SCHOOL OF ENGINEERING - STI
ELECTRICAL ENGINEERING INSTITUTE
SIGNAL PROCESSING LABORATORY

Pascal Frossard

EPFL - FSTI - IEL - LTS
Station 11
Switzerland-1015 LAUSANNE

Phone: +4121 6932601
Fax: +4121 6937600
e-mail: `pascal.frossard@epfl.ch`

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

SYMMETRIC DISTRIBUTED CODING OF STEREO OMNIDIRECTIONAL IMAGES

Vijayaraghavan Thirumalai, Ivana Tosic and Pascal Frossard

Ecole Polytechnique Fédérale de Lausanne (EPFL)
Signal Processing Laboratory

Technical Report LTS-2008-002

February 15, 2008

Symmetric Distributed Coding of Stereo Omnidirectional Images

Vijayaraghavan Thirumalai, Ivana Tosic and Pascal Frossard

Abstract

This paper presents a distributed coding scheme for the representation of 3D scenes captured by omnidirectional cameras. We consider a scenario with a pair of similar cameras that benefit from equivalent bandwidth and computational resources. The images are captured at different viewpoints and encoded independently, while a joint decoder exploits the correlation between images for improved decoding quality. The distributed coding is built on the multiresolution representation of spherical images, whose information is split into two partitions. The encoder then transmits one partition after entropy coding, as well as the syndrome bits resulting from the channel encoding of the other partition. The joint decoder exploits the intra-view correlation by predicting one partition from the other partition. At the same time, it exploits the inter-view correlation by using motion estimation between images from different cameras. Experiments demonstrate that the distributed coding solution performs better than a scheme where images are handled independently. Furthermore, the coding rate stays balanced between the different cameras, which interestingly permits to avoid hierarchical relations between vision sensors in camera networks.

I. INTRODUCTION

Camera networks find widespread usage in several applications that rely on the effective representation of scenes or the analysis of 3D information. These networks normally consist of several cameras distributed in the scene, and pose several problems like the coding of multi-view images, the reconstruction of the 3D structure from multiple views, or the multi-view object recognition, for example. This paper focuses on the compression of multi-view images and particularly stereo omnidirectional images. The images captured from different viewpoints are usually correlated, which permits to reduce the coding rate by exploiting efficiently the redundancy between the different views. Instead of joint encoding that requires communication between cameras, we rely on the Slepian-Wolf theorem [1] and design a distributed coding scheme where images are encoded independently, but decoded jointly in order to exploit the correlation between the sources, as illustrated in Fig. 1.
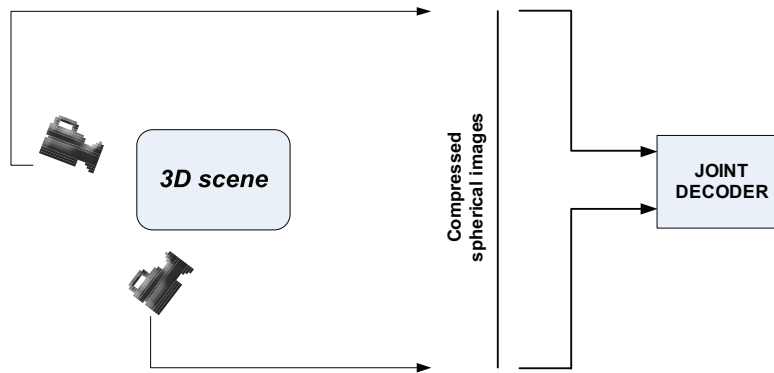

Fig. 1. Distributed source coding of the 3D scenes. The correlated images are compressed independently and are decoded jointly

Most of the research carried out on distributed coding for multi-view images or videos propose solutions based on coding with side information. One of the cameras is chosen as the primary source and its output is encoded independently. The other cameras represent secondary sources whose rate can be drastically reduced if the joint decoder uses the primary source as side information. Such a coding scheme obviously does not balance the transmission rate between the encoders. However, it is often interesting in practice to avoid hierarchical relations between sensors and to distribute the coding and transmission cost equally among the sensors. We therefore concentrate on symmetric coding schemes in this paper, where all cameras are equally important in the representation of the 3D scenes.

