Compression of CT Images with Modified Inverse Pyramidal Decomposition

ROUMEN KOUNTCHEV
Department of Radio Communications
Technical University of Sofia
Bul. Kl. Ohridsky 8, Sofia 1000
BULGARIA
rkountch@tu-sofia.bg
http://www.tu-sofia.bg

ROUMIANA KOUNTCHEVA
T&K Engineering
Mladost 3
Pob 12
Sofia 1712
BULGARIA
kountcheva_r@yahoo.com

Abstract: In the paper is presented one novel approach for adaptive compression of groups of computer tomography (CT) images, based on modified branched inverse pyramid decomposition. In result is obtained high compression ratio with retained image quality. To achieve this, was used the high correlation between sequences of CT images, representing same object(s). The experimental results confirm the efficiency of the presented method.

Key-Words: Image processing, Archiving of CT images, Compression of sequences of medical images; Image group coding, Modified branched pyramidal image decomposition.

1 Introduction

Electronic healthcare records are important part of medical diagnostics and healthcare. As it is well-known, the old practice was based on paper rectors, which have significant disadvantages, related to legibility, resistance to humidity, accessibility, etc., etc. The contemporary approach is based on electronic rectors, which offer significant advantages, especially when visual medical data is concerned. This comprises various X-ray or ultrasound images, electrocardiograms, and many more. Computed tomography (CT) is one of the main contemporary diagnostic procedures. It uses special x-ray equipment to obtain cross-sectional pictures of the body, and for this, usually sequences of images for every patient are created and used. Important stage in CT is archiving the images obtained in an efficient manner concerning the data volume occupied and the image quality. A vast number of medical image compression techniques already exist [1-3,7], which can be divided into two large groups – lossless [4,9] and lossy [5,6] depending on the ability to restore the image fully or not. In both groups some type of image decomposition is used (for example, linear orthogonal transform or a wavelet one combined with spectral coefficients rearrangement and entropy coding).

The most famous file format, used for medical images archiving is DICOM [8], which is based on the JPEG standard. In [14] is proposed one approach based on adaptive sampling of DCT coefficients. The quality of the restored images after this compression is comparable to that of the JPEG2000 as the author shows while the JPEG coder produces images with PSNR between 31 and 40 dB for same levels (i.e. visually lossless). In [5] authors confirm that wavelet decomposition assures better quality for the compressed images being than the JPEG-based coder.

Some resent publications emphasize the advantages of the wavelet decomposition for medical image compression combined with other techniques in order to construct more efficient coders (for example, using joint statistical characterization [1], by linear prediction of the spectral coefficients [13], introducing region of interest (ROI) [6], incorporating planar coding [12], etc). Nevertheless the higher compression levels achieved, there is also reported significant reduction of the visual quality of these images [9,10]: while cumulative quality measures such as PSNR stay high the smoothing of vast image areas due to the wavelet coefficients quantization becomes intolerable for higher compression ratios.

In this paper is proposed one new approach for lossy compression of sequences of CT images with
2 Basics of the Branched Inverse Pyramid Decomposition (BIDP)

The Inverse Pyramid decomposition is aimed at the efficient compression of still digital images. The digital image is represented by a matrix of size \((2^m) \times (2^m)\). For the processing, the matrix is first divided into blocks of size \(2^n \times 2^n\) and on each is applied Inverse Pyramid Decomposition (IPD) \([11]\).

It is performed as follows: on the matrix \([B]\) of each block is applied pre-selected “truncated” orthogonal transform (TOT) and are calculated the values of relatively small number of “retained” coefficients, located in the high-energy area of the so obtained transformed (spectrum) matrix \([S_0]\), for example, these are usually the coefficients with spatial frequencies \(0,0\), \(0,1\), \((1,0)\) and \((1,1)\). After inverse orthogonal transform (IOT) of the "truncated" spectrum matrix \([\hat{S}_0]\), which contains the retained coefficients only, is obtained the matrix \([\hat{B}_0]\) for the initial (zero) IPD level \((p=0)\), which approximates the matrix \([B]\). The accuracy of the approximation depends on: the positions of the retained coefficients in the matrix \([S_0]\); the values, used to substitute the missing coefficients from the approximating matrix \([\hat{S}_0]\) for the zero level, and the selected orthogonal transform (DCT, WHT, CHT, KLT, etc.).

