7

Novel Prediction Based Technique for Efficient Compression of Medical Imaging Data

Josip Knezović, Mario Kovač, Martin Žagar,
Hrvoje Mlinarić and Daniel Hofman
University of Zagreb
Croatia

1. Introduction

The annual volume of imaging data in modern paperless hospitals can approach up to 10 terabytes, heavily pressing the storage and transmission requirements (Choong et al., 2007). Utilizing efficient compression techniques for those data in order to reduce associated costs is very attractive from both viewpoints: financial and organizational (Sanchez, Abugharbieh & Nasiopoulos, 2009; Sanchez et al., 2008). Although lossy techniques can yield better compression results, due to possible compression artifacts in the compressed image, they are less favored compared to lossless compression techniques in certain medical applications such as image-based diagnosis, archival etc. Compression itself helps in alleviating storage requirements for medical imaging system. Additionally, it also helps in accommodating the on-line transmission and availability of patient diagnostic imaging data which is essential for future electronic health frameworks.

Moreover, new approaches in medical imaging such as 3D and 4D imaging and bio–modeling produce even greater amounts of image data. For efficient storage and transmission of those data and utilization of systems that exploit 3D and 4D imaging technologies, compression is inevitable. In this field, at least certain parts of images are required to be stored and transmitted without any loss of information. The lossless compression algorithm that we propose can also be efficiently employed for at least those vital parts of interest in this kind of applications (Zagar et al., 2007).

Important property of image data is high degree of correlation among neighboring pixels which is crucial for any compression technique since it makes it possible to decorrelate the samples using some sort of prediction–based modeling. If employed modeling technique effectively models the spatial correlation among neighboring pixels, remaining data will be mostly decorrelated and easily coded with an entropy coder. On the other hand, it is well known that image data are nonstationary, i.e. properties of image regions vary all over the image (Memon & Wu, 1999). Accordingly, it is necessary to adapt the model to the changing image characteristics. Another assumption of local stationariness is very well applicable to the image data. This means that for arbitrarily small image regions, the model adapted to the dominant local property will be effective inside the region. Predictive image coding in which the prediction error of the current pixel is coded has shown to be the most effective technique in lossless image compression. Using prediction, image data are decorrelated prior

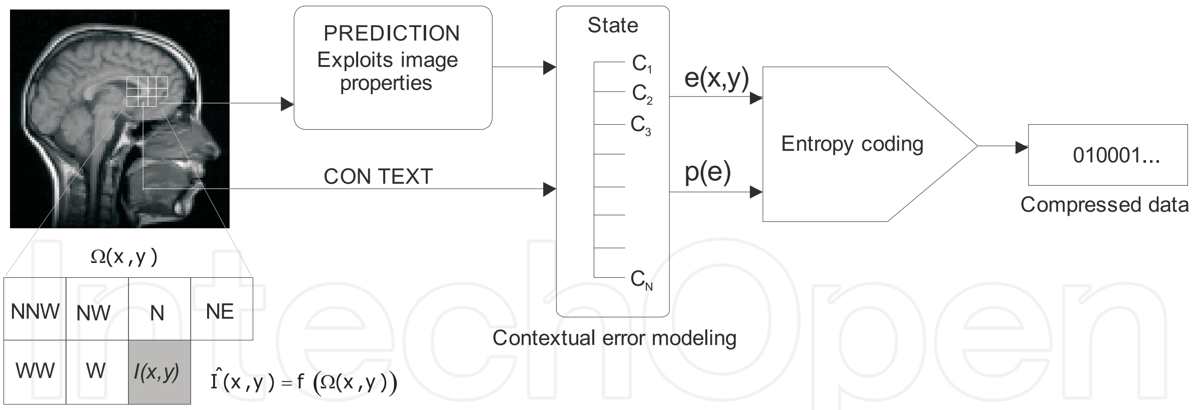


Fig. 1. Block scheme of predictive image coding

to the entropy coding so that better compression is achieved. In the framework of sequential, backward adaptive lossless image compression, predictive image coding can be formulated as composed of the following steps as shown in Figure 1:

1. Prediction of the current pixel based on the casual set of surrounding pixels - pixel prediction.
2. Contextual error modeling: Determination of the conditional probabilistic context in which the current prediction error occurs (Memon & Wu, 1999).
3. Entropy coding of the prediction error in the detected probabilistic context - entropy coding.

In this work a new predictive model for lossless image compression is proposed. The model is based on classification and blending of static predictors which is followed by heuristic contextual error modeling. The classification is performed in order to capture and model higher order redundancies inside the local image region. Then, on the causal set of classified neighboring pixels, the set of selected static predictors is dynamically blended to produce the prediction. The idea behind the blending of predictors is to find a dominant property inside the current image region while taking other properties into account. The dominant property will have the biggest impact on the final, blended predictor. Based on our proposed predictor, we developed two lossless image compression codecs characterized with high compression efficiency. First codec neglects computational complexity while the second proposal employs an effective technique which reduces the computational requirements of the first proposal while maintaining compression efficiency.

2. Proposed Predictive Coding Method

We treat image as a two-dimensional array $I(x,y)$ of pixel grey intensity values of width W and height H, where $0 \leq x < W$ and $0 \leq y < H$. Pixels are observed sample by sample in raster scan order, from top to bottom, left to right. In the assumed backward adaptive approach, the encoder is allowed to use only past information that is also available to decoder. This means that for forming the prediction only previously observed pixels are used, as shown in Figure 1. In fact, only a small subset of previously encoded pixels is used to form the causal template. Predictor from Figure 1 uses causal context of surrounding pixels for the prediction of the current pixel :

$$\hat{I}(x,y) = f(\Omega(x,y)) \tag{1}$$

In simple prediction schemes that use small causal context the compass point notation for surrounding pixels is used which is also illustrated in Figure 1. For example N denotes North pixel, W denotes the West pixel from the current pixel, NE North–East, NW North–West etc. After the prediction is performed, the model outputs the prediction error. This means that instead of coding the real pixel value, the pixel prediction $\hat{I}(x,y) = f(\Omega(x,y))$ is performed and the prediction error $e(x,y) = I(x,y) - \hat{I}(x,y)$ is further encoded by the statistical encoder. As previously observed, a typical image can be treated as composed of regions with varying dominant properties such as edges, textures, smooth regions, noisy regions etc. Those properties pose different and conflicting constraints on the prediction function if it is required to adapt to the region properties. The main properties of image regions and their requirements on the predictor are given in (Seemann & Tischer, 1997). If we consider a linear predictor:

