# Distributed single source coding with side information

J. E. Vila-Forcen, O. Koval and S. Voloshynovskiy*

University of Geneva, Department of Computer Science, 24 rue Général-Dufour,
CH 1211, Geneva, Switzerland

## ABSTRACT

In the paper we advocate image compression technique in the scope of distributed source coding framework. The novelty of the proposed approach is twofold: classical image compression is considered from the positions of source coding with side information and, contrarily to the existing scenarios, where side information is given explicitly, side information is created based on deterministic approximation of local image features. We consider an image in the transform domain as a realization of a source with a bounded codebook of symbols where each symbol represents a particular edge shape. The codebook is image independent and plays the role of auxiliary source. Due to the partial availability of side information at both encoder and decoder we treat our problem as a modification of Berger-Flynn-Gray problem and investigate a possible gain over the solutions when side information is either unavailable or available only at decoder. Finally, we present a practical compression algorithm for passport photo images based on our concept that demonstrates the superior performance in very low bit rate regime.

Keywords: image compression, wavelets, distributed source coding, source coding with side information, image enhancement, rate-distortion function, Markov chain, random coding argument, Slepian-Wolf coding, Wyner-Ziv coding, Berger-Flynn-Gray coding, Estimation-Quantization.

## 1. INTRODUCTION

The urgent demand of efficient image representation is recognized by the industry and research community. The necessity in such a technology is highly increased due to the novel requirements to the security documents such as passports, ID cards, and visas as well as recent extended functionalities of wireless communication devices. The document, ticket, entry pass or even device personalization are often required in many application protocols. Therefore, more than often, classical compression techniques developed for generic applications are not suitable for these purposes.

Wavelet transform [2, 3] based lossy image compression techniques [4-7] are proved to be the most efficient from the rate-distortion point of view for the rate range of 0.2÷1 bit per pixel (bpp). Superior performance of this class of algorithms is justified by decorrelation and energy compaction properties of wavelet transform and by efficient adaptive interband and intraband models that describe the data in wavelet subbands (zero trees [4], Estimation-Quantization (EQ) [8,9]). Recent results in wavelet-based image compression show that some modest performance improvement (in terms of peak signal to noise ratio (PSNR) up to 0.3 dB) could be achieved taking into account non-orthogonality of the transform [10] or using more complex higher-order context models of wavelet coefficients [11].

During years, a standard benchmark for wavelet-based compression algorithm evaluation was used. It includes several 512×512 greyscale test images (like Lena, Barbara, Goldhill) and verification is performed for the rates 0.2÷1 bpp. In some applications that include person authentication data like photo images or fingerprint images operational conditions might be different. In this case, especially for the case of strong compression (below 0.15 bpp), resulting image quality of state-of-the-art algorithms is not high enough (Fig. 1). Therefore, for this kind of applications more advanced techniques are necessary to satisfy the fidelity constrains.

In this paper we address the problem of enhancement of classical wavelet-based image compression using side information within framework of distributed coding of correlated sources. Recently, it was practically shown that it is possible to achieve significant performance gain when side information is available even only at the decoder while the encoder has not access to the side information [12]. Using side information from auxiliary analog Additive White Gaussian Noise (AWGN) channel in form of the noisy copy of the input image at the decoder, it was reported the PSNR enhancement in the range of 1÷2 dB depending on the test image and compression rate. It could be noted that the performance of this scheme strongly depends on the state of the auxiliary channel, which should be known in advance at the encoding stage. Moreover, it is assumed that the noisy (analog) copy of the original image should be directly available at the decoder. This situation is typical for distributed coding in remote sensing applications or can be simulated like in the case of simulcast of analog and digital television. In the

Further author information: (Send correspondence to S. Voloshynovskiy): E-mail: svolos@cui.unige.ch

case of single source compression, the side information is not directly available at the decoder.

a) b) c)

**Figure 1.** Results of compression of the (a) 256×256 8-bit test image "Slava" with rate 0.073 bits per pixel (bpp) using (b) JPEG 2000 standard software (PSNR is 25.33 dB); (c) state-of-the-art compression EQ (PSNR is 25.96 dB).

Therefore, the main goal of this paper consists in development of a concept of single source compression within distributed coding framework using "virtually" created side information. This concept is based on the accurate approximation of a source data using structured codebook, which is shared by the encoder and the decoder, and communication of the residual approximation term within the classical wavelet-based compression paradigm.

The paper is organized as following. In Section II fundamentals of source coding with side information are presented. In Section III approach for single source distributed lossy coding will be introduced. Practical algorithm for very low bit rate compression of passport photo images is developed in Section IV. Section V contains experimental results and Section VI concludes the paper.

**Notation.** Scalar random variables are denoted by capital letters $X$, bold capital letters $\mathbf{X}$ denote vector random variables, letters $x$ and $\mathbf{x}$ are reserved to denote realization of scalar and vector random variables, respectively. The superscript $N$ is used to denote $N$ – length vectors $\mathbf{x} = \{x_1, x_2, ..., x_N\}$ where $i$-th element is denoted $x_i$. $X \sim p_X(x)$ or $X \sim p(x)$ indicate that a random variable $X$ is distributed according to $p_X(x)$. The mathematical expectation of a random variable $X \sim p(x)$ is denoted by $E_X[X]$ or $E[X]$. $H(X), H(X,Y), H(X|Y)$ denote entropy of random variable $X$, joint entropy of random variables $X$ and $Y$, and conditional entropy of random variables $X$ given $Y$. By $I(X;Y)$ and $I(X;Y|Z)$ mutual information between random variables $X$ and $Y$ and conditional mutual information between random variables $X$ and $Y$ given random variable $Z$ are denoted, respectively. Calligraphic font $\mathcal{X}$ is used to indicate sets $X \in \mathcal{X}$ and $|\mathcal{X}|$ indicates a cardinality of set. $R^+$ is used to represent a set of positive real numbers.