In the next (first) IPD level \((p=1)\) is calculated the difference matrix \([E_0] = [B] - [\hat{B}_0]\). The result is then split into 4 similar sub-matrices of size \(2^{n-1} \times 2^{n-1}\) and on each is applied the corresponding TOT. The total number of retained coefficients for the level \(p=1\) is 4 times larger than that in the zero level. In case, that the for this level is used the Walsh-Hadamard transform, the values of coefficients \((0,0)\) in the IPD decomposition levels 1 and higher are always equal to zero, which permits to reduce the number of retained coefficients with \(\frac{1}{4}\). On each of the 4 spectrum matrices \([\hat{S}_1]\) for the IPD level \(p=1\) is applied IOT and in result are obtained 4 corresponding sub-matrices, which build the approximating difference matrix \([E_0]\). In the next IPD level \((p=2)\) is calculated the difference matrix \([E_1] = [E_0] - [\hat{E}_0]\). After that each difference sub-matrix is divided in similar way as in level 1, into 4 matrices of size \(2^{n-2} \times 2^{n-2}\), for each is performed TOT, etc. The maximum possible number of decomposition levels for one image is \(n\) (for \(p=n-1\)). In this case the total number of "retained" coefficients for all levels \((4m^2 + 3.4m^2 + ... + 3.4^{n-1}m^2 = 4^n m^2)\) is equal to the number of pixels in the image, and hence, the IPD is not "overcomplete". In the last (highest) IPD level is obtained the “residual” difference matrix. In case that the image should be losslessly coded, each block of the residual matrix is processed with full orthogonal transform and no coefficients are omitted.

The Branched IPD is enhanced version of IDP. It is represented by the Block diagram shown on Fig. 1. The IPD for each block, called “Main Pyramid” is of 3 levels \((n = 3, \text{for } p = 0, 1, 2)\). The values of coefficients, calculated for these 3 levels, compose the inverse pyramid, whose sections are of different color each. The coefficients \((0,0)\), \((0,1)\), \((1,0)\) and \((1,1)\) in the level \(p=0\) from all blocks compose corresponding matrices of size \(m \times m\) each, colored in yellow. These 4 matrices build the “Branch for level 0" of the Main Pyramid. Each is then divided into blocks of size \(2^{n-1} \times 2^{n-1}\), on which in similar way are build corresponding 3-level IPDs \((p=0,01,02)\). The retained coefficients \((0,1), (1,0)\) and \((1,1)\) in the level \(p=1\) of the Main Pyramid from all blocks build matrices of size \(2m \times 2m\) (colored in pink). Each such matrix is divided into blocks of size \(2^{n-1} \times 2^{n-1}\), on which in similar way are build corresponding 3-level IPDs \((p=10,11,12)\). The retained coefficients, calculated after TOT from the blocks of the Residual Difference in the last level \((p=2)\) of the Main Pyramid, build matrices of size \(4m \times 4m\); from the first level \((p=00)\) of the "Branch Pyramid 0" - matrices of size \((m/2)^2 \times m/2^3)\); and from the first level \((p=10)\) of the "Branch Pyramid 1" - matrices of size \((m/2)^2 \times m/2^3)\). In order to reduce the correlation between elements from the so obtained matrices, on each group of 4 spatially neighboring elements is applied the following transform: the first one is substituted by their average value, and each of the remaining 3 – by its difference to next elements, scanned counter clockwise. The coefficients, obtained this way from all levels of the Main and Branch Pyramids are arranged in one-dimensional sequences in accordance to Peano-Hilbert scan and
after that are quantized and entropy coded using Adaptive RLC and Huffman. The values of the spectrum coefficients are quantized only in case that the image coding is lossy. In order to retain the visual quality of the restored images, the quantization values are related to the sensibility of the human vision to errors in different spatial frequencies. For the reduction of these errors, together with retained compression efficiency, in the consecutive BIDP levels could be used various fast orthogonal transforms: for example, in the zero level could be used DCT, and in the next levels - WHT.

