ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Accurate and Scalable Correlation Analysis approach for decoding of distributed encode multi view images

V.Pratyusha
MTech Student
Department of CSE
CMR College of Engineering & Technology
Kandlakoya (V), Medchal Road, Hyderabad

B.Rajani
Asst. Professor
Department of CSE
CMR College of Engineering & Technology
Kandlakoya (V), Medchal Road, Hyderabad

Abstract: ***Large data volumes that are produced during multi view image generation have to be efficiently compressed in order to be stored or transmitted. Two main classes of encoders use transform coding techniques either by utilizing spatial prediction methods or by using higher degree transforms. Here we device an accurate fast correlation estimation strategy for joint reconstruction approach of multi view image decoding. We propose a distributed compression and joint reconstruction enabled decompression strategy for an array of multiview images, which is to jointly improve the quality of multiple compressed correlated (multiview) images and not to increase the spatial resolution of the compressed images or to extract a single high quality image.***

Index Terms: Depth estimation, distributed compression, joint reconstruction, multiview images, optimization, proximal splitting.

INTRODUCTION

Compressive sampling (CS) can be used to reconstruct image signals from a "small" number of (random or deterministic) linear combinations, referred to as measurements or samples, of the original image pixels without collecting the entire frame [1], [2]. CS-based imaging and video coding has been recently discussed as the basis for a clean-slate approach to low-power wireless video streaming systems based on simple encoder and high-complexity decoder, with applications to wireless multimedia sensor networks [3], [4].

In this context, we propose and study a multi-view video encoding and decoding architecture based on compressive imaging principles, designed to acquire multiple correlated images from the same area of interest from different views. The architecture is motivated by wireless video sensing applications with low-complexity, independent encoders with minimal inter-sensor communication; and a potentially more complex joint decoder, which can lead to substantial rate savings and

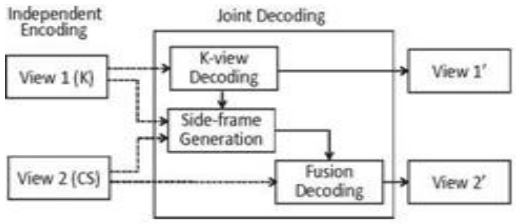

Fig. 1: Muti-view encoding/decoding architecture.

energy savings on battery-powered wireless sensors. Theoretical results for noiseless [5] and noisy [6] distributed source coding have been available since the late seventies. Several video coding schemes based on CS have been proposed in the literature [4], [7], [8], [9]. However, they mainly focus on performing CS reconstruction by exploiting correlation among successive frames [8], [9] without considering interview correlation; or consider rate allocation with traditional CS reconstruction methods [4]. In [10], a distributed multi-view video coding scheme based on CS is proposed, which however assumes the same measurement rates for different views, and can only be applied together with specific structured dictionaries as sparse representation matrix. Differently, in this work we consider multi-view video sequences encoded at different rates and with more general sparsifying matrices, e.g., Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT). The authors of [11] propose a CS-based joint reconstruction method for multi-view images, which uses two images from the two nearest views of the current image (the right and left neighbors) to calculate a prediction frame.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

However, in our work, only one reference view (not necessarily the nearest one) is selected to reconstruct the side frame for the joint reconstruction process.

Related Work

In this paper, we consider a scenario where a set of cameras are distributed in a 3D scene. In most practical deployments of such systems, the images captured by the different cameras are likely to be correlated. The captured images are encoded independently using standard encoding solutions and are transmitted to the central decoder. Here, we assume that the images are compressed with a balanced rate allocation; this permits to share the transmission and the computational costs equally among the sensors. It thus prevents the necessity for hierarchical relationships among the sensors. The central decoder builds a correlation model from the compressed images which is used to jointly decode [6] the multi-view images. The joint reconstruction is formulated as a convex optimization problem. It reconstructs the multi-view images that are consistent with the underlying correlation information and at the same time close to the compressed images information[12,13]. Furthermore, we effectively handle the occlusions that commonly arise in multi-view imaging. We propose to solve the joint reconstruction problem using effective parallel proximal algorithms.

