Simulated Annealing Algorithm for 2D Image compression

Pedro Moreno-Bernal[1], Marco Antonio Cruz-Chávez[2], Abelardo Rodríguez-León[3], O. López[4], M.P. Malumbres[4], Martín G. Martínez-Rangel[5], Alina Martínez-Oropeza[1], Beatriz Martínez-Bahena[1], Jazmín Yanel Juárez-Chávez[1]

[1]Posgrado en Ingeniería y Ciencias Aplicadas, [2]CIICAp-UAEM, [5]FCAeI,
Avenida Universidad 1001. Col. Chamilpa, C.P. 62209. Cuernavaca, Morelos, México,
[3]ITVer, Calzada Miguel Ángel de Quevedo 2779, Veracruz, México
[4]Miguel Hernandez University, Av. Universidad s/n, 03202 Elche, Spain

{pmoreno,mcruz,alinam}@uaem.mx

Abstract

In this paper a new sign coding approximation method for the wavelet coefficients in a 2D image codec based on a simulated annealing metaheuristic is presented. The efficiency of the proposed algorithm versus a genetic algorithm using benchmarks of Kodak is compared and showing that the proposed sign prediction algorithm is efficient and provides a significant reduction of wavelet coefficients sign information in the final bit-stream. The results show that, by including sign coding capabilities to a non-embedded encoder, the sign compression gain is up to 17.35%, being the rate-distortion (R/D) performance improvement up to 0.25 dB.

1. Introduction

Many image compression algorithms, including the standard JPEG2000 [1], employ the Discrete Wavelet transform (DWT) [2] into their algorithms. One of the most valuable advantages of the wavelet transform is the provision of both frequency and spatial localization of image energy. An image can be represented by a two-dimensional function *f(x,y)* where the value of the *f* in spatial coordinates *(x,y)* gives the image intensity in that point. When the image is generated from a physical process, their values are proportional to the intensity of the energy radiated by the physical source resulting *f(x,y)* different from zero and finite. A digital image has been discretized in spatial coordinates and brightness, this involves sampling and quantization process that creates an array where each point identifies the level of light intensity called pixel. Therefore, the image energy is compacted into a small fraction of the transform coefficients, where the values have been represented by magnitude and sign, and the compression can be achieved by coding these coefficients. The energy of a wavelet transform coefficient is restricted to non-negative real numbers, but the coefficients themselves are not, and they are defined by both a magnitude and a sign.

Shapiro stated in [3] that a transform coefficient is equally likely to be positive or negative and thus one bit should be used to encode the sign (raw coding). However, in recent years, several authors have begun to use context modeling for sign coding [4][5][6]. A context model allows defining different schemes for grouping subsets based on experience. Schwartz, Zandi and Boliek were the first authors to consider sign coding, using one neighboring pixel in their context modeling algorithm [7]. The main idea behind this approach is to find correlations along and across edges.

In [4], X. Wu presents a high order context modeling encoder. In this coder, the sign and the textures share the same context modeling. This model is based on a different neighborhood for the HL, LH and HH wavelet subbands. For the HL subband, the information of North, North-West, North-East, North-North and South sign is used to predict the current coefficient sign. The neighbors sign information used for the LH subband is North, North-West, North-East, West-West and East. Finally, for the HH subband, an inter-band prediction is used besides the intra-band prediction used by the HL and LH subbands. In [5] the Embedded Block Coding with Optimized Truncation of the embedded bit-streams (EBCOT), core coding tool of the JPEG 2000 standard, encodes the sign of wavelet coefficients using context information from the sign of horizontal and vertical neighbor coefficients (North, South, East, West directions). Five contexts are used to model the sign coding stage.

In [6], A. Deever and S. Hemami examine sign coding in detail in the context of an embedded wavelet image coder. The paper shows that a Peak Signal to Noise Ratio (PSNR) improvement up to 0.7 dB is possible when sign entropy coding and a new extrapolation technique based on the mutual information that biorthogonal basis vectors provide to improve the estimation of insignificant coefficients are combined. However, the contribution of sign coding by itself to the PSNR improvement is only up to 0.4 dB.

This work uses a similar context modeling than in [5], adapting the context neighborhood to the LTW encoder [11]. LTW encoder is a non-embedded tree-based wavelet image encoder. As other tree-based wavelet coders, it is based on the construction and efficient coding of wavelet coefficient trees. Nevertheless, it does not use an iterative loop in order to determine the significant coefficients and to assign bits to them. It builds the significance map in only one step by using two symbols for pruning tree branches, and also codes the significant coefficients in one step. Previous studies have verified that there is a strong correlation between the sign of a wavelet coefficient and the signs of their neighbors. This correlation opens the possibility of using a sign predictor in order to improve the image compression process. However, this relationship between signs is not uniform and constant for any image, or even consistent within the same image. To obtain an efficient sign prediction scheme a simulated annealing algorithm was developed to explore the solution space that can not be exhaustively tested due to its huge computational and time costs.

Simulated annealing (SA) is an optimization method that implements an iterative local search in an intelligent way to avoid terminating in local optima through stochastic process [8], and originally has been design to find the minimal cost of the objective function derived from complex nonlinear systems. The technique was first introduced by Kirkpatrick [9]. The method is an adaptation of the Metropolis-Hastings algorithm [10] to generate sample states of a thermodynamic system.

This paper explores the convenience of employing simulated annealing to efficiently predict the wavelet coefficient signs based on the correlation found in a given neighborhood set (context). If simulated annealing algorithm helps to define a good wavelet sign predictor, then, instead of coding the sign, encoding the result of the prediction (i.e success or failure). A binary entropy encoder will be able to get significant compression rates if the sign prediction is really good. The proposed simulated annealing algorithm maximizes the cost function in order to obtain the best successful sign predictions, taking into account the given neighborhood.

In order to test the impact of the sign coding module in the behavior of an image wavelet encoder, employs the non-embedded wavelet-based encoder LTW proposed by J. Oliver in [11], to perform the experiments and to help us to determine the advantages of sign compression and to quantify the bit-rate savings.

The remainder of the paper is organized as follows: Section 2 describes the sign coding approximation used and the simulated annealing algorithm that gives an optimized sign prediction. Section 3, shown the experimental results of the sign encoding proposal and its impact on the overall performance of the LTW image encoder. Finally, in Section 4 some conclusions are drawn.

2. Wavelet Coefficient Sign Coding

In a Discrete Wavelet Transform applies a low-pass and a high-pass filter over the image and gets four subbands called LL(Low-Low), HL (High-Low), LH (Low-High) and HH (High-High). The LL subband represents the lower frequencies in the image, while the HL, LH and HH represent the vertical, horizontal and diagonal frequency details. Further decompositions could be applied over the LL subband obtaining again four frequency subbands which represent the same frequencies at a different scale. Typically, for image coding applications, five or six wavelet decompositions are applied. Figure 1, shows typical wavelet decomposition on an image for one and two wavelet decompositions.

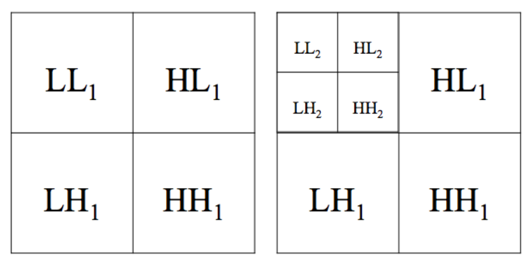

Figure 1. Wavelet descomposition for one and two levels

As Deever explained in [6], given a vertical edge in an HL subband, it is reasonable to expect that neighboring coefficients along the edge have the same sign as the coefficient being coded. This is because vertical sign correlation often remains very high along vertical edges in images. When a low-pass filter is applied along the image columns, it results in a series of similar rows, as elements in a row tend to be very similar to elements directly above or below due to the high vertical correlation. Subsequent high-pass filtering along similar rows is expected to yield vertically correlated transform coefficients.

It is also important to consider correlation across edges, being the nature of the correlation directly affected by the structure of the high pass filter. For the popular Daubechies' 9/7 filter, wavelet coefficient signs are strongly negatively correlated across edges because this filter is very similar to a second derivative of a Gaussian, so, it is expected that wavelet coefficients will change sign as the edge is crossed. Although the discrete wavelet transform involves sub sampling, the sub sampled coefficients remain strongly negatively correlated across edges. In this manner, when a wavelet coefficient is optimally predicted as a function of its across-edge neighbors (e.g. left and right neighbors in HL subbands), the optimal prediction coefficients are negative, indicating an expected sign change. This conclusion is general for any wavelet with a shape similar to a second derivative of a Gaussian.

To estimate sign correlation in a practical way, have been applied a 6-level DWT decomposition of the source image. As a first approach and taking into account that the sign neighborhood correlation depends on the subband type (HL, LH, HH), have been defined three different neighborhoods, one for each subband type. So, the neighborhood of HL subband is composed by the coefficients located at N (North), NN (North of North) and W (West). Taking into account symmetry, for the LH subband, those neighbors are W, WW, and N. For the HH subband they are N, W, and NW, exploiting the correlation along and across the diagonal edges. So, for each neighborhood there are three neighbors whose sign value may be positive (+), negative (-) or zero (*). This leads to a maximum of 27 different Neighborhood Sign Patterns (NSP) for each subband type.

The selection neighborhood criterion has been chosen as an example, being available other configurations that will be more appropriate or not depending on the encoder features (i.e. coefficient scanning order, bit-plane coding or one pass coding, etc). Figure 2, shows the neighborhood to be analyzed when coding the current coefficient ‘x’ for each subband type. Using the previously mentioned neighborhood for each subband type, a simulated annealing algorithm (SA) has been developed in order to find an accurate sign predictor.

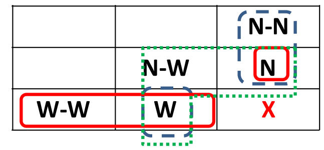

Figure 2. Neighborhood selection criteria for each subband type. Solid line (red) for LH subband, dotted line (green) for HH subband and slashed line (blue) for HL subband

2.1. Simulated Annealing Algorithm for Wavelet sign Prediction

The goal of the proposed algorithm is to find a sign prediction that is accurate for each Neighborhood Sign Pattern (NSP_k, $K=0\ldots\ldots3^3\text{-}1$). Although for a particular NSP_k the current coefficient sign is not always positive (+) or negative (-), this can be not significant (*), it is posible to determinate the probability which would help the optimization algorithm to establish the apropiate prediction (Table 1). But, the problem is still more complex, because a sign prediction for a neighbor sign pattern could fit well for an image and not for others. Therefore, the idea is to find a sign pattern prediction of the wavelet coefficients that better fit for a representative set of images.

The SA algorithm to compress the sign of wavelet coefficients note that if the number of neighbors used to analyze the sign correlation grows or when is a great number of images to be used to the analysis, the search space is excessively wide. Therefore, it is not intuitive to find a way of combining the predictions obtained for several images. The SA algorithm is based on a stochastic method for deciding to accept or not a solution, whether good or bad, based on a probability of acceptance to involve cost-evaluation of the objective function, with this, SA is able to avoid local optimum. The acceptance probability is based in Boltzmann function [8].

Figure 3, presents the SA pseudo code for wavelet sign prediction. First of all defined an initial solution *(s)*, which is formed by a data structure containing an integer element for the cost of the solution and a prediction vector to containing a sign prediction for each 3^3 NSP, then each NSP sign prediction is randomly initialized as a positive or negative, subsequently evaluated the quality of the solution using the objective function (Eq. 1) where N,M are the image dimensions, $S^{\wedge}C_{i,j}\ [k]$ is the sign prediction for NSP(k) and $SC_{i,j}$ is the sign of wavelet coefficient $C_{i,j}$.

$$\sum_{i=0,j=0}^{N,M} \sum_{k=1}^{3^3 NSP} \hat{S}C_{i,j}[k].SC_{i,j} \ \forall \ image \qquad (1)$$

```
function SignPrediction (SubbandType,ImageFiles)
  Initialize(s,T0,Tf)
  EvaluateCostFunction(SubbbandType,ImageFiles,s)
  Repeat
    s' = GenerateRandomSolution(s)
    EvaluateCostFunction(SubbbandType,ImageFiles,s')
    if (f(s') > f(s))
      s = s'
    else
      ProbabilityAcceptance = exp^( (f(s') > f(s) / T0 )
      α = Random [0,1)
      if (α < ProbabilityAcceptance )
        s = s'
      End
    //Temperature cooling
    T0 = T0 * β
    until T0 ≤ Tf
End of function
```

Figure 3. Simulated annealing for the sign prediction

Then, during cooling process, SA algorithm attempts to replace the current solution *(s)* by a new solution *(s')* in which one of the NSP sign prediction is changed. The new solution *(s')* will be accepted if it´s cost of the objective function *f(s')* is better than cost of the current solution *f(s)* one. Also, in case the cost of the new solution *f(s') < f(s)*, the quality of the solution *s'* is checked with the Boltzmann function, for possible acceptance, despite being a bad solution. Remember that SA algorithm minimizes cost-function originally, to solve this problem the following adjustments were made to maximize. Eliminates the negative sign of the exponential and the inequality that verifies the cost difference is reversed in Metropolis cycle, to compare if *f(s') > f(s).* When the algorithm finishes, a quality prediction for each NSP is provided. This prediction is not known if the global optimum for the problem is.

The tuning of the parameters of SA algorithm is obtained by performing sensitive analysis based on [14]. The parameters for sign prediction of the wavelet coefficients are: cooling control parameter β = 0.965, the initial temperature parameter T_0 = 5, the stop criterion or frozen known as final temperature T_f = 2 and Markov chain length is 27.

After running the SA algorithm for each subband type and the representative image set, is obtained a solution containing the prediction of the current coefficient sign $S^{\wedge}C_{i,j}$ *[k]*, that will be shared for both encoder and decoder. In table 1 shown the sign prediction table for the HL subband. The next step is to encode the sign prediction found for each subband. The will be lighter image with a lower number of bits maintaining image quality.

NSP(k)	N	NN	W	Prediction $(S\hat{C}_{i,j}[k])$
0	*	*	*	-
…				
13	+	+	+	+
14	+	+	-	+
…				
26	-	-	-	+

Table 1. Sign prediction for HL subband for some NSPs

Figure 4, shows the integration of the prediction vector for sign prediction obtained for the SA algorithm within the encoder. The original image is applied the discrete wavelet transform with a decomposition level, subsequently applies quantification process to reduce the values of the discrete wavelet transform matrix divided by any value. In this way reduces the magnitude of the wavelet coefficients. Then a sign matrix is created with symbols for the coefficients sign using 0 for positive sign (+), 1 for negative sign (-) and 2 for non-significant (*). So, SA algorithm works with the sign matrix to find the sign prediction for each NSP and the encoder is the sum of the compression of the magnitude and the sign compression to obtain the bit stream ratio.

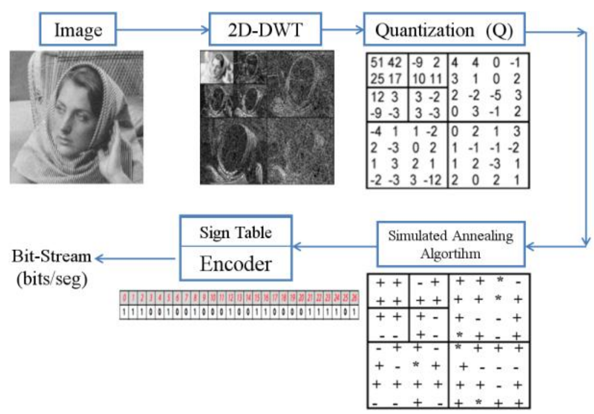

Figure 4. Integration of sign in the encoder

3. Performance Evaluation

This section analyzes the behavior of the sign coding proposal. After developing and properly tuning the SA algorithm, the SA algorithm runs over Kodak test image set [17] in order to obtain the sign prediction tables for HL, LH and HH subbands. Then, these tables (sign prediction for each NSP) were included in the encoding and decoding modules of the reference wavelet image encoder, LTW [11]. The new encoder version including the sign coding module is called S-LTW [12].

In general, to perform the evaluation of sign coding proposal includes results from the original LTW, JPEG2000 (Jasper 1.701.0), and SPIHT (Spiht 8.01) encoders. The test images used in the evaluation were: Lena (512x512), figure 5, Barbara (512x512), figure 6, and Bike (2560x2048), figure 7.

Figure 5. Lena

Figure 6. Barbara

Figure 7. Bike

Figure 8, shows the relative compression gain with respect to the original LTW due only to the sign coding capability for several test images. The maximum sign compression gain is up to 17.35% for Barbara image at 1 bpp and 9% on average for all tested images, being the improvement greater at low compression rates and for high textured images. This behavior is mainly due to the higher number of significant coefficients in the neighborhood that lead to increase the sign prediction success rate. On the other side, the lowest compression gain is achieved on low textured images like Lena.

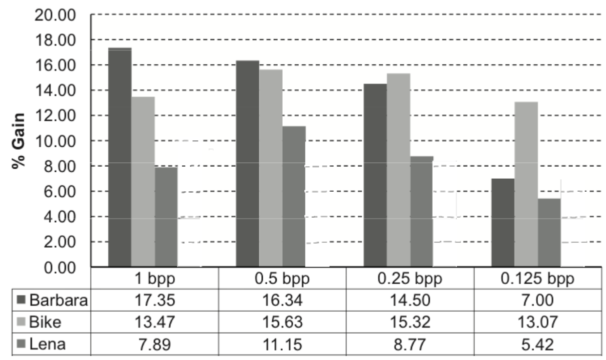

Figure 8. Sign compression performance at different bit-rates.

In [12], authors performed estimation of the bit-rate savings for the SPIHT encoder if the sign coding proposal is applied. They show that up to 9,482 bits could be saved for Barbara image and up to 184,711 for Bike image.

Figure 9, shows the R/D improvement when comparing original LTW versus JPEG2000/SPIHT and S-LTW versus JPEG2000/SPIHT. As shown, there is an increase in the PSNR difference between SPIHT and the new S-LTW encoder, and regarding JPEG2000, there can be observed that S-LTW has a minor loss in PSNR than original LTW. This behavior is similar for the rest of tested images and the use of our sign coding proposal represents, in general, a PSNR increase of 0.25 dB.

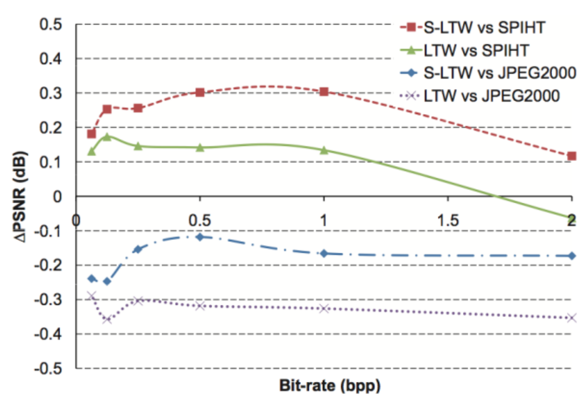

Figure 9. PSNR-Gain for Bike image

Finally, remark that this improvement is exclusively due to our sign coding proposal, being unaltered the rest of the encoder modules.

4. Conclusions

The simulated annealing optimization algorithm that is able to find a good sign predictor for wavelet coefficients sign. The sign prediction result (success or failure) will be highly compacted in the final bit-stream using an entropy encoder. Have been Included the sign prediction tables provided by the SA algorithm (one for each subband type) using a three-neighborhood configuration into a non-embedded encoder. The new encoder version (S-LTW) exhibits better R/D performance (up to 0.25 dB), or in terms of bit-stream, it is able to reduce it up to 17% the sign information for the same quality level.

5. References

[1] González, R. C., Woods R. E., *Tratamiento digital de imágenes*, Addison Wesley/Díaz de Santos, 1996.

[2] Antonini, M., Barlaud, M., Mathieu, P., & Daubechies, I., "Image Coding Using Wavelet Transform", *IEEE transaction on Image Processing, 1992, pp.* 205–220.

[3] Shapiro, J. M., "A Fast Technique for Identifying Zero Trees in the EZW Algorithm", *Proc. IEEE Int. Conf. Acoustic, Speech, Signal Processing*, vol. 3, 1996, pp. 1455-1458.

[4] Wu, X., "High-Order Context Modeling and Embedded Conditional Entropy Coding of Wavelet Coefficients for Image Compression", *Proc. Of 31 St. Asilomar Conf. On Signals, Systems, and Computers*, 1997, pp. 1378-1382.

[5] Taubman, D., "High Performance Scalable Image Compression with EBCOT", *IEEE Transactions on Image Processing*, vol. 9, no. 7, 2000, pp. 1158-1170.

[6] Deever, A, Hemami, S.S., "What's your sign?: Efficient Sign Coding for Embedded Wavelet Image Coding", *Proc. IEEE Data Compression Conf.*, Snowbird, UT, 2000, pp. 273-282.

[7] Schawrtz, E. L., Ahmad, Z., Boilek, M., "CREW: Compression with Reversible Embedded Wavelets", *In Proc. SPIE*, 1995, pp. 212-221.

[8] E. Aarts and J. Korts, *Simulated Annealing and Boltzmann Machines*, John Wiley, New York, 1989.

[9] Vecchi, M. P. and Kirkpatrick, S., "Global wiring by simulated Annealing", *IEEE Trans. Computer-Aided Design*, vol. CAD-2, 1983, pp. 215-222.

[10] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N. , Teller, A. H. and Teller, E., "Equations of State Calculations by Fast Computing Machines", *The Journal of Chemical Physics*, vol. 21, no. 6, 1953, pp. 1087-1092.

[11] Oliver, J. and Malumbres, M. P., "Low-Complexity Multiresolution Image Compression using Wavelet Lower Trees", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 11, 2006, pp. 1437-1444.

[12] López, O., Martínez, M. , Piñol, P., Malumbres, M. P., Oliver, J., **"**E-LTW: an Enhanced LTW Encoder with Sign Coding and Precise Rate Control", *ICIP'09 Proceedings of the 16th IEEE international conference on Image processing*, 2009, pp. 2821-2824.

[13] Otoniel M. López Granado, *Fast and Efficient Coding Tools for Digital Image and Video Signals*, Tesis Doctoral, Alicante España, 2010.

[14] Cruz-Chávez, Marco Antonio, Martínez-Oropeza, Alina, Serna Barquera, Sergio A., "Neighborhood Hybrid Structure for Discrete Optimization Problems", *Electronics, Robotics and Automotive Mechanics Conference,* CERMA 2010, IEEE-Computer Society, México, 2010, pp. 108-113.

[15] Said, A. and Pearlman, W., "A new, Fast and Efficient Image Codec based on Set Partitioning in Hierarchical Trees", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, 1996, pp. 243–250.

[16] R. García Gutiérrez, *Implementación de un algoritmo Genético para la Búsqueda de una Solución Óptima o Subóptima, Asociada a la Codificación del Signo de los Coeficientes de la Transformada Wavelet 2D,* Trabajo para obtención del Diploma de Estudios Avanzados por la Universidad Miguel Hernández dentro del programa de doctorado Tecnologías Industriales y de Telecomunicación, del Departamento de Ciencia de Materiales, Óptica y Tecnología Electrónica, 2010

[17]http://www.kodak.com/digitalImaging/samples/classic.shtml