DISTRIBUTED SOURCE CODING OF STILL IMAGES

Çagatay Dikici, Radhouane Guermazi, Khalid Idrissi, and Atilla Baskurt

INSA de Lyon, Laboratoire d'InfoRmatique en Images et Systèmes d'information,
LIRIS, UMR 5205 CNRS, France
{cdikici,rguermaz,kidrissi,abaskurt}@liris.cnrs.fr
http://liris.cnrs.fr

ABSTRACT

We propose a compression scheme for still images, by exploiting the theory of Distributed Coding of correlated multi-sources. Two corrupted versions of an image are encoded separately but decoding jointly. Our approach results in twofold. i) use of decomposition of low-pass wavelet coefficients for creating the Side Information, and ii) variable-length coset creation by estimating the bit-rate of the cosets on the encoder using the joint distribution statistics of the original image and the side info. In the case of coding for mobile terminals, the proposed codec exploits the channel coding principles in order to have a simple encoder with a low transmission rate and high PSNR. Experimental results are given for lossy encoding case.

1. INTRODUCTION

According to the Slepian-Wolf Theorem, the achievable lossless compression rate of two independent sources X_1 and X_2 are same even they encoded separately but decoding jointly [1]. The theorem was extended to the lossy case by Wyner-Ziv [2], while the input source X_1 is lossy encoded, the second source X_2 called Side Info is available lossless at the decoder. The rate distortion function of X_1 is same even the Side Information is known only at the decoder. However in the case of both X_1 and X_2 are lossy encoded, the rate distortion limits have not been solved yet [3]. This paper focuses on the Rate-Distortion performance if both X_1 and X_2 are lossy encoded. The systems that exploit the distributed source coding are presented recently [4, 5]. The duality between channel coding and the distributed source coding is used for the separate compression of the correlated sources. One application of the distributed source coding is video encoding for low-power handhelds. In [6], an inter-frame encoder but intra-frame decoder is proposed by encoding the key frames as Side Info and the P frames are encoded only with a syndrome based method such that the correlation between the Side Info frames and syndrome based coded frames are exploited only at the decoder. We realized that this syndrome based approach can also be used within the key frames. Hence we propose a distributed source coding approach within a single frame.

The application of distributed source coding techniques on still images are not trivial, because the image should be divided into two sources X_1, X_2 which are encoded separately. In this paper, the low-pass component of the discrete wavelet composition of the image is used as X_2. For X_1, a modulo based binning that has error correcting capabilities on edge boundaries is used. Instead of classical source encoding of X_1, the pixel values are grouped into bins based on a modulo operation, and decoder finds the value of the syndrome that is closest to the X_2. We introduce an efficient distributed coding technique for still images, using discrete wavelet transformation as the side information and modulo generation as the generation of the cosets. Consider the sensor communication [7] in Fig. 1. The input sensors observe the noisy version of the true data X, which are X_1 and X_2 respectively. After the separately encoding of the two noisy observation, the central receiver decodes the two sources jointly. The inputs are encoded with R_1 and R_2 such that the total rate should satisfy the conditions $R_1 \geq H(X_1|X_2)$, $R_2 \geq H(X_2|X_1)$ and $R_1 + R_2 \geq H(X_1, X_2)$, where $H(X_i|X_j)$ is the conditional entropy of X_i given X_j and $H(X_1, X_2)$ is the joint entropy of X_1 and X_2.

In Section 2, the architecture of the coding system is presented. Section 3 contains the details of the algorithm such as creation of the Side Info and cosets, estimation of the parameters, and iterative coding. The experimental setup and the results are given in Section 4.

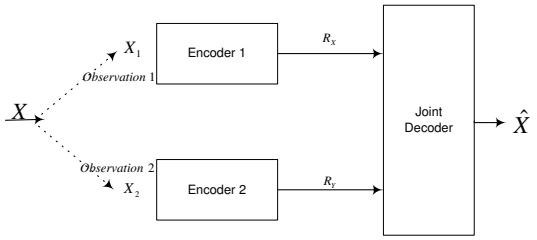

Figure 1: Distributed Source Coding Scheme of the proposed method. Two noisy version of the image are encoded independently and decoded jointly.

