

IMAGE RECONSTRUCTION FROM COMPRESSED LINEAR MEASUREMENTS WITH SIDE INFORMATION

Vijayaraghavan Thirumalai and Pascal Frossard

Ecole Polytechnique Fédérale de Lausanne (EPFL)
Signal Processing Laboratory - LTS4, Lausanne, 1015 - Switzerland.
{vijayaraghavan.thirumalai, pascal.frossard}@epfl.ch

ABSTRACT

This paper proposes a joint reconstruction algorithm for compressed correlated images that are given under the form of linear measurements. We consider the particular problem where one image is selected as the reference image and it is used as a side information for decoding the compressed correlated images. These compressed images are given under the form of random measurements that are further quantized and entropy coded. The joint decoder estimates the correlation model based on the geometric transformation of features captured by a structured dictionary. We observe that the high frequency components are not efficiently captured in the estimated image when the correlation information is used alone for image prediction. Hence, we propose a reconstruction strategy that uses the information in the measurements to recover the missing visual information in the predicted image. The reconstruction is based on an optimization algorithm that enforces the reconstructed image to be consistent with the quantized measurements. We further add additional constraints to ensure that the reconstructed image is close to the image predicted from the correlation estimation. The non-linearities introduced due to quantization are considered on both correlation and reconstruction algorithms in order to improve the performance. Experimental results demonstrate the benefit of the reconstruction algorithm as it brings improved coding performance especially at high rate and outperforms independent coding solutions based on JPEG 2000.

1. INTRODUCTION

Distributed source coding (DSC) usually refers to the independent encoding and joint decoding of correlated sources. It permits to design low complexity acquisition systems and to shift the computational burden to the decoder. DSC typically finds applications in vision sensor networks where low-power cameras perform a spatio-temporal sampling of the visual information and send the resulting images to a central decoder. While most common encoders in DSC systems acquires the entire image before compression, the complexity of the encoders can be further reduced if the sensors directly acquires the compressed image in the form of random projections [1, 2]. Such a solution computes only few linear projections at the encoder and thereby significantly reduces the computational cost and the power requirements at the encoder. A joint decoder eventually reconstructs the visual information from the compressed images by exploiting the correlation between the samples, which permits to achieve a good rate-distortion (RD) tradeoff in the representation of video or multi-view information.

This work has been partly supported by the Swiss National Science Foundation, under grant 200021-118230.

The concept of random projections in distributed scenarios has been previously studied in [3] where three joint sparsity models (JSM) are proposed for joint signal reconstruction algorithms. These simple JSM's are however not ideal for multi-view images or video sequences as the correlation model in such scenarios is usually given in the form of disparity or motion vectors. Later distributed algorithms have been developed for multi-view images [4] and video sequences [5, 6] based on linear measurements. These schemes first build the correlation model that is eventually used to construct the side information for joint signal recovery. The joint reconstruction stage then assumes that the prediction error between the original and side information is sparse in an orthonormal basis [6] or a dual-tree wavelet basis [4], and recovers the missing values in the side information by solving an optimization problem. In an other approach, the authors in [7] derived the significant position of the source using the side information and use this as a priori knowledge for joint signal recovery. However, these distributed schemes based on linear measurements usually fail to consider the effect of quantization and hence cannot be applied directly in practical coding applications.

In this paper we build on our previous work [8] and propose a joint reconstruction algorithm from quantized linear measurements. We first compute the most prominent visual features in the reference image and approximate them with geometric functions drawn from a parametric dictionary. Then the correlation model is constructed by solving a regularized optimization problem that computes the corresponding features in the compressed image along with the relative geometric transformation. When the correlation model is used for estimating the compressed image we lose the visual information mostly along the edges and high frequency components. We therefore propose a reconstruction algorithm to capture the missing details and texture information in the predicted image from the information provided by the quantized measurements. We consider the predicted image as a side information and formulate a convex optimization problem to reconstruct an image that is as close as possible to the side information. At the same time we use additional constraints to enforce that the reconstructed image is consistent with the quantized measurements, where the consistency is measured using the l_p norm in order to take into account the effect of quantization nonlinearities [9]. Experimental results confirm that the proposed reconstruction scheme improves the quality of the predicted image and thereby improves the RD performance of the proposed DSC scheme.

2. PROPOSED SCHEME

We consider a framework (see Fig. 1) where a pair of image I_1 and I_2 represent a scene at different time instants or from different viewpoints; these images are correlated through the motion of visual ob-

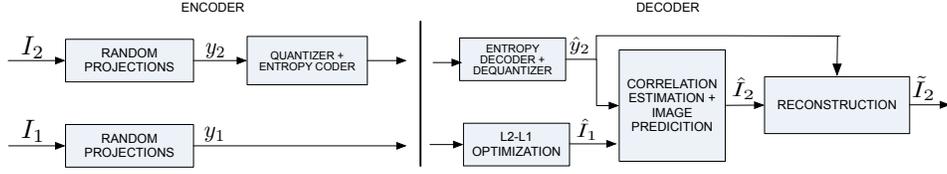


Fig. 1. Schematic representation of the proposed scheme. The images I_1 and I_2 are correlated through displacement of scene objects, due to view point change or motion of scene objects.

jects. These images are represented by linear measurements that corresponds to the projection of the image pixel values on a random set of coding vectors. They are then transmitted to a joint decoder that estimates the relative motion or disparity between the received signals and jointly reconstructs the images.

We focus on the particular problem where the first image serves as a reference for the correlation estimation and the reconstruction of the second image. While this image could be encoded with any coding algorithm (e.g., JPEG 2000) we choose in this work to represent the reference image I_1 by random linear measurements $y_1 = \psi I_1$ with a projection matrix ψ . The measurements are used by the decoder to reconstruct an approximation \hat{I}_1 using a convex optimization algorithm under the assumption that I_1 is sparse in particular basis (e.g., a Wavelet basis). Next the second image I_2 is also projected on the random matrix ψ to generate the measurements $y_2 = \psi I_2$. The generated measurements y_2 are quantized using an uniform quantizer and are further entropy coded (e.g., Arithmetic encoder). The decoder performs the reverse operations (dequantization and entropy decoding) to form the measurement vector \hat{y}_2 (see Fig. 1). This measurement vector is finally used by the joint decoder to estimate the relative transformation between the images I_1 and I_2 and eventually predicts the second image \hat{I}_2 based on image warping. However, the image \hat{I}_2 thus estimated based on warping fails to capture the details or texture information as the motion or disparity compensation predicts only the low frequency components and not the high frequency. Hence in this work we propose a reconstruction algorithm that approximates the second image \tilde{I}_2 from \hat{I}_2 by enforcing consistency with the quantized measurements \hat{y}_2 (see Fig. 1).

3. CORRELATION ESTIMATION FROM QUANTIZED MEASUREMENTS

In this section we briefly review our correlation estimation scheme based on quantized measurements and for more details we refer the reader to [8]. We propose to model the correlation between the images by relative transformations between prominent visual features in both images. In particular, the joint decoder first computes the sparse approximation of the image \hat{I}_1 using vectors in a parametric dictionary of geometric functions. Such an approximation captures the most prominent geometrical features in the image \hat{I}_1 . Given these features we estimate the corresponding features in the second image I_2 that is given in terms of quantized linear measurements \hat{y}_2 . We propose to select this set of atoms in a regularized energy minimization framework with energy model E expressed as,

$$E(\Lambda) = E_d(\Lambda) + \alpha_1 E_s(\Lambda) + \alpha_2 E_t(\Lambda) \quad (1)$$

where E_d , E_s and E_t represent the data term, smoothness term and consistency term respectively and Λ represents the set of N atom parameters. The regularization constants α_1 and α_2 balance the data,

smoothness and consistency terms. The data cost term E_d picks the set of atoms that agrees best with the quantized measurements \hat{y}_2 . To provide robustness to quantization errors we have additional constraints that consider all the measurement values in the quantized interval while picking the atoms. The smoothness cost E_s enforces the estimated dense disparity or motion field to be coherent. The consistency term E_t is used to enforce consistency of the estimated image \hat{I}_2 with the quantized measurements \hat{y}_2 , where the image \hat{I}_2 is estimated through image warping. We propose an iterative optimization algorithm to solve Eq. 1 that picks a solution in a search space formed by perturbing the N atom parameters by one increment at a time in the parameter space. In every iteration a search space is constructed based on the solution of the previous iteration and a (minima) solution is estimated. The estimated solution is used for constructing the search space in the next iteration and the algorithm is iterated till convergence is reached. Finally the image \hat{I}_2 is reconstructed by warping the reference image \hat{I}_1 using the estimated correlation model. We observed experimentally that the quality of \hat{I}_2 saturates at high rate as the high frequency components are not captured efficiently using the warping operator. In the next section we describe a new reconstruction algorithm based on an optimization framework that captures the high frequency components with the help of measurement information to improve the quality of \hat{I}_2 .

4. CONSISTENT RECONSTRUCTION

The key idea in the proposed reconstruction algorithm is to consider the warped image \hat{I}_2 as a side information and combine it with the information from the measurements for the reconstruction of the second image \tilde{I}_2 . We propose to reconstruct an image \tilde{I}_2 that is not only consistent with the measurements y_2 but also close to the image \hat{I}_2 . By merging these two constraints the proposed convex optimization problem is given as,

$$\min \| I_2 \|_{TV} \text{ s.t. } \| y_2 - \psi I_2 \|_2 = 0, \| I_2 - \hat{I}_2 \|_2 \leq \epsilon_2. \quad (2)$$

In the above optimization we use the prior based on total variation (TV) norm [2] that works well for natural images. Nevertheless one could also use a sparsity prior, i.e., the image is sparse in a particular basis. When the measurements are quantized it is well known that the optimization problem given in Eq. (2) fails to meet the quantization consistency, i.e., the reconstructed image \tilde{I}_2 is not consistent with the quantized measurements. Jacques *et al.* [9] showed that quantization consistency is enforced with l_p norm with $p > 2$ and not using l_2 norm in the first constraint. Inspired by Jacques *et al.* [9], when the measurements y_2 are quantized the above optimization problem can be modified as,

$$\min \| I_2 \|_{TV} \text{ s.t. } \| \hat{y}_2 - \psi I_2 \|_p \leq \epsilon_1, \| I_2 - \hat{I}_2 \|_2 \leq \epsilon_2 \quad (3)$$

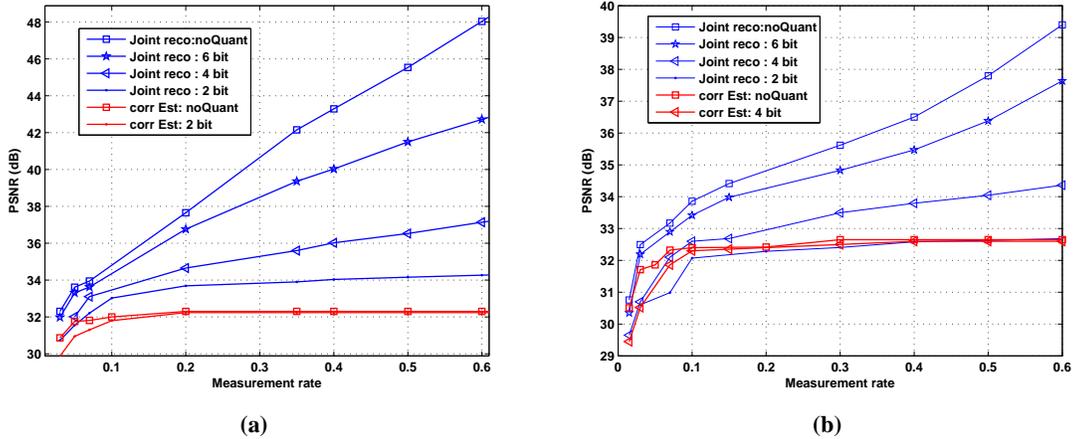


Fig. 2. Performance comparison of the proposed scheme at various quantization bitrates for (a) Plastic and (b) Foreman datasets. The PSNR is computed between I_2 and \tilde{I}_2 in the joint reconstruction and between I_2 and \hat{I}_2 in the correlation estimation.

where the measurement consistency is measured using the l_p norm with $p > 2$ instead of using $p = 2$.

Now we describe an optimization methodology to solve the problem given in Eq. (3). The optimization problem can be visualized as the special case of general convex problem given as,

$$\min_{x \in \mathcal{H}} f_1(x) + f_2(x) + f_3(x) \quad (4)$$

where $\mathcal{H} = \mathbb{R}^N$ is the Hilbert space and the functions f_1, f_2 and $f_3 \in \Gamma_0(\mathbb{R}^N)$ [9, 10]. $\Gamma_0(\mathbb{R}^N)$ is the class of lower semicontinuous convex functions from \mathbb{R}^N to $]-\infty, +\infty]$ such that the convex function f is not infinity everywhere, i.e., $\text{dom } f \neq \emptyset$. For the optimization problem given in Eq. (3) the functions are (i) $f_1(x) = \|x\|_{TV}$ (ii) $f_2(x) = i_{T^p(\epsilon_1)}(x) = 0$ if $x \in T^p(\epsilon_1)$ and ∞ otherwise, i.e., the indicator function of the closed convex set $T^p(\epsilon_1) = \{x \in \mathbb{R}^N : \|\hat{y}_2 - \psi x\|_p \leq \epsilon_1\}$, (iii) $f_3(x) = i_{C(\epsilon_2)} = 0$ if $x \in C(\epsilon_2)$ and ∞ otherwise, where C is the closed convex set given as $C(\epsilon_2) = \{x \in \mathbb{R}^N : \|x - \hat{I}_2\|_2 \leq \epsilon_2\}$.

The solution to the problem in Eq. (4) can be found by generating the recursive sequence $x^{(t+1)} = \text{prox}_{\beta f}(x^{(t)})$, $\beta > 0$, where the function f is given as $f = f_1 + f_2 + f_3$. The proximity operator is defined as the $\text{prox}_f(x) = \min_{x \in \mathcal{H}} f(y) + \frac{1}{2} \|x - y\|_2$. The main difficulty with these iterations are the computations of the $\text{prox}_{\beta f}(x)$ operator as there is no closed form expression to compute the $\text{prox}_f(x)$, especially when the function f is the cumulative sum of three functions. In such cases, instead of the computing the $\text{prox}_f(x)$ directly for the combined function f one can perform a sequence of calculations involving separately the individual operators $\text{prox}_{f_1}(x)$, $\text{prox}_{f_2}(x)$ and $\text{prox}_{f_3}(x)$. This class of algorithm is popularly known as *splitting methods* as these methods proceed by splitting the combined f into f_1, f_2 and f_3 , and hence allow for an easily implementable algorithm [10]. For the function $f_1(x) = \|x\|_{TV}$ the operator $\text{prox}_{f_1}(x)$ can be computed iteratively using Chambolle's algorithm [11]. The $\text{prox}_{f_2}(x)$ for function $f_2(x) = i_{T^p(\epsilon_1)}(x)$ with $T^p(\epsilon_1) = \{x \in \mathbb{R}^N : \|\hat{y}_2 - \psi x\|_p \leq \epsilon_1\}$ can be computed for $p \geq 2$ using Newton's iterations [9]. The proximity operator $\text{prox}_{f_3}(x)$ for function $f_3 = i_{C(\epsilon_2)}$ with $C(\epsilon_2) = \{x \in \mathbb{R}^N : \|x - \hat{I}_2\|_2 \leq \epsilon_2\}$ can be found out using radial projection.

In our work, we use the parallel proximal algorithm (PPXA) proposed by Combettes *et al.* [10] to solve Eq. (4) as the algorithm can be easily implementable on multicore architectures due to the parallel structure. PPXA algorithm starts with an initial solution $x^{(0)}$ and computes the $\text{prox}_{f_1}(x)$, $\text{prox}_{f_2}(x)$ and $\text{prox}_{f_3}(x)$ in each iteration and the result is used to update the current solution $x^{(0)}$. The iterative procedure of computing the *prox* and updating the solution is repeated until convergence is reached. The authors have shown that the sequence $(x^{(l)})_{l \geq 1}$ generated by PPXA algorithm is guaranteed to converge to the solution of the problem given in Eq. (4).

5. EXPERIMENTAL RESULTS

The performance of the joint reconstruction has been studied on three natural datasets namely the stereo images Sawtooth and Plastic, and Foreman video sequence (frames 2 and 3). In our experiments we set the quality of the reference image \hat{I}_1 to 33 dB for Sawtooth and Plastic datasets and to 45 dB for Foreman dataset. The measurements are generated using a block scrambled Hadamard transform with block size 8 [12]. For a given measurement rate we first estimate the correlation model and then the image \hat{I}_2 is predicted based on disparity or motion compensation respectively. We then reconstruct the second image \tilde{I}_2 from the compressed measurements \hat{y}_2 and the predicted image \hat{I}_2 by solving Eq. (3).

Fig. 2 shows the benefit of using the additional reconstruction stage when the measurements are not quantized for Plastic and Foreman datasets. It should be noted that when the measurements are not quantized we solve the optimization problem in Eq. (3) with $p = 2$. The parameter ϵ_1 is set to $1e-4$, and parameter ϵ_2 is selected based on trial and error experiments such that the quality of reconstructed image \tilde{I}_2 is maximized. The PSNR value corresponding to the predicted image \hat{I}_2 and the reconstructed image \tilde{I}_2 is marked in red and blue respectively. It is clear from the plots that the quality of the predicted image \hat{I}_2 saturates around a measurement rate 0.2. Then by activating the reconstruction stage the quality of the reconstructed image \tilde{I}_2 improves as the measurement rate increases; this proves that the reconstruction stage captures the details and the texture components. Similar experimental finding is observed on the Sawtooth

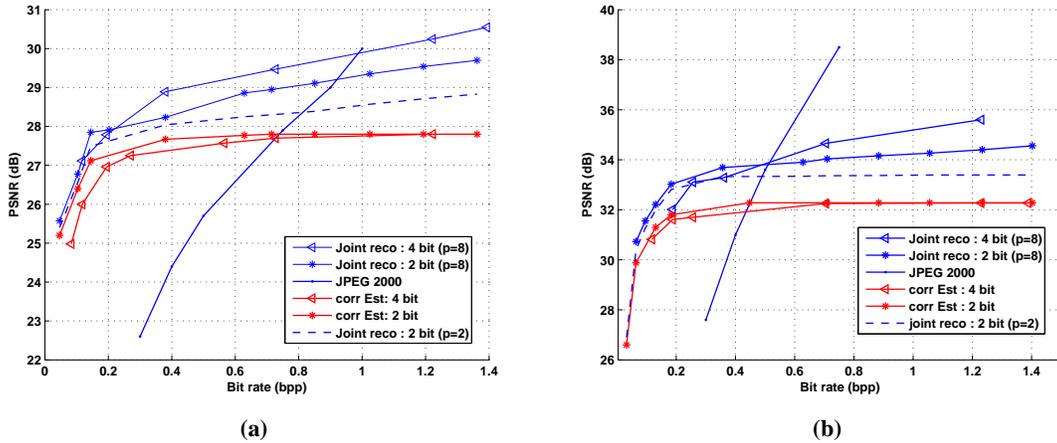


Fig. 3. Performance of the proposed scheme at 2 and 4 quantization bitrates and comparison with JPEG 2000. (a) Sawtooth and (b) Plastic datasets. The PSNR is computed between I_2 and \hat{I}_2 in the joint reconstruction and between I_2 and \hat{I}_2 in the correlation estimation.

dataset. Now we analyze the behavior of the proposed joint reconstruction scheme when the measurements are quantized. We solve the optimization problem with $p = 8$ (selected based on trial and error experiments), and the parameter ϵ_1 is calculated based on the quantization bitrate and the value of p [9]. Fig. 2 compares the reconstruction quality of \tilde{I}_2 when the measurements are quantized using 2, 4 and 6 bits. As expected for a given measurement rate the quality of the reconstructed image \tilde{I}_2 degrades as the quantization bitrate decreases. We further notice that the quality of the image \tilde{I}_2 saturates above rates > 0.2 when the measurements are quantized using a 2-bit quantizer.

We finally analyze the coding performance of the proposed scheme where the bitrate is computed by entropy coding the quantized measurements. Fig. 3 shows the coding performance of the proposed scheme when the measurements are quantized using 2 and 4 bits. It is clear from the plot that the proposed reconstruction scheme improves the quality of the predicted image \hat{I}_2 . Also, as expected the quality of the reconstructed image \hat{I}_2 increases with the bitrate and thus corrects the saturation behavior of the predicted image \hat{I}_2 especially at high rate. In addition, our solution shows significant improvement over independent coding solutions based on JPEG 2000 between low to medium rate. Finally the choice of using $p > 2$ is proven effective in Fig. 3, when measurements are quantized using a 2-bit quantizer. It is clear that the coding performance is better when $p = 8$ (rather than $p = 2$) as it better handles the quantization error while reconstructing the image \hat{I}_2 .

6. CONCLUSIONS

In this paper we present a framework for the distributed representation and joint reconstruction of an image pair with quantized linear measurements. Building on our correlation estimation and image prediction algorithm [8], we propose a reconstruction algorithm that captures the texture and details in the predicted image. We consider the predicted image as a side information and we formulate a convex optimization algorithm that ensures that the reconstructed image is consistent with the measurements. Also the reconstruction algorithm efficiently handles the quantized measurements in order to provide robustness to quantization noise. We show by experiments that the

performance is improved especially at high rate by activating the reconstruction algorithm. Finally we show that the performance of our scheme is far superior to independent coding solutions based on JPEG 2000. This illustrates the potential of our scheme in distributed multi-view or video coding applications.

7. REFERENCES

- [1] D. Donoho, "Compressed sensing," *IEEE Trans. on Infor. Theo.*, vol. 52, pp. 1289–1306, 2006.
- [2] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Infor. Theo.*, vol. 52, pp. 489–509, 2006.
- [3] M. F. Duarte, S. Sarvotham, D. Baron, M. B. Wakin, and R. G. Baraniuk, "Distributed compressed sensing of jointly sparse signals," in *Proc. Asilo. Conf. on Sig. Sys. and Comp.*, 2005.
- [4] M. Trocan, T. Maugey, E. W. Tramel, J. E. Fowler, and B. P. Popescu, "Multistage compressed-sensing reconstruction of multi-view images," *Proc. IEEE ICME*, 2010.
- [5] L. W. Kang and C. S. Lu, "Distributed compressive video sensing," *Proc. IEEE ICASSP*, 2009.
- [6] T. T. Do, Y. Chen, D. T. Nguyen, N. Nguyen, L. Gan, and T. D. Tran, "Distributed compressed video sensing," *Proc. IEEE ICIP*, 2009.
- [7] V. Stankovic, L. Stankovic, and S. Cheng, "Compressive image sampling with side information," *Proc. IEEE ICIP*, 2010.
- [8] V. Thirumalai and P. Frossard, "Distributed Representation of Geometrically Correlated Images with Compressed Linear Measurements," *EPFL Technical Report*, 2010. [Available] <http://infoscience.epfl.ch/record/161685>.
- [9] L. Jacques, D. K. Hammond, and M. J. Fadili, "Dequantizing compressed sensing: When oversampling and non-gaussian constraints combine," *IEEE Trans. on Infor. Theo.*, vol. 57, pp. 559–571, 2011.
- [10] P. L. Combettes and J-C. Pesquet, "Proximal splitting methods in signal processing," *Fixed Point Alg. for Inv. Probl. in Sci. and Engi.*, Springer, 2010.
- [11] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Math. Imag. and Vis.*, pp. 89–97, 2004.
- [12] L. Gan, T. T. Do, and T. D. Tran, "Fast compressive imaging using scrambled hadamard ensemble," *Proc. EUSIPCO*, 2008.