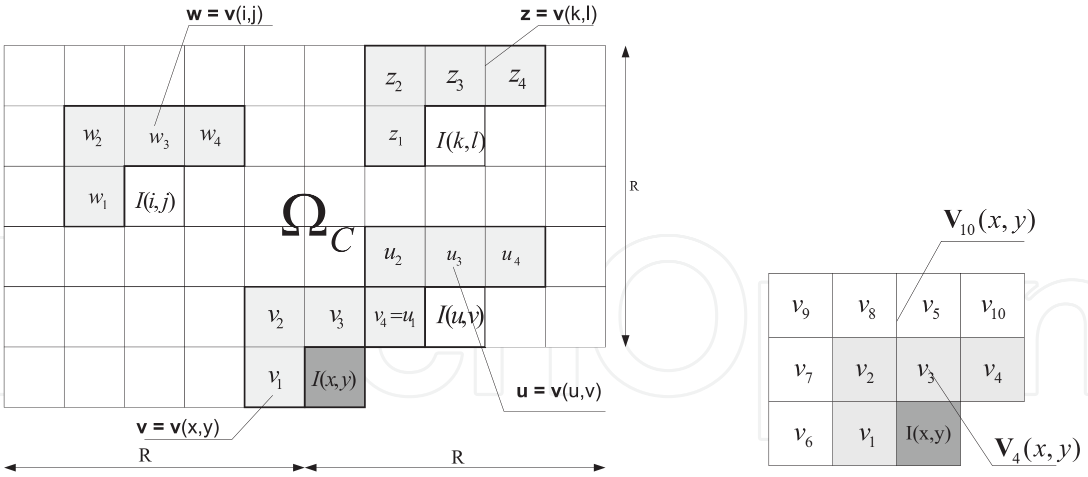
$$\hat{I}(x,y) = f\big(\Omega(x,y)\big) = \sum_{I(i,j)\in\Omega} a_{i,j} \cdot I(i,j), \tag{2}$$

then we can formulate the following constraints on the predictor coefficients $\{a_{i,j}\}$ depending on the dominant property of the current region:

- **Smooth regions** in which the intensity of pixel doesn't change require that $\sum a_{i,j} = 1$. For planar region it is required for at least one of $a_{i,j}$ to be negative so that the gradient can be estimated.
- **Noisy regions** require the minimum magnitude of noise is introduced into prediction which implies that $\sum |a_{i,j}|$ should be as small as possible. Therefore, the best approach for noisy regions is the use of averaging prediction function.
- **Edges and textured regions** constitute the most important visual part of images (Seemann & Tischer, 1997). Edges require some kind of adaptation mechanism in the predictor to provide the detection and orientation of the edge. Textures are the most difficult to model and they can be considered as combination of noise and edges.

In general, image regions can be viewed as composed of structures mentioned before with one or two dominant properties; therefore the choice for predictor to satisfy given constraints is very often in conflict. The best choice for the most cases is to assume the noisy property and corresponding constraint as suggested in (Seemann & Tischer, 1997) where the blending of static predictors is proposed which practically gives the final predictor to be an averaging predictor.

The effectiveness of any prediction scheme depends on its ability to adapt to different image regions. This precludes the use of static predictors if efficient prediction is required. Typical, heuristically tuned switching predictors use a set of static prediction functions and heuristics to determine which function will be used for the prediction of the current pixel. Such predictors include GAP from the CALIC algorithm (X. Wu and N. Memon and K. Sayood, 1995), MED predictor from the LOCO-I and JPEG–LS standard (M.J. Weinberger and G. Seroussi and G. Sapiro, 1998), etc. The main drawback of switching predictors is the lack of robustness in the presence of nontrivial image structures. Another approach is to use adaptive predictor which we also propose. There is a large spectrum of adaptive predictors with various mechanisms of adaptation and complexities such as LS–based predictors (G. Motta and J.A. Storer and B. Carpentieri, 2000; X. Li and M.T. Orchard, 2001) and blending predictors (Seemann & Tischer, 1997;



(a) Causal context

(b) Vector template

Fig. 2. Contextual elements of proposed predictor

Seemann et al., 1997). The aim of the prediction scheme proposed in this paper is to predict well in all various image regions with moderate computational complexity.

2.1 Proposed predictor

In order to efficiently model different image structures we propose adaptive predictor based on the idea of predictor blends (Seemann & Tischer, 1997). The blending predictor is extended with dynamic determination of blending context on a pixel–by–pixel basis. The set of predictors

$$\mathcal{F} = \{f_1, f_2, \cdots, f_N\}$$

is composed of N static prediction functions adjusted to predict well in the presence of specific property. For example simple predictor $f_W = I(x-1, y)$ is known to predict well in the presence of sharp horizontal edge. The classification process determines the set of neighboring pixels on which the blending of $\mathcal{F}$ is performed. It is similar to initial step of vector quantization design substantially simplified in order to be usable in symmetric, backward adaptive algorithm (M.J. Slyz and D.L. Neuhoff, 1994). The process of blending of selected predictors will finally produce an averaging predictor which is known to predict well in the noisy regions. This way, the final predictor will try to adapt to the most relevant image properties while detecting and adjusting to the dominant property in the current local region, whether it is an edge, a planar region or even a texture with nontrivial coarseness and period. Figure 2 depicts basic elements of proposed prediction scheme. Ω_C denotes the causal context used by the predictor. It is a rectangular window of radius R composed of previously encoded pixels on which the search procedure for classification is performed. Each pixel from the Ω_C and currently unknown pixel $I(x, y)$ is associated with its vector template $\mathbf{v}(x, y)$ that is composed of d closest causal neighboring pixels as shown in Figure 2(b). As an example, the vectors of size $d = 4$ ($\mathbf{v}_4(x, y)$) and $d = 10$ ($\mathbf{v}_{10}(x, y)$) are shown. In Figure 2(a) the vectors of size four are used. The Euclidean distance between associated vectors will be used for classifications of pixels into the *current* cell of pixels *similar* to the current pixel $I(x, y)$, just like in VQ design (A. Gersho, 1993; A. Gersho and V.M. Gray, n.d.). In order to reduce the complexity of proposed scheme some basic simplifications are introduced. First, as shown by

