JOINT RECONSTRUCTION OF CORRELATED IMAGES FROM COMPRESSED IMAGES

Vijayaraghavan Thirumalai and Pascal Frossard

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

ABSTRACT

This paper proposes a novel joint reconstruction algorithm to decode sets of correlated images from distributively compressed images. We consider a scenario where the images captured at different viewpoints are encoded independently using transform-based coding solutions (e.g., SPIHT) with a balanced rate distribution among different cameras. A central decoder jointly processes the compressed images and reconstructs an image pair by exploiting the inter-view correlation. The central decoder first estimates the underlying correlation model from the independently decoded images and it is eventually used for the joint signal recovery. The joint reconstruction is cast as a constrained convex optimization problem that reconstructs a total-variation (TV) smooth image pair that satisfies with the estimated correlation model. At the same time, we add constraints that force the reconstructed images to be as close as possible to the compressed views. We show by experiments that the proposed joint reconstruction scheme outperforms independent reconstruction in terms of image quality, for a given target bit rate.

Index Terms— Distributed representation, Joint reconstruction, Convex optimization, Disparity estimation.

1. INTRODUCTION

Distributed source coding (DSC) usually refers to the independent encoding and joint decoding of correlated sources [1]. 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 the low-power cameras perform a spatio-temporal sampling of the visual information and send the resulting compressed images to a central decoder. 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 tradeoff in the representation of multi-view information.

Several distributed coding schemes for compressing the video and the multi-view images have been proposed in the literature [2, 3]. In such schemes, a feedback channel is generally used for accurately controlling the Slepian-Wolf coding rate. Unfortunately, this results in high latency and bandwidth usage due to the multiple requests from the joint decoder. These schemes can thus hardly be used in real time applications. One solution to avoid the feedback channel is to use a separate encoding rate control module in order to precisely control the Slepian-Wolf coding rate [4]. The overall computational complexity at the encoder becomes non-negligible due to this rate control module. In this paper, we build a symmetric distributed coding scheme, where the correlated compressed images are directly transmitted to the joint decoder without implementing any Slepian-Wolf coding; this avoids the necessity for complex estimation of the statistical correlation estimation and of the coding rate at the encoder. In a similar framework, Schenkel *et al.* [5] have proposed a distributed joint representation of image pairs. In particular, they have proposed an optimization framework to enhance the quality of the JPEG compressed images. This work, however, considered an asymmetric scenario that requires a reference image for joint decoding.

In this paper, we propose a symmetric distributed joint representation scheme for compressing a pair of correlated images captured in stereo camera networks. We consider a scenario where the captured images are compressed independently using standard encoding solutions (e.g., SPIHT [6]) and are transmitted to a central decoder. The central decoder builds a correlation model from the compressed images, which is used to jointly decode a pair of images. The joint reconstruction is formulated as a convex optimization problem; it reconstructs a pair of images that are consistent with the underlying correlation information and with the compressed images information. We solve this joint reconstruction problem using parallel proximal algorithms [7]. Experimental results demonstrate that the proposed distributed coding solution improves the rate-distortion performance of the separate coding results by taking advantage of the inter-view correlation. Also, the quality of the decoded images is quite similar for a given bit rate; this confirms the symmetrical nature of the proposed scheme. Thus, our framework certainly provides an interesting alternative to the most classical DSC solutions [2, 3, 4], since it does not require any statistical correlation information at the encoder nor any feedback channel.

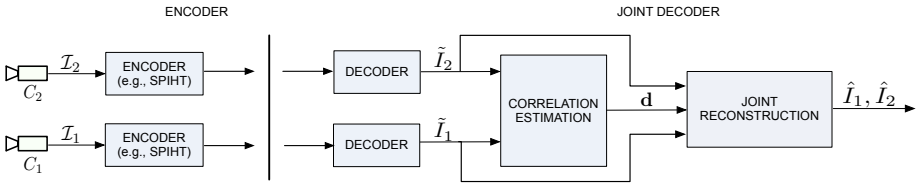


Fig. 1. Schematic representation of the proposed scheme. The images $\mathcal{I}_1$ and $\mathcal{I}_2$ are correlated through displacement of scene objects due to positioning of the cameras C_1 and C_2.

