Wyner-Ziv Image Coding from Random Projections

Shoulie Xie, Susanto Rahardja, and Zhengguo Li
Institute for Infocomm Research
21 Heng Mui Keng Terrace, Singapore 119613
Email: {slxie, rsusanto, ezgli}@i2r.a-star.edu.sg

Abstract— **In this paper, we present a Wyner-Ziv coding based on random projections for image compression with side information at the decoder. The proposed coder consists of random projections (RPs), nested scalar quantization (NSQ), and Slepian-Wolf coding (SWC). Most of natural images are compressible or sparse in the sense that they are well-approximated by a linear combination of a few coefficients taken from a known basis, e.g., FFT or Wavelet basis. Recent results show that it is surprisingly possible to reconstruct compressible signal to within very high accuracy from limited random projections by solving a simple convex optimization program. Nested quantization provides a practical scheme for lossy source coding with side information at the decoder to achieve further compression. SWC is lossless source coding with side information at the decoder. In this paper, ideal SWC is assumed, thus rates are conditional entropies of NSQ quantization indices. Recently theoretical analysis shows that for the quadratic Gaussian case and at high rate, NSQ with ideal SWC performs the same as conventional entropy-coded quantization with side information available at both the encoder and decoder. We note that the measurements of random projects for a natural large-size image can behave like Gaussian random variables because most of random measurement matrices behave like Gaussian ones if their sizes are large. Hence, by combining random projections with NSQ and SWC, the tradeoff between compression rate and distortion will be improved. Simulation results support the proposed joint codec design and demonstrate considerable performance of the proposed compression systems.**

I. Introduction

Wyner-Ziv coding refers to lossy source coding with side information at the decoder. Wyner-Ziv coding in general suffers rate loss when compared to lossy coding with side information available both at the encoder and decoder [12], [13]. However, there is no rate loss with Wyner-Ziv coding in a special quadratic Gaussian sources. The quadratic Gaussian corresponds to a case when the source and information are zero mean and jointly Gaussian with mean-square error (MSE) distortion. Recently, Wyner-Ziv coding results have been used in image/video compression [7], [10], [15]. A practical encoder was proposed in [10] that consists of DCT, uniform quantization and trellis coding for SWC, while turbo code was used for SWC in [7]. In [15], a layered Wyner-Ziv video coder was presented to enhance the video quality, which consists of DCT, NSQ, and low-density parity-check (LDPC) code based SWC [9]. NSQ is a binning scheme that partitions the input into cosets and outputs only the coset indices.

In the above typical paradigms of Wyner-Ziv image/video coding, transforms such as DCT and wavelet are used to provide a more sparse decomposition of an image so that a more efficient method of image coding is possible in the later stages. However, this kind of transform coding has at least two weaknesses: 1) Computing all coefficients of the basis and keeping only the fewer significant ones in the encoder end. This is not practicable if the encoder has limited computational capability and/or the image is presented at high rate. 2) The locations of nonzero coefficients vary from one image to another, and must be known in the decoder, this needs more bits to be transmitted or saved. In this sense, the transform coding should be considered as adaptive.

Random projections have emerged recently as a useful tool in signal reconstruction [3], [5]. This has intrigued a lively area of research called "compressed sensing," [6] which offers an alternative approach to conventional image acquisition and compression. This approach shows that only a few number of linear random projections can be used to reconstruct the signal at a certain distortion level by solving a convex programming if it is sparse or compressible (i.e., the reordered entries of its coefficients in some orthonormal basis decay like a power-law). This compression process actually computes only a few linear measurements and can proceed without exploiting any information (aside from sparsity or compressibility), in this sense random projection method is a universal measurement tool and non-adaptive. Thus the computations can be significantly simplified. The foundation of these results is built on infinite precisions. Unfortunately, in today's digital world, all measurements are performed with analog-to-digital conversion, the quantization noise must be considered in the performance analysis [2], [4], [5].