## 2. DISTRIBUTED CODING OF CORRELATED SOURCES

### 2.1. Slepian-Wolf encoding

Assume that it is necessary to encode two discrete-alphabet independent identically distributed (i.i.d.) random variables $\mathbf{X}$ and $\mathbf{Y}$ with joint distribution $p_{\mathbf{XY}}(\mathbf{x},\mathbf{y}) = \prod_{k=1}^{N} p_{X_k Y_k}(x_k, y_k)$, where each pair $\{X_k, Y_k\}$ has a joint pdf $p_{X_k Y_k}(x_k, y_k)$. A Slepian-Wolf [13, 14] code allows performing lossless encoding of $\mathbf{X}$ and $\mathbf{Y}$ individually using two separate encoders. Decoding is performed jointly (Fig. 2). Using random coding argument, it was shown that the efficiency of such a code is the same as in case when joint encoding is used. It means that the pair of bit rates of encoders $R_X, R_Y$ of random variables $\mathbf{X}$ and $\mathbf{Y}$ respectively is achievable when the following relationships hold:

$$R_X \geq H(\mathbf{X}|\mathbf{Y}), \tag{1}$$

$$R_Y \geq H(\mathbf{Y}|\mathbf{X}), \tag{2}$$

$$R_X + R_Y \geq H(\mathbf{X},\mathbf{Y}). \tag{3}$$

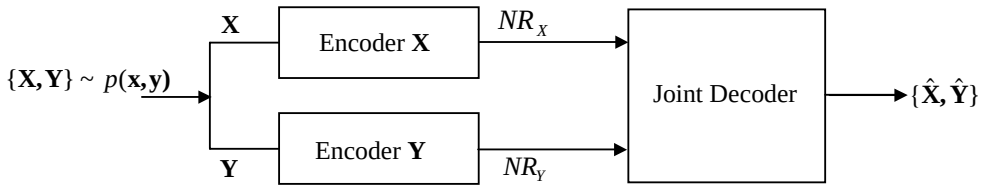



**Figure 2.** Slepian-Wolf coding.

## 2.2. Lossy compression with side information

In lossy compression set-up it is necessary to achieve target coding rate for minimal possible distortions. Depending on the availability of side information several possible scenarios exist [15].

*No side information is available*: Imagine that we need to represent an i.i.d. source sequence $\mathbf{X} \sim p_X(\mathbf{x})$, $\mathbf{X} \in \mathcal{X}^N$, using coding mapping $\mathcal{X}^N \to \{1,2,...,2^{RN}\}$ and decoding mapping $\{1,2,...,2^{RN}\} \to \hat{\mathcal{X}}^N$ with average bit rate $R$ bits per element. Fidelity of representation is evaluated using average distortion $D = \frac{1}{N}\sum_{k=1}^{N} E[d(X_k, \hat{X}_k)]$ where distortion measure $d(x,\hat{x})$ in general is determined as a mapping $\hat{\mathcal{X}}^N \times \mathcal{X}^N \to \mathcal{R}^+$. Due to Shannon [14, 16] it is well known that optimal performance of such compression system (Fig. 3) (the minimal achievable rate for certain distortion level) is determined by rate-distortion function:

$$R(D) = \min_{p(\hat{x}|x):\ \sum_{\hat{x},x} p(x)p(\hat{x}|x)d(x,\hat{x}) \le D} I(X;\hat{X}). \tag{4}$$

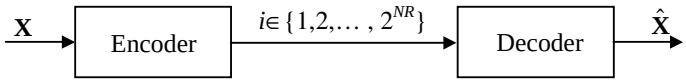



**Figure 3.** Lossy source coding system without side information.

*Side information is available at encoder*: In this case performance limits coincide with the previous case and rate-distortion function could be determined using (4) [17].

*Wyner-Ziv coding*: Fundamental performance limits of source coding systems with side information available only at the decoder (Fig. 4) were established by Wyner and Ziv [14, 18]. Wyner-Ziv problem could be formulated in the following way: given analog side information only at the decoder, what will be the minimum rate *R* necessary to reconstruct source **X** with average distortion less or equal some *D*. By other words, assume that we have a sequence of independent drawings of a pair $\{X_k, Y_k\}$ of dependent random variables, $\mathbf{\{X,Y\}} \sim p(\mathbf{x},\mathbf{y})$, $\mathbf{(X,Y)} \in \mathcal{X}^N \times \mathcal{Y}^N$. Our goal is to construct $R$ bit per element coder $f_E : \mathcal{X}^N \to \{1,2,...,2^{RN}\}$ and joint decoder $f_D : \{1,2,...,2^{RN}\} \times \mathcal{Y}^N \to \hat{\mathcal{X}}^N$ such that average distortions satisfies the constrain: $E[d(\mathbf{X}, f_D(\mathbf{Y}, f_E(\mathbf{X})))] = \sum_{x,\hat{x},u,y} d(x,\hat{x})p(x,y)p(u|x)p(\hat{x}|x,y) \le D$. Using asymptotic properties of random codes it was shown in [18] that the set of achievable rate-distortion pairs of such a coding system will be bounded by Wyner-Ziv rate-distortion function:

$$R(D)_{X|Y}^{WZ} = \min_{p(u|x)\,p(\hat{x}|x,y)} [I(U;X) - I(U;Y)]. \tag{5}$$

Minimization is performed over all $p(u|x)\,p(\hat{x}|x,y)$ and all decoder functions satisfying fidelity constrain and $U$ is an auxiliary random variable such that $|\mathcal{U}| \le |\mathcal{X}| + 1$ and $Y \to X \to U$ form a Markov chain. Hence, (5) could be rewritten as follows:

$$R(D)_{X|Y}^{WZ} = \min_{p(u|x)\,p(\hat{x}|x,y)} I(U;X|Y). \tag{6}$$

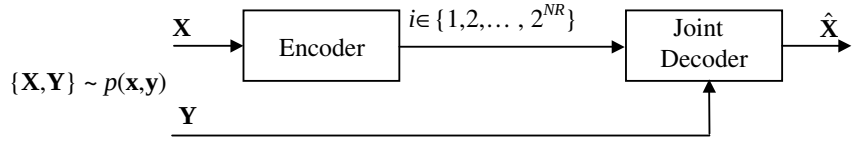



**Figure 4.** Wyner-Ziv coding.

It is worth to note that for the case of zero distortions Wyner-Ziv problem corresponds to Slepian-Wolf problem, i.e.: $R(0)_{X|\mathbf{Y}}^{WZ} = H(X|Y)$.