2. PROPOSED FRAMEWORK

In this section, we give an overview of our distributed coding framework. We consider the scenario illustrated in Fig. 1, where a pair of cameras C_1 and C_2 sample a 3D scene in different viewpoints. Without loss of generality, we assume that the images $\mathcal{I}_1$ and $\mathcal{I}_2$ (with resolution $N = N_1 \times N_2$) are rectified, so that correlation between images is effectively described by a disparity field. The correlated images $\mathcal{I}_1$ and $\mathcal{I}_2$ are compressed independently with b bits per view using common coding solutions (e.g., SPIHT [6]). Balanced rate allocation permits to share the transmission and the computational costs equally among the sensors. It prevents the necessity for hierarchical relationship among the sensors. The compressed information is transmitted to a central decoder that exploits the underlying correlation between views for joint decoding. In particular, as shown in Fig. 1, the joint decoder estimates the correlation between images in terms of a dense disparity image $\mathbf{d}$ from the decoded images $\tilde{I}_1$ and $\tilde{I}_2$. Several algorithms have been proposed in the literature to compute the dense disparity images [8]. In this work, we estimate the disparity image $\mathbf{d}$ using graph-based optimization techniques, due to better performance over other techniques [8, 9]. A joint reconstruction stage eventually uses the disparity information $\mathbf{d}$ and enhances the quality of the independently decoded views $\tilde{I}_1$ and $\tilde{I}_2$. Finally, note that our framework can also be extended to the joint decoding of unrectified multiview images. The details are available in [10].

3. JOINT RECONSTRUCTION IN OPTIMIZATION FRAMEWORK

In this section, we describe the proposed optimization-based joint reconstruction algorithm. We propose to reconstruct an image pair $(\hat{I}_1, \hat{I}_2)$ as a solution to the following optimization problem:

$$(\hat{I}_1, \hat{I}_2) = \underset{I_1, I_2}{\operatorname{argmin}} \left(\|I_1\|_{TV} + \|I_2\|_{TV}\right) \quad (1)$$
$$\text{s.t. } \|\mathcal{R}(I_1) - \mathcal{R}(\tilde{I}_1)\|_2 \leq \epsilon_1, \|\mathcal{R}(I_2) - \mathcal{R}(\tilde{I}_2)\|_2 \leq \epsilon_1,$$
$$\|I_2(m,n) - I_1(m + \mathbf{d}(m,n), n)\|_2^2 \leq \epsilon_2,$$

where $\tilde{I}_1$ and $\tilde{I}_2$ represent the compressed views and $\|.\|_{TV}$ represents the total-variation (TV) norm. The reshaping operator $\mathcal{R} : I_{N_1 \times N_2} \rightarrow X_{N_1 N_2 \times 1}$ produces a vector $X = \mathcal{R}(I) = [I_{.,1}^T \; I_{.,2}^T \ldots I_{.,N_1}^T]^T$, where $I_{.,m}$ represents the m^{th} row of the matrix I. The first two constraints of Eq. (1) forces the reconstructed images $\hat{I}_1$ and $\hat{I}_2$ to be consistent or close to the respective decoded images $\tilde{I}_1$ and $\tilde{I}_2$. The last constraint imposes the reconstructed images to fit with the correlation information $\mathbf{d}$. Finally, the TV prior term ensures that the reconstructed images are smooth. In general, inclusion of the prior knowledge brings effective reduction in the search space, which leads to efficient optimization solutions. Therefore, optimization problem of Eq. (1) reconstructs a pair of TV smooth images that is consistent with both the compressed images and the correlation information. In our framework, we use the TV prior on the reconstructed images, however, one could also use a sparsity prior that minimizes the l_1 norm of the coefficients in a sparse image representation.