Slyz and Neuhoff, we do not perform full vector quantization. Rather, only the *current* cell in which the current pixel lies needs to be calculated (M.J. Slyz and D.L. Neuhoff, 1994). Next, the cell population, i.e. number of pixels that go into the cell together with the current pixel's vector, is set as constant M at the beginning of the coding process.
Proposed prediction scheme operates as follows:

Classification and Blending Predictor CBP

1. *Iteration:* Iterate for every pixel $I(x, y)$ in the image.
2. *Classification:* For each pixel $I(i, j) \in \Omega_C$ compute the Euclidean distance $D(i, j)$ between its corresponding vector $\mathbf{v}(i, j)$=$\mathbf{w}$ and the current pixel's vector $\mathbf{v}(x, y)$=$\mathbf{v}$:

$$\begin{aligned} D(i,j) &= ||\mathbf{v}(i,j) - \mathbf{v}(x,y)|| \\ &= ||\mathbf{w} - \mathbf{v}|| \\ &= \textstyle\sum_{k=1}^{d} |w_k - v_k|^2 . \end{aligned} \tag{3}$$

   Based on the computed distances, determine M pixels from Ω_C that belong to the current cell, i.e. with the smallest vector distances from the current pixel's vector. The current cell will be used as blending context Ω_B for $\mathcal{F}$. This step is similar to nearest neighbor selection in VQ design.
3. *Predictor penalties:* For every predictor f_k the penalty G_k is calculated by the following equation:

$$G_k = \sum_{I(i,j) \in \Omega_B} (\hat{I}_k(i,j) - I(i,j))^2, \tag{4}$$

   where $\hat{I}_k = f_k(i, j)$ is the prediction of $f_k \in \mathcal{F}$ for the pixel $I(i, j) \in \Omega_B$.
4. *Final predictor blending:* Based on the penalties we form the prediction for the current pixel $\hat{I}(x, y)$ as:

$$\hat{I}(x,y) = F(x,y) = \left[\frac{\left(\sum_{k=1}^{N} \frac{1}{G_k} \cdot \hat{I}_k(x,y) \right)}{\sum_{k=1}^{N} 1/G_k} \right]. \tag{5}$$

   The prediction for the current pixel is the weighted sum of predictions of all the predictors from $\mathcal{F}$ with weights inversely proportional to corresponding penalties. The penalty of predictor reflects its prediction accuracy on the blending context. If the predictor predicts well, its contribution, i.e. its weight in the final prediction will be higher. The predictors that do not predict well on the current blending context will eventually be blended out by associated large penalties. The denominator in (5) normalizes the final prediction so that the sum of weights equals to 1.
5. *Error correction:* On the blending context Ω_B calculate *typical* error of the final predictor as:

$$\bar{e}(\Omega_B) = \frac{1}{M} \sum_{I(i,j) \in \Omega_B} (F(i,j) - I(i,j)). \tag{6}$$

   Based on the typical error of blending predictor F the final prediction for the current pixel is further refined as:

$$\dot{I}(x,y) = F(x,y) + \bar{e} = \hat{I}(x,y) + \bar{e}. \tag{7}$$

   This final step of proposed predictor captures typical bias of the blending predictor F on the classified set of pixels Ω_B that are the part of similar structure as current pixel.

Through the classification and blending process, proposed predictor adjusts itself to the dominant local property. The blending allows other non–dominant properties to be included in the final predictor, although with less contribution. This is crucial difference compared with switching predictors that don't have the capability to model nontrivial image structures with mixture of properties. Note that pixels from the search window that do not belong to the region with the same dominant property as the region in which current pixel resides will not be included in the current cell and thus they will not be part of the blending context.

2.2 Contextual error modeling

Although the prediction step removes statistical redundancies within image data, there are remaining structures in the error image which cannot be completely removed using only previously applied prediction step (Memon & Wu, 1999). Those structures are removed using contextual modeling of prediction error, where the context or the state is the function of previously observed pixels, errors or any other relevant variables. As reported by Wu, the heuristic method that uses both, previous pixels template and causal error energy estimate is best suited for this purpose (Wu, 1997). Wu's contextual model is composed of two different submodels: (1) Model with large number of states that is used for prediction error feedback; and (2) Model with low number of states used for error probability estimation. On the other hand, Wu's predictor is a heuristic predictor with low degree of adaptation and our proposal is highly adaptive predictor with already built in error feedback mechanism (error correction step in the prediction). This implies that our mechanism needs smaller and less complex contextual model for estimation of symbol probabilities. Therefore it is built as follows: Besides of the high correlation with texture pattern, current prediction error is also highly correlated with the errors on neighboring pixels. This is modeled with the error discriminant

$$\Delta = d_h + d_v + 2|e_w|, \tag{8}$$

where

$$d_h = |W - WW| + |N - NN| + |N - NE|, \tag{9}$$

and

$$d_v = |W - NW| + |N - NN| + |NE - NNE|, \tag{10}$$

are horizontal and vertical gradients around the current pixel, and e_w is the prediction error on the west pixel W from the current pixel. Δ is uniformly quantized into eight levels to produce the state of the model (Wu, 1997). Every state contains the histogram table which is used for probability estimation of the prediction error in the current state. Because of the context dilution effect, this contextual model is required to have small number of states.

2.3 Entropy coding of prediction error

The final step of proposed image compression algorithm is entropy coding of the resulting prediction error. For the given error symbol and given probability estimate obtained from the contextual state, the codeword is computed by the entropy coder. This codeword is output as the final result of predictive image coding algorithm. Our proposal uses highly efficient implementation of adaptive arithmetic coding (P.G. Howard and J.S. Vitter, 1994).

3. Selective computation of Predictor Blends

In this section we extend our Classification and Blending Predictor CBP with the goal to reduce the computational requirements of the method while maintaining its compression efficiency. We achieve this by implementing an effective heuristic approach in the predictive part of the compression method.

