is less sensitive to changes in textured areas than in smooth areas, WPSNR has another parameter that takes into account the texture of the image. The WPSNR formula is shown below:

$$WPSNR = 10\ log_{10}\left(\frac{\left(2^n - 1\right)^2}{NVF \times MSE}\right) \tag{14}$$

The noise visibility function (NVF) uses a Gaussian model to estimate the amount of texture content in all parts of the image. In textured regions with edges, NVF will have a value greater than 0, while in smooth regions the value of NVF will be greater than 1.

$$NVF = NORM\left\{\frac{1}{1 + \delta_{bloc}^2}\right\} \tag{15}$$

where δ_{bloc} is the luminance variance for a block and NORM is the normalization function.

7.6. The visual signal to noise ratio (VSNR)

The principle of the parameter VSNR is based first on the difference image computation E between the original image I and the distorted image $\widehat{I}$. The images I and E are then subjected to a discrete wavelet transformation (2D-DWT, M decomposition levels).

Within each sub-band $\{S_I\}$ and $\{S_E\}$, the VSNR then calculates the distortion visibility, comparing the contrast of the distortion to the detection threshold, and then calculates the RMS contrast of the error signal (d_{pc}).

$$\begin{aligned} d_{pc} &= C(E) \\ &= \frac{1}{\mu_{L(I)}}\left(\frac{1}{N}\sum_{i=1}^{N}\left[L(E_i + \mu_I) - \mu_{L(E+\mu_I)}\right]^2\right)^{1/2} \end{aligned} \tag{16}$$

where $\mu_I = \frac{1}{N}\sum_{i=1}^{N} I_i$, $\mu_{L(I)} = \frac{1}{N}\sum_{i=1}^{N} L(I_i)$, $L(E_i + \mu_I)$ and $\mu_{L(E+\mu_I)} = \frac{1}{N}\sum_{i=1}^{N} L(E_i + \mu_I)$ respectively denote the average pixel value, the average luminance of I, the distortion of the i-pixel luminance of the image, and the distortion of the average luminance.

Based on the strategy of the global priority idea, in the HVS, the VSNR calculates a global priority, preserving the contrast (d_{gp}).

$$d_{gp} = \left(\sum_{m=1}^{M}\left[C^*\left(E_{f_m}\right) - C\left(E_{f_m}\right)\right]^2\right)^{1/2} \tag{17}$$

where $C(E_{f_m})$ denotes the current contrast of distortion.

The final index is a linear combination of d_{pc} and d_{gp}. We define a visual distortion VD, as the linear combination of d_{pc} and d_{gp}.

$$VD = \alpha\, d_{pc} + (1 - \alpha)\frac{d_{gp}}{\sqrt{2}} \tag{18}$$

where the parameter $\alpha \in [0,1]$.

Table 2
The characteristics of tested medical video.

Tested medical video	Number of Frames	Number of Frames/ Second	Width × Height	Video Format
CORONARY	594	30	360 × 360	AVI
MRI1	145	30	240 × 240	AVI
MRI2	296	30	340 × 352	AVI
KERATINOCYTES AND NEURONAL CELLS	1260	30	960 × 720	AVI
BLADER CANCER	662	30	480 × 320	AVI

The quantity d_{pc} is about this distance from the origin. The quantity d_{gp} is the distance between the two points; so, in general, $d_{gp} \in [0, \sqrt{2}\, d_{pc}]$ and so, $VD \in [0,\ d_{pc}]$.

The amount is d_{pc} necessary to account for differences in perceived fidelity when two images are of different total distortion contrasts, but the two images have $d_{gp} = 0$. If both images have $d_{gp} = 0$, the image with the higher total distortion contrast (d_{pc}) will generally be less ranked in the perceived fidelity (assuming the additional distortion contrast is visible). The quantity d_{pc} is necessary to account for this condition. The VSNR, in dB, is therefore given by:

$$VSNR = 10 \log_{10}\left(\frac{C^2(I)}{(VD)^2}\right) = 20 \log_{10}\left(\frac{C(I)}{\alpha\, d_{pc} + \left(1 - \alpha\right)\frac{d_{gp}}{\sqrt{2}}}\right) \tag{19}$$

where $C(I)$ indicates the RMS contrast of the original image I, given by $C(I) = \sigma_{L(I)}/\mu_{L(I)}$ ($\sigma_{L(I)}$ is the standard deviation, $\mu_{L(I)}$ denotes the average luminance of I). Note that when the global priority is at most disturbed for a given data $C(E)$, at most $d_{gp} = \sqrt{2}\, d_{pc}$, $VD = d_{pc}$ therefore $VSNR = 20 \log_{10}(C(I)/d_{pc}) = 20 \log_{10}(C(I)/C(E))$; in this case the VSNR is given by the SNR contrast from I to E.