*Lossy compression of correlated sources*: This problem was investigated by Berger [19] and Flynn and Gray [20] (Fig. 5).

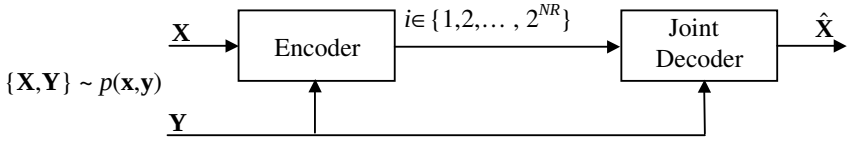



**Figure 5.** Berger-Flynn-Gray coding.

As in the previous case, Berger-Flynn-Gray coding refers to the sequence of pairs $\{\mathbf{X},\mathbf{Y}\} \sim p(\mathbf{x},\mathbf{y})$, $(\mathbf{X},\mathbf{Y}) \in \mathcal{X}^N \times \mathcal{Y}^N$, where $\mathbf{Y}$ is available at the encoder and at the decoder. It is necessary to construct $R$ bit per element joint coder $f_E : \mathcal{X}^N \times \mathcal{Y}^N \to \{1,2,\ldots,2^{RN}\}$ and joint decoder $f_D : \{1,2,\ldots,2^{RN}\} \times \mathcal{Y}^N \to \hat{\mathcal{X}}^N$ such that it will be possible to reconstruct source $\mathbf{X}$ with average distortions satisfying the following inequality: $E[d(\mathbf{X}, f_D(\mathbf{Y}, f_E(\mathbf{X})))] \le D$. Performance limits in this case are determined by conditional rate-distortion function:

$$R(D)_{X|Y} = \min_{p(\hat{x}|x,y)} I(\hat{X};X|Y), \tag{7}$$

where minimization is performed over all $p(\hat{x}|x,y)$ with the fidelity constraint $E[d(\mathbf{X}, f_D(\mathbf{Y}, f_E(\mathbf{X})))] = \sum_{x,\hat{x}} p(x,y)p(\hat{x}|x,y) \le D$.

Comparing rate-distortion performance of different coding scenarios with side information, it should be noted that in general case the following inequalities hold [21]:

$$R(D) \ge R(D)_{X|Y}^{WZ} \ge R(D)_{X|Y}. \tag{8}$$

The last inequality becomes equality, i.e., $R(D)_{X|Y}^{WZ} = R(D)_{X|Y}$, only for the case of Gaussian distribution of the source $\mathbf{X}$ and

mean square error (MSE) distortion measure. For any other pdf performance loss exists in Wyner-Ziv coding. It was shown in [21] that this loss is upper-bounded by 0.5 bit:

$$R(D)_{X|Y}^{WZ} - R(D)_{X|Y} \leq 0.5. \tag{9}$$

Therefore, due to the fact that natural images have highly non-Gaussian statistics, compression of this data using Wyner-Ziv strategy will always leads to the performance loss. Another open issue targeting compression improvement using side information consists in availability of the output of auxiliary channel either only at decoder or at both encoder and decoder in classical set-up. Moreover, it is assumed that global correlation is significant between source data and side information. Therefore the main goal of subsequent sections consists in extension of classical distributed coding set-up to the case of single source coding scenario.

## 3. PRACTICAL APPROACH: DISTRIBUTED SOURCE CODING WITH SIDE INFORMATION OF A SINGLE SOURCE

The block-diagram of practical single source distributed coding system with side information is presented in Fig. 6. The system consists of two main functional parts. The first part includes the main encoder that is working as a classical quantization-based lossy coder with varying rates. The second part includes the block of transition detection and shape approximation that creates some auxiliary image **Y**, as a very close approximation of **X**. Index encoder communicates the parameters of approximation model to the decoder. The shape codebook is shared by both the approximation block and the decoder.

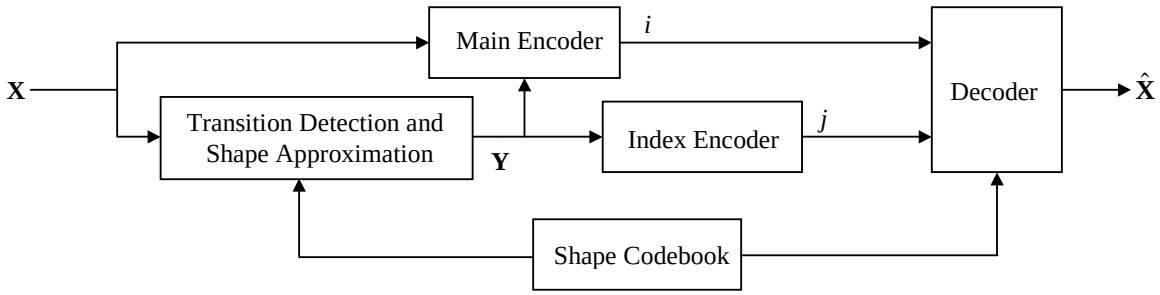



**Figure 6.** Compression system of a single source distributed coding with side information.

The intuition behind our approach is based on the assumption that natural images in the coordinate domain can be represented as a union of several statistically homogeneous regions of different intensity levels or in the non-decimated wavelet transform domain using edge process (EP) model (Fig. 7). This assumption and the EP model have been used in our previous work in image denoising where promising results have been reported [22].