Before solving Eq. (1), we represent the last constraint $\|I_2(m,n) - I_1(m + \mathbf{d}(m,n), n)\|_2^2$ in the matrix format as $\|\mathcal{R}(I_2) - A \cdot \mathcal{R}(I_1)\|_2^2$. That is, we represent the disparity compensation $I_1(m + \mathbf{d}(m,n), n)$ as a linear transformation $A \cdot \mathcal{R}(I_1)$ given as

$$\underbrace{\begin{bmatrix} \bar{I}_{2,1}^T \\ \bar{I}_{2,2}^T \\ \vdots \\ \bar{I}_{2,N_1}^T \end{bmatrix}}_{\mathcal{R}(\bar{I}_2)} = \underbrace{\begin{bmatrix} A^1 & 0 & \ldots & 0 \\ 0 & A^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & A^{N_1} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} I_{1,1}^T \\ I_{1,2}^T \\ \vdots \\ I_{1,N_1}^T \end{bmatrix}}_{\mathcal{R}(I_1)}, \quad (2)$$

where $\bar{I}_2$ is the predicted image. The sub-matrix A^m is of dimensions $N_2 \times N_2$ that is computed as

$$A^m(p, min(p + \beta, N_2)) = \begin{cases} 1 & \mathbf{d}(m,p) = \beta, \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

where $\mathbf{d}(m,p)$ represents the disparity value at the p^{th} location in the m^{th} row. If the value of $p + \beta > N_2$ (which might happen at the boundaries) we set $p + \beta = N_2$, so that the dimensions of the matrix A^m is $N_2 \times N_2$. It is easy to check that the matrix A^m formed using Eq. (3) contains only one non-zero value in each row. For example, the matrix A^m corresponding to $\mathbf{d}(m,.) = [2\ 2\ 1\ 1]$ is given as

$$A^m = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (4)$$

Since the matrix A^m contains only one non-zero value in each row it is evident that $\bar{I}_2(m,i) = I_1(m,j)$ if $A^m(i,j) = 1$. Thus, it is clear that the matrix A^m shifts the pixels in $I_{1,m}$ by the corresponding disparity vector $\mathbf{d}(m,.)$, to form $\bar{I}_{2,m}$. Using this linear relationship of disparity compensation, Eq. (1) can be rewritten as

$$(\hat{I}_1, \hat{I}_2) = \underset{I_1, I_2}{\operatorname{argmin}} \left(\|I_1\|_{TV} + \|I_2\|_{TV}\right) \quad \text{(OPT-1)}$$
$$\text{s.t. } \|\mathcal{R}(I_1) - \mathcal{R}(\tilde{I}_1)\|_2 \leq \epsilon_1, \|\mathcal{R}(I_2) - \mathcal{R}(\tilde{I}_2)\|_2 \leq \epsilon_1,$$
$$\|\mathcal{R}(I_2) - A \cdot \mathcal{R}(I_1)\|_2^2 \leq \epsilon_2.$$

Finally, note that the joint reconstruction of compressed images has been considered in other applications such as super-resolution, where multiple compressed images are fused to enhance the resolution [11]. Such techniques usually target reconstruction of a *single* high resolution image from multiple compressed images. Alternatively, in [12] several encoded versions of the single image are fused together to extract a *single* high quality image. Our main target in this paper is to jointly improve the quality of *multiple* compressed views and not to increase the spatial resolution of the compressed images or to extract a single high quality image.

4. OPTIMIZATION METHODOLOGY

We propose now a solution for the joint reconstruction problem OPT-1. We first show that the problem OPT-1 is convex. Then, we propose a solution based on proximal methods.

Proposition 1. *The optimization problem OPT-1 is convex.*

Proof. Our objective is to show that all the functions in the problem OPT-1 are convex. However, it is quite easy to check that the functions $\|I_j\|_{TV}$ and $\|\mathcal{R}(I_j) - \mathcal{R}(\tilde{I}_j)\|_2$, $\forall j \in \{1,2\}$ are convex [13]. So, we have to show that the last constraint $\|\mathcal{R}(I_2) - A \cdot \mathcal{R}(I_1)\|_2^2$ is a convex function. Let $g(\grave{I}_1, \grave{I}_2) = \|\grave{I}_2 - A\grave{I}_1\|_2^2$, where $\grave{I}_2 = \mathcal{R}(I_2)$ and $\grave{I}_1 = \mathcal{R}(I_1)$. The function g can be represented as

$$g(\grave{I}_1, \grave{I}_2) = \grave{I}_2^T\grave{I}_2 - \grave{I}_2^T A\grave{I}_1 - \grave{I}_1^T A^T\grave{I}_2 + \grave{I}_1^T A^T A\grave{I}_1.$$

The second derivative $\nabla^2 g$ of the function g is given as

$$\nabla^2 g = \begin{bmatrix} 2AA^T & -2A \\ -2A^T & 2 \end{bmatrix} = 2C^T C \succeq 0,$$

where $C = [A^T \quad -\mathbb{1}]$ and $\mathbb{1}$ represents the identity matrix. $2C^TC \succeq 0$ follows from $2x^TC^TCx = 2\|Cx\|_2^2 \geq 0$ for any x. This means that $\nabla^2 g$ is positive semi-definite and thus $g(\grave{I}_1, \grave{I}_2)$ is convex. □

We now propose an optimization methodology to solve OPT-1 with proximal splitting methods [7]. For mathematical convenience, we rewrite OPT-1 as

$$\begin{aligned} \underset{X \in \mathbb{R}^{2N}}{\operatorname{argmin}} \quad & \{\|\mathcal{R}^{-1}(S_1X)\|_{TV} + \|\mathcal{R}^{-1}(S_2X)\|_{TV}\} \\ \text{s.t.} \quad & \|S_1(Y-X)\|_2 \leq \epsilon_1, \|S_2(Y-X)\|_2 \leq \epsilon_1, \\ & \|[-A \quad \mathbb{1}]X\|_2^2 \leq \epsilon_2, \end{aligned} \quad (5)$$

where $X = [\mathcal{R}(I_1)\ ; \mathcal{R}(I_2)]$, $Y = [\mathcal{R}(\tilde{I}_1)\ ; \mathcal{R}(\tilde{I}_2)]$, $S_1 = [\mathbb{1}\ 0]$ and $S_2 = [0\ \mathbb{1}]$. The operator $\mathcal{R}^{-1}_{N_1 \times N_2}$ outputs a matrix of dimensions $N_1 \times N_2$ from a column vector of dimensions $N = N_1N_2$, i.e., it performs the inverse operations corresponding to $\mathcal{R}$. The optimization problem of Eq. (5) can be visualized as a special case of general convex problem as

$$\underset{X \in \mathbb{R}^{2N}}{\operatorname{argmin}} \{f_1(X) + f_2(X) + f_3(X) + f_4(X) + f_5(X)\}. \quad (6)$$

The functions $f_1, f_2, f_3, f_4, f_5 \in \Gamma_0(\mathbb{R}^{2N})$, where $\Gamma_0(\mathbb{R}^{2N})$ is the class of lower semicontinuous convex functions from $\mathbb{R}^{2N}$ to $(-\infty\ +\infty]$ that are not infinity everywhere [7]. For the optimization problem given in Eq. (5) the functions in the representation of Eq. (6) are (i) $f_1(X) = \|\mathcal{R}^{-1}(S_1X)\|_{TV}$; (ii) $f_2(X) = \|\mathcal{R}^{-1}(S_2X)\|_{TV}$; (iii) $f_3(X) = i_{c_1}(X) = 0$ if $X \in c_1$ and ∞ otherwise, i.e., $f_3(X)$ is the indicator function of the convex set $c_1 = \{X : \|S_1(Y-X)\|_2 \leq \epsilon_1\}$; (iv) $f_4(X) = i_{c_2}(X) = 0$ if $X \in c_2$ and ∞ otherwise, where the convex set $c_2 = \{X : \|S_2(Y-X)\|_2 \leq \epsilon_1\}$; and (v) $f_5(X) = i_{c_3}(X) = 0$ if $X \in c_3$ and ∞ otherwise, where the convex set $c_3 = \{X : \|[-A \quad \mathbb{1}]X\|_2^2 \leq \epsilon_2\}$.

The solution to the problem of Eq. (6) can be estimated by generating the recursive sequence $X^{(t+1)} = prox_f(X^{(t)})$, where the function f is given as $f = \sum_{i=1}^5 f_i$. The proximity operator is defined as $prox_f(X) = min_X \{f(X) + \frac{1}{2}\|X - Z\|_2^2\}$. The main difficulty with these iterations is the computation of the $prox_f(X)$ operator. There is no closed form expression to compute the $prox_f(X)$, especially when the function f is the cumulative sum of two or more functions. In such cases, instead of computing the $prox_f(X)$ directly for the combined function f, one can perform a sequence of calculations involving separately the individual operators $prox_{f_i}(X), \forall i \in \{1,\ldots,5\}$. The algorithms in this class are known as *splitting methods* [7], which lead to an easily implementable algorithm.

We describe in more details the methodology to compute the *prox* for the functions $f_i, \forall i \in \{1,\ldots,5\}$. For the function $f_1(X) = \|\mathcal{R}^{-1}(S_1X)\|_{TV}$, the $prox_{f_1}(X)$ operator can be computed using Chambolle's algorithm [14]. A similar approach can be used to compute the $prox_{f_2}(X)$. The function f_3 can be represented as $f_3 = F \circ G$, where $F = i_{d(\epsilon_1)}$ and $G = S_1Y - S_1X$. The set $d(\epsilon_1)$ represents the l_2-ball defined as $d(\epsilon) = \{y \in \mathbb{R}^{2N} : \|y\|_2 \leq \epsilon_1\}$. Then, the $prox_{f_3}$ can be computed using the following closed form expression:

$$prox_{f_3}(X) = prox_{F \circ G}(X) = X + (S_1)^*(prox_F - \mathbb{1})(G(X)) \quad (7)$$

[15], where $(S_1)^*$ represents the conjugate transpose of S_1. The $prox_F(y)$ with $F = i_{d(\epsilon_1)}$ can be computed using radial projection [7] as $prox_F(y) = y$ if $\|y\|_2 \leq \epsilon_1$ and $\frac{y}{\|y\|_2}$ otherwise. The $prox$ for the function f_4 can also be solved using Eq. (7) with $F = i_{d(\epsilon_1)}$ and $G = S_2Y - S_2X$. Finally, the function f_5 can be represented with $F = i_{d(\sqrt{\epsilon_2})}$ and an affine operator $G = [-A \quad \mathbb{1}]X = \Omega X$, i.e.,

$f_5 = F \circ G$. As the operator Ω is not a tight frame, the $prox_{f_5}$ can be computed using an iterative scheme [15]. Let $\mu_t \in (0, 2/\gamma_2)$, and γ_1 and γ_2 be the frame constants with $\gamma_1 \mathbb{1} \leq \Omega\Omega^* \leq \gamma_2 \mathbb{1}$. The $prox_{f_5}$ can be calculated iteratively [15] as

$$u^{(t+1)} = \mu_t(\mathbb{1} - prox_{\mu_t^{-1} F})(\mu_t^{-1} u^{(t)} + Gp^{(t)}) \quad (8)$$

$$p^{(t+1)} = X - \Omega^* u^{(t+1)}, \quad (9)$$

where $u^{(t)} \to u$ and $p^{(t)} \to prox_{F \circ G} = prox_{f_5} = X - \Omega^* u$. It has been shown that both $u^{(t)}$ and $p^{(t)}$ converge linearly and the best convergence rate is attained when $\mu_t = 2/(\gamma_1 + \gamma_2)$.

In our work, we use the parallel proximal algorithm (PPXA) proposed by Combettes *et al.* [7] to solve Eq. (6), as this algorithm can be easily implementable on multicore architectures due to its parallel structure. The PPXA algorithm starts with an initial solution $X^{(0)}$ and computes the $prox_{f_i}, \forall i \in \{1, \ldots, 5\}$ in each iteration and the result is used to update the current solution $X^{(0)}$. The iterative procedure for computing the *prox* of functions $f_i, \forall i \in \{1, \ldots, 5\}$, and the updating steps are repeated until convergence is reached. The authors have shown that the sequence $(X^{(t)})_{t \geq 1}$ generated by the PPXA algorithm is guaranteed to converge to the solution of problems such as the one given in Eq. (6).