We propose a rate-balanced distributed coding scheme for stereo omnidirectional cameras, whose images can be appropriately mapped to spherical images [33]. Compared to perspective stereo images, spherical images certainly provide great advantage and increased flexibility in representing the 3D scene due to the wider field of view of the sensors. The images captured by omnidirectional cameras distributed in the 3D scene are likely correlated, and we design a distributed coding scheme that is able to exploit the redundancy between spherical images. The omnidirectional images initially undergo a multi-resolution decomposition based on the Spherical Laplacian Pyramid, which brings the advantage of shift invariance. The resulting sets

of coefficients are split into two partitions. The coefficients of the first partition are quantized and entropy coded, and sent to the decoder. The second partition is encoded using the Nested Scalar Quantization (NSQ) [9], which is a binning scheme that encompasses a scalar quantizer and a coset encoder. It outputs the coset bin indexes and permits to reduce the coding rate compared to encoding the quantized coefficients directly. The coset bin indexes are further encoded using a Slepian-Wolf encoder [1] based on multi-level LDPC codes [10], [11], in order to achieve further compression. The resulting syndrome bits are finally transmitted to the joint decoder.

The joint decoder uses the syndrome bits to decode the missing information, and further implements motion estimation on the sphere to take benefit of the correlation between views. The proposed scheme therefore efficiently combines the intra and inter Wyner-Ziv image coding, which allows for a balanced coding rate between cameras. Such a strategy proves to be beneficial with respect to independent processing of the omnidirectional images and shows only a small performance loss compared to joint encoding of the different views. Finally, exploiting the inter-view correlation by motion estimation is highly advantageous in camera networks where no knowledge of the camera parameters is given or when camera network calibration is not achievable in practice. The proposed scheme therefore provides a low-complexity coding solution for the representation of 3D scenes, which does not require complex setup nor hierarchical organization between vision sensors.

The rest of the paper is organized as follows. Section 2 overviews the related work in distributed coding with a special focus on camera networks. Section 3 presents the distributed coding algorithm adapted to omnidirectional images. Section 4 presents in more details the Wyner-Ziv coding strategy, while Section 5 describes the joint decoding scheme. Section 6 finally presents the experimental results that show the benefits of the proposed solution.

II. Related Work

The first information-theoretical results on distributed source coding appeared already in the late seventies. In particular, it has been shown that independent coding of correlated sources can achieve the same rate-distortion bound as joint encoding if a joint decoder can efficiently exploit the correlation between the sources [1]. Rate-distortion bounds have been established later for the particular case of coding with side information [2]. However, most results presented in [1], [2] have remained non-constructive for about three decades. Practical DSC schemes have been designed only recently, by establishing a relation between the Slepian-Wolf theorem and channel coding [3]. Subsequently, several practical DSC systems have been presented using different channel codes, e.g., Turbo codes [4], [5] or LDPC codes [6]. These results have been used to construct low complexity video coders by assuming that adjacent frames represent correlated sources. These frames are categorized into key frames and WynerZiv frames, where the key frames are encoded independently and the Wyner-Ziv frames are Slepian-Wolf encoded [12], [13], [14].

Only a few studies have been reported however about the application of distributed coding principles to camera networks. The few works reported in the literature generally consider the particular case of coding with side information, where one camera is used a reference to decode the information from the other cameras. For example, in [15], [20], [17] the cameras are categorized into reference and Wyner-Ziv cameras and the correlation among views is exploited at the joint encoder with disparity estimation. When camera parameters are not available and calibration is not possible, the joint decoder can rather use motion estimation to exploit the redundancy between images [18], [19]. These schemes however introduce a hierarchical relation among the sensors and the coding rate is therefore not balanced.

In practice however, it is often interesting to avoid hierarchical relations between sensors and to further balance the coding and transmission costs among the sensors. One of the first works that addresses balanced rate allocation in distributed coding is based on time sharing mechanisms [27], which are however hard to implement due to node synchronization issues. The first practical scheme for symmetric coding based on channel code partitioning has been proposed in [29]. This scheme has been later extended to multiple sources using systematic channel codes by Stankovic *et al.* [30]. It is based on horizontally splitting the generator matrix of the channel code into two sub-generator matrices. Codewords are then generated using the sub-matrixes and assigned to each encoder. The compression rate of each encoder is determined by the number of rows retained in the corresponding sub-matrix. The advantage of this system is the need for only one channel code. However, this framework is limited to systematic channel codes. The authors in [35] have developed a symmetric DSC using a general linear channel code (which includes both systematic and non systematic channel codes) and their framework is based on algebraic binning concept of the channel code. Simulation results have shown that almost the entire Slepian-Wolf region can be covered with this coding algorithm.

Symmetric distributed coding can also be achieved by information partitioning. Sartipi *et.al.* [31] have considered the compression of two sources at the symmetric rate by information partitioning, where half of the source bits are transmitted directly while the corresponding syndrome bits are generated on the other half (complementary part) of the source bits. Similarly to [35], the authors show that they can approach the entire SW region and thus the decoding error is insensitive to arbitrary rate allocation among the encoders. However, both schemes are based on capacity approaching channel codes, and they approach the Slepian-Wolf bound only for long source length (typically 10^4). Grangetto *et. al* [36] have proposed a balanced coding scheme for small block length binary sources. The algorithm is based on time sharing and distributed arithmetic codes that perform better than the turbo based DSC scheme in the considered framework.
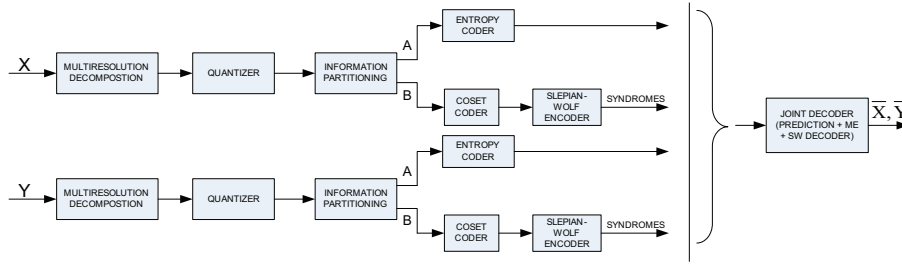
Fig. 2. Overview of the symmetric distributed coding scheme.

Symmetric coding has also been proposed for coding image sequences. In [16], the authors propose to divide each frame into two partitions and to transmit directly one partition. In addition, each frame is Wyner-Ziv encoded and the side information is eventually generated using disparity estimation. This scheme permits to avoid hierarchical relations between frames. However, it results in high coding rates, since one of the partitions in each frame is encoded using both Wyner-Ziv and independent coding. Finally, a balanced distributed coding scheme for camera networks has been proposed in [32], based on linear channel code construction that can achieve any point in the Slepian-Wolf region. The developed linear codes were not however applied to the practical coding of images in camera networks.

In this paper, we present a balanced distributed coding scheme for multi-view image coding. We focus on omnidirectional images and partition multiresolution information in different subsets. The coefficients are Slepian-Wolf encoded, and both intra- and inter-view correlation is exploited at the joint decoder. Motion estimation is used to generate the side information, which does not require explicit knowledge of the camera parameters nor precise calibration, in contrary to disparity estimation.

III. Distributed coding scheme

A. Overview

This section presents an overview of the symmetric distributed coding scheme, illustrated in Fig.2. We consider omnidirectional images that can be exactly mapped on the sphere, as those captured by catadioptric mirrors [38]. Since these images have a single center of projection, they permit to capture the plenoptic function [40] in its radial form without discrepancies that usually appear due to Euclidian assumptions for planar images. The wide field of view is another advantage of omnidirectional cameras for the representation of 3D scenes. The stereo omnidirectional images undergo a multiresolution decomposition, whose coefficients are then quantized. As we target a balanced rate between cameras, the information of both images are partitioned in a similar way. The quantized coefficients are split into two subsets or partitions: A and B. These partitions are generally correlated due to the simple partitioning process, which puts coefficients alternatively in both partitions. The first partition (e.g., partition A) is transmitted directly to the joint decoder after entropy coding. The second partition (B) is coset encoded and the resulting coset indexes are Slepian-Wolf encoded with a multilevel LDPC code. Hence, each encoder transmits one half of the quantized coefficients, and only the syndrome bits for the coefficients of the second partition.

The joint decoder tries to exploit intra- and inter-view correlation for improved decoding performance. It estimates the coefficients of the partition B by using the coefficients of the partition A. Under the assumption that the images X and Y are correlated, the predicted coefficients are further refined using motion estimation (ME) on the sphere [23]. Motion estimation permits to compensate for the displacement of the objects captured from different viewpoints. Prediction and motion estimation together lead to effective side information, which permits to reduce the channel rate of the Slepian-Wolf encoder. The coset indexes for the coefficients in each subband are further recovered after correcting the virtual channel noise in the side information using the corresponding syndromes. The subband coefficients are then estimated from the recovered coset index, and the images are finally reconstructed by inverse transform.

B. Spherical Laplacian Pyramid

Multiresolution analysis is an efficient tool that permits to decompose a signal at progressive resolutions and perform coarse to fine computations on the data. The two most successful embodiments of this paradigm are the various wavelet decompositions [24] and the Laplacian Pyramid (LP) [25]. As shift invariance represents an interesting property for distributed coding in camera networks, we have chosen to use the Laplacian Pyramid on our scheme. It proves to be beneficial for predictive coding based on motion estimation. Furthermore, since we work with omnidirectional images, we propose to use the Laplacian Pyramid on the sphere (SLP), which is presented below.

Let first $L^2(S^2, d\mu)$ denote the Hilbert space of the square integrable signals on the 2D sphere S^2, where $d\mu(\theta, \phi) = sin\theta \, d\theta \, d\phi$ represents the rotation invariant Lebesgue measure on S^2. Any spherical signal $F \in L^2(S^2)$ can be expanded

using the spherical harmonics $W_{m,n}$, whose Fourier coefficients are given by

$$\check{F}(m,n) = \int_{S^2} d\mu(\theta,\phi) W_{m,n}^*(\theta,\phi) F(\theta,\phi) \ , \tag{1}$$

where $W_{m,n}^*$ is the complex conjugate of the spherical harmonic of order (m,n). The omnidirectional images are sampled on the nested equi-angular grids on the sphere described as:

$$\mathcal{G}_j = \{(\theta_{jp},\varphi_{jq}) \in S^2 : \theta_{jp} = \tfrac{(2p+1)\pi}{4B_j}, \varphi_{jq} = \tfrac{q\pi}{B_j}\}, \tag{2}$$

$p,q \in \mathcal{N}_j \equiv \{n \in \mathbb{N} : n < 2B_j\}$, for a range of bandwidth $B = \{B_j \in 2\mathbb{N}, j \in \mathbb{Z}\}$. These grids permit to perfectly sample any band-limited function $F \in L^2(S^2)$ of bandwidth B_j, i.e., such that $\check{F}(m,n) = 0$ for all $m > B_j$. This class of sampling grids are advantageously associated to a Fast Spherical Fourier Transform [39], which permits rapid transformation.

Similar to classical Laplacian Pyramid decomposition, the Spherical Laplacian Pyramid proceeds first by low pass filtering of the spherical signal in the Fourier domain for speeding up the computations. Suppose the original data F_0 is bandlimited, i.e, $\check{F}_0(m,n) = 0$, $\forall m > B_0$, and sampled on $\mathcal{G}_0$. We capture the low frequency information by using an axisymmetric half-band axisymmetric filter $\check{H}_{\sigma_0}$. The bandwidth parameter σ_0 is chosen so that the filter is numerically close to a perfect half-band filter $\check{H}_{\sigma_0}(m) = 0$, $\forall m > B_0/2$. The low pass filtered data is then downsampled on the nested sub-grid $\mathcal{G}_1$, which gives the low-pass channel of our spherical laplacian pyramid F_1. The high-pass channel of the pyramid is computed as usual, that is by first upsampling F_1 on the finer grid $\mathcal{G}_0$, low-pass filtering it with H_{σ_0} and taking the difference with F_0. Coarser resolutions are computed by iterating this algorithm on the low-pass channel F_l and scaling the filter bandwidth accordingly, i.e., $\sigma_l = 2^l \sigma_0$.

The coefficients of the Spherical Laplacian Pyramid need to be quantized with efficient rate distribution among the subbands. We follow the algorithm proposed for the Laplacian Pyramid in [8]. The rate allocation can be computed by Lagrange's multipliers method when the quantizers are uniform. Unsurprisingly, the rate in the different subbands is chosen to be proportional to the variance of the coefficients.

IV. SLEPIAN-WOLF ENCODER

A. Coefficient partitioning

We describe now in more details the symmetric distributed coding scheme illustrated in Fig. 4. The omnidirectional images undergo a multiresolution decomposition with the Spherical Laplacian Pyramid, whose coefficients are then quantized uniformly with proper rate allocation among the subbands, as described above. The quantized coefficients of each image are then distributed alternatively into two correlated partitions that form a kind of checkerboard pattern. Let (θ,ϕ) denote the coordinate of the coefficients in the subband U_j. The coefficients are put in the partition A if $((\phi \mod 2) \; XOR \; (\theta \mod 2)) = 1$. Otherwise, they are put in the partition B.

The quantized indexes in the partition A are compressed using the Arithmetic encoder and transmitted directly to the joint decoder. The partition B however further undergoes coset and Slepian-Wolf coding to save bit rate in the distributed coding scheme. The quantized coefficients in partition B are put in different cosets. The cosets group coefficients from the jth subband that are separated by a distance d_j. Only the coset indexes are eventually encoded, which provides some significant rate savings. The coset distance is estimated as $d_j = 2^{\lceil \log_2(2E_j+1) \rceil}$, where E_j is the maximum error between the original and the side information images in the jth subband. Scalar quantization and coset encoding together behave similarly to nested scalar quantization.

B. Multilevel LDPC coding

Even after coset encoding, some correlation still exists between the coefficients and the side information available at the decoder. We propose to achieve further compression by encoding the coset indexes with multilevel LDPC codes [22]. In other words, instead of sending the coset indexes to the decoder, the encoder only transmits the syndrome bits resulting from LDPC encoding. We propose to use irregular LDPC codes and we follow the procedure described in [11], [21] in order to construct the parity check matrix. The decoder is eventually able to reconstruct the coset indexes from the syndrome bits and the side information. We describe now in more details the multilevel LDPC coding, and the channel rate estimation.

The coset indexes from each subband are first decomposed into n bit planes b_0, b_1, ...b_{n-1}, where b_0 represents the most significant bit plane (MSB) and b_{n-1} represents the least significant bit plane (LSB). Each bitplane b_i $(0 \le i \le n-1)$ is encoded by the LDPC encoder, starting from the bit plane b_0, as illustrated in Fig.3. The code rate is chosen by assuming that the error between the SLP subbands and the side information follows a Laplacian distribution. The subsequent bit planes are also encoded with LDPC codes, where the code rate is however adapted based on the previously encoded bitplanes.