Our optimization is based on the observation that the changes in the predictor, i.e. the changing of predictor weights, i.e. coefficients, mainly happens on the boundaries of image regions. For example, once computed prediction parameters in the CBP predictor remain mostly unchanged while the coder is in the local region with dominant property such as smooth or planar region. Noticeable change in the CBP predictor parameters happens when the coder reaches out of the one local region and enters into another with a new dominant property. Shortly, this change happens on the edge areas. Based on this fact, we propose a novel way of reducing the number of required computations in the predictor by preserving the predictor coefficients as long as we are inside a local region. This idea is implemented in the CBP predictor in a way that when the magnitude of the prediction error on the current pixel is beyond a predefined threshold, we trigger the compute-intensive calculation of the predictor weights in the CBP. As long as the prediction error magnitude is below the threshold, previously calculated predictor weights are used for the prediction of the current pixel. This way, we evict the pixel-by-pixel computation of predictor weights, and perform it only on pixels on which the current predictor fails beyond a predefined threshold. This technique is motivated by the Edge Directed Property of LS–based predictor reported in (X. Li and M.T. Orchard, 2001).

Modified prediction scheme operates as follows:

Selective Classification and Blending Predictor SCBP

1. *Initialization:* Set the previous prediction error to zero.
2. *Iteration:* Iterate for every pixel $I(x,y)$ in the image.
3. *Selective computation:* Check the previous prediction error. If it is less than the predefined prediction error threshold T_e, skip the computation and go to step 6; otherwise go to step 4.
4. *Classification:* For each pixel $I(i,j) \in \Omega_C$ compute the Euclidean distance $D(i,j)$ between its corresponding vector $\mathbf{v}(i,j)$=$\mathbf{w}$ and the current pixel's vector $\mathbf{v}(x,y)$=$\mathbf{v}$:

$$\begin{aligned} D(i,j) &= ||\mathbf{v}(i,j) - \mathbf{v}(x,y)|| \\ &= ||\mathbf{w} - \mathbf{v}|| \\ &= \textstyle\sum_{k=1}^{d} |w_k - v_k|^2 . \end{aligned} \tag{11}$$

   Based on the computed distances, determine M pixels from Ω_C that belong to the current cell, i.e. with the smallest vector distances from the current pixel's vector. The current cell will be used as blending context Ω_B for $\mathcal{F}$. This step is similar to nearest neighbor selection in VQ design.

5. *Predictor penalties:* For every predictor f_k the penalty G_k is calculated by the following equation:

$$G_k = \sum_{I(i,j)\in\Omega_B} (\hat{I}_k(i,j) - I(i,j))^2, \tag{12}$$

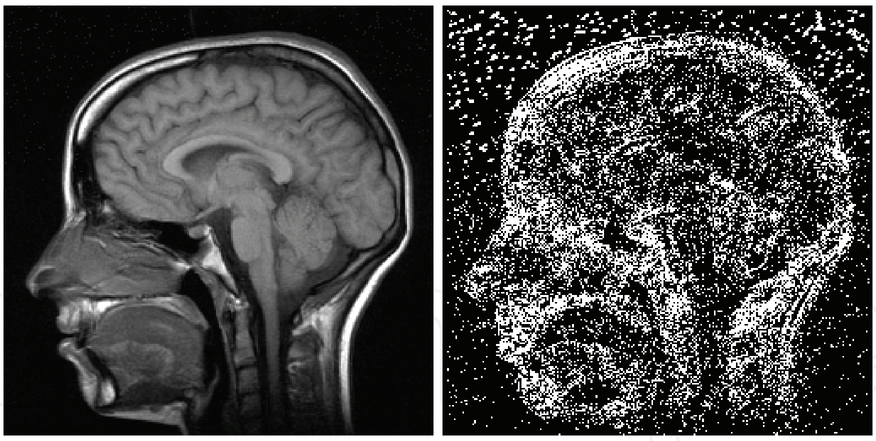


Fig. 3. Effects of selective computations of predictor weights

where $\hat{I}_k = f_k(i,j)$ is the prediction of $f_k \in \mathcal{F}$ for the pixel $I(i,j) \in \Omega_B$.

6. *Final predictor blending:* Based on the blend penalties we form the prediction for the current pixel $\hat{I}(x,y)$ as:

$$\hat{I}(x,y) = F(x,y) = \left[\frac{\left(\sum_{k=1}^{N} \frac{1}{G_k} \cdot \hat{I}_k(x,y) \right)}{\sum_{k=1}^{N} 1/G_k} \right]. \tag{13}$$

The prediction for the current pixel is the weighted sum of predictions of all the predictors from $\mathcal{F}$ with weights inversely proportional to corresponding penalties.

7. *Error correction:* On the blending context Ω_B calculate *typical* error of the final predictor as:

$$\bar{e}(\Omega_B) = \frac{1}{M} \sum_{I(i,j) \in \Omega_B} (F(i,j) - I(i,j)). \tag{14}$$

Based on the typical error of blending predictor F the final prediction for the current pixel is further refined as:

$$\dot{I}(x,y) = F(x,y) + \bar{e} = \hat{I}(x,y) + \bar{e}. \tag{15}$$

This final step of proposed predictor captures typical bias of the blending predictor F on the classified set of pixels Ω_B that are the part of similar structure as current pixel.

Based on the step of selective computation in which the algorithm checks whether the current predictor penalties form the prediction which is reasonably accurate, we omit a number of computations of predictor weights which do not contribute substantially to the overall prediction efficiency and therefore the compression. On the other side, this simple, yet effective technique reduces the computational complexity of our first proposal, a CBP predictor.

As an illustration, Figure 3 shows the effects of applying Selective Classification and Blending Predictor SCBP on a sample medical image. Left image is the original while on the right side, white pixels denote the positions for which the recomputation of the predictor penalties was performed. More specifically, we set the recomputation threshold T_e to zero, meaning that only the perfect prediction on the previous pixel will not trigger the predictor recomputation. For this particular case, this resulted in reduction of predictor blend calculations by a factor of 2.05 and reduced the computation time by 39% of the original CBP predictor. At the

Predictor	Prediction function	Description
f_1	N	North pixel to the current pixel
f_2	W	West pixel to the current pixel
f_3	NW	Northwest pixel to the current pixel
f_4	NE	Northeast pixel to the current pixel
f_5	$N+W-NE$	Planar region predictor
f_6	$2N-NN$	Planar region predictor horizontal
f_7	$2W-WW$	Planar region predictor vertical

Table 1. Set of static predictors

same time the compression efficiency measured by zero–order entropy of prediction error was unnoticeably increased by approximately 2%. More details are given in the Section 4 where we demonstrate our experimental results.

4. Experimental results

In this section we show the experimental results obtained with our proposed predictors. We also show the results obtained while implementing our predictors in complete lossless image codecs. Proposed predictors have several parameters that can be varied in order to balance between the compression ratio and the computational complexity. These are, as illustrated in Figure 2:

1. Radius R of the classification context which defines the size of the classification template Ω_C on which the current cell is estimated. Increasing R, the algorithm has more chance to capture the structure of the current region, but at the cost of requiring more computations.
2. Population size M of the current cell. Captures the similar pixels to the current pixel for blending of predictors. If this parameter is too small the resulting predictor is overspecialized and having too few samples to effectively model the local region. On the other hand, higher value of would result in pixels in the current cell that are not similar enough to the current pixel. This will also jeopardize the final predictor efficiency to predict well in the current region.
3. The size of the vector template d. Vector template is composed of d closest already observed pixels. In proposed predictor we experimented with the vector size 1 to 10 as shown in Figure 2(b). The higher vector size, the higher capability to capture more complex structures. Also higher vector size demands more computations.
4. Set of predictors $\mathcal{F}$. Exhaustive tests on a large number of medical test images had shown that the set $\mathcal{F}$ should contain simple static predictors that are suited to various regions, such as oriented edges, planar regions, smooth regions etc. Our experiments resulted in best choice for set $\mathcal{F}$ to be:

$$\mathcal{F} = \{N, W, NW, NE, N+W-NW, G_W, G_N\}, \tag{16}$$

   where $G_W = 2N - NN$ and $G_N = 2W - WW$. Table 1 gives the details of the static predictors used in our experiments. We use compass point notation for surrounding pixels,

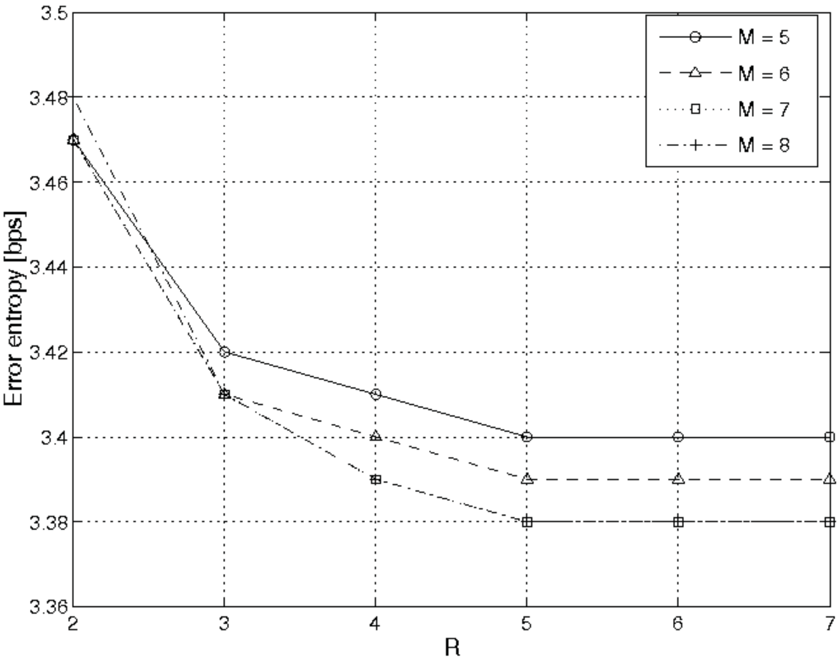


Fig. 4. Average zero order entropy of prediction error

> i.e. W denotes west pixel, N denotes a pixel placed on the North side of the current pixel, NN denotes pixel placed two pixels to the North of the current pixels and so on. We selected a set of seven predictors that include horizontal and vertical edge predictors and planar predictors. The process of blending of selected predictors will finally produce an averaging predictor which is known to predict well in the noisy regions. This way, final predictor will try to adapt to most relevant image properties while detecting and adjusting to the dominant property in the current local region, whether it is an edge, a planar region or even a texture with nontrivial coarseness and period The set includes simple predictors suited for edges, planar predictor and predictors with gradient modeling for planar regions. Through the blending process the final predictor has the form of averaging predictor which is suited for noisy regions and for modeling higher structures such as textures.

Figure 4 shows the average zero-order entropy of the Classification and Blending Predictor CBP obtained on a test set of images shown in Figure 5. The zero-order entropy is shown in bits per symbol (bps) versus radius of the classification template and with the cell population M as a parameter. The vector size was fixed to four. As can be seen, the entropy saturates for $R > 6$ and $M > 7$. That clearly shows that increasing the radius and cell size beyond these limits will result in diminishing returns in terms of compression while increasing computational time. Consequently, in Table 2 we show experimentally chosen working parameters of the CBP and SCBP predictors that we will further use in the comparison with other popular predictors for lossless image compression. We used these settings in experiments in which we employ our predictors as parts of the complete image compression algorithms. Notice that due to the decreased computational requirements, SCBP predictor can afford larger classification template Ω_C with higher radius $R = 6$ as opposed to CBP with the $R = 5$.

Predictor	Radius R	Cell size M	Vector size d	Threshold T_e
CBP	5	7	4	–
SCBP	6	7	4	0