However, contrarily to the image denoising set-up, in the case of lossy wavelet-based image compression we are interested to consider not the behaviour of edge profile along the direction of edge propagation but the different edge shapes. Due to high variability of edge shapes in real images and corresponding complexity of approximation problem, we will utilize a *structured codebook* for shape representation. It means that we will use one or several types of shapes, to construct a codebook where each codeword represents one edge of some magnitude (Fig.8).

The structured codebook of edge shapes can be schematically ordered as cosets where each coset contains the approximation structures of basic shapes, e.g., three basic shapes as it is shown in Fig. 8. More formally, the structured codebook can be represented as in Fig. 9.

Important points about the codebook are: (a) it is image invariant, (b) shapes of codewords could be expressed analytically, for instance, using apparatus of splines, and (c) the dimensionality of codebook is determined by the type of used transform and compression regime. Therefore, a concept of successive refinement construction [23] of the codebook might be used. The

intuition behind this approach could be explained using coarse-fine quantization framework presented in Fig. 10. It means that for the case of low rate compression when we do not have much rate to code shape index, a single shape profile will be used (like a coarse quantizer). In other regimes (medium and high rates) we could improve the fidelity of approximation adding more edge shapes (in this case we could assume that high rate quantization assumption becomes valid).

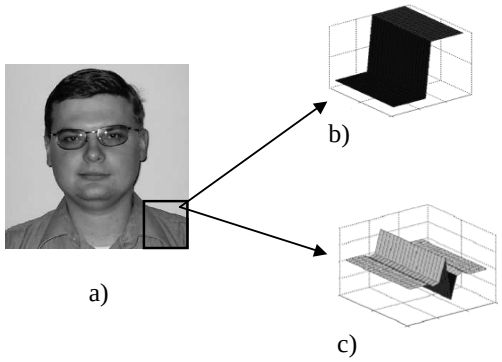



**Figure 7.** (a) Test image "Slava" and its fragment (marked by square): two-region modelling of the fragment (b) in the coordinate domain and (c) in the non-decimated wavelet transform domain.

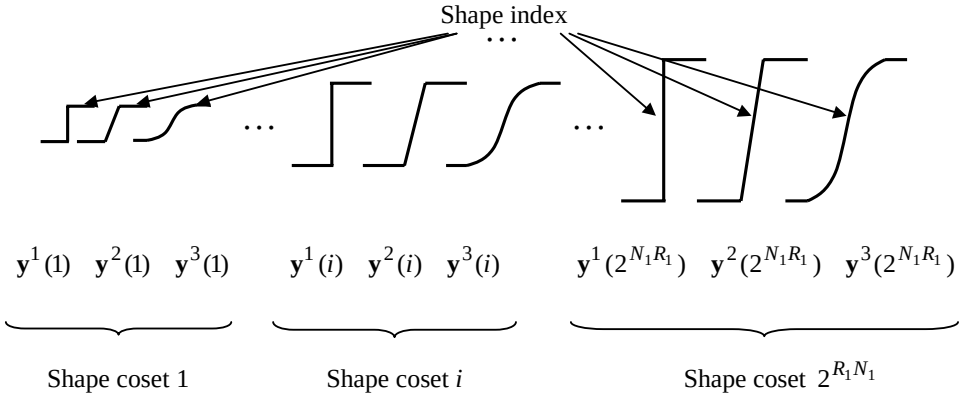



**Figure 8.** Three shape cosets from shape codebook **Y**.

The task of real edge approximation according to the shape codebook can be formulated, for instance, like a classical $L^2$ norm approximation problem:

$$\tilde{\mathbf{y}}^j(i) = \underset{\substack{\{\mathbf{y}^j(i)\},\, 1\le i\le 2^{R_1N_1},\\ 1\le j\le N_2}}{\arg\min} \left\| \mathbf{x} - \mathbf{y}^j(i) \right\|^2, \tag{10}$$

where the minimisation is performed over the whole codebook in each image point, and the coset index $i$ is in the range of $1 \le i \le 2^{R_1N_1}$ and index of edge shape $1 \le j \le N_2$, where $N_2$ is a number of edge shapes in one coset ( $N_2 = 3$ in our case).

It is clear that in the presented set-up computational complexity of approximation will be significant, which can be unacceptable in some real-time application scenarios. To simplify the situation searching space dimensionality might be significantly reduced using techniques that simplify the edge localization. Canny edge detector [24] can be used.

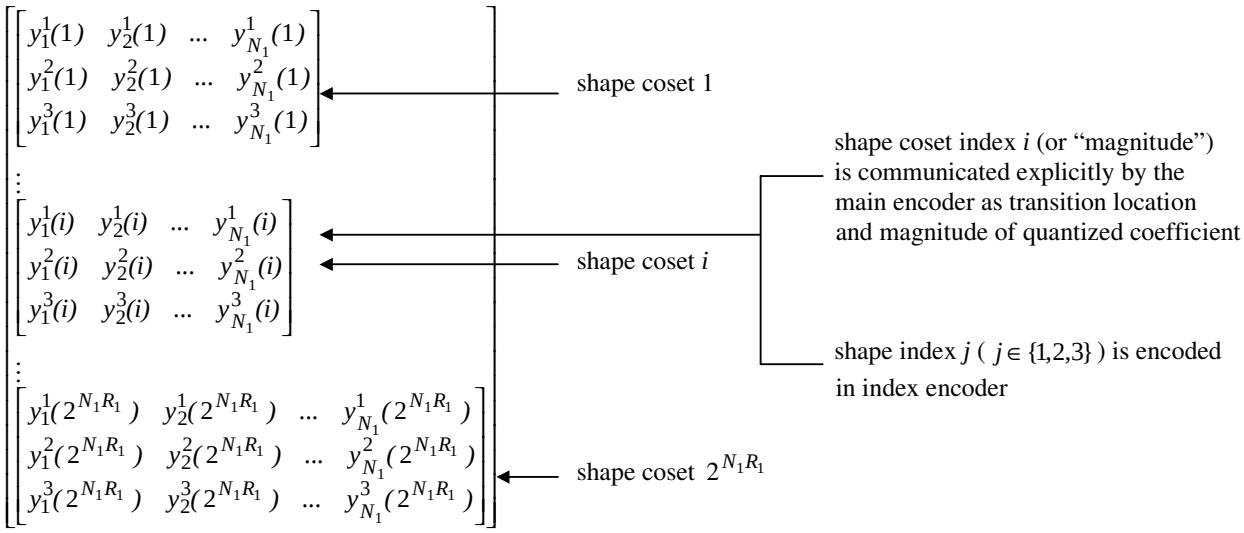



**Figure 9.** Structured codebook: shape coset index *i* is communicated by main encoder and shape index *j* is communicated by index encoder.

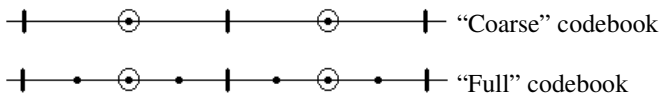



**Figure 10.** Successive refinement codebook construction.

The edge of real image could be considered as a noisy or distorted version of corresponding codeword (edge shape) with respect to the codebook **Y,** i.e., some correlation assumption between original edge and a codeword can be introduced. Therefore, structure of the codebook is similar to the structure of channel coset code [25], meaning that the distance between codewords of equal magnitude (Fig. 8) in the transform domain should be enough to perform correct shape approximation.

The coding strategy can be performed in the distributed manner. In general, the main encoder, working in the approximation quantization regime with the specified rate performs the quantization of the edge and communicates corresponding indices of reconstruction levels to the decoder. This information suffices to determine the shape coset index *i* at the decoder for different compression regimes including even low bit rate regime (besides the case when quantization to zero is performed). The index *j* of edge shape within a coset is communicated by the index encoder or is estimated based on the main encoder output. Having the index of coset and shape index, the decoder looks in coset bin *i* for a $\mathbf{y}^{j}(i)$ and generates the reproduction sequence $\hat{\mathbf{x}} = f_D(\hat{\mathbf{x}}'(i), \tilde{\mathbf{y}}^{j}(i))$, where $\hat{\mathbf{x}}'(i)$ is the data reproduced at the decoder based on index *i* only.

In particular, in the case of high rate compression the main encoder performs high rate (high accuracy) approximation of image edges. It means that the index encoder does not produce any output, i.e., both the edge magnitude and the edge shape could be reconstructed directly from information contained in the main decoder bit stream. Therefore, the role of side information, represented by the codebook, consists in compensation of quantization noise influence.

For middle rates, prediction of the edge magnitude is still possible using main encoder bitstream. However, accuracy of edge approximation for this regime is not high enough to estimate edge shape and the edge shape index should be communicated to the decoder by the index encoder. One can note that in such a way we end up with vector-like edge quantization using off-line designed 1-D edge codebook. The role of side information remains similar to the previous case and targets the compensation of quantization effect.

At low rates, the index encoder is idle due to the restrictions of the target rate. A single codeword, optimal in the MSE sense, is chosen to represent all shapes within given image. In more general case, one can choose a single shape codeword that is the same for all images. This is a valid assumption for the compression of image databases with the same type of images. Contrarily to the above case of middle rates, the decoder operates with a single edge codeword that can be applied to all cases where the edge coefficients are partially preserved in the corresponding subbands. Moreover, the edge reconstruction is even possible when the edge coefficients in some subbands are completely discarded by the dead-zone quantization.

Practical aspects of implementation of the presented framework of single source distributed coding with side information at high and middle rates are out of the scope of this work. The very low bit rates are of particular importance for many practical applications where the traditional compression technologies do not produce satisfactory results. The next section presents a particular scenario in which the proposed framework is applied to the compression of photo images at very low bit rates.

## 4. DISTRIBUTED CODING OF IMAGE WITH SYMMETRIC SIDE INFORMATION: COMPRESSION OF PASSPORT PHOTOS AT VERY LOW BIT RATES

In this section, the case of single source distributed coding system with side information presented in section 3 is discussed for the case of very low bit rate (less than 0.1 bpp) compression of passport photo images. The importance of this task is justified by the urgent necessity to store personal information on the memory restricted media identification documents that includes passports, visas, ID cards, driver licenses and credit cards using digital watermarks, bar-codes or magnetic strips. In this paper we assume that the images of the interest have 8-bit and 256x256 format. As it was shown in Fig. 1, standard compression tools are unable to provide the satisfactory quality results.

The scheme presented in Fig. 6 is used as a basic set-up for this application. As it was discussed, for the case of very low bit rate regime, only one (simple step edge) setup profile is used. Therefore, index encoder is not used in this particular case. Certainly, better performance can be expected if one approximates transitions using complete image codeword (Fig. 8). The price to pay for that is additional $\log_2 N_2$ bits of side information per shape, where $N_2$ is the number of edge shapes within the coset.

In next sections, we discuss in details the particularities of encoding/decoding for very low bit rates.

### 4.1. Transition detection and shape approximation

*Encoding*: On the first step, the detection of positions of the principal edges (the edges with the highest contrast) is performed using Canny edge detector. Due to the fact that detection result is not always precise (some position detection deviation is possible), actual transition detector is performed using correlation detector. The output of detection is used to take into account the nature of passport images meaning that the most important information (human face) is placed in the central part of the photo. Therefore, the rate allocating to the face part in the Main Encoder of the photo is increased according to the region-of-interest compression paradigm. Hence, a region of interest within a rectangle is determined using three additional bytes of side information (Fig. 11, a). Extra four bytes are used to store the mean brightness of the background of the photo image in four quadrants (Fig. 11, b).

a) b)

**Figure 11.** Test image “Slava”: (a) region of interest and (b) background four quadrant splitting.

*Decoding*: For this mode it is assumed that low resolution version of the original data obtained using Main Coder Bitstream is already available. Detection of the coarse positions of main edges is performed on the interpolated image

analogically to the encoder case. To adjust detection results a new concept of "zero-crossing" detection is used.

This concept is based on the fact that in non-decimated wavelet transform domain (algorithm "a trios" [27] is used for its implementation) all the representations of an ideal step edge in the same spatial orientation cross the horizontal axe in the same point that we call "zero-crossing" point. This point coincides with a spatial position of the transition in the coordinate domain. Besides, the magnitudes of principle peaks (maximum and minimum values of the data in the vicinity of transition in non-decimated transform domain) of the components are related pairwise from high to low frequencies with certain fixed ratios (Fig. 12). Therefore, when the position of the "zero-crossing" point is given and at least one of the component peak magnitudes is known from original step edge, it is possible to completely reconstruct the missing data components with no error.

At the targeted very low bit rate compression scenario, the data are severely degraded by quantization. To make zero crossing detection more reliable in this case more levels of overcomplete decomposition can be used. The additional reliability is coming from the fact that the data at very low frequency levels almost do not suffer from quantization. The gain in this case is limited by the information that is still presented at these low frequencies: only the edges propagating in all the subbands in the same orientation could be detected in such a way.

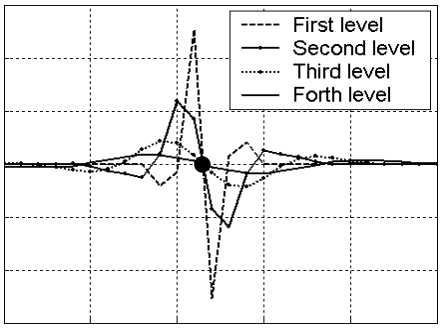



**Figure 12.** Step edge representation in the non-decimated wavelet transform domain.

Besides the principal edges represented by the contour of human face, important information is mainly localized inside the region-of-interest (Fig. 11,a). To improve the performance of the system in this area zero-crossing detection is applied to each point within this region. If "zero-crossing" is detected, magnitudes of the disregarded high frequency components are estimated and localized in the corresponding high frequency non-decimated wavelet transform subbands.

## 4.2. Main encoder

To justify Main Encoder structure we would like to point out that the main gain achieved recently in wavelet-based lossy transform image coding is due to the accuracy of the stochastic image model. It was shown by Ramstad *et. al.* [28] that stationary zero mean i.i.d. Laplacian pdf, approximating the data in wavelet subbands [29], could be without loss according to Kullback-Leibler divergence represented using non-stationary mixture of zero-mean Gaussian components with local variances drawn from exponential distribution:

$$\frac{\lambda}{2}\exp(-\lambda|x|) = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{x^2}{2\sigma^2}\right)\lambda\exp(-\lambda\sigma^2)d\sigma^2, \tag{10}$$

where $\lambda$ is a parameter of Laplacian distribution.

If one assumes that side information (local variances) is symmetrically available at decoder and encoder, the gain in rate-distortion sense coding the Gaussian mixture instead of i.i.d. Laplacian source is equal to:

$$R_L(D) - R_{MD}(D) \approx 0.312 \text{ bit/sample}, \tag{11}$$

where $R_L(D), R_{MD}(D)$ denote rate-distortion functions for i.i.d. Laplacian source and infinite Gaussian mixture source, respectively. The practical problem of side information for MG model consists in the communication of MG model parameters (local variances $\{\sigma^2_{x_i}\}$) to the decoder.

The problem of backward local variances estimation was elegantly solved by S. LoPresto *et. al.* [8, 9]. They marginally modelled the wavelet subband data using Generalized Gaussian distribution [3] and assumed that approximation (10) is still valid. They supposed that these local variances are slowly varying in wavelet subbands and could be reliably predicted using small dimensionality causal neighbourhood using even only quantized past coefficients. The developed practical image Estimation Quantization (EQ) coder shows the performance that is among the state-of-the-art for the case of lossy image compression [30].

Motivated by the EQ coder performance, we constructed our Main Encoder using the same principles with several modifications:

- at very low bit rate regime most of the information on the first and on the second wavelet decomposition levels is quantized to zero. We assume that all the data about strong edges could be reconstructed with some precision using side information and do not allocate any rate to these subbands;
- high frequency subbands of the third decomposition level are compressed using region of interest strategy;
- only 3×3 causal window is applied for local variance estimation;
- no interscale dependencies are assumed and wavelet coefficients are modelled based on intraband stochastic models.

The actual bitstream from the encoder constitutes the data from the EQ encoder, three bytes determining the position of the rectangular region-of-interest and four bytes characterizing background brightness.

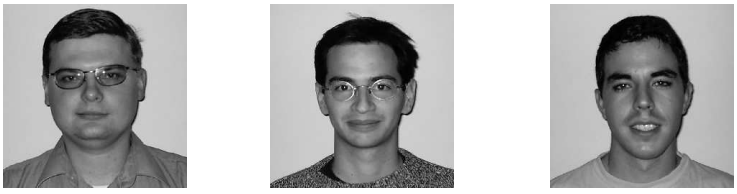
### 4.3. Decoder

The decoder has to perform reconstruction of compressed data using Main Encoder output and available side information. Bitstream of the main encoder is decompressed by EQ decoder. The fourth wavelet transform decomposition level is decompressed using classical algorithm version and the third level is reconstructed using region of interest EQ decoding.