We evaluate the performance [12] of our novel joint decoding scheme in several multi-view datasets. Experimental results demonstrate that the proposed distributed coding solution improves the rate-distortion performance of the separate decoding results by taking advantage of the inter-view correlation. We show that the quality of the decoded images is quite balanced for a given bit rate, as expected from symmetric coding solutions. We observe that our scheme, at low bit rate, performs close to the joint encoding solutions based on H.264, when the block size used for motion compensation is set to 4 x 4. Finally [6,8], we show that our framework outperforms state-of-the-art distributed coding solutions based on motion learning and on the DISCOVER algorithm in terms of rate-distortion performance. Our scheme certainly provides an interesting alternative to most classical DSC solutions since it does not require statistical correlation information at the encoder.

JOINT DECODING OF IMAGE PAIRS

We consider the scenario illustrated in Fig. 1, where a pair of cameras C_1 and C_2 project the 3D visual information on the 2D plane I_1 and I_2 (with resolution $N = N_1 \times N_2$), respectively. The images I_1 and I_2 are compressed independently using standard encoding solutions (e.g., JPEG, H.264 intra)and are transmitted to a central decoder. The joint decoder has access to the compressed versions of the correlated images and its main objective is to improve the quality of all the compressed views by exploiting the inter-view correlation. We first propose to estimate the correlation between images from the decoded images $\tilde{I}_1$ and $\tilde{I}_2$, which is effectively modeled by a dense depth image D. The joint reconstruction stage then uses the depth information D in order to enhance the quality of the decoded images $\tilde{I}_1$ and $\tilde{I}_2$. Note that one could solve a joint problem to simultaneously estimate the correlation information D, and the enhanced quality images $\tilde{I}_1$ and $\tilde{I}_2$. However, such a joint optimization problem would be hard to solve with a complex objective function. Therefore, we propose to split the problem in two steps: (i) we estimate the correlation information from the decoded images; and (ii) we perform joint reconstruction using the estimated correlation information. These two steps are detailed in the rest of this section.

A. Depth Estimation

The first task is to estimate the correlation between multi-view images, which typically consists in a dense depth image. It is generally estimated by matching the corresponding pixels between correlated images. A comprehensive overview of several depth estimation algorithms has been presented. In this work, we estimate a dense depth image from the decoded images in a regularized energy minimization framework, where the energy E is composed of a data term E_d, and a smoothness term E_s. A dense depth image D is obtained by minimizing the energy function E as

$$D = \arg\min_{D_c} \mathrm{E}(\mathrm{D}_c) = \arg\min_{D_c} \{\mathrm{E}_d(D_c) + \lambda E_s(D_c)\}, \quad (1)$$

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

where λ balances the importance of the data and smoothness terms, and D_c represents the candidate depth images. The candidate depth values $D_c(m,n)$ for every pixel position (m,n) are discrete; this is constructed by uniformly sampling the inverse depth in the range $[1/D_{\max}, 1/D_{\min}]$, where $D_{\min}$ and $D_{\max}$ are the minimum and maximum depth values in the scene, respectively.

We now discuss in more details the components of the energy function of Eq. (1). The data term E_d is used to match the pixels across views by assuming that the intensity is consistent irrespective of the viewpoints. It is computed as

$$E_d(D_c) = \sum_{m=1}^{N_1}\sum_{n=1}^{N_2} \| \tilde{I}_2(m,n) - \omega \tilde{I}_1(m,n), D_c(m,n)) \|_2^2, \quad (2)$$

where N_1 and N_2 represent the image dimensions. W represents a warping function that warps the image $\tilde{I}_1$ using the depth information D_c. This warping, in general, is a two-step process [23]. First the pixel position (m, n) in the image $\tilde{I}_1$ is projected to the world coordinate system. This projection step is represented as

$$[u,v,w]^T = R_1 P_1^{-1}[m,n,1]^T D_c(m,n) + T_1, \quad (3)$$

where P_1 is the intrinsic camera matrix of the camera C_1, and (R_1, T_1) represent the extrinsic camera parameters with respect to the global coordinate system. Then, the 3D point $[u,v,w]^T$ is projected on the coordinates of the camera C_2 with the internal and external camera parameters, respectively as P_2 and (R_2, T_2). This projection step can be described as

$$[x',y',z']^T = P_2 R_2^{-1}\{[u,v,w]^T - T_2\}. \quad (4)$$

Finally, the pixel location of the warped image is taken as $(m',n') = (round(x'/z'), (round(y'/z'))$ where $round(x)$ rounds x to the nearest integer. The smoothness term E_s is used to enforce consistent depth values at neighboring pixel locations. It is measured as

$$E_s(D_c) = \sum_{(m,n),(\tilde{m},\tilde{n})\in N} \min(|D_c(m,n) - D_c(\tilde{m},\tilde{n})|, \tau), \quad (5)$$

Where N represents the usual four-pixel neighborhood. The parameter τ sets an upper level on the smoothness penalty such that discontinuities can be preserved. Finally, a depth image can be estimated by solving the optimization problem of Eq. (1), which is usually a non-convex problem. Several minimization algorithms exist in the literature to solve Eq. (1), e.g., Simulated Annealing, Belief Propagation Graph Cuts. Among these solutions, the optimization techniques based on Graph Cuts algorithm compute the minimum energy in polynomial time and they generally give better results than the other techniques. Therefore, in our work, we solve the minimization problem of Eq. (1) using Graph Cuts techniques.

B. *Image Warping as Linear Transformation*

Before describing our joint reconstruction problem, we show that the image warping operation $W(\tilde{I}_1, D)$ in Eq. (2) can be written as matrix multiplication of the form $A.R(\tilde{I}_1)^1$; this linear representation offers a more flexible formulation of our joint reconstruction problem. The reshaping operator $R: I_{N_1\times N_2} \rightarrow X_{N_1N_2\times 1}$ produces a vector $X = R(I) = [I_{.,1}^T I_{.,2}^T \ldots I_{.,N_1}^T]^T$ from the matrix I, where $I_{.,m}$ represents the row m of the matrix I, and $(.)^T$ denotes the usual transpose operator. For our convenience, we also define another operator $R^{-1}_{N_1\times N_2} : X_{N_1N_2\times 1} \rightarrow I_{N_1\times N_2}$ that takes the vector $[R(I)]_{N_1N_2\times 1}$ and gives back the matrix $I_{N_1\times N_2}$, i.e., this operator R^{-1} performs the inverse operations corresponding to R. The matrix A describes the warping by re-arranging the elements of $R(\tilde{I}_1)$. Its construction is described in this section.

Recall that the warping function W in Eq. (2) shifts the intensity at the pixel position (m,n) in the reference image to the position (m',n') in the target image. Alternatively, this pixel shift between images can be represented using a

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

horizontal component m^h and a vertical component m^v of the motion field as $(m', n') = (m + m^h(m,n), n + m^v(m,n))$. Note that this motion field (m^h, m^v) can be easily computed from Eqs. (3) and (4), once the depth information and the camera parameters are known. By knowing this motion field (m^h, m^v), we can represent the image warping step $W(\tilde{I}_1, D)$ as a linear transformation of the form

$$\underbrace{\begin{bmatrix} \tilde{I}_{2,1}^T \\ \tilde{I}_{2,2}^T \\ \vdots \\ \tilde{I}_{2,N_1}^T \end{bmatrix}}_{R(\tilde{I}_2)} = \underbrace{\begin{bmatrix} A^1 \\ A^2 \\ \vdots \\ A^{N_1} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \tilde{I}_{1,1}^T \\ \tilde{I}_{1,2}^T \\ \vdots \\ \tilde{I}_{1,N_1}^T \end{bmatrix}}_{R(\tilde{I}_1)}. \quad (6)$$

Here, $\tilde{I}_2 = W(\tilde{I}_1, D)$ represents the warped image and A^m is a matrix of dimensions $N_2 \times N_1 N_2$ whose entries are determined by the horizontal and vertical components of the motion field in the row m, i.e., $m^h(m, .)$ and $m^v(m, .)$.

In general, the matrix A^m can be constructed in two ways: (i) forward warping; and (ii) inverse warping. In this work, we propose to construct the matrix A^m based on forward warping; this permits easier handling of the occluded pixels, as shown later. The elements of the matrix A^m are given as

$$A^m(n - \beta_1 - \beta_2 N_2, n) = \begin{cases} 1 & \text{if } m^h(m,n) = \beta_1, \\ & \text{and } m^v(m,n) = \beta_2, \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

If $n - \beta_1 - \beta_2 N_2 < 0$ (e.g., at image boundaries), we set $n - \beta_1 - \beta_2 N_2 = 1$, so that the dimensions of the matrix A^m remains $N_2 \times N_1 N_2$. The matrix A^m formed using Eq. (7) may contain multiple entries with values of '1' in each row; this is because several pixels in the source image can be mapped to the same location in the destination image during forward warping. In such cases, for a given row index m, we keep only the last '1' entry, while the remaining ones in the row m are set to zero. This is motivated by the fact that, during forward warping when multiple source pixels are mapped to the same destination point (m', n'), the intensity value of the last source pixel is assigned to the destination pixel (m', n')[2]. Furthermore, some of the rows in the matrix A^m do not contain any entry with value of '1', i.e., all entries in row m of the matrix A^m are zeros. This means that the set of pixel locations $\{j : j \in J^m\}$ in the warped image $\tilde{I}_{2,m}(j)$ has zero value, where J^m is the set of row indexes in the matrix A^m that do not contain any entry with value of '1'. These pixel positions represent holes in the warped image and correspond to the occluded regions.

Based on the above arguments, it is easy to check that the constructed matrix A^m shifts the pixels in $\tilde{I}_1$ in order to form $\tilde{I}_{2,m}$, whose entries are given as

$$\tilde{I}_{2,m}(j) = \begin{cases} 0 & \text{if } j \in J^m \\ \tilde{I}_1(k,n) & \text{if } A^m(j, (k-m)N_2 + n) = 1. \end{cases} \quad (8)$$

By constructing the matrix $A^m, \forall m \in \{1, 2, \ldots, N_1\}$ and using it in Eq. (6), the image warping can be represented as a linear transformation. Finally, note that similar operations can also be performed with an inverse mapping.

C. Joint Reconstruction

We now discuss our novel joint reconstruction algorithm that takes benefit of the estimated correlation information given by the matrix A (or D) in order to reconstruct the images. We propose to reconstruct an image pair $(\tilde{I}_1, \tilde{I}_2)$ as a solution to the following optimization problem:

$$(\tilde{I}_1, \tilde{I}_2) = \underset{I_1, I_2 \in R^{N_1 \times N_2}}{\arg\min} \; (\|I_1\|_{TV} + \|I_2\|_{TV})$$
$$\text{s.t. } (\|R(I_1) - R(\tilde{I}_1)\|_2 \le \epsilon_1,$$
$$(\|R(I_2) - R(\tilde{I}_2)\|_2 \le \epsilon_1,$$
$$(\|R(I_2) - A \cdot R(I_1)\|_2^2 \le \epsilon_2. \quad (9)$$

Similarity Weights by Feature Correlation

The increasing order series of the features and their happening in the set of multi-view images set is known as concept weight

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

(cw). The ordered series is originally considered with lone feature and then it rises by adding every feature per iteration. The sequence is terminated once the concept weight is found to be less than the said threshold. In this process if the series s1 is subset of sequence s2 and concept weight of s1 is less than or equal to concept weight of s2 then s1 can be trimmed [46]. This procedure results set of hypothesis as feature set 'CFS'

Feature pruning by similarity score

Feature with many terms will be got in this phase and the obtain the parallel score in each selected feature x and other each feature x' with lesser number of terms than x. in case the semblance between x and x1 are found to be more than the said threshold and bigger than all of the similarity scores between x' and rest of bigger length features selected then x1 will be grouped to the x.

The algorithmic approach of the feature set optimization in pseudo code format

Let ts be the terms set selected from CFS.

Let $c\omega_\tau$ be the concept weight threshold

Order the terms belongs to 'ts' in descending by their frequency score.

1. For each term $\{t_i \mid t_i \in ts\}$:

Begin

a. Let $\{c_i \mid c_i \in cs\}$

b. $c_i \leftarrow t_i$ (add t_i to c_i)

c. For each term $\{t_i^{'} \mid t_i^{'} \in ts, t_i^{'} \neq t_i\}$

Begin

Project concept $c_i^{'}$ by adding $t_i^{'}$ to c_i in sequence

If ($cw(c_i^{'}) \geq c\omega_\tau$)

If ($cw(c_i^{'}) \cong cw(c_i)$ then

Discard c_i ;

Set $c_i \leftarrow c_i^{'}$ continue step c.

Else

$cs \leftarrow c_i^{'}$ (add $c_i^{'}$ to cs)

Continue step c;

Else

Discard $c_i^{'}$

Continue step c;

End of step c;

End of step 1.

Order cs in descending order by the length of the concepts (here after concepts referred as features)

Let ml be the maximal length of the feature in cs

For each $\{c_i \mid c_i \in cs; tc(c_i) \equiv ml\}$ move c_i to label set ls

2. For each $\{l_i \mid l_i \in ls; tc(l_i) \cong ml\}$;Here $tc(c_i)$ indicates the term count of the feature c_i

Begin

a. For each $\{c_i \mid c_i \in cs; c_i \notin ls\}$

Begin

Find similarity score $ss_{(l_i \leftrightarrow c_i)} = \frac{tc(l_i \cap c_i)}{tc(l_i \cup c_i)}$

End of Step a;

End of Step 2;

3. For each $\{c_i \mid c_i \in cs; c_i \notin ls\}$

Begin

a. For each $\{l_j \mid l_j \in ls\}$ set $\bigcup_{j=1}^{|ls|} ss_{(l_j \leftrightarrow c_i)}$ in descending order and select first element as $ss_{l \leftrightarrow c_i}$

b. If ($ss_{l \leftrightarrow c_i} \geq ss_\tau$) then consider c_i as feature of the group represented by the label l

Else move c_i to ls

End of step 3;

If ls got updated then go to step 2 else

Return ls as set of class labels

Further the classification of the multi-view images set is initiated that performs supervised learning by using concept labels as the labels of the categories

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Finding Correlation of the semantics

This stage of supervised learning estimates the correlation between activities that extracted from the given multi-view images set dataset. In this regard the activities found are considered to be categorical as they associate with divergent arguments. Henceforth here we use mean-square contingency coefficient [22] to estimate the correlation between attributes. Any given two activities A and B such that $\{a_1, a_2, a_3, \ldots..a_m\}$, $\{b_1, b_2, b_3, \ldots.b_n\}$ are categorical arguments found to be associated to A and B respectively. The size of the set of arguments associated with activity A is m and activity B is n. Then the mean square contingency coefficient between activities A and B can be measured as follows:

$$\rho_{ij} = \sum_{i=1}^{m}\sum_{j=1}^{n} 1 - \frac{1}{o(a_i, b_j)}$$

Here in this equation ρ_{ij} is the fraction of co occurrence of a_i, b_j

$$\rho_i = \sum_{i=1}^{m} 1 - \frac{1}{o(a_i)}$$

Here in this equation ρ_i is the fraction of occurrence of a_i

$$\rho_j = \sum_{j=1}^{n} 1 - \frac{1}{o(b_j)}$$

Here in this equation ρ_j is the fraction of occurrence of b_j

$$\chi^2_{(A \leftrightarrow B)} = \frac{1}{\min(m,n) - 1} * \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{\left(\rho_{ij} - (\rho_i . \rho_j)\right)^2}{\rho_i . \rho_j}$$

Here in this equation $\chi^2_{(A \leftrightarrow B)}$ is the mean square contingency coefficient that indicates the correlation between activities A and B. According to the correlation estimation process explored here, the activities that are highly correlated will be grouped. Further each group of activities will be used as class label for second level of supervised learning.

EXPERIMENTS AND EXPLORATION OF THE RESULTS

Number ofsequence images taken as input are 771, which are related to a football match. In the preprocessing step, around 750 multi-view image groups were formed. The groups and their correlation thresholds are listed in table 1. The performance analysis of the proposal was done by comparing its compression ratio with the model devised in [14]. The results are indicating that the compression ratio is significantly in proposed model, which due to the process of correlation analysis (see figure 2). The figure 3 indicates the group correlation thresholds observed during preprocess.

groups	mean correlation
1-100	0.238281615
101-200	0.219613847
201-300	0.20793224
301-400	0.156296504
401-500	0.170067523
501-600	0.162341092
601-700	0.170406263
700-750	0.163233233

Table 1: The mean correlation of each 100 groups of multiview images

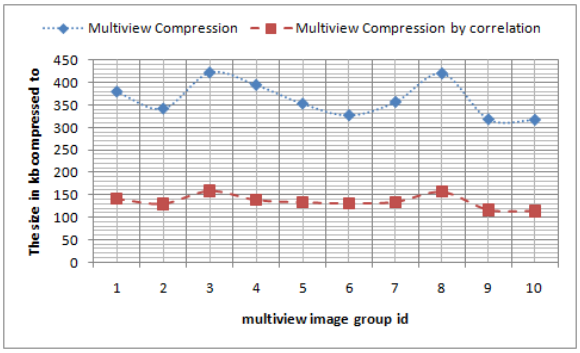

Figure 2: The multiview image group compressed to the size in KB

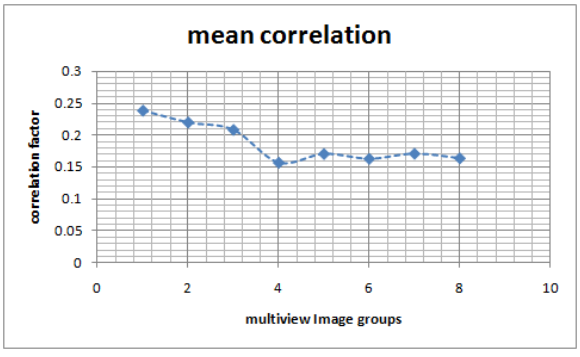

Figure 3: image correlation considered to form the multi-view image groups

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

CONCLUSIONS

We have designed a new scheme for jointly decoding independently- and compressively-sampled multi-view video streams. Simulation results showed that the proposed joint decoder outperforms the independent CS-decoder in the case of both fast and moderate motion levels. The accuracy of a newly-proposed blind video quality estimation method was also verified.

REFERENCES

[1] E. J. Candes and M. B. Wakin, "An Introduction to Compressive Sampling," IEEE Signal Processing Magazine, vol. 25, no. 2, pp. 21-30, March 2008.

[2] D. L. Donoho, "Compressed Sensing," IEEE Transactions on Informa¬tion Theory, vol. 52, no. 4, pp. 1289-1306, April 2006.

[3] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, "A Survey on Wireless Multimedia Sensor Networks," Computer Networks, vol. 51, no. 4, pp. 921-960, March 2007.

[4] S. Pudlewski, T. Melodia, and A. Prasanna, "Compressed-sensing En¬abled Video Streaming for Wireless Multimedia Sensor Networks," IEEE Transactions on Mobile Computing, vol. 11, no. 6, pp. 1060-1072, June 2012.

[5] D. Slepian and J. K. Wolf, "Noiseless Coding of Correlated Information Sources," IEEE Transactions on Information Theory, vol. 19, no. 4, pp. 471-480, July 1973.

[6] A. D. Wyner and J. Ziv, "The Rate-distortion Function for Source Coding with Side-information at the Decoder," IEEE Transactions on Information Theory, vol. 22, no. 1, pp. 1-10, Jan. 1976.

[7] S. Pudlewski and T. Melodia, "A Tutorial on Encoding and Wireless Transmission of Compressively Sampled Videos," IEEE Communica¬tions Surveys & Tutorials, vol. 15, no. 2, pp. 754-767, Second Quarter 2013.

[8] H. W. Chen, L. W. Kang, and C. S. Lu, "Dynamic Measurement Rate Allocation for Distributed Compressive Video Sensing," Visual Communications and Image Processing, vol. 7744, pp. 1-10, July 2010.

[9] Y. Liu, M. Li, and D. A. Pados, "Motion-aware Decoding of Compressed-sensed Video," IEEE Transactions on Circuits System Video Technology, vol. 23, no. 3, pp. 438-444, March 2013.

[10] X. Chen and P. Frossard, "Joint Reconstruction of Compressed Multi¬view Images," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Taipei, Taiwan, April 2009.

[11] M. Trocan, T. Maugey, E. W. Tramel, J. E. Fowler, and B. Pesquet- Popescu, "Compressed Sensing of Multiview Images Using Disparity Compensation," in Proc. IEEE International Conference on Image Processing (ICIP), pp. 3345-3348, Hong Kong, Sep. 2010.

[12] F. H. Jamil, R. R. Porle, A. Chekima, R. A. Lee, H. Ali, and S. M. Rasat, "Preliminary Study of Block Matching Algorithm (BMA) for Video Coding," in Proc. International Conference On Mechatronics (ICOM), Istanbul, Turkey, May 2011.

[13] A. M. Huang and T. Nguyen, "Motion Vector Processing Using Bidirec¬tional Frame Difference in Motion Compensated Frame Interpolation," in Proc. IEEE International Symposium on A World of Wireless, Mobile and Multimedia Networks, Newport Beach, CA, USA, June 2008.

[14] VijayaraghavanThirumalai and Pascal Frossard; Joint Reconstruction of Multiview Compressed Images; IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 22, NO. 5, MAY 2013