In more details, the coding rate for encoding the mth bitplane of the jth subband U_j is estimated as follows. First, the conditional probability $p(i) = Pr(b_m(i) = 1 | \tilde{U}_j(i), b_0(i), b_1(i), .., b_{m-1}(i))$ is calculated for each bit i, where $\tilde{U}_j$ denotes the
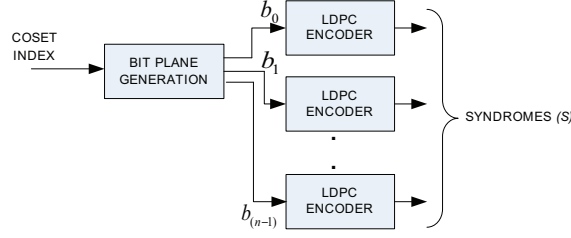
Fig. 3. Block scheme of the LDPC encoder bank

side information for decoding the subband U_j. The rate of the LDPC encoder is then chosen to be equal to the following conditional entropy:

$$H(b_m|\tilde{U}_j, b_0, b_1, ...b_{m-1}) \simeq \frac{1}{N} \sum_{i=1}^{N} H(p(i)), \tag{3}$$

where N denotes the number of bits in the bit plane [28]. However, the side information is unfortunately not available at the encoder, and the conditional probability p has generally to be estimated as described below.

C. Noise models

One of the main difficulty in distributed coding is the estimation of the correlation between the sources, or equivalently the construction of noise models for the proper design of the Slepian-Wolf encoder. The encoder has to estimate the noise distribution in order to determine the coset distance, and the LPDC coding rate. Unfortunately, the side information that is used for joint decoding is only present at the decoder, and the encoder can only predict the noise distribution.

We assume that the error E_j between the jth SLP subbands U_j and the corresponding subband $\tilde{U}_j$ in the side information image follows a Laplacian distribution, of the form $f_{E_j}(e) = \frac{1}{2\lambda_j} exp(-\frac{|e|}{\lambda_j})$. The Laplacian distribution is a common assumption in such a scenario, and it provides a good approximation of the actual distribution of the error, as shown later. In this case, the rate R_j necessary to code the error E_j is equivalent to the conditional entropy $H(U_j|\tilde{U}_j)$. When the quantization is uniform with stepsize δ, the rate depends only on the variance of the Laplacian distribution [41]. It can be written as:

$$R_j = H(U_j|\tilde{U}_j) = \alpha_j \ log_2(\frac{\lambda_j}{\delta}) + \beta_j, \tag{4}$$

where α_j, β_j are constants that can be estimated offline on test image sets. The construction of the noise model for a proper choice of the coding parameters therefore consists in estimating the parameter λ_j of the Laplacian distribution.

In our scheme, the side information is actually built on two forms of prediction. We can thus model separately the effect of spatial prediction of the coefficients, and the benefit of motion estimation. In the first case, the encoder can estimate the rate R_j^p that is necessary to correct the error due to spatial prediction of the coefficients. The encoder can implement the coefficient prediction step, since it does not depend on the information from the other sensors. The residual error between the coefficients in the subband U_j and the corresponding subband computed by coefficient prediction $\hat{U}_j$ can then be modeled with a Laplacian distribution. The parameter of the distribution λ_j^p is finally estimated from the prediction error. The rate R_j^p needed to code the prediction error in the jth subband can also be computed by the conditional entropy, as $R_j^p = H(U_j|\hat{U}_j) = \alpha_j \ log_2(\frac{\lambda_j^p}{\delta}) + \beta_j$.