2. SYSTEM ARCHITECTURE

The proposed encoding and decoding system for still images, exploiting Distributed Source Coding can be seen in Fig. 2. A similar system is presented in [8] as a syndrome based distributed coding of image. Note, however, that the proposed system does not use the true pixel values for encoding and decoding of the Side Information (SI). Instead, the low-pass frequency components of the image are utilized. Moreover, we proposed a variable-length coset creation in order to improve the Peak Signal to Noise Ratio (PSNR) of the decoded image, especially on the edge contours.

Let $I(M,N)$ is an $M \times N$ gray level image matrix which has integer pixel values within the range of 0 and 255. $\vec{I}$ is a vector version of the image length MN. The $\vec{I}$ is constructed

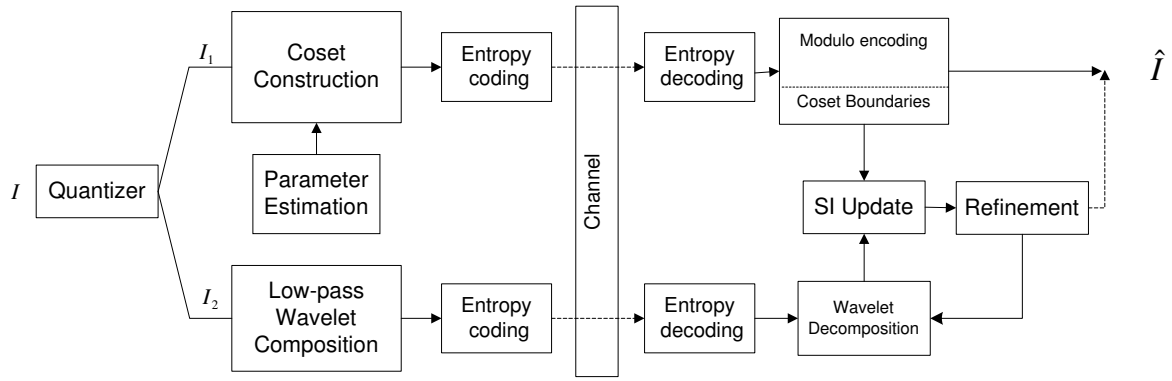

Figure 2: Encoder and decoder structure. Two sources are compressed using modulo coding and wavelet composition, and decoded iteratively.

by a snake scanning of the image matrix as

$$\vec{I}(z) = \begin{cases} \text{I(x,y)} & x \text{ is odd} \\ \text{I(x,N+1-y)} & \text{otherwise} \end{cases} \quad (1)$$

Where $z = N \times x + y$, that z, x, and y are all integer numbers. Hence the discontinuity of the image vector $\vec{I}$ is reduced by the snake scan. In the encoding process, a scalar quantizer is followed by the low-pass wavelet decomposition of the resultant image. In our system, the *Enc*. 1 does not have an access to the values of the *Enc*. 2, but we assumed that the joint distribution statistics of the inputs are available at the encoder in order to estimate the bit-rate of the *Enc*. 1. Hence the encoding parameters of the modulo coding are calculated on the encoder. The drawback of this process is it adds some complexity on the encoder, but the error performance will be superior to that of the estimation on the decoder. For the coset construction process, the possible input values are grouped into bins and the corresponding index of the input (syndrome) is sent to the decoder. Then the decoder predicts the output such that the values of the received syndrome that is closest to the SI. The proposed decoding scheme contains an iterative update of the SI by using the received wavelet decomposition and the correct information coded within the cosets.