The decoding of the compressed image data is done, performing the already described operations in inverse order.

The basic quality of the BIPD is that it permits to achieve significant decorrelation of the processed image data. In result, the BIPD permits the following:

- To achieve highly efficient compression with retained visual quality of the restored image (i.e. visually lossless coding), or efficient lossless coding, depending on the application requirements;
- Layered coding and transfer of the image data, in result of which is obtained low transfer bit-rate with gradually increased quality of the decoded image;
- Lower computational complexity than that of the wavelet decompositions;
- Easy adaptation of the coder parameters, so that to ensure the needed concordance of the so obtained data stream, to the ability of the communication channel;
- Resistance to noises in the communication channel, or due to compression/decompression. The reason for this is the use of TOT in the decoding of each image block;
- Retaining the quality of the decoded image unchanged after multiple codings/decodings;

BIDP could be further developed and modified in accordance to requirements of various possible applications. One of these applications for processing of groups of similar images is given below.

3 Representation of CT image sequences through Modified BIDP

For the Modified BIDP coding is used the high similarity between group of CT images of same object. Based on this, images in one group are coded together (Group coding). In order to make the information redundancy in the sequence of matrices [Bn] for n=0, ±1, ±2, … , ±N smaller, here is offered to use the following modification of the image decomposition (the relations below are for decomposition of 2 levels only). First, one of the images in the group is selected to be used as a reference. For this is evaluated the mutual correlation between any 2 images in the group and the image with highest correlation is chosen to be the reference. The decomposition starts with the calculation of the lower decomposition level for the reference image [B0], which corresponds to its coarse approximation. The following steps are performed:

- For the IDP level p = 0 of the reference image in the Group coding is calculated the transform [S0] by applying the direct orthogonal transform:

\[ [S_0^n] = [T_0][B_0][T_0]^\dagger, \]  

where [T0] is the matrix of the selected 2D direct orthogonal transform.

- The matrix of the approximated transform of the reference image is calculated:

\[ [\hat{S}_0^n] = [m_0(u,v)]S_0^n(u,v), \]  

where m0(u,v) is the element of the matrix-mask [M0] which defines the set of retained transform coefficients:

\[ m_0(u,v) = \begin{cases} 1, & \text{if } S_0^n(u,v) - \text{retained coefficient } t, \\ 0, & \text{in all other cases}, \end{cases} \]  

- The first approximation [ˆB0] of the reference image is obtained after inverse orthogonal transform:

\[ [\hat{B}_0] = [T_0][\hat{S}_0^n][T_0]^\dagger, \]  

where [T0] = [T0]−1 is the matrix of the inverse orthogonal transform of size 2^m×2^m.

- The difference matrix is calculated:

\[ [E_0^n] = [B_0] - [\hat{B}_0]. \]  

- The difference matrix is then split into 4 sub-matrices:

\[ [E_0^n] = \begin{bmatrix} [E_{0,1}^1] & [E_{0,1}^2] \\ [E_{0,2}^1] & [E_{0,2}^2] \end{bmatrix}, \]  

where [E_{0,i}^j] for i = 1, 2, 3, 4 are sub-matrices of size 2^{m-1}×2^{m-1}.

- For the next decomposition level p = 1 of the reference image is calculated the transform [S0] of the i-th sub-matrix of the difference [E0], using direct orthogonal transform:

\[ [S_i] = [T_i][E_0][T_i]^\dagger \]  

where [T_i] is the matrix of the direct orthogonal transform, of size 2^{m-1}×2^{m-1}. 