Table 2. Working parameters for CBP and SCBP

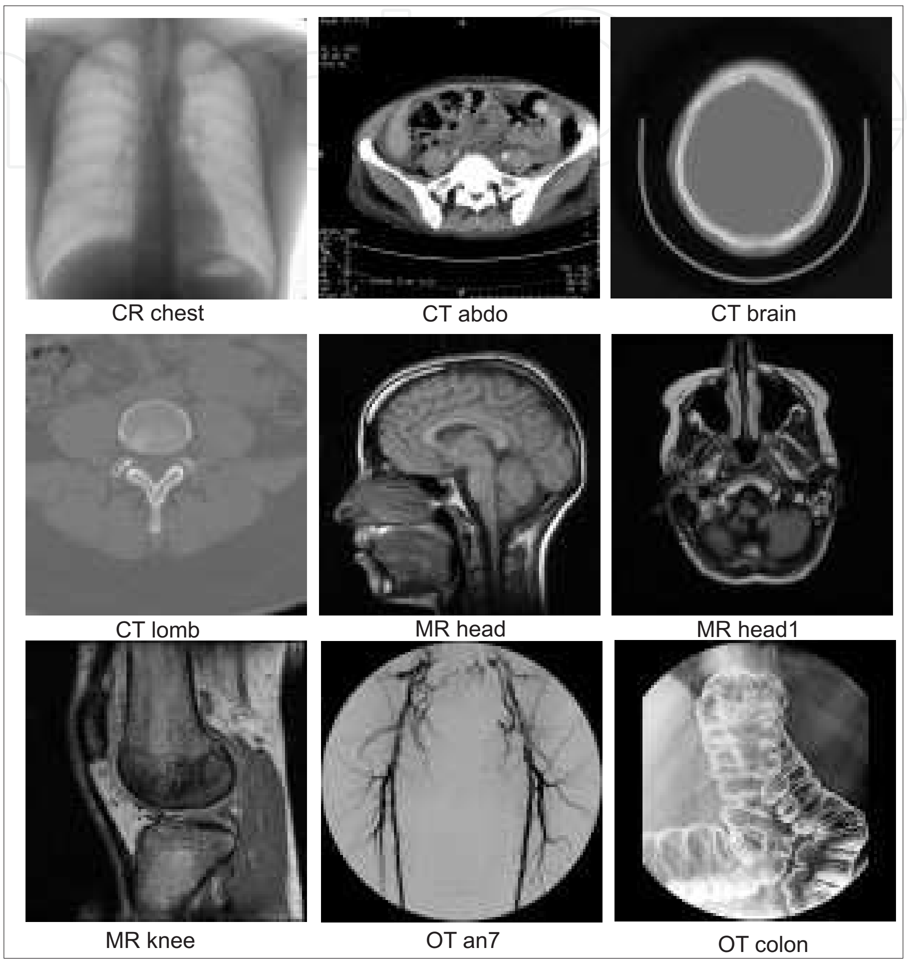

CR chest
CT abdo
CT brain
CT lomb
MR head
MR head1
MR knee
OT an7
OT colon

Fig. 5. Test set of medical images

In order to demonstrate the performance of proposed predictors we will show their results on the set of nine medical images obtained with varying modalities (Meyer, 2005; S. Barré Medical Image Samples, 2004). Those images are grey level images with 8 bits per pixel precision and their thumbnails are shown in Figure 5.

We compared our predictors CBP and SCBP with several popular predictors used in predictive lossless image compression algorithms in Table 3. The results show the zero-order entropy of the prediction error in bits per symbol for the same test set of medical images. GAP predictors is static heuristically tuned predictor used in the CALIC algorithm (Wu, 1997), and MED is the predictors from LOCO-I algorithm which is used as JPEG-LS standard (M.J. Weinberger and G. Seroussi and G. Sapiro, 1998). Our predictors outperform other predictors for all images in the set except for *CT abdo* where MED predictor performed best. CBP was slightly better than SCBP for all test images.

Image	**GAP(CALIC)**	**MED(JPEG–LS)**	**CBP**	**SCBP**
CR Chest	2.49	2.52	**2.28**	2.29
CT Abdo	2.73	2.67	**2.72**	2.75
CT Brain	1.93	1.79	**1.17**	1.19
CT Lomb	2.59	2.46	**2.19**	2.19
MR Head	4.75	4.80	**4.55**	4.65
MR Head1	4.82	4.85	**4.51**	4.55
MR Knee	5.20	5.24	**5.12**	5.16
OT An7	4.28	4.20	**4.17**	4.19
OT Colon	3.91	3.84	**3.75**	3.77
Average	3.63	3.60	**3.38**	3.42

Table 3. Zero–order entropy of prediction error (bps)

We built two complete lossless image codecs that incorporate CBP and SCBP predictors, contextual error modeling described section 2.2, and arithmetic coding. We call them Classification and Blending Predictive Coder CBPC and Selective Classification and Blending Predictive Coder SCBPC, respectively. To compare both coders we show their compression efficiency in the Figure 6. We show the compression ratios obtained on the test set. From the figure, we can see that the SCBPC encoder closely follows the CBP although with reduced computational complexity. Geometric mean of compression ratios is within 2% of the CBP results. Additionally, Figure 7 illustrates the time savings of SCBPC encoder compared to CBPC encoder on the test set. On average, SCBPC encoder has reduced the execution time by more than 35% compared to CBPC method.

Finally, Table 4 shows the results of our CBPC and SCBPC coders compared with the results of popular lossless coders for our test set of medical images. The first column shows the compression ratios of the CALIC algorithm (Memon & Wu, 1999), the second column the JPEG-LS results (M.J. Weinberger and G. Seroussi and G. Sapiro, 1998) and the third column shows the results of JPEG 2000 lossless compression that uses reversible wavelet transform (Santa-Cruz & Ebrahimi, 1997). Last two columns show the results of CBPC and SCBPC encoders respectively. Our both coders outperform other coders. CBPC obtained the best compression ratios on all test images except for *CT abdo* where it was slightly outperformed by JPEG–LS. SCBPC closely followed CBPC encoder beating other encoders on all images except on *CT abdo*.

5. Conclusions

The results obtained with our proposed compression algorithms are encouraging. Proposed approach of modeling the image as composed of regions with the mixture of dominant and non–dominant properties has shown to be useful for lossless image compression. On the other hand, it should be noted that the increase in compression performances comes with the increase in computational complexity. This was extensively studied and resulted in the heuristic approach described in this work. Our predictors are moderately more complex than predictors used in contemporary lossless compression algorithms such as CALIC and