5. EXPERIMENTAL RESULTS

In this section, we study the performance of our proposed joint reconstruction algorithm on two correlated datasets, *Tsukuba* and *Venus* [8]. In our experiments, the images $\mathcal{I}_1$ and $\mathcal{I}_2$ are encoded independently using the SPIHT algorithm [6]. Note that one can also use standard encoding techniques like JPEG for compressing the images (see [10] for details). We estimate a dense disparity image from the decoded images $\tilde{I}_1$ and $\tilde{I}_2$ using α-expansion algorithm in Graph Cuts [9]. We solve the OPT-1 problem with $\epsilon_1 = 2$ and $\epsilon_2 = 3$ that are selected based on trial and error methods such that the quality of the reconstructed images $\hat{I}_1$ and $\hat{I}_2$ is maximized.

Fig. 2(a) and Fig. 2(b) compare the quality of the decoded images for the independent and joint decoding solutions, respectively for the Tsukuba and Venus datasets. For bit rates $b > 0.2$, we see from Fig. 2 that the proposed joint reconstruction scheme outperforms the independent reconstruction by a margin of about 0.7-0.8 dB. This confirms that the proposed joint reconstruction framework effectively exploits the inter-view correlation. We also see that the reconstruction quality of the images $\hat{I}_1$ and $\hat{I}_2$ is quite similar for a given bit rate b. We have also observed substantial coding gains when compared to the performance of a DSC scheme based on disparity learning [16]. Unfortunately, due to page restrictions we omit the discussion here. More details are available in [10].

From Fig. 2, we further see that the joint reconstruction fails to improve the quality of the compressed images at low rates; this is due to the poor quality of the estimated disparity images. For the Venus dataset, the disparity images estimated from the decoded images $\tilde{I}_1, \tilde{I}_2$ that are encoded at bit rates $b = 0.1$ and $b = 0.4$ are shown in Fig. 3(b) and Fig. 3(d), respectively. Comparing the respective disparity images with respect to the groundtruth information $\mathbf{D}$ in Fig. 3(a), we observe poor quality disparity results for a bit rate of 0.1. Quantitatively, the respective disparity errors are found to be 28% and 10.4%, when it is measured as the percentage of pixels with an absolute error greater than one. This confirms that the quantization noise in the compressed images are not properly handled while estimating the correlation information. Developing robust correlation estimation techniques to alleviate this problem is the target of our future works.

6. CONCLUSION

In this paper, we have proposed a novel rate balanced distributed representation scheme for compressing a pair of correlated images captured in camera networks. Contrary to the classical DSC schemes, our scheme compresses the images independently without knowing the inter-view geometrical or statistical relationship at the encoder. We have proposed a new joint decoding algorithm based on a constrained optimization problem that permits to improve the reconstruction quality by exploiting the correlation between images. We have shown that our joint reconstruction problem is convex so that it can be efficiently solved using proximal methods. Simulation results confirm that the proposed joint representation algorithm is successful in improving the reconstruction quality of the compressed images at medium to high coding rates, with a balanced quality between images. Our future work focuses on developing robust techniques to estimate an accurate correlation information from highly compressed images.

7. REFERENCES

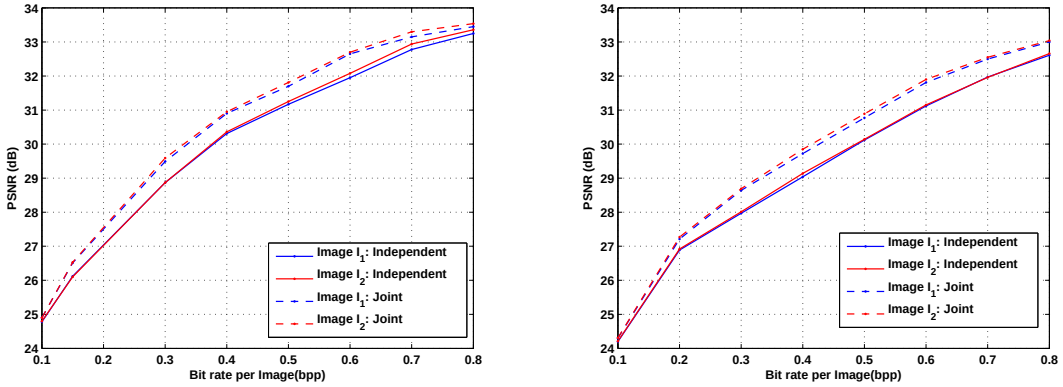
[1] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. on Infor. Theo.*, vol. 19, pp. 471 – 480, 1973.

[2] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proc. of the IEEE*, vol. 93, pp. 71 – 83, 2005.

[3] C. Guillemot, F. Pereira, L. Torres, T. Ebrahimi, R. Leonardi, and J. Ostermann, "Distributed monoview and multiview video coding," *IEEE Sig. Proc. Mag.*, vol. 24, no. 5, pp. 67–76, 2007.

[4] R. Puri, A. Majumdar, and K. Ramachandran, "PRISM: A video coding paradigm with motion estimation at the decoder," *IEEE Trans. on Image Proc.*, vol. 16, no. 10, pp. 2436–2448, 2007.

[5] M. B. Schenkel, C. Luo, P. Frossard, and F. Wu, "Joint decoding of stereo image pairs," in *Proc. ICIP*, 2010.



(a) (b)

Fig. 2. Comparison of the rate-distortion performance between the independent and proposed joint decoding schemes: (a) Tsukuba dataset; and (b) Venus dataset.

(a) $\mathbf{D}$ (b) $\mathbf{d}$ (c) $|\mathbf{D}-\mathbf{d}|>1$ (d) $\mathbf{d}$ (e) $|\mathbf{D}-\mathbf{d}|>1$

Fig. 3. Comparison of the estimated disparity image from the compressed images with respect to groundtruth information in the Venus dataset. (a) Groundtruth disparity image $\mathbf{D}$; (b) computed disparity image $\mathbf{d}$ at bit rate $b = 0.1$; (c) disparity error at rate $b = 0.1$. The pixels with absolute error greater than one is marked in white. The percentage of white pixels is 28%. (d) Computed disparity image $\mathbf{d}$ at bit rate $b = 0.4$; (e) disparity error at rate $b = 0.4$. The percentage of white pixels is 10.4%.

[6] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec using set partitioning in hierarchical trees," *IEEE Trans. on Cir. and Sys. for Video Tech.*, vol. 6, no. 3, pp. 243–250, 1996.

[7] P. L. Combettes and J-C. Pesquet, "Proximal splitting methods in signal processing," *In Fixed Point Algo. for Inv. Prob. in Sci. and Eng., Springer*, 2010.

[8] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense stereo," *Inter. Jour. on Comp. Vis.*, vol. 47, pp. 7–42, 2002.

[9] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. on Patt. Anal. and Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, Jan 2002.

[10] V. Thirumalai and P. Frossard, "Joint reconstruction of multi-view compressed images," *arXiv e-prints*, 2012. [Online]. Available: http://arxiv.org/abs/1206.4326

[11] S. Park, M. Park, and M. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Sig. Proc. Mag.*, vol. 20, no. 3, pp. 21–36, 2003.

[12] M. Dalai, S. Malavasi, and R. Leonardi, "Consistent image decoding from multiple lossy versions," in *Proc. ACM Multimedia workshop*, 2010.

[13] S. P. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge University Press, New York, 2004.

[14] A. Chambolle, "An algorithm for total variation minimization and applications," *Jour. Math. Imag. Vis.*, pp. 89–97, 2004.

[15] M. J. Fadili and J. L. Starck, "Monotone operator splitting for optimization problems in sparse recovery," *Proc. ICIP*, 2010.

[16] D. Varodayan, Y. C. Lin, A. Mavlankar, M. Flierl, and B. Girod, "Wyner-Ziv coding of stereo images with unsupervised learning of disparity," in *Proc. PCS*, 2007.