ISBN: 978-1-61804-115-9  76
The approximated $i$th transform is then calculated:

$$[\hat{S}_i] = [m_i(u,v)s_i^n(u,v)],$$  \hspace{1cm} (8)

where $m_i(u,v)$ is the element of the matrix-mask $[M_i]$ which defines the retained transform coefficients for the next level (similar to Eq. 3):

For the decomposition level $p=1$ of the image $[B_n]$ of the processed group of images is calculated the difference:

$$[E_n] = [B_n] - [\hat{B}_n] \text{ for } n=0, \pm 1, \pm 2, \pm 3, \pm 4. \hspace{1cm} (9)$$

The difference matrix is split into 4 sub-matrices:

$$[E_n] = \begin{bmatrix} [E_1^n] & [E_2^n] & [E_3^n] & [E_4^n] \end{bmatrix},$$ \hspace{1cm} (10)

$[E_i^n]$ for $i=1,2,3,4$ are sub-matrices of size $2^{m-1} \times 2^{m-1}$.

The $i$th transform $[S_i^n]$ of the sub-matrix of the difference $[E_i^n]$ is obtained using direct orthogonal transform:

$$[S_i^n] = [T_i][E_i^n][T_i^T] \text{ for } i=1,2,3,4. \hspace{1cm} (11)$$

The approximated $i$th transform is:

$$[\hat{S}_i^n] = [m_i(u,v)s_i^n(u,v)],$$ \hspace{1cm} (12)

where $m_i(u,v)$ is the element of the matrix-mask, which defines the retained coefficients (similar to Eq. 3).

The difference matrices of the approximated transforms are calculated:

$$[\Delta \hat{S}_n] = [\hat{S}_n] - [\hat{S}_{n-1}] \text{ for } n=\pm 1, \pm 2, \ldots, \pm N. \hspace{1cm} (13)$$

The values of coefficients of matrices $[\hat{S}_i^n]$ and $[\Delta \hat{S}_n]$ for $i=1,2,3,4$ and $n=0, \pm 1, \pm 2, \ldots, \pm N$ in decomposition levels $p=0,1$ are losslessly coded.

Eqs. (9) and (13) represent the main difference between the basic IDP decomposition, and the modification, used for the Group image coding. In the basic IDP decomposition each image has its own approximation for the consecutive decomposition levels. Besides, in the modified approach, presented here, all images use the same coarse approximation, and in the second decomposition level each image is processed individually. Another significant difference is that the basic IDP decomposition usually comprises 3 or 4 levels, starting with large image sub-blocks. The modification, used for the group representation, is based on 2-level decomposition, built for relatively small sub-blocks, usually of size 8 x 8 pixels for the lower level and 4 x 4 pixels – for the higher one.

The block diagram of the Modified Branched 2-level IDP, used for coding of one sub-group of CT images is shown on Fig. 2. The decoding is performed in reverse order.

In a group of medical images the first in the sequence is usually used as a reference. In case that the sequence contains large number of images, best results are obtained if the sequence is divided into groups of 5 or 6 consecutive images (the mutual correlation becomes lower for larger number of consecutive images in the group).

Fig. 2. Block diagram of the Modified BIPD coder

Here is used two-dimensional (2D) TOT, whose retained sets of coefficients are defined by the “ones” in the binary matrices-masks $[M_0]$ and $[M_1]$ used respectively for decomposition levels $p = 0,1$. For each sub-group of CT images the lossless compression in the last stage of the processing comprises run-length coding (RLC), Huffman coding (HC) and arithmetic coding (AC).