3. DISTRIBUTED SOURCE CODING

3.1 Side Information

Side Information(SI) is known as the information encoded parallel with the syndromes, and be used at the decoder in order to estimate the $\hat{I}$ with the help of received syndromes. In the proposed scheme, the information encoded as parallel will be refined at the decoder by using the extra information coded within the cosets. Hence we defined the encoding process as Pre-SI. The quantized version of the image I_q is composed into its wavelet coefficients. In our case, a 2 level of $7/9$ filter set is used [9]. However, note that in each level of wavelet composition, only the lowest part of the coefficients are calculated. The rest of the composition coefficients need not to be calculated, hence reduces the computational complexity of the encoder. The visualization of the SI that is computed in the encoder can be seen in Fig. 3, such that the number of Low-Low coefficients of the Pre-SI (I_{LL}^{ζ}) is only $1/16$th of the size of the original image. The values of I_{LL}^{ζ} is then encoded by arithmetic coding in order to reduce the cost of the transmission rate R_2. In the decoding process of the Pre-SI, an image I_{Pre-SI} is constructed with the same size of the original image by using the transmitted coefficients I_{LL}^{ζ} and the rest of the coefficients setting to 0.

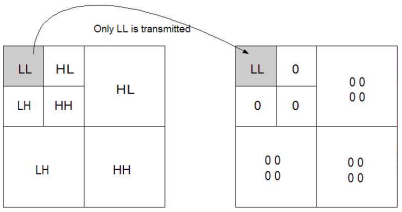

Figure 3: Construction of the Side Information. The Low-Low wavelet composition of the second level is transmitted only. Decoder reconstructs the side information by setting all other coefficients to 0.

3.2 Estimation of the Parameters

In this paper, we modelled the error between the Original Image I, and the Pre-SI I_{Pre-SI} as a laplacian distribution $f(I, I_{Pre-SI}) = \frac{\alpha}{2} e^{-\alpha |I - I_{Pre-SI}|}$ where α can be estimated at the encoder using the residual error between the LL wavelet decomposition of the first level and that of second level. We observed that using the estimate of α instead of calculating the real distance values does not significantly degrade the performance of the system. A distance D_1 is calculated in order to be cover the nth percentage of the pixel values, and we will use this value in order to code our cosets. The values of the cosets are created by using the modulo of $2 \times D_1 + 1$ that will be explained in the next section.

In our experiments we observed that the estimation error will be located mostly on the edge contours, hence we propose a coset creation with variable length. A difference vector is defined such that $\vec{f_{diff}}(\vec{I_q}) = \vec{I_q}(z) - \vec{I_q}(z-1)$. The boundary of each coset is dependent on the difference vector $\vec{f_{diff}}(\vec{I_q})$. We estimate a D_2 such that the m percentage of the values of $\vec{f_{diff}}(\vec{I_q}) < D_2$.

The estimation of the D_1 and D_2 can be visualized in Fig. 4. While the estimation of D_1 depends on the difference between the wavelet decomposition of the first and second low-pass level, the estimation of D_2 depends on the histogram of the absolute difference vector $|\vec{f_{diff}}|$.

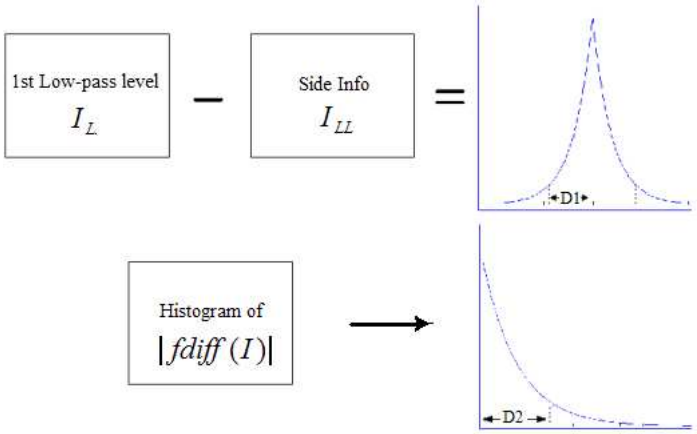

Figure 4: Estimation of the parameters D_1 and D_2

3.3 Coset Creation

In distributed coding approach, the possible input values are grouped into bins and instead of coding the input values, their corresponding bins are transmitted. Receiver chooses the value that is closest to the value of the SI. The proposed coset creation is similar to the [8]. However, in the proposed method, a variable-length cosets are created to improve the quality of the image at the edge boundaries. The creation of the cosets is as follows: The quantized image matrix I_q is vectorized to $\vec{I_p}$ which is defined in Sec. 2.

The parameters D_1 and D_2 that are estimated on Sec.3.2 and $f_{\vec{diff}}(\vec{I_p})$ are used for the coset creation. The vector $\vec{I}_{coset}$ is divided into non-overlapping coset vectors $[\vec{C_1};\vec{C_2};\vec{C_3};..;\vec{C_t}]$ such that whenever the condition $f_{\vec{diff}}(I_q) > D_2$ the current coset vector $\vec{C_i}$ is terminated and the new coset vector $\vec{C}_{i+1}$ is started. The length of a coset $\vec{C_i}$ may vary from 1 to MN depending on the $f_{\vec{diff}}(I_q)$ vector and the value of D_2.

Let the length of the coset $\vec{C_i}$ is ℓ. The values of the coset $\vec{C_i}$ is created as the first element $\vec{C_i}(1)$ is the gray value of the quantized image corresponding to that position $\vec{I_q}(x)$. And the rest of the values are $\vec{C_i}(n) = mod(f_{\vec{diff}}(I_q(x)), 2 \times D_1 + 1)$ which is the modulo of the difference value with respect to D_1. After creating all the cosets, a unique separator string is added at the end of each coset so that the coset boundaries can be easily detected at the decoder. Hence the total length of the cosets are $(M*N)+t$ where t is the number of cosets created. Finally the values of the cosets are encoded by an entropy coder and sent through the channel. A sample of coset creation can be seen in Fig. 5 where $D_1 = 10$, $D_2 = 30$.

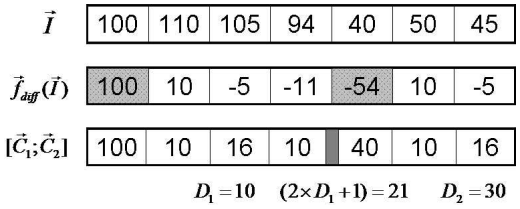

Figure 5: A coset creation example.

Hence $f_{\vec{diff}}(\vec{I})$ is created by $\vec{I}(z) - \vec{I}(z-1)$ and if the absolute difference is not superior then D_2 the modulo $2 \times D_1 +$ 1 value is sent. And if the difference is more than D_2 the real pixel value is coded. In the decoding process, the first elements of each coset are used for updating the SI with a recursive manner. The update process of SI is described below.

3.4 Iterative Decoding

The receiver decodes the Pre-SI I^{ζ}_{LL} and the coset vectors $[\vec{C_1};\vec{C_2};\vec{C_3};..;\vec{C_t}]$. In the first iteration, an estimate of the SI, which is $M \times N$ matrix I_{SI}, is created by using the decoded wavelet coefficients I^{ζ}_{LL} and construct the wavelet decomposition by setting the higher frequencies to 0. Then, the first elements of each coset $C_i(1)$ are substituted on to the corresponding position of I_{SI}. Since the I^{ζ}_{LL} of the updated image must be same as the received I^{ζ}_{LL} of the Pre-SI, all the wavelet coefficients of the updated SI is constructed at the decoder. In this case, the values of LH, HL, and HH could be different than 0. On the new SI, the LL coefficients are replaced by the received I^{ζ}_{LL} of the Pre-SI, and inverse decomposition of the updated SI is calculated. By using the updated SI, the process will continue by replacing the first elements of the cosets $C_i(1)$ to the new SI.

According to our experimental results, 3 iterations on average, the PSNR of the decoded image $\hat{I}$ is improved 1.5DB and will be constant with the consecutive iterations. Thus we chose the number of refinement of the SI as 3. The improvement of the visual quality of the decoded image as the result of iteratively refinement of the Side info is given in Fig. 8.