Compressed sensing moves large amount of computations at traditional encoder side into the decoder side. This is also the idea of Wynzer-Ziv coding. In this paper, random projection technique is applied in image compression with side information at the decoder side. It is very natural that existing image compression technologies such as JPEG and JPEG2000 can also be used to provide side information. This kind of existing side information can assist on increasing the compression ratios and/or decreasing distortion. Here, our proposed encoder will consist of random projections, nested scalar quantization, and Selepian-Wolf coding. For simplicity, ideal SWC is assumed, thus the coding rates are conditional entropies of quantization indices given the side information. NSQ is used for further compression after random projections of an image. The measurements of random projections of an image can behave like Gaussian random variables because almost all random measurement matrices behave like Gaussian matrix if their sizes are large. Thus, the measurement vector after random projections is more suitable for later practical Wyner-Ziv coding with NSQ and SWC due to theoretical rate lossless result of Wyner-Ziv coding in quadratic Gaussian sources and practical good Wyner-Ziv coding with NSQ and SWC [8] (compared to DCT transform coding because DCT coefficients of images are better modeled as Laplacian distributed [11]). Hence, our proposed practical Wynzer-Ziv image coding based on random projections and SWC-NSQ will improve compression efficiency and distortion-rate functions.

II. Background of Signal Recovery Using RPs

Recent papers [2]–[6] have shown that a finite discrete signal $x_0 \in \mathbb{R}^N$ can be exactly recovered from a very limited number of observations if x_0 is sparse in spike basis or other fixed basis. A finite signal x_0 is sparse in that there are very few significant components; that is, whose support $T_0 = \{t : x_0(t) \neq 0\}$ is assumed to have small cardinality. Instead of observing x_0 directly, all that we know about x_0 are the very small number $K \ll N$ of linear measurements

$$y_k =< x_0,\ \phi_k >, \quad k = 1, 2, \cdots, K \qquad \text{or} \quad y = \Phi x_0, \quad (1)$$

where Φ is measurement matrix whose rows are the test function $\phi_k \in \mathbb{R}^N$. The results in [3], [5], [6] show that one can recover x_0 with very high probability by solving the convex optimization problem

$$(P_1) \qquad \min \|x\|_1 \quad \text{subject to} \quad \Phi x = y \qquad (2)$$

provided that Φ is random matrix with i.i.d zero-mean Gaussian entries and $K \geq C \cdot M \log N$ with small constant C and $M = |T_0|$. Similar results also can be derived for other measurement ensembles such as Fourier random matrix obtained by randomly selecting K rows from the $N \times N$ discrete Fourier transform [3], [5].

Real-world signal is not exactly sparse in any orthogonal basis in general. An important model is a class of compressible signals [4], [5]. Vector x_0 is compressible if its entries (or its coefficients in a fixed basis) obey a power law $|x_0|_{(k)} \leq C_r k^{-r}$, where $|x_0|_{(k)}$ is the kth largest value of x_0 ($|x_0|_{(1)} \geq |x_0|_{(2)} \geq \cdots \geq |x_0|_{(N)}$), $r > 1$, and C_r is a constant which depends only on r. In most practical situations, we cannot also assume that the measurements are known with arbitrary precision. Instead, we assume that one is given noisy observation $y = \Phi x + e$, where e is an unknown perturbation whose size can be bounded by a known amount $\|e\|_2 \leq \epsilon$. In place of problem (P_1), one can solve the relaxed problem

$$(P_2) \qquad \min \|x\|_1 \quad \text{subject to} \quad \|\Phi x - y\|_2 \leq \epsilon. \qquad (3)$$

In [4], [5], it was shown that the solution $\hat{x}$ to (P_2) for compressive signal x_0 has error that can be bounded by

$$\|\hat{x} - x_0\|_2 \leq C_1 \cdot \epsilon + C_2 \cdot S^{-\alpha} \qquad (4)$$

provided that Φ is Gaussian random matrix and K is at least proportional to $S \log N$ with the S largest values of x_0, where $\alpha = r - \frac{1}{2}$, C_1 and C_2 are small constants. Similar results can be also derived for other random measurement ensembles.

The above result says that minimizing ℓ_1 problem (P_2) recovers the S-largest entries of vector x_0 from only a small portion of the measurements, whose approximation error is almost as good as that obtained by knowing everything about the signal x_0 and selecting its S-largest entries.

III. Proposed WZ Image Coding Based on RPs

A. Encoder

The encoder here consists of three components: Random projections, NSQ and SWC shown in Fig. 1.

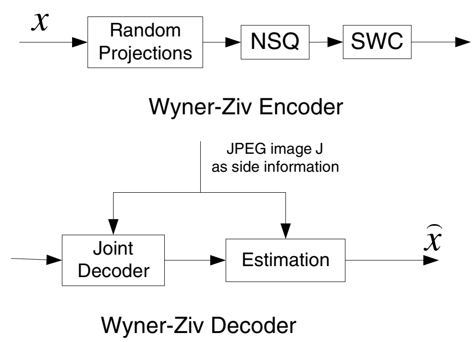

Fig. 1. Random projections based WZ image coding.

1) Random Projections: We can first acquire K measurements of an image x by multiplying x with a random measurement matrix $\Phi \in \mathbb{R}^{K \times N}$; that is, $y = \Phi x$. Therefore, the statistics of the measurement random projections are fully characterized by specifying the measurement matrix Φ known both at the encoder and decoder sides. Here we will follow the approach in [4] to design the random measurement matrix Φ. In general, an image is a large dimensional vector, this makes the standard Gaussian random measurement matrix be very large dimensional and too unwieldy. Instead, an implicit random matrix is produced by using a scrambled real Fourier ensemble. That is, this random matrix is obtained from the real-valued sines and cosines of Fourier ensemble by a random permutation of the columns. This has a computational advantage because of existence of fast Fourier transform (FFT).

On the other hand, we treat a standard JPEG decoded image as side information J at the decoder side, which is also randomly measured by using $p = \Phi J$, We assume that y and p are jointly Gaussian with $y = p + \eta$ where η is zero-mean Gaussian and independent of p. This assumption is reasonable because the random measurement matrices behave like Gaussian random matrix if their sizes are large. Thus, the measurement vector after random projections is more suitable for later practical Wyner-Ziv coding with NSQ and SWC due to theoretical rate lossless result of Wyner-Ziv coding in quadratic Gaussian sources (compared to DCT transform coding because DCT coefficients of images are better modeled as Laplacian distributed [11]).

2) Nested Scalar Quantization: In this step, NSQ is applied, which consists of a coarse coset channel code nested in a fine uniform scalar quantizer. The fine source code employs a uniform scalar quantizer with stepsize q and the coarse channel code with minimum distance $d_{min} = nq$. To encode, each entry in the measurement vector y is first quantized by the fine source code (uniform scalar quantization), resulting the "high resolution" with "good" distortion, which is an average quantization error of $q^2/12$ at high rate. Then the index a $(0 \le a \le n-1)$ of the coset in the coarse channel code that the quantized y belongs to is coded to further save rate after random projections.

3) SWC: Due to the correlation between y and p, there still remains correlation between the coset index a and the side information p. SWC can be used to compress a to the rate of conditional entropy $R = H(a|p)$. Here we assume that ideal SWC coders are employed. In practical implementation, we employ multilevel LDPC codes to compress a based on the syndrome-based method [8], [9], [15] to approach the Slepian-Wolf limit $R = H(a|p)$. The role of SWC is thus to exploit the correlation between a and p for further compression.

B. Decoder

We perform optimal estimation at joint decoder, the decoded coset index $\hat{a}$ specifies the uncertainty region of y. The side information essentially supplies the conditional pdf of y given p, which is a Gaussian with mean p and variance proportional to the correlation between y and p. The optimal estimation of y is computed as the conditional centroid $\hat{y} = E(y|\hat{a}, p)$. In high rate case, the estimator is linear [8]; that is, $\hat{y} = \hat{a} + Q_{nq}(p - \hat{a})$ where Q_{nq} is quantization with stepsize nq.