The block diagram represents the processing of one sub-block of the processed image only. The group coding for color images is performed in a similar way, but it requires each color component to be processed individually. Depending on the color format (RGB, YUV, YCbCr, KLT, etc.), and the color sampling format (4:4:4, 4:2:0, 4:1:1, etc.) for each component is built an individual pyramid. The approach based on the processing of the reference image and the remaining ones in the group, is retained.

4 Experimental results

For the experiments were used images from the image database of the Medical University and Technical University in Sofia. The CT test images
used for the experiments given below, were a group of 576 grayscale slices in DICOM format, of size 512x512 pixels each, with intensity depth of 16 bpp.

On Fig. 3.a is presented graphically the correlation coefficient between each two images from the group of 560 images, and on Fig. 3.b – the correlation coefficient between the first image and all the others. As suggested before, a strong variation of the correlation exists inside a candidate group around a proper reference and outside it asymptotically goes to a constant value.

One test group consisting of 9 images is shown on Fig.4: the first image was selected to be used as a reference. The size of the initial sub-block was 16x16 (n = 4). In the zero and first level of the branch, the retained coefficients were 4, placed in the low-frequency area. For the main branch of the inverse pyramid in the first and second level, the same 4 low-frequency coefficients were retained.

In Table 1 are given the results obtained for 5 test images with MBIDP coding (GC) and comparison with the JPEG standard. For the experiments was used the software implementation of the method in Visual C. It is easy to notice, that for approximately same quality the compression ratio (CR) is higher for the MBIDP (Group coding). Besides, for similar quality obtained through JPEG 2000, the visual quality of MBIDP is better, because the JPEG 2000 decreases the small details (in result of the use of wavelet transform).

Table 1. Results obtained for 5 test images.

<table>
<thead>
<tr>
<th>Image</th>
<th>CR&lt;sub&gt;GC&lt;/sub&gt;</th>
<th>PSNR&lt;sub&gt;GC&lt;/sub&gt;</th>
<th>CR&lt;sub&gt;JPEG&lt;/sub&gt;</th>
<th>PSNR&lt;sub&gt;JPEG&lt;/sub&gt; [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>68.30</td>
<td>31.84</td>
<td>47</td>
<td>31.88</td>
</tr>
<tr>
<td>b</td>
<td>51.89</td>
<td>31.88</td>
<td>46</td>
<td>31.76</td>
</tr>
<tr>
<td>c</td>
<td>47.28</td>
<td>31.76</td>
<td>46</td>
<td>31.78</td>
</tr>
<tr>
<td>d</td>
<td>46.28</td>
<td>31.51</td>
<td>46</td>
<td>31.91</td>
</tr>
<tr>
<td>e</td>
<td>43.10</td>
<td>31.37</td>
<td>43</td>
<td>32.44</td>
</tr>
<tr>
<td>Mean</td>
<td>51.37</td>
<td>31.67</td>
<td>45.6</td>
<td>31.95</td>
</tr>
</tbody>
</table>

5 Conclusions

The main advantages of the presented new method for compression of sequences of medical CT images are:

- The higher visual quality obtained for same compression ratios;
- The comparatively low computational complexity (the computational complexity of the new method is comparable to that of the IDP method which was presented and evaluated in detail in [15]).

The general characteristics of the Modified BIDP are a reliable basis for its successful use in various application areas. As it was proved by the experiments, this compression is very efficient for archiving of sequences of CT images. It could also be used for compression of still multispectral, hyperspectral or multi-view images and video sequences, obtained from surveillance video cameras, supersound scanners, thermo-vision systems, scanning microscopes, etc.

The investigation will be further developed through modeling and experiments. For this will be used large image databases with CT images of various kinds. The so obtained results will be evaluated and compared to other similar algorithms.
and will be investigated possible new application areas: remote investigation of the earth surface, medical diagnostic, automatic manufacturing control, defense, etc.

Acknowledgement
This paper was supported by the Joint Research Project: Electronic Health Records for the Next Generation Medical Decision Support in Romanian and Bulgarian National Healthcare Systems, DNTS 02/09.

References: