Wavelet Image Compression Method Combined With The GPCA

Indrit Enesi
Department of Electronic and Telecommunication
Polytechnic University of Tirana
Sheshi “Nene Tereza”, Nr. 4, Tirana, Albania
ienesi@fti.edu.al

Abstract—In the most common standard coding schema of the image, coder has greater complexity than the decoder, usually 5-10 times higher. In wireless sensor networks is the opposite, the coder is implemented in a battery feed device, having so very limited power resource, while the decoder is typically implemented in a powerful computer. The paper will focus on a combination of wavelet technique with algebraic GPCA method, realizing an improvement of traditional wavelet technique, providing a suitable compression of multimedia information without reducing its quality. Simulations in MATLAB provide desired compression levels. The proposed method, unlike other existing methods, takes into account the structure of multi-modal data, except the correlation in 1 D and 2 D it takes in consideration correlation between color channels. The method gives a significant improvement in the PSNR values reaching an average of 15% compared with traditional wavelet compression.

Index Term—wavelet, GPCA, multi-modal, linear, hybrid

I. INTRODUCTION

Nowadays wireless sensor networks are being used more and more. In such networks, nodes have typically very limited power resources [1]. Images use a large size in memory, so the compression is required. The transmission of data, especially images, is a procedure that consumes more energy from a node. It is clear that compressing the image followed by its transmission, is generally more efficient than direct transmission of the image without compression. Transmitting the compressed data, the energy will be reduced [2]. A lot of research has been done in this field, and the techniques are implemented in various ways. Depending on the specific use and the network design, the best featured method is selected.

II. WAVELET TRANSFORM

The first step in image compression is converting a RGB image in YUV. Wavelet transform admits a series of colors plan YUV9 [3]. It operates in multiple steps in any plan of colors, dividing them into a series of subbands, as shown in Fig. 1. Wavelet algorithm extracts both high and low frequency data on each step, separating them into two regions. The process is repeated successfully in the low frequency half, dividing it further into smaller subbands, and stops when the smaller subbands are not broken anymore. In many cases, six steps are sufficiently adequate. The subband algorithm forms a pyramid, collecting data on low-frequency branch on the left, leaving the DC coefficients in the corner [3]. The final transformation is presented using 16-bit whole numbers. Fig. 2 shows how coloring plans are transformed in wavelet subbands. It is clear that many subbands contain very little information, especially in regions of high frequency. This is revealed by applying a quantization filter in these regions, leading to zero most of the high frequency coefficients.

Redundant data will favor the tree-zero coder for better compression of data [3]. Fig. 3 shows the same wavelet data set presented through a colored coding map. This highlights the extent in bits of each wavelet coefficient. Black areas correspond to 0-bit information, while the blue gradient represents increasing sized bits. Redundant zero coefficients are outlined in Fig. 4. Wavelet coefficients are currently 16 bits, but most of them have values mostly located in the 5 most significant bits. About 1% of the coefficients have values larger than 8-bits-those corresponding to low frequency and DC coefficients.
III.  ENTHROPY CODING

If the set of wavelet coefficients is divided into sixteen bit individual plans, we notice that the plan that corresponds to the least significant bit rate would be rare and generally zero, while the plan that corresponds to the most significant bit rate will be more populated [4]. Moreover, every bit plan will become rarer being directed towards regions of high frequency. Using this feature, a zero-tree compression algorithm is implemented, which operates on these bit plans independently. Fig. 5 shows how the zero-tree bits is arranged in a single plan. The diagram on the right shows the tetrahedron-tree structure [3]. The root node corresponds to the position of the DC coefficients. The following values expand, spreading to cover areas with high frequency.

![Diagram of a wavelet bit plan structure](image)

The tree is constructed by connecting each parent node with four children that contain one or more non-zero bits. Parent nodes, whose children contain only zero bits, return to a leaf node. So, all branches will be limited to a node, whose following values are all zero. Since the coefficients in the high frequency data mainly contain zero, the technique of restriction makes the tree pretty extensive with branches that grow levels even more, thus making the data compression possible [3][4]. We note that in areas where there is very little high frequency information, most of the corresponding wavelet coefficients are quantized to zero. Fig. 6 shows the lack of corresponding parent node, because all subsequent branches are zero. Having built the zero-tree bits, a series of codes are generated to indirectly present the entire tree. These bit codes are stored in a file holder as compressed data, which would allow a decoder to reconstruct a zero-tree.

![Diagram showing parent node maps for zero-tree Y, U, and V respectively. Black regions show that there is no parent node in any of the bit plans.](image)

IV. HYBRID LINEAR MODEL

An image with width, length, and color channels lies in a multidimensional space $R^{WHx}$ [5]. We initially can reduce the size by dividing the image into a set of blocks (do not expect each other). Each block with dimensions $bxb$ will be collected in a vector $x_i \in R^K$, where $K=bxc$ is space dimension. In this way, the image $I$ will be converted in a vector set $x_i \in R^K, i = 1...M$, where $M = WH/b^2$ is the vectors’ total number. It is assumed that $x_i$ are random samples from a probability distribution (not necessarily the only). Since the distribution can be tricky, a common solution is to be achieved for a better approximation between a simple class models. The “optimal” model is one that minimizes a certain distance from the true model. Different choices of models and measurements of the distance classes have led to many algorithms implemented in test machines, structure’s recognition, prediction software and image processing. The most common size for measuring the image compression is “Mean Square Error” (MSE) between the original image and the reconstructed image $\hat{I} : e^2 = \frac{1}{WHc} \| I - \hat{I} \|^2$ [5]. Since we will approximate vectors (blocks) instead of pixels of the image, in the following derivation is better to define MSE for an array that is different from $e^2$ during each step:

$$e^2 = \frac{1}{M} \sum_{i=1}^{M} \| x_i - \bar{x}_i \|^2 = b^2 \epsilon_i^2 = K \epsilon_i^2$$

(1)

The “Peak Signal to Noise Ratio” (PSNR) of the approximated image is defined as:

$$PSNR = -10 \log e^2 = -10 \sqrt{\epsilon_i^2}$$

(2)

![Diagram showing linear models data lie in a common subspace](image)

Linear Models. If we assume that vectors $x$ are derived from a Gaussian distribution or a linear subspace, the optimal model for a given PSNR can be taken from Key Components Analysis (PCA) [6] or equivalently by Karhunen-Loève Transform (KLT).

Lets $\bar{x} = \frac{1}{M} \sum_{i=1}^{M} x_i$ average vector data, and

$$X = [x_1 - \bar{x}, x_2 - \bar{x}, ..., x_M - \bar{x}] = U \in V^T$$

be the SDV of the subtracted average matrix. Then all the vectors $x_i$ can be represented as a linear superposition:

$$x_i = \bar{x} + \sum_{j=1}^{K} \alpha_j \phi_j, i = 1,...,M$$

(3)

where: $\phi_j, j = 1,..., K$ are simply the columns of the matrix $U$. 

123105-4747-IJIPNS-IJENS © October 2012 IJENS
The matrix $\sum = \text{diag}(\sigma_1, \sigma_2, ..., \sigma_K)$ contains the sorted values $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_K$ [7]. It is known that the optimal linear presentation $\hat{x}_i$ by the MSE$^2$ is obtained by keeping the first (fundamental) components:

$$\hat{x}_i = x + \sum_{j=1}^{k} \alpha_i^j \varphi_j, i = 1, ..., M$$

(4)

where: $k$ is chosen to be:

$$k = \min(n), \frac{1}{M} \sum_{i=n+1}^{K} \sigma_i^2 \leq \varepsilon^2$$

(5)

The complexity of the model marked with $\Omega$ is the total number of coefficients to be presented to model $\{\alpha_i^j, \varphi_j, \hat{x}_i\}$ and following by the loss of approximation $\hat{I}$ of image I. This is provided by:

$$\Omega(M, k) = Mk + k(K - k + 1)$$

(6)

where the first term is the number of coefficients $\{\alpha_i^j\}$ denoting $\{x_i - \hat{x}_i; i = 1, ..., M\}$ regarding to the basic principle $\Phi = \{\varphi_j; j = 1, ..., k\}$ and the second term the number of coordinates that need to present $\Phi$ and the average vector $\bar{x}$. Second term is often called the overhead. Note that the original set of vectors $x_i$ containing MxK entry coordinates. If $\Omega \ll MK$, the new presentation, even though with loss, is more compact [8]. The search for such a compact representation is the core of any method of image compression (with loss). However, in order to make fair comparisons with other methods, in subsequent discussions and experiments always we will count the total number of coefficients required for the presentation, including the redundancy.

A. Hybrid Linear Models.

Linear model is more efficient when the function is one-modal distribution [9]. Since the image I contains several heterogeneous regions $\{I_n\}_{n=1}^{N}$, the data vectors $x_i$ may be samples from a subspace set of different size or a combination of multiple Gaussian distributions. Fig. 8 shows three main components of the vector data $x_i$ (as points in $R^3$) of the image. Here we can notice the multi-modal characteristic in the data [9]. It is assumed that the natural image I can be segmented in $N$ separate regions $I = \bigcup_{n=1}^{N} I_n$ with $I_n \cap I_n = \emptyset$ for $n \neq n'$. At every region $I_n$, we can assume that the linear model is valid for the vectors subset $\{x_{n,i}\}_{i=1}^{M_n}$ in $I_n$:

$$\hat{x}_{n,i} = \bar{x}_n + \sum_{j=1}^{k_n} \alpha_i^j \varphi_{n,j}, i = 1, ..., M_n$$

(7)

As in the linear model, dimensions $k_n$ of every subspace are decided from the common $\varepsilon^2$ MSE using equation (2). The complexity of the model, expressed by the total number of coefficients needed for presentation of the hybrid linear model $\{\phi_{n,j}, \hat{x}_{n,i}\}$ is [10]:

$$\Omega = \Omega(M_1, k_1) + ... + \Omega(M_k, k_k) = \sum_{n=1}^{N} M_n k_n + k_n(K - k_n + 1)$$

(8)

Fig. 8. In linear hybrid models, the data vectors are located in multiple subspaces.

Fig. 9. Left: Figure of the Palace of Congresses, Tirana. Right: The coordinates of each point are the three main components of vectors $x_i$.

It is noticed that $\Omega$ is similar to the effective dimensions (ED) of the hybrid linear presentation specified in (8). In this way, finding a presentation that reduces $\Omega$ is the same as minimizing the effective dimensions of image data [11]. If we model the set of all vectors $\{x_{n,i}\}_{i=1}^{M_n}$ with a single subspace (of the same MSE), the subspace dimensions in general should be $k = \min(k_1 + ... + k_N, K)$. It is easy to prove by the equation (3) that in some reasonable conditions (for e.g. when $N$ is not too large), we have $\Omega(M, k) > \Omega(M_1, k_1) + ... + \Omega(M_k, k_k)$. So, if we can define a linear hybrid model for an image, the resultative presentation will be generally more compressed than that of a single subspace[12]. This will also be verified by experiments on real images.

Identification of a linear model is generally a difficult problem: if the segmentation of the vectors $x_i$ would be known, optimal subspaces for every subset can be found easily using PCA; and vice-versa if the subspaces are known, vectors $x_i$ (and in this way the image) can be easily segmented in their most approximate subspace. This kind of problem can be treated
with one of several methods of grouping developed in statistics or in testing machine [13]. However these techniques are iterative and incremental in nature and in this way they tend to converge to local minima if the initialization is too far. Also these methods do not fit directly with the hybrid models, the components of which may have different and unknown dimensions. One method, for e.g is the “Generalized Principal Component Analysis” (GPCA) [14]. This method doesn’t require preliminary knowledge of the subspace dimensions and may estimate in the mean time the multiple subspaces and segment the data in them [15], which is suitable for our purpose here.

Here we assumed that an image can be divided into several separate regions. In every region a valid linear model is applied. We are interested in evaluating such a division in subspaces. To realize this we are going to use GPCA. This method evaluates the multiple subspaces and divides data into segments. In this paper we propose a new linear hybrid model depending on MSE that we need.

If we can find a linear hybrid model for an image, the result presentation in general will be more compressed comparing with the one which has only one subspace.

V. THE IMPLEMENTATION OF THE PROPOSED METHOD

The wavelet image will be decomposed into a subband tree. Each of subband levels LH, HL, HH will contain information about high frequency contours and LL subband is compressed further to next subband level. Fig. 11 shows such a structure for the decomposition of level 2. Vectors are reconstructed by grouping wavelet coefficients in subbands LH, HL, HH. Let assume that the color is constant along a corner. If the edge is along the horizontal, vertical or diagonal direction it will have an edge in subbands coefficients in LH, HL or HH, respectively [16]. Two other subbands will be zero. The size of data vector associated with this edge will be 1.

If the edge is not exactly in one of these areas, it will be noticed in one of the three subbands coefficients. If the color varies along the edge of the subset, the proportions will be higher. The wavelet method does not need the under-sampling pyramid [17]. The decomposition in many wavelet levels gives a multi-scale structure. Fig. 12 shows a structure of the decomposition of the level 3 in wavelet 4.4.

In the nonlinear wavelet approximation, coefficients that are below an error threshold will be ignored. Similarly, in our model, not all the data vectors of will be modeled and approximated. We should only approximated (coefficients) vectors \( x_i \) satisfying the following condition:

\[
\|x_i\|^2 > \varepsilon^2
\]  

(9)

We do not need to scale error tolerance at different levels because the wavelet bases are orthonormal by its construction. Practically, energy of multiple vectors is very close to zero. Only a small amount of vectors at every level should be modeled. The evaluation process of such a model can be summarized as following:

**TABLE 1: Hybrid Linear Multi-Step Model: Wavelet Domain Algorithm**

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>function ( \hat{I} = \text{Multi_Step_Model}(I, \text{level}, \varepsilon^2) )</td>
</tr>
<tr>
<td>2.</td>
<td>( \hat{I} = \text{WaveletTransform}(I, \text{level}) )</td>
</tr>
<tr>
<td>3.</td>
<td>for each level do</td>
</tr>
<tr>
<td>4.</td>
<td>( \hat{I} = \text{HybridLinearModel}(\hat{I}_{\text{nivel}}, \varepsilon^2) )</td>
</tr>
<tr>
<td>5.</td>
<td>end for</td>
</tr>
<tr>
<td>6.</td>
<td>( \hat{I} = \text{InversWaveletTransform}(\hat{I}_{\text{nivel}}, \varepsilon^2) )</td>
</tr>
<tr>
<td>7.</td>
<td>Return ( \hat{I} )</td>
</tr>
</tbody>
</table>

VI. EXPERIMENTAL RESULTS

The comparison of a multi-scale hybrid linear in wavelet domain with the classic one is treated. The experiment is performed with two standard images: 1200x1600 images of the “Palace of Congresses” and a photo near it as shown in Fig. 13.
Simulations are done in Matlab [18]. The number of levels for the chosen model is 3. In Fig. 15 and 16 we see the comparison of our method with standard wavelet. Hybrid multi-scale model in the linear wavelet domain achieves a better PSNR than classic wavelet model.

From the comparison of multi scale hybrid method with the classic wavelet for the two images, we see that our model achieves better PSNR than wavelet. In Fig. 17 is shown a visual comparison of files with a multi-scale wavelet hybrid using the same amount of coefficients for both models. We notice that in the area around the contours, the wavelet transform blurs both contours and fine details of them. Hybrid linear model in the wavelet domain eliminates the effects of blocks, holds and preserves the sharp contours and more details in the background. In gives a better PSNR, and it also gives a better visual effects.
While studying some types of images, we distinguish some specific cases. Normally the wavelet methods blurs the sharp edges in 2D, but represents best-frequency regions/low entropy. To make the comparison of our method, we select specific cases as in Fig. 21.
Below are some graphic above comparative figures because we applied both models to an average of 2.3% of the approximate coefficients. Fig. 21 shows the average performance of the model multi-scale hybrid linear for all images taken in consideration. Hybrid linear model in wavelet domain does not produce good results in black and white images since the size of the vectors are only 3 and there is not much room for further reduction. For black and white images should be seen other adaptations and modifications of this model, but this is beyond the scope of this paper. Our goal was to show the great potential of a new image presentation obtained by the combination of the subspaces methods with the traditional approximation schemes for presenting the image.

VII. CONCLUSIONS

In this paper we propose a simple but effective class of mathematical model for the presentation and approximation of natural images to be used in wireless sensor networks. In these networks, due to the constraints arising from the use of sensory devices, we need to provide efficient compressing models and at the same time with less calculation and complexity. The model presented in this paper handle the heterogeneous and hierarchical structure of the image and exploits the correlation between color channels. Experiments show that the performance of the proposed method is comparable or even better than the classic wavelet method. Furthermore our scheme complements the existing schemes and together they achieve a higher performance. PSNR-values achieved are 15% larger than the existing traditional wavelet method. In the future it remains to be worked on how can be realized schemes for the quantization and coding the entropy, such a scheme can be developed as a complete package for image compression. The next step will be adapting the code in MATLAB and TinyOS language for the implementation of the software in the sensor nodes.

VIII. REFERENCES


123105-4747-IJVIPNS-IJENS © October 2012 IJENS