Next, ℓ_1 minimum optimal approach is applied to $\hat{y}$ to obtain $\hat{x}$ with high accuracy and probability. In [4], it was shown that the recovered images by solving (P_2) tend to contain visually displeasing high frequency oscillatory artifacts. To address this problem, in practical, this optimal problem is replaced by

$$(TV) \qquad \min \|x\|_{TV} \quad \text{subject to} \quad \|\Phi x - \hat{y}\|_2 \le \epsilon, \quad (5)$$

where

$$\|x\|_{TV} = \sum_{i,j} \sqrt{(x_{i+1,j} - x_{i,j})^2 + (x_{i,j+1} - x_{i,j})^2} \quad (6)$$

is the total variation of the image x: the sum of the magnitude of the gradient. Instead of looking for an image with a sparse transform that explains the observations, TV searches for an image with a sparse gradient (without spurious high frequency oscillations). In fact, it is shown in [4] that the reconstructions by solving (TV) have smaller error than that by solving (P_2) and do not contain visually displeasing artifacts.

In (5), $\epsilon^2 = Kq^2/12$ which is selected from the fact that the probability of $\|\Phi x - \hat{y}\|_2^2$ exceeding this value is small. In fact,

$$\|\Phi x - \hat{y}\|_2^2 \le \|\Phi x - y_q\|_2^2 + \|y_q - \hat{y}\|_2^2,$$

where y_q is the fine quantized vector of measurement vector y. It was shown in [8], [14] that at high rate and for the quadratic Gaussian case, NSQ with ideal SWC performs the same as classic entropy-coded quantization with the side information available at both the encoder and decoder; that is, the probability of $\|y_q - \hat{y}\|_2^2 \ne 0$ is very small, and the total distortion $\|\Phi x - \hat{y}\|_2^2$ (or distortion-rate function) with practical SWC-NSQ is 1.53dB away from the theoretical Wyner-Ziv DR function. This result of high-rate performance for Wyner-Ziv coding of quadratic Gaussian sources with NSQ is relatively accurate, but in general, average quantization errors of any nonuniform quantizers taking advantage of statistics properties of sources are less than the average quantization error of uniform scalar quantizer. Hence, it is reasonable that $\epsilon^2 = Kq^2/12$ is chosen as crude quantization error of the measurement vector regardless of whether Wyner-Ziv coding with SWC is used or not, but it is universal in the sense that it is not related with Wyner-Ziv coding.

Finally, it is obvious that by combing random projections and NSQ techniques for Wyner-Ziv image coding, the transmitted (or saved) rate can be reduced largely and the distortion is small enough to ensure sufficient accuracy of image recovery.

IV. Simulation Results

This section shows the simulation results to demonstrate the validity of our approach. First, the convex programs (P_2) and (TV) can be solved efficiently by interior-point methods [1].

We apply our Wyner-Ziv image coding approach to a representative natural image—Lena. However, the observations and conclusions drawn here apply to most real natural images. We use Lena image with $N = 256 \times 256 = 65536$ pixels shown in Fig. 2 as original one, and the standard JPEG coded Lena image with 32K bits shown in Fig. 3 as side information at the decoder. The proposed Wyner-Ziv image coding approach consisting of random projections, NSQ and SWC is used to generate a Wyner-Ziv bitstream. We first make $K = 12500$ measurements of the image using a scrambled real Fourier matrix. Then fine image quantization is performed with stepsize $q = 10$. Next NSQ is used to compute the coset for nesting ratio $n = 16$ and output the corresponding coset index a $(0 \le a \le n-1)$. Then ideal SWC is assumed in our simulations for simplicity, so the transmitted rate $R = H(a|p)$. At the decoder side, the joint linear NSQ decoding is carried out first, and then the reconstructed image is completed by solving (TV) problem. The simulation result for the proposed Wyner-Ziv image coding with $K = 12500$ is shown in Fig. 4. Similar, for $K = 25000$ measurements, the simulation result is shown in Fig. 5. We see that the image quality using our proposed Wyner-Ziv coding with random projections $K = 12500$ is better than that obtained using JPEG technique, and in our results, there are no visually displeasing high frequency oscillatory artifacts, this benefits from solving (TV) problem. Also, we note that the image quality for $K = 25000$ is better than that for $K = 12500$. In fact, in our simulations, we obtained that using our proposed Wyner-Ziv with random projections image coding $\frac{1}{N}\|\hat{x} - x_0\|_2 = 0.0272$ for $K = 25000$, and $\frac{1}{N}\|\hat{x} - x_0\|_2 = 0.0398$ for $K = 12500$, respectively, while $\frac{1}{N}\|\hat{x} - x_0\|_2 = 0.0410$ for JPEG coded Lena image. Hence, we observe from our simulations that the image quality improves as the number K of random projections (measurements) increases. In addition, the bit rate is of course lower than that of directly fine quantization after random projections without Wyner-Ziv coding.

V. Conclusion

We have proposed a Wyner-Ziv image coding method based on random projections and solving convex optimal problems. The advantages of this coding method are manifold. First, only a few of measurements from random projections are computed. Second, the resulted measurement vector behaves like Gaussian random that is more suitable for later stages of our image coding (compared to Laplacian distributed DCT coefficients). Third, NSQ and SWC are used to further save the rate. Fourth, the computational burden is moved to the decoder side compared to conventional standard image coding, it is suitable for uplink multimedia transmission from low-computation power devices. However, most of reconstruction algorithms based on random projections involves linear programming or second-order cone programming. These decoding methods have high computation cost, which are expected to be cubic in the length of signal although they are not NP problems. In the future, we will seek more fast and efficient computation techniques in order to reduce the computational burden of decoder.

References

[1] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
[2] E. J. Candès and J. Romberg, "Encoding the ℓ_p ball from limited measurements," in *Data Compression Conf. (DCC'06)*, Snowbird, UT, 2006, pp. 33–42.
[3] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction form highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, 2006.
[4] ——, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.
[5] E. J. Candès and T. Tao, "Near optimal signal recovery form random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.
[6] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
[7] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proc. IEEE*, vol. 93, no. 1, pp. 71–83, 2005.
[8] Z. Liu, S. Cheng, A. D. Liveris, and Z. Xiong, "Slepian-Wolf coded nested lattice quantization for Wyner-Ziv coding: High-rate performance analysis and code design," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4358–4379, 2006.
[9] A. D. Liveris, Z. Xiong, and C. Georghiades, "Compression of binary sources with side information at the decoder using LDPC codes," *IEEE Commun. Lett.*, vol. 6, no. 10, pp. 440–442, 2002.
[10] S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 626–643, 2003.
[11] D. Reininger and J. Gibson, "Distribution of the two-dimensional DCT coefficients," *IEEE Trans. Commun.*, vol. 31, no. 6, pp. 835–839, 1983.
[12] A. Wyner, "The rate-distortion function for source coding with side information at the decoder-II: General sources," *Inf. Contr.*, vol. 38, pp. 60–80, 1978.
[13] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, 1976.
[14] Z. Xiong, A. D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Process. Mag.*, vol. 21, no. 9, pp. 80–94, 2004.
[15] Q. Xu and Z. Xiong, "Layered Wyner-Ziv video coding," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3791–3803, 2006.

Fig. 2. Original Lena image with 256×256 pixels.

Fig. 3. JPEG coded image.

Fig. 4. Random projections based Wyner-Ziv coded image ($K = 12500$).

Fig. 5. Random projections based Wyner-Ziv coded image ($K = 25000$).