8. Results and discussion

In order to evaluate the efficiency of the proposed algorithm, the CDF9/7 + SPIHT, DBT(CDF9/7)+EZW, DBT(GALL5/3)+SPIHT, DBT (CDF9/7)+SPIHT and DBT(LIFTING)+SPIHT algorithms are respectively applied to the 8-bit, 512 × 512 grayscale medical video. Estimates and judgments of the compressed medical video quality are given by the PSNR, the MSSIM similarity index and VIF [28,29] evaluation parameters. An input medical video is decomposed into four levels(See. Table 2).

The original medical videos used in the compression test are shown in Fig. 4.

- **CORONARY**: Coronary artery disease (CAD), also known as ischemic heart disease (IHD), involves the reduction of blood flow to the heart muscle due to build-up of plaque in the arteries of the heart.
- **MRI1, MRI2**: Magnetic resonance imaging (MRI) is a medical imaging technique used in radiology to form pictures of the anatomy and the physiological processes of the body. MRI scanners use strong magnetic fields, magnetic field gradients, and radio waves to generate images of the organs in the body. MRI does not involve X-rays or the use of ionizing radiation, which distinguishes it from CT or CAT scans and PET scans. Magnetic resonance imaging is a medical application of nuclear magnetic resonance (NMR). NMR can also be used for imaging in other applications such as NMR spectroscopy.
- **KERATINOCYTES AND NEURONAL CELLS**: Keratinocytes are cells constituting 90% of the superficial layer of the skin (epidermis) and integuments (nails, hair, hairs, feathers, scales). They synthesize keratin (keratinization), a fibrous protein that is insoluble in water, which gives the skin its impermeability and external protection properties.
- **BLADER CANCER**: Bladder cancer is a growth of abnormal tissue, known as a tumor, that develops in the bladder lining. In some cases, the tumor spreads into the bladder muscle.

The all tested medical data are from:

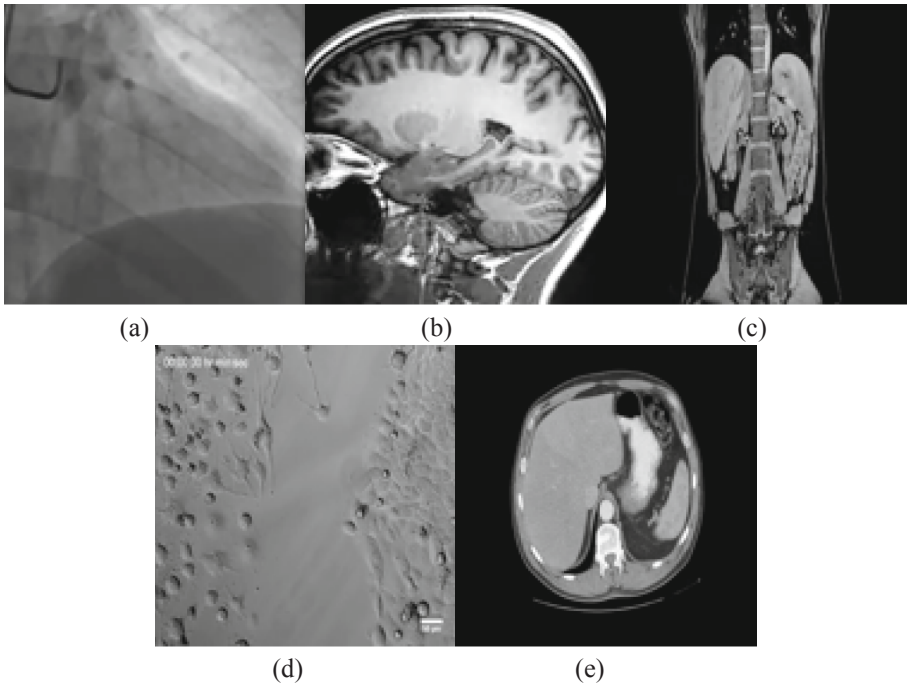

Fig. 4. Used medical video for evaluation, (a). CORONARY, (b). MRI1, (c). MRI2, (d). KERATINOCYTES AND NEURONAL CELLS, (e). BLADER CANCER.

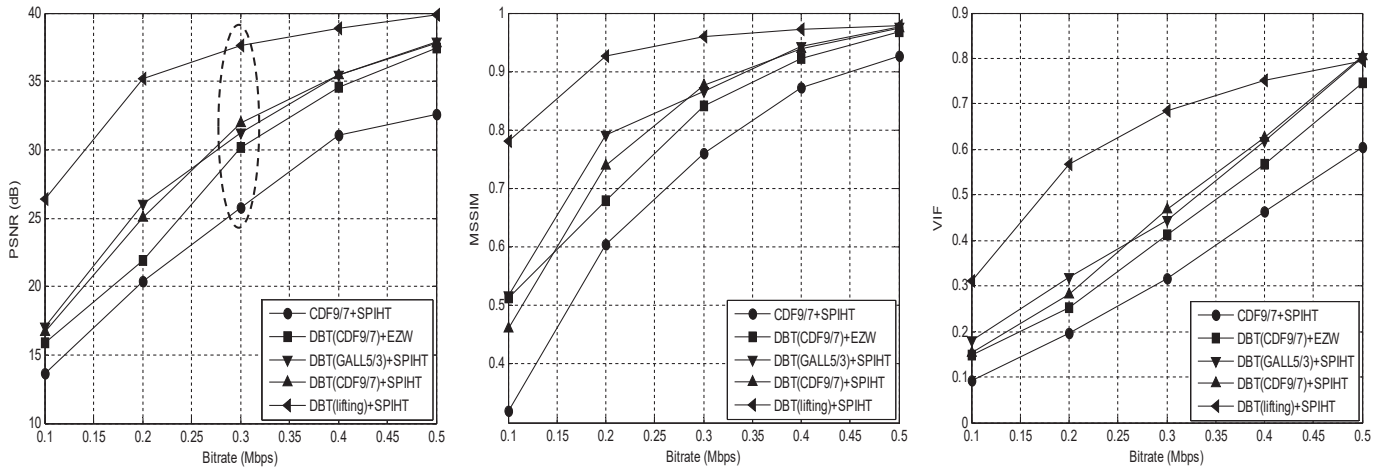

Fig. 5. PSNR, MSSIM and VIF coding results comparison between CDF9/7 + SPIHT, DBT(CDF9/7)+EZW, DBT(GALL5/3)+SPIHT, DBT(CDF9/7)+SPIHT and DBT (Lifting)+SPIHT algorithms for CORONARY-ANGIOGRAPHY video.

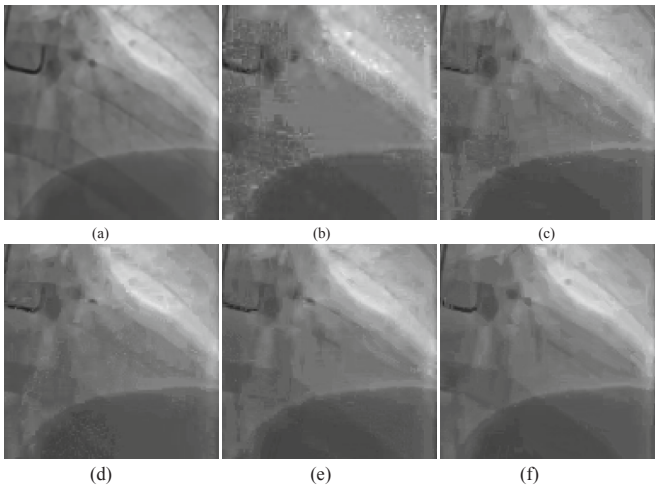

Fig. 6. Performance of all tested algorithms for CORONARY-ANGIOGRAPHY video at 0.3 Mbps. (a). Original frame, (b). CDF9/7 + SPIHT, (c). DBT(CDF9/7)+EZW, (d). DBT(GALL5/3)+SPIHT, (e). DBT(CDF9/7)+SPIHT.

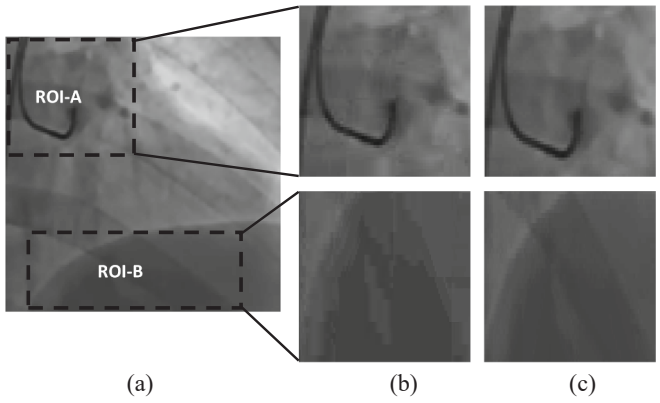

Fig. 7. The visual region of interest (ROI) comparison results for CORONARY-ANGIOGRAPHY video at 0.3 Mbps: (a) Original video. (b) ROI by DBT(CDF9/7)+SPIHT. (c) ROI by DBT(LIFTING)+SPIHT.

- INSERM (French National Institute of Health and Medical Research), including large databases used for the comparative analysis.
- www.GEMedicalSystem.com

Fig. 5 shows PSNR, MSSIM and VIF at several bit rates. As we can notice form this figure, our proposed method outperforms baseline CDF9/7 + SPIHT, DBT(CDF9/7)+EZW, DBT(GALL5/3)+SPIHT and DBT(CDF9/7)+SPIHT. At the same bit rate, the PSNR, MSSIM and VIF values of the reconstructed medical videos for DBT(LIFTING)+SPIHT algorithm are higher than those for the other tested algorithms. The gain improvement in terms of PSNR, for example at 0.3Mbps for CORONARY ANGIOGRAPHY video, reaches 7,8701 dB, 4,3162 dB, 3,3902 dB, and 3,4364 dB over the algorithms of CDF9/7 + SPIHT, DBT(CDF9/7)+EZW, DBT(GALL5/3)+SPIHT and DBT(CDF9/7)+SPIHT respectively. Moreover, it can be seen that DBT(LIFTING)+SPIHT has better performance for all bit rates (from 0.1 Mbps until 0.5Mbps).

Fig. 6 shows the decoded CORONARY ANGIOGRAPHY video for the proposed algorithm at 0.3 Mbps, where we compare the perceptual quality between the all tested algorithms. It can be observed that the subjective quality of medical video in Fig. 6 f is improved over Fig. 6 b, Fig. 6 c, Fig. 6 d, and Fig. 6 e. The overview of visual comparison results of the local regions are shown in Fig. 7.

As a second evaluation, the proposed algorithm is applied to the other medical sequences. The compression is performed to select sequences in which our algorithm outperforms preceding ones. From Fig. 8, we note that the DBT (LIFTING)+SPIHT algorithm performs better for the CORONARY video. As a final evaluation, the proposed algorithm is also intended for comparison with another compression standard known as H.264 (see Fig. 9).

The H.264 is the newest international video coding standard. However this standard is based on the use of the popular transform termed the Discrete Cosine Transform (DCT) which suffers from blocking artifacts, limited block size, and is unable to capture complex geometries, so it kills all contours existing in the video and therefore reduces directly the visual quality, and data is lost in the reconstructed video sequence [30]. The obtained results show that bandelet based on the lifting scheme and SPIHT algorithm provide high visual quality and important PSNR gain values, especially for low bit rate (for example the PSNR gain (Fig. 8) is 6.69 dB, 4.58 dB and 1.54 dB at 0.25 Mbps, 0.30 Mbps and 0.35Mbps respectively).

9. Conclusion

In this paper, an improved lossy medical video coding algorithm based on the modified bandelet basis and SPIHT coder is proposed. The simulation results indicate that the proposed algorithm can produce a better reconstructed medical video in terms of visual quality. The objective parameters (PSNR, MSSIM, etc.) of the proposed algorithm DBT(lifting)+SPIHT are better than the objective parameters of DBT +

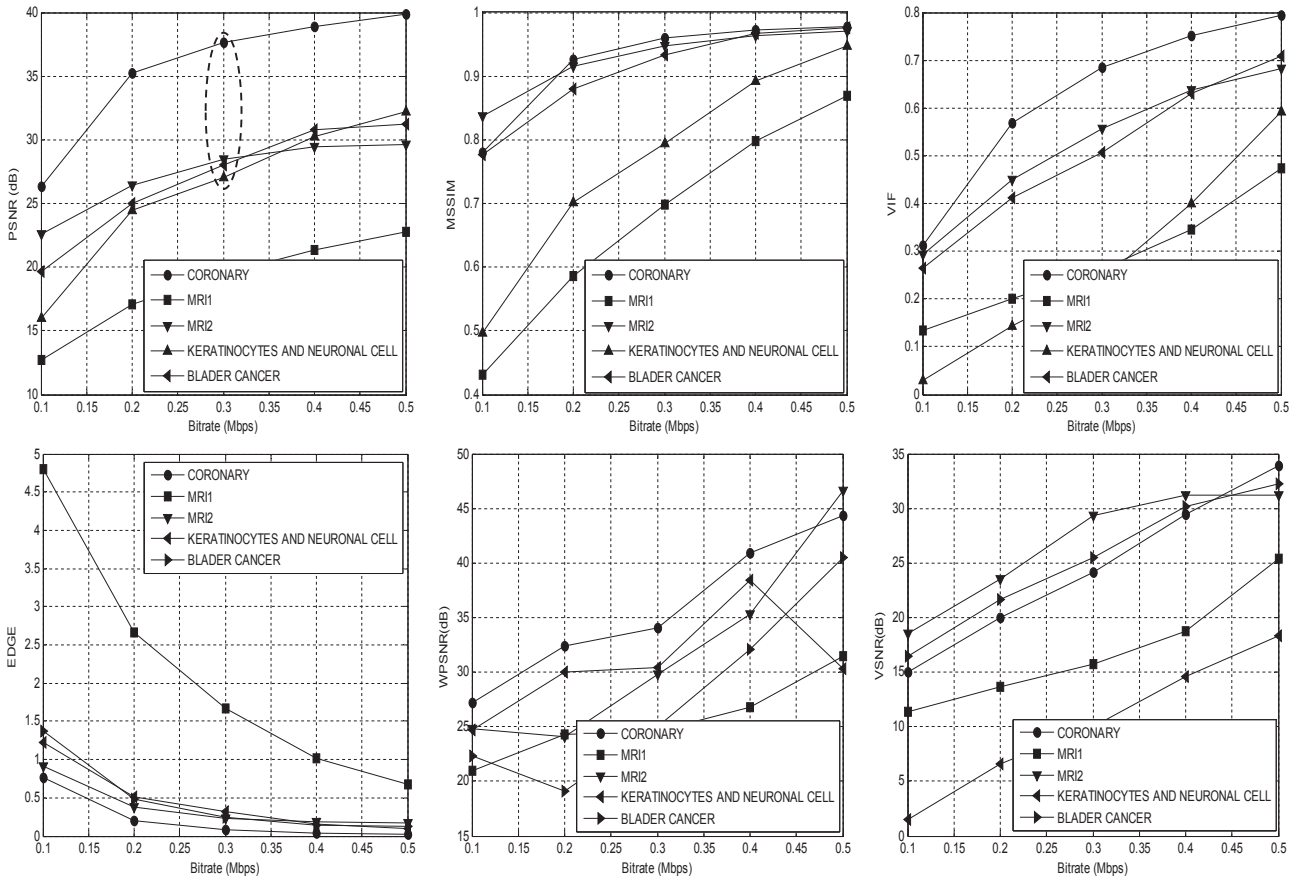

Fig. 8. PSNR, MSSIM, VIF, EDGE, WPSNR, and VSNR coding results comparison between all tested medical video using DBT(LIFTING)+SPIHT algorithm.

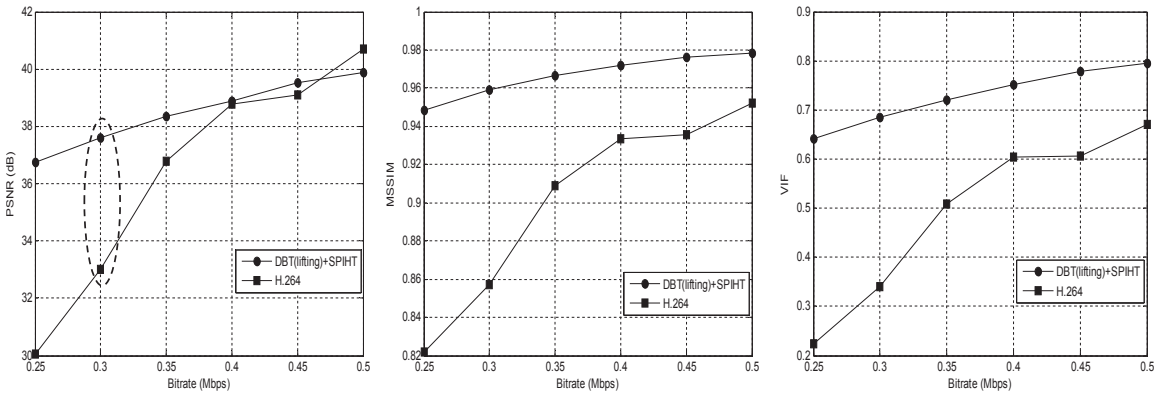

Fig. 9. PSNR, MSSIM and VIF coding results comparison for CORONARY-ANGIOGRAPHY video between DBT(Lifting)+SPIHT algorithm and H.264 standard.