Having two low pass levels of decomposition, low-resolution reconstruction (with two high frequency decomposition levels equal to zero) of the original photo using wavelet transform is obtained. Final reconstruction of high quality data is performed based on the interpolated image and Translation Detection and Shape Approximation block.

## 5. EXPERIMENTAL RESULTS

In this section we present experimental results of very low bit rate passport photo compression based on the proposed framework of distributed single source coding with symmetrical side information. For the experiments a set of 11 face images was used. Due to the restrictions of paper size, we show the results only for three images shown in Fig. 13.



a) b) c)

**Figure 13**. Test images: (a) “Slava”, (b) “Julien“, and (c) “Jose”.

The comparison of performance in terms of $\mathrm{PSNR} = 10\log_{10}\left(\frac{255^2}{\|\mathbf{x}-\hat{\mathbf{x}}\|^2}\right)$ of the proposed technique versus the EQ coder is shown in Table 1.

The results of compression are also shown in Fig. 14. The comparison of visual quality and PSNR proves a significant performance improvement using developed algorithm in comparison to the EQ.

**Table 1.** Comparison of the developed approach versus the EQ.

| Rate, bytes (Bitrate, bpp) | Slava | | Julien | | Jose | |
|---|---|---|---|---|---|---|
| | EQ | Developed approach | EQ | Developed approach | EQ | Developed approach |
| 500 (0.061) | 22.36 dB | 27.03 dB | 21.86 dB | 23.86 dB | 25.09 dB | 28.65 dB |
| 600 (0.073) | 25.96 dB | 27.85 dB | 22.81 dB | 24.28 dB | 27.09 dB | 29.10 dB |
| 700 (0.085) | 27.12 dB | 28.09 dB | 23.61 dB | 24.50 dB | 28.37 dB | 29.35 dB |
| 800 (0.099) | 27.71 dB | 28.16 dB | 24.24 dB | 24.56 dB | 29.31 dB | 29.46 dB |

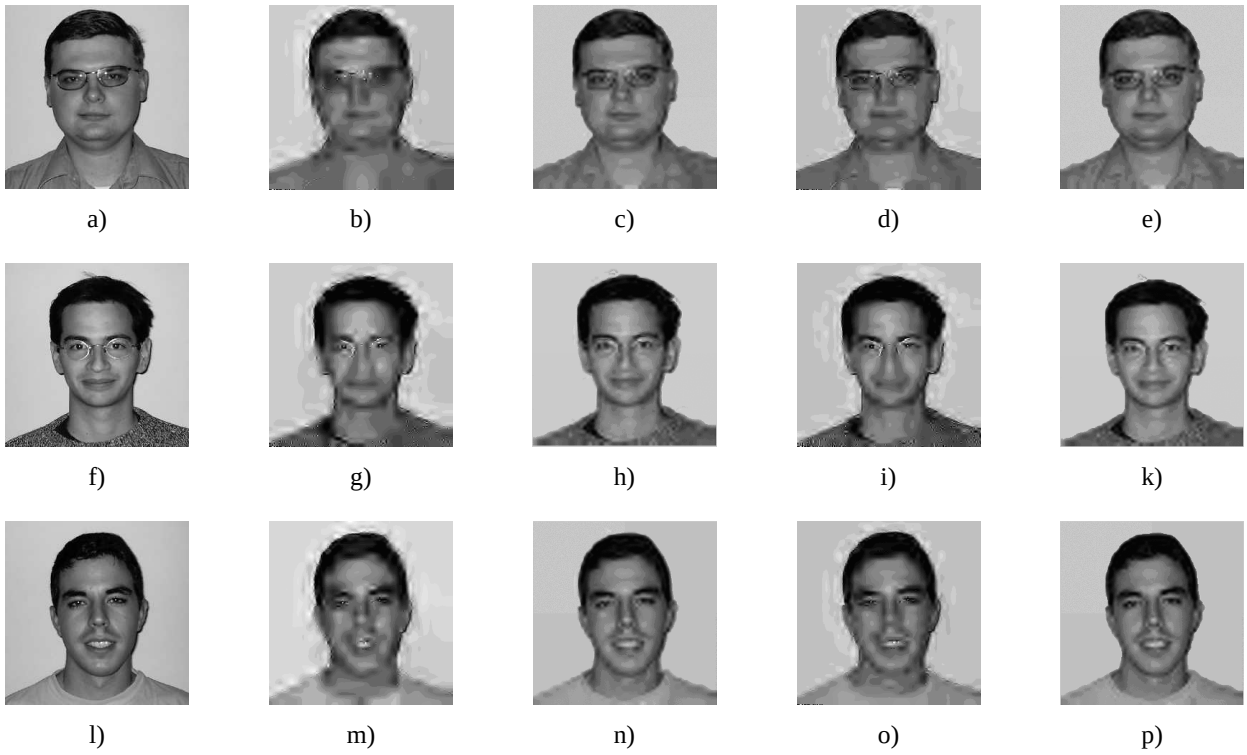



**Figure 14.** Experimental results: the first column, (a), (f), (l) original test images; the second column, (b), (g), (m) EQ compression for the rate 500 bytes; the third column, (c), (h), (n) proposed algorithm compression for the rate 500 bytes; the fourth column, (d), (i), (o) EQ compression for the rate 700 bytes; the fifth column, (e), (k), (p) proposed algorithm compression for the rate 700 bytes.