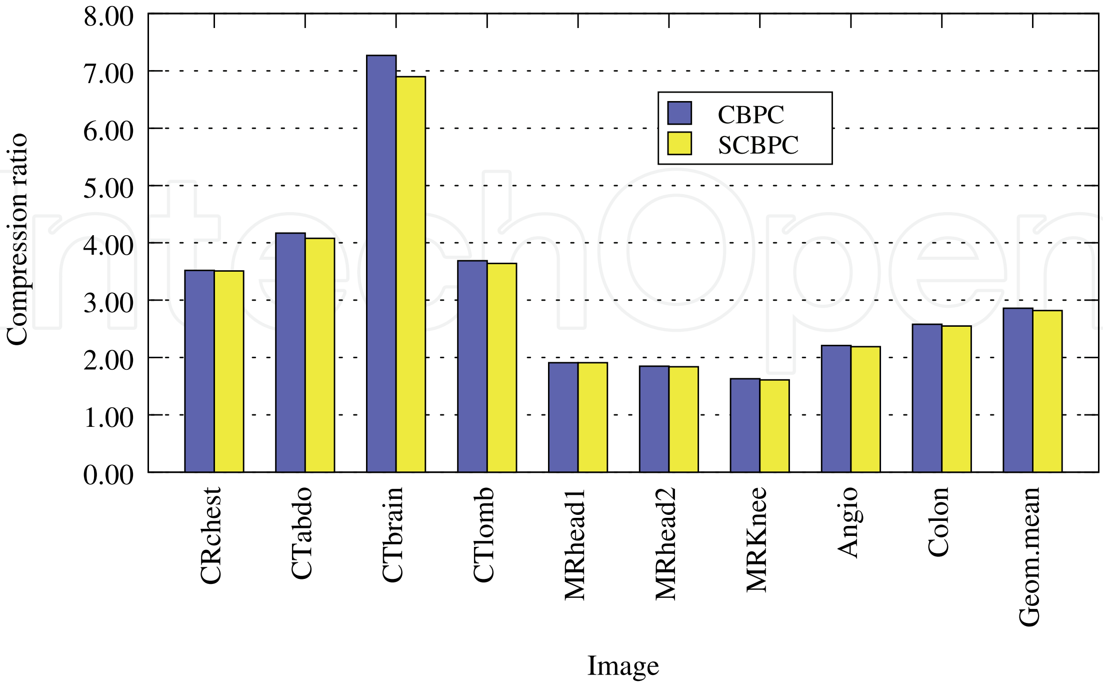


Fig. 6. Compression efficiency of SCBPC and CBPC

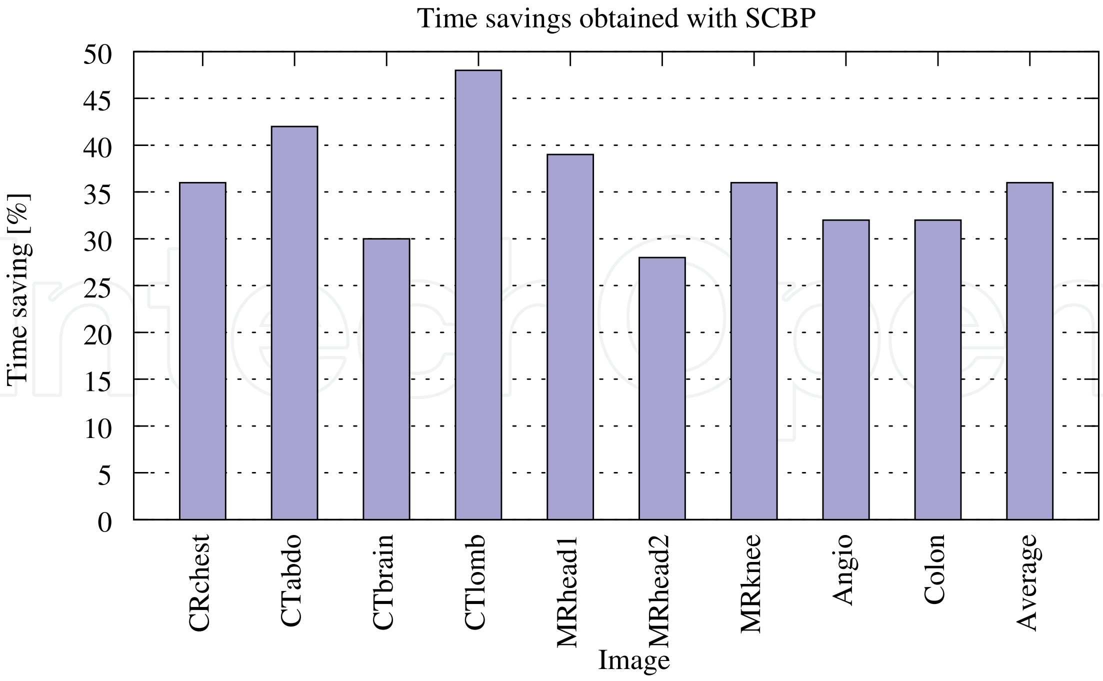


Fig. 7. SCBP and CBP: Time savings

Image	CALIC	JPEG–LS	JPEG2000$_R$	CBPC	SCBPC
CR Chest	3.40	3.35	3.17	3.52	3.51
CT Abdo	3.52	4.23	3.09	4.17	4.08
CT Brain	6.45	6.20	5.67	7.27	6.90
CT Lomb	3.62	3.42	3.36	3.69	3.64
MR Head	1.87	1.80	1.79	1.91	1.91
MR Head1	1.80	1.73	1.70	1.85	1.84
MR Knee	1.61	1.57	1.57	1.63	1.61
OT An7	2.16	2.18	2.02	2.21	2.19
OT Colon	2.49	2.50	2.32	2.58	2.55
Geom.mean	2.72	2.71	2.52	2.86	2.82

Table 4. Obtained compression ratios

JPEG–LS, but less complex than highly–adaptive predictors that can be found in proposals based on least squares approach (X. Li and M.T. Orchard, 2001). Also, the complexity of proposed predictors can be tuned for both goals: better compression and faster time by changing its parameters which can be set on a image basis.

As part of the future work, we plan to investigate possible extension of our image compression techniques to 3D and volumetric medical data. In this framework, we would propose to have regions of data that will be stored using lossless compression, while other, less relevant parts could be compressed with lossy algorithms. Additionally, prediction template can be extended by incorporating three dimensions rather than using two dimensions as is currently the case.

6. References

A. Gersho (1993). Optimal Vector Quantized Nonlinear Estimation, *Proc. 1993 IEEE Int. Symp. on Information Theory*, San Antonio, p. 170.

A. Gersho and V.M. Gray (n.d.). *Vector Quantization and Signal Compression*, Kluwer Academic Publishers.

Choong, M., Rajasvaran, L. & Bister, M. (2007). Cost–effective handling of digital medical images in the telemedicine environment, *International Journal of Medical Informatics* 76: 646–654.