SPIHT algorithm and H.264 standard at a given bit rate for all tested medical videos. The proposed improvement simplifies the coding and improves the visual quality of medical video. In perspective, we endeavor to apply this algorithm to color medical video and compare it with HEVC (H.265).

References

[1] Vetterli M. Wavelets, approximation and compression. IEEE Signal Process Mag 2001;18(5):59–73. https://doi.org/10.1109/79.952805.

[2] Changsheng L, Hong L, Guangzheng L, Xiujuan Z. Combined sparse representation based on curvelet transform and local DCT for multi-layered image compression. In: IEEE 3rd international conference on communication software and networks; 2011. p. 316–20. https://doi.org/10.1109/ICCSN.2011.6013723.

[3] Belbachir AN, Goebel PM. The contourlet transform for image compression, physics in signal and image process. 2005.

[4] Donoho DL. Wedgelets: nearly minimax estimation of edges. Ann Stat 1999;27(3): 859–97. https://doi.org/10.1214/aos/1018031261.

[5] Do MN, Vetterli M. The finite ridgelet transform for image representation. IEEE Trans Image Process 2003;12(1):16–28. https://doi.org/10.1109/TIP.2002.806252.

[6] Yen-Yu C, Shen-Chuan T. Embedded medical image compression using DCT based subband decomposition and modified SPIHT data organization. In: Proceedings. Fourth IEEE symposium on bioinformatics and bioengineering; 2004. p. 167–74. https://doi.org/10.1109/BIBE.2004.1317339.

[7] Pennec EL, Mallat S. Image compression with geometrical wavelets. In: Proceedings 2000 international conference on image processing (cat. No.00CH37101); 2000. p. 661–4. https://doi.org/10.1109/ICIP.2000.901045.

[8] Pennec EL, Mallat S. Sparse geometric image representation with bandelets. IEEE Trans Image Process 2005;14(4):423–38. https://doi.org/10.1109/TIP.2005.843753.

[9] Yanga S, Qi W, Wanga Z, Jiao L. SAR image compression based on multi-bandelets and geometric flow optimization, Multispectral Image Acquisition and Processing. Proc SPIE Int Soc Opt Eng 2009:7494. https://doi.org/10.1117/12.832427.