However, the side information is not only built on coefficient prediction, since motion estimation is used at the decoder in order to exploit the correlation between the images from different sensors. We propose to compute a conservative approximation of the gain due to motion estimation, expressed as $\gamma_j = R_j^p/R_j$. It can be computed by offline encodings of several test images, where the complete side information is made available at the encoder. The offline estimation of γ_j finally permits to estimate at the encoders the parameter of the complete noise model. The parameter λ_j of the Laplacian distribution can be expressed as

$$\lambda_j = \lambda_j^p \ 2^{(-\frac{R_j^p(1-\frac{1}{\gamma_j})}{\alpha})}, \tag{5}$$

by combinations of the equations of the rates R_j and R_j^p. As we have now an approximation of the distribution of the error induced by the side information, we can estimate the side information for the jth subband $\tilde{U}_j$ at the encoder. It permits to estimate the error E_j and hence the coset distance for coding the quantized coefficients. Finally, we can estimate the LDPC coding rate by computed the probability $p(i)$, and the conditional entropy given in Eq. (3). The complete encoder is illustrated in Figure 4.
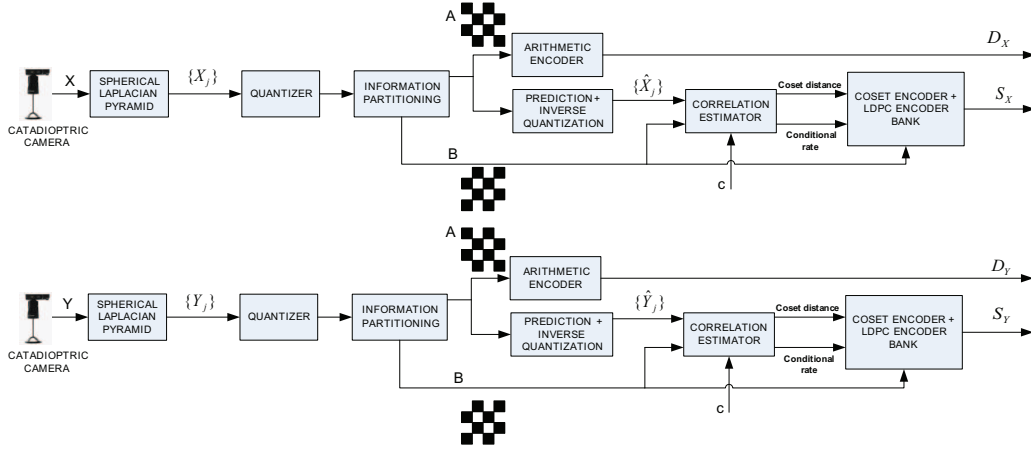
Fig. 4. Detailed block scheme of the Slepian-Wolf encoder with correlation estimation.
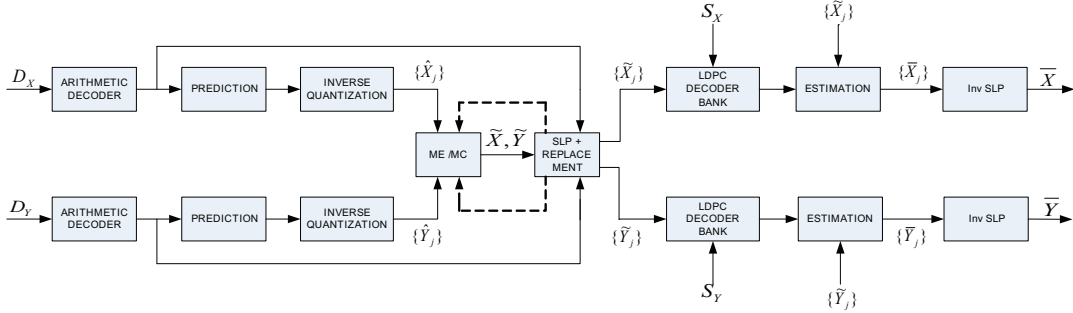

Fig. 5. Block scheme of the Slepian-Wolf Decoder.

V. JOINT DECODING

A. Overview

The joint decoder exploits the correlation