Dikbas, S. & Zhai, F. (2010). Lossless image compression using adjustable fractional line-buffer, *Sig. Proc.: Image Comm.* 25(5): 345–351.

G. Motta and J.A. Storer and B. Carpentieri (2000). Lossless Image Coding via Adaptive Linear Prediction and Classification, *Proceedings of The IEEE* 88(11): 1790–1796.

Knezović, J. & Kovač, M. (2003). Gradient Based Selective Weighting of Neighboring Pixels for Predictive Lossless Image Coding, *Proc. 25th Int. Conf. Information Technology Interfaces ITI 2003*, Cavtat, Croatia, pp. 483–488.

M. Boliek and A. Zandi (1995). CREW: Lossless/lossy Image Compression – Contribution to ISO/IEC JTC 1.29.12, *ISO Working Document ISO/IEC JTC1/SC29/WG1 N196* .

Memon, N. & Wu, X. (1999). Recent developments in context-based predictive techniques for lossless image compression, *IEEE Computer* 10: 127–136.

Meyer, B. (2005). Tmw page, http://www.csse.monash.edu.au/ bmeyer/tmw/index.html.

Miaou, S.-G., Ke, F.-S. & Chen, S.-C. (2009). A lossless compression method for medical image sequences using jpeg-ls and interframe coding, *Information Technology in Biomedicine, IEEE Transactions on* 13(5): 818 –821.

M.J. Slyz and D.L. Neuhoff (1994). A Nonlinear VQ–based Lossless Image Coder, *Proc. 1994 Data Compression Conference*, IEEE, Snowbird, UT, pp. 491–500.

M.J. Weinberger and G. Seroussi and G. Sapiro (1998). The LOCO–I Lossless Image Compression Algorithm: Principles and Standardization into JPEG–LS, *HP Laboratories Technical Report HPL-98-198* .

Neves, A. J. R. & Pinho, A. J. (2009). Lossless compression of microarray images using image-dependent finite-context models, *IEEE Trans. Med. Imaging* 28(2): 194–201.

P.G. Howard and J.S. Vitter (1994). Fast and Efficient Lossless Image Compression, *Proc. 1994 Data Compression Conference*, IEEE, Snowbird, UT, pp. 351–360.

S. Barré Medical Image Samples (2004). http://www.barre.nom.fr/medical/samples/, WWW.

S. Forchhammer and X. Wu and J.D. Andersen (2004). Optimal Context Quantization in Lossless Compression of Image Data Sequence, *IEEE Transactions on Image Processing* 13(4): 509–517.

Sanchez, V., Abugharbieh, R. & Nasiopoulos, P. (2009). Symmetry-based scalable lossless compression of 3d medical image data, *IEEE Trans. Med. Imaging* 28(7): 1062–1072.

Sanchez, V., Nasiopoulos, P. & Abugharbieh, R. (2008). Efficient lossless compression of 4-d medical images based on the advanced video coding scheme, *IEEE Transactions on Information Technology in Biomedicine* 12(4): 442–446.

Sanchez, V., Nasiopoulos, P. & Abugharbieh, R. (2009). Novel lossless fmri image compression based on motion compensation and customized entropy coding, *Trans. Info. Tech. Biomed.* 13: 645–655.

Santa-Cruz, D. & Ebrahimi, T. (1997). An analytical study of jpeg 2000 functionalities, *Proceeding of IEEE Data Compression Conference*, IEEE, pp. 49–52.

Seemann, T. & Tischer, P. (1997). Genralized Locally Adaptive DPCM, *Proc. 1997 Data Compression Conference*, IEEE, Snowbird, UT, pp. 473–488.

Seemann, T., Tischer, P. & Meyer, B. (1997). History–Based Blending of Image Sub-Predictors, *Proceedings of PCS'97 Picture Coding Symposium*, VDE–Verlag, Berlin, Germany, pp. 147–151.

Sriraam, N. & Eswaran, C. (2008). An adaptive error modeling scheme for the lossless compression of eeg signals, *IEEE Transactions on Information Technology in Biomedicine* 12(5): 587–594.

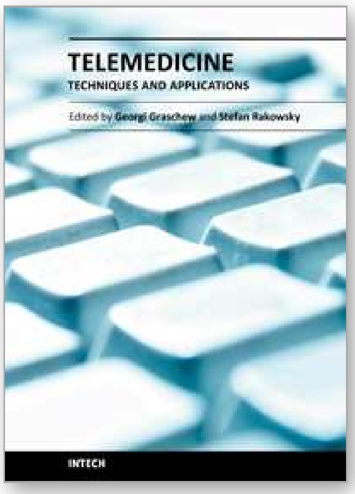
Wu, X. (1997). Lossless Compression of Continuous-tone Images via Context Selection, Quantization, and Modeling, *IEEE Transactions On Image Processing* 6(5): 656–664.

X. Li and M.T. Orchard (2001). Edge–Directed Prediction for Lossless Compression of Natural Images, *IEEE Transactions on Image Processing* 10(6): 813–817.

X. Wu and N. Memon and K. Sayood (1995). A Context–Based, Adaptive, Lossless/nearly–lossless Coding Scheme for Continuous–Tone Images, ISO Working Document ISO/IEC/ SC29/WG1/N256.

Zagar, M., Kovac, M. & Bosnic, I. (2007). Lossless and lossy compression in 4d bio-modeling, *Proceeding of the First International Conference in Information and Communication Technology & Accessibility*, pp. 271–276.

Zhang, Y. & Adjeroh, D. A. (2008). Prediction by partial approximate matching for lossless image compression, *IEEE Transactions on Image Processing* 17(6): 924–935.


TELEMEDICINE
TECHNIQUES AND APPLICATIONS
Edited by Georgi Graschew and Stefan Rakowsky
INTECH

Telemedicine Techniques and Applications
Edited by Prof. Georgi Graschew

ISBN 978-953-307-354-5
Hard cover, 514 pages
Publisher InTech
Published online 20, June, 2011
Published in print edition June, 2011

Telemedicine is a rapidly evolving field as new technologies are implemented for example for the development of wireless sensors, quality data transmission. Using the Internet applications such as counseling, clinical consultation support and home care monitoring and management are more and more realized, which improves access to high level medical care in underserved areas. The 23 chapters of this book present manifold examples of telemedicine treating both theoretical and practical foundations and application scenarios.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821