[10] Wenge Z, Fang L, Licheng J, Xinbo G. SAR image despeckling based on edge detection and non subsampled second generation bandelets. J Syst Eng Electron 2009;20(3):519–26.
[11] Qu X, Yan J, Xie G, Zhu Z, Chen B. A novel image fusion algorithm based on bandelet transform. Chin Optic Lett 2007;5(10):569–72.
[12] Sweldens W. The lifting scheme: a custom-design construction of biorthogonal wavelets. Appl Comput Harmon Anal 1996;3(2):186–200. https://doi.org/10.1006/acha.1996.0015.
[13] Shapiro JM. Embedded image coding using zerotrees of wavelet coefficients. IEEE Trans Signal Process 1993;41(12):3445–62. https://doi.org/10.1109/78.258085.
[14] Said A, Pearlman W. A new fast and efficient image codec based on set partitioning in hierarchical trees. IEEE Trans Circuits Syst Video Technol 1996;6(3):243–50. https://doi.org/10.1109/76.499834.
[15] Ria, Subhash C Y, Mapping Three Step Search algorithm onto a systolic array processor in H.264, 2016 3rd international conference on recent advances in information technology (RAIT) 2016:561-563. DOI: 10.1109/RAIT.2016.7507960.
[16] Zandi A, Allen JD, Schwartz EL, Boliek M. CREW: compression with reversible embedded wavelets. In: Proceedings DCC '95 data compression conference; 1995. p. 212–21. https://doi.org/10.1109/DCC.1995.515511.
[17] Calderbanka AR, Daubechies I, Sweldens W, Boon-Lock Y. Wavelet transforms that map integers to integers. Appl Comput Harmon Anal 1998;5(3):332–69. https://doi.org/10.1006/acha.1997.0238.
[18] Beladgham M, Bessaid A, Moulay Lakhdar A, Taleb-ahmed A. Improving quality of medical image compression using biorthogonal CDF wavelet based on lifting scheme and SPIHT coding. Serb J Electr Eng 2011;8(2):163–79. https://doi.org/10.2298/SJEE1102163B.
[19] Hao C, Weiliang X. Secure and robust color image watermarking for copyright protection based on lifting wavelet transform. In: 2018 25th international conference on mechatronics and machine vision in practice (M2VIP); 2018. https://doi.org/10.1109/M2VIP.2018.8600889.
[20] Abhishek KG, Archit A, Akhilesh S, Divay V, Dharmendra K. Blind audio watermarking using adaptive quantization and Lifting wavelet transform. In: 2018 5th international conference on signal processing and integrated networks (SPIN); 2018. https://doi.org/10.1109/SPIN.2018.8474250.
[21] Binqiang C, Yang L, Nianyin Z, Wangpeng H. Fractal lifting wavelets for machine Fault diagnosis. IEEE Adv Prognostics Syst Health Manag 2019;(7). https://doi.org/10.1109/ACCESS.2019.2908213.
[22] Prakash O, MinPark C, Khare A, Jeon M, Gwak J, Multiscale fusion of multimodal medical images using lifting scheme based biorthogonal wavelet transform, Optik, 182: 995-1014, DOI: https://doi.org/10.1016/j.ijleo.2018.12.028.
[23] Antonini M, Barlaud M, Mathieu P, Daubechies I. Image coding using wavelet transform. IEEE Trans Image Process 1992;1(2):205–20. https://doi.org/10.1109/83.136597.
[24] Illasenor JD, Belzer B, Liao J. Wavelet filter evaluation for image compression. IEEE Trans Image Process 1995;4(8):1053–60. https://doi.org/10.1109/83.403412.
[25] Haouam I, Beladgham M, Bendjillali RI, Habchi Y. MRI image compression using level set method and biorthogonal CDF wavelet based on lifting scheme. In: 2018 international conference on signal, image, vision and their applications (SIVA); 2018. https://doi.org/10.1109/SIVA.2018.8661068.
[26] Unser M, Blu T. Mathematical properties of the JPEG2000 wavelet filters. IEEE Trans Image Process 2003;12(9):1080–90. https://doi.org/10.1109/TIP.2003.812329.
[27] Dabhole SH, Jadhav ST. An efficient codec of 2D adaptive directional lifting based on CDF9/7 with improved SPIHT algorithm for lossy to lossless image coding. In: 2nd international conference on electronics and communication systems (ICECS); 2015. p. 84–91. https://doi.org/10.1109/ECS.2015.7125031.
[28] Wang S, Gu K, Zeng Kai, Wang Z, Lin W. Objective quality assessment and perceptual compression of screen content images. IEEE Comput Graph Appl. 2016; 38(1):47–58. https://doi.org/10.1109/MCG.2016.46.
[29] Sheikh HR, Bovik AC. Image information and visual quality. IEEE Trans Image Process ;15(2):430–44. https://doi.org/10.1109/TIP.2005.859378.
[30] Alotaibi RA, Elrefaei LA. Text-image watermarking based on integer wavelet transform (IWT) and discrete cosine transform (DCT). Appl Comput Inf 2018. https://doi.org/10.1016/j.aci.2018.06.003.

Update

Informatics in Medicine Unlocked

Volume 21, Issue , 2020, Page

DOI: https://doi.org/10.1016/j.imu.2020.100474

Informatics in Medicine Unlocked 21 (2020) 100474

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: http://www.elsevier.com/locate/imu

Erratum regarding missing Declaration of Competing Interest statements in previously published articles

Declaration of Competing Interest statements were not included in published version of the articles that appeared in previous volumes of Informatics in Medicine Unlocked. Hence, the authors of the below articles were contacted after publication to request a Declaration of Interest statement:

1. "Automated scraping of structured data records from health discussion forums using semantic analysis" [Informatics in Medicine Unlocked, 2018; 10C: Pages: 149–158] https://doi.org/10.1016/j.imu.2018.01.003
2. "Molecular dynamics simulation approach to explore atomistic molecular mechanism of peroxidase activity of apoptotic cytochrome c mutants" [Informatics in Medicine Unlocked, 2018; 11C; Pages: 51–60] https://doi.org/10.1016/j.imu.2018.04.003
3. "An efficient and secure remote user mutual authentication scheme using smart cards for Telecare medical information systems" [Informatics in Medicine Unlocked, 2018; 16C: Article Number: 100092] https://doi.org/10.1016/j.imu.2018.02.003
4. "A numerical modelling of an amperometric-enzymatic based uric acid biosensor for GOUT arthritis diseases" [Informatics in Medicine Unlocked, 2019; 12C: Pages: 143–147] https://doi.org/10.1016/j.imu.2018.03.001
5. "Automated heartbeat classification and detection of arrhythmia using optimal orthogonal wavelet filters" [Informatics in Medicine Unlocked, 2019, 16C; Article number: 100221] https://doi.org/10.1016/j.imu.2019.100221
6. "CHROMATOGRAPHIC ANALYSIS OF PHYTOCHEMICALS IN COSTUS IGNEUS AND COMPUTATIONAL STUDIES OF FLAVONOIDS" [Informatics in Medicine Unlocked, 2018; 13C: page range: 34–40] https://doi.org/10.1016/j.imu.2018.10.004
7. "Sperm motility analysis system implemented on a hybrid architecture to produce an intelligent analyzer" [Informatics in Medicine Unlocked, 2020; 19C; Article number: 100324] https://doi.org/10.1016/j.imu.2020.100324
8. "Medical video compression using bandelet based on lifting scheme and SPIHT coding: in search of high visual quality" [Informatics in Medicine Unlocked, 2019; 17C: Article number 100244] https://doi.org/10.1016/j.imu.2019.100244
9. "A histopathological image dataset for grading breast invasive ductal carcinomas" [Informatics in Medicine Unlocked, 2020; 19C: Article number 100341] https://doi.org/10.1016/j.imu.2020.100341
10. "Cancer diagnosis in histopathological image: CNN based approach" [Informatics in Medicine Unlocked, 2019; 16C: Article number 100231] https://doi.org/10.1016/j.imu.2019.100231
11. "The open D1NAMO dataset: A multi-modal dataset for research on non-invasive type 1 diabetes management" [Informatics in Medicine Unlocked, 2018; 13C: Pages: 92–100] https://doi.org/10.1016/j.imu.2018.09.003
12. "The prediction of good physicians for prospJective diagnosis using data mining" [Informatics in Medicine Unlocked, 2018; 12C: Pages 120–127] https://doi.org/10.1016/j.imu.2018.07.005
13. "Prediction of Pathological Complete Response after Neoadjuvant Chemotherapy for Breast Cancer using Ensemble Machine Learning" [Informatics in Medicine Unlocked, 2019; 16C: Article number: 100219] https://doi.org/10.1016/j.imu.2019.100219
14. "Classification of intra-genomic helitrons based on features extracted from different orders of FCGS" [Informatics in Medicine Unlocked, 2019; 18C: Article number: 100271] https://doi.org/10.1016/j.imu.2019.100271
15. "Visual feedback framework for rehabilitation of stroke patients" [Informatics in Medicine Unlocked, 2018; 13C: Pages: 41–50] https://doi.org/10.1016/j.imu.2018.10.002
16. "Performance of source imaging techniques of spatially extended generators of uterine activity" [Informatics in Medicine Unlocked, 2019; 16C Article number: 100167] https://doi.org/10.1016/j.imu.2019.100167
17. "Automated classification of benign and malignant cells from lung cytological images using deep convolutional neural network" [Informatics in Medicine Unlocked, 2019; Volume 16C: Article number: 100205] https://doi.org/10.1016/j.imu.2019.100205

DOIs of original article: https://doi.org/10.1016/j.imu.2019.100205, https://doi.org/10.1016/j.imu.2019.100221, https://doi.org/10.1016/j.imu.2019.100167, https://doi.org/10.1016/j.imu.2018.04.003, https://doi.org/10.1016/j.imu.2020.100341, https://doi.org/10.1016/j.imu.2018.01.003, https://doi.org/10.1016/j.imu.2018.03.001, https://doi.org/10.1016/j.imu.2018.10.002, https://doi.org/10.1016/j.imu.2020.100324, https://doi.org/10.1016/j.imu.2018.09.003, https://doi.org/10.1016/j.imu.2019.100219, https://doi.org/10.1016/j.imu.2018.02.003, https://doi.org/10.1016/j.imu.2019.100244, https://doi.org/10.1016/j.imu.2018.07.005, https://doi.org/10.1016/j.imu.2019.100271, https://doi.org/10.1016/j.imu.2018.10.004, https://doi.org/10.1016/j.imu.2019.100231.

https://doi.org/10.1016/j.imu.2020.100474

Available online 4 November 2020