4. EXPERIMENTAL RESULTS

The proposed algorithm is applied on the image 'Lena'. In our experimental setup, we examine the effects of the quantization and the selection of the encoder parameters D_1 and D_2 respectively. The image Lena is processed as the following steps:

- The input image is first linearly quantized at 256, 128, 64, 32 and 16 levels respectively. Please recall that 256 level quantization corresponds to lossless quantization because of the input image pixels are 8 bits depth.
- For each quantized image, the parameter D_1 is calculated such that Nth percentage of the pixel differences are less than D_1. For our experiments the percentage values for calculating the D_1 are 60%, 70%, 80%, 85%, 90%, 95%.
- For each quantized image, the second encoding parameter D_2 is calculated as described in Sec. 3.2. The D_2 values are estimated according to the 80%, 85%, 90%, 95%, 98%, 99%, 100% of the difference pixel values $f_{\vec{diff}}(\vec{I})$ are within the range of D_2.
- For each D_1, D_2, and the quantization level, the proposed distributed image coding algorithm is applied. the Peak Signal to Noise Ratio(PSNR) and the compression ratio of the output are computed.

The PSNR versus compression ratio of the image which is calculated without quantization can be shown in Fig. 6. In this figure, the dashed lines have the same D_1 value but the values of D_2 are ranging from 80% to 100% of the pixels. As the percentage of $f_{\vec{diff}}(\vec{I_q}) < D_2$ decreases, the number of cosets that are encoded are increased. Hence while the PSNR values are improved, the compression ratio will be less in that case. Moreover each dashed lines has unique D_1

values. While the rightmost line corresponds to the 60% of the pixels, the PSNR values are better if we increase the percentage value. Please note that 95% line cuts the other lines for D_2 covers the 90%. iteration versus PSNR.

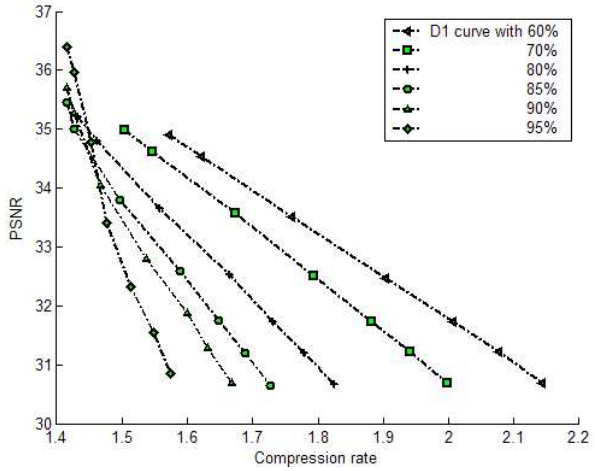

Figure 6: RD Curve Without quantizer.

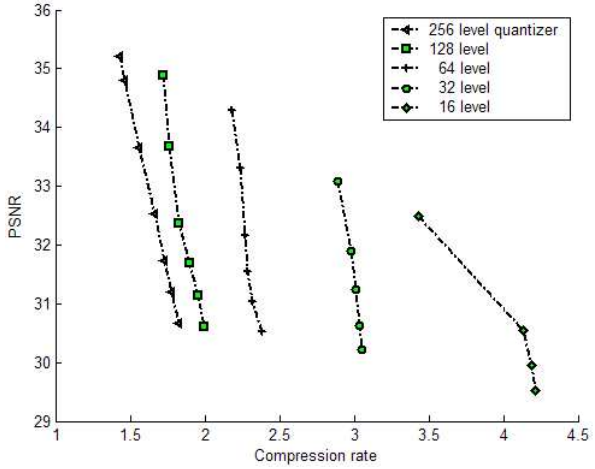

Figure 7: RD Curve comparing the level of quantizers.

The effects of the linear scalar quantization of the input image can be shown in Fig. 7. In order to plot this figure, we used a fixed D_1 value which the percentage of the pixels are less than 80%. The values of D_2 is computed as described above. Each dashed line corresponds to a quantization level. While the left-most line is the non-quantized case, the following lines corresponds to the next quantization level. As you can see from the figure, the quantization process decreases the PSNR value a little bit, while it has great advantage on compression performance.

Moreover the effects of the refinement of the Side Info at the decoder can be shown in Fig. 8. In this figure, a 130×160 pixel subset of the outputs with a compression rate 1.43 are given. The leftmost picture is the pre SI that is reconstructed such that all of the wavelet coefficients except the ones that are LL received by the encoder are set to 0. The decoding of the cosets based on the pre SI is given in the center. And the rightmost one is the output of the decision based on the refinement of the Pre SI after 3 iterations. On each iteration the true values of the first elements of the cosets are substituted on the same location of the Pre-SI, and a wavelet decomposition and reconstruction is applied such that the LL decomposition of the Pre-SI is always preserved same. The quality improvement on the edges such as the face, shoulder and hair region can be seen. Moreover the PSNR values of these three images are 28.9DB 34.16DB and 34.97DB respectively.

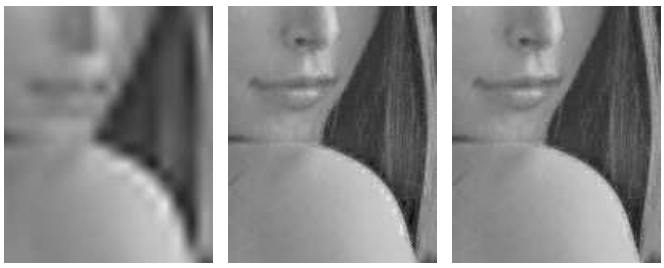

Figure 8: Left: Side Info at the receiver; Center: First iteration output of the decoded image; Right: Final decoding output after 3 iterations on Pre-Side Info update.

5. CONCLUSIONS

This study shows the feasibility of such a multi-source coding scheme for still images. It is clear that the performance of the proposed algorithm is not compatible to those obtained with actual coding schemes. The Rate Distortion performance can be increased and several ideas have been given below. Furthermore, that the real interest of this kind of multi-source applications can be initialized in new functionalities for multimedia services and uses; transmission on ad-hoc networks, combining the compression with image indexing such that the Side Information can represent local and global descriptors in order to enrich the transmitted data. In this work, the performance of the proposed image coding algorithm is examined, which the Low-pass wavelet coefficients and modulo binning of the image are encoded on parallel, and decoded jointly at the receiver. Based on the edge features of the image, the algorithm use more bits on these regions for recovering the high frequencies. Moreover, since the algorithm has a parameter estimation unit on the encoder, it enables us to determine the quality and the rate of the decoded image. This algorithm can also be used for coding a single frame within a video data that exploits distributed source coding idea.

For the future work, the performance of the system can e increased as follows: For the Pre-SI, the second level LL composition of the image is encoded however the effects of sending the third or fourth level instead can be observed. Moreover, for preserving the simplicity of the encoder, the coset creation depends only on the horizontal line features of the image, hence a block based approach can also be implemented for the future work.

REFERENCES

[1] J. D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources", *IEEE Transactions on Information Theory*, vol. IT-19, pp. 471–480, July 1973.

[2] A. D. Wyner and J. Ziv, "The Rate-Distortion Function for Source Coding with Side Information at the Decoder", *IEEE Transactions on Information Theory*, vol. IT-22, no. 1, pp. 110, Jan. 1976.

[3] T. M. Cover and J. A. Thomas, “Elements of Information Theory”, *Wiley*, New York, 1991.

[4] R. Puri and K. Ramchandran, “PRISM: A new robust video coding architecture based on distributed compression principles”, *Proc. Allerton Conference on Communication, Control, and Computing*, Allerton, Illinois, Oct. 2002.

[5] A. Aaron and B. Girod, “Compression with side information using turbo codes”, *Proc. IEEE Data Compression Conf.*, pp. 252–261, 2002.

[6] A. Aaron, E. Setton and B. Girod,“Towards practical Wyner–Ziv coding of video”, *Proc. IEEE International Conference on Image Processing*, Barcelona, Sept. 2003.

[7] T.J. Flynn and R.M. Gray, “Encoding of correlated observations”, *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 773–787, Nov. 1987.

[8] K. Ozonat, “Lossless distributed source coding for highly correlated still images”, *Electrical Eng. Dept., Stanford Univ., Stanford, CA, Tech. Rep.*, 2000.

[9] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, “Image coding using wavelet transform”, *IEEE Transactions on Image Processing*, Vol. pp. 205-220, 1992.