## 6. CONCLUSIONS

In the paper the problem of distributed source coding of a single source with side information was considered. It was shown that the optimal performance for non-Gaussian sources can be achieved using Berger-Flynn-Gray coding set-up. A practical very low bit rate compression algorithm based on this set-up was proposed for coding of passport photo images. It was shown on a set of test images that the algorithm demonstrates superior performance in comparison to one of the state-of-the-art techniques with respect to both visual quality and PSNR (PSNR gain up to 4.5 dB for different images and different target rates). Some performance saturation in PSNR gain is observed for compression regime close to 0.1 bpp. This phenomenon could be justified by non-optimality of quantizing to zero of two complete levels of wavelet decomposition in rate-distortion sense and poor approximation properties of single letter codebook. To improve the performance for these rates the modification to the Main Coder should be performed and finer codebook should be used. These extensions are a subject of our ongoing research.

## ACKNOWLEDGEMENTS

This paper was partially supported by Swiss SNF grant No 21-064837.01 and SNF Professorship grant No PP002-68653/1 and IM2 project. The authors are thankful to Pierre Vandergheynst (EPFL, Lausanne), Thierry Pun, Yuriy Rytsar, Frédéric Deguillaume and Emre Topak from University of Geneva for many helpful and interesting discussions.

## REFERENCES

1. X. Wu and N. Memon, "Calic – a context based adaptive lossless image codec," *in ICASSP'96*, 1996.

2. I. Daubechies, *Ten lectures on wavelets*. SIAM, Philadelphia, 1992.

3. S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7): 674-693, July 1989.

4. J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing*, vol. 41, pp. 3445--3462, Dec. 1993.

5. A. Said and W.A. Pearlman, "A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 6, pp. 243-250, June 1996.

6. Z. Xiong, K. Ramchandran and M. T. Orchard, "Space-frequency Quantization for Wavelet Image Coding," *IEEE Trans. Image Processing*, vol. 6, pp. 677-693, May 1997.

7. C. Chrysafis and A. Ortega, "Efficient Context-based Entropy Coding for Lossy Wavelet Image Compression," In Proc. *IEEE Data Compression Conference*, pages 241-250, Snowbird, Utah, 1997.

8. S. M. LoPresto, K. Ramchandran, and M. T. Orchard, "Image Coding based on Mixture Modeling of Wavelet Coefficients and a Fast Estimation-Quantization Framework," in Proc. *IEEE Data Compression Conference*, Snowbird, Utah, 1997.

9. S. Lopresto, K. Ramchandran and M. T. Orchard, "Wavelet image coding using rate-distortion optimized backward adaptive classification," *SPIE Proc. of Visual Communications and Image Processing (VCIP)*, San Jose, Feb. 1997.

10. A. Deever, S.S. Hemami, "What's Your Sign?: Efficient Sign Coding for Embedded Wavelet Image Coding," *Proceedings of Data Compression Conference 2000,* Snowbird, Utah, March 2000.

11. X. Wu, "*Compression of Wavelet Transform Coefficients*," *The Transform and Data Compression Handbook, (K.R. Rao et. al, ed.)*, Boca Raton, CRC Press LLC,2001.

12. S. S. Pradhan and K. Ramchandran, "Enhancing analog image transmission systems using digital side information: a new wavelet based image coding paradigm," *IEEE Data Compression Conference*, March, 2001.

13. D. Slepian and J. K. Wolf, "Noiseless Coding of Correlated Information Sources, " *IEEE Transactions on Information Theory,* vol. IT-19, pp. 471–480, Mar. 1973.

14. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, 1991.

15. T. M. Cover and M. Chiang, "Duality between Channel Capacity and Rate Distortion with Two-Sided Side Information," *IEEE Transactions on Information Theory*, vol. 48, no.6, pp.1629-1638, June 2002.

16. C. E. Shannon**.** "Coding Theorems for a Discrete Source With a Fidelity Criterion**,**" *Institute of Radio Engineers, International Convention Record*, Vol. 7 (Part 4, 1959), pp. 142-163.

17. T. Berger, *Rate-distortion theory: A mathematical basis for data compression.* Englewood Cliffs, NJ: Prentice-Hall, 1971.

18. A. Wyner and Y. Ziv, "The Rate Distortion Function for Source Coding with Side Information**,**" *IEEE Transactions on Information Theory*, vol. 22, pp.1-10, Jan. 1976.

19. T. Berger, *The information theory approach to communications (G.Liongo ed.,) chapter Multiterminal Source Coding*, Springer-Verlag, 1978.

20. T. J. Flynn and R. M. Gray, "Encoding of the Correlated Observations," *IEEE Transactions on Information Theory*, vol. 33, pp.773-787, Nov. 1987.

21. R. Zamir, "The Rate Loss in the Wyner-Ziv Problem," *IEEE Transactions on Information Theory*, vol. 42, pp. 2073-2084, Nov. 1996.

22. S. Voloshynovskiy, O. Koval, and T. Pun, "Wavelet-based image denoising using non-stationary stochastic geometrical image priors," in *ISJT/SPIE's Annual Symposium, Electronic Imaging 2003: Image and Video Communications and Processing V*, (San Clara, California USA), 20-24 January 2003.

23. I. Kozintsev and K. Ramchandran, "Multiresolution joint source-channel coding using embedded constellations for power-constrained time-varying channels," *Proc. of IEEE ICASSP-96*, Atlanta, May 1996.

24. J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, Nov. 1986.

25. J. G. Proakis. *Digital Communications*, 3rd Ed., McGraw-Hill, Inc., 1995.

26. http://www.jpeg.org/JPEG2000.html.

27. S. G. Mallat, *A wavelet tour of signal processing*, Academic Press, 1997.

28. A. Hjørungnes, J. M. Lervik, and T. A. Ramstad**,** Entropy Coding of Composite Sources Modeled by Infinite Mixture Gaussian Distributions, In Proc. *1996 IEEE Digital Signal Processing Workshop (Loen, Norway)*, September 1996.

29. Y. Yoo, A. Ortega, and B. Yu, "Image subband coding using context-based classification and adaptive quantization," *IEEE Transactions on Image Processing*, vol.8, 1702-1215.

30. http://www.icsl.ucla.edu/~ipl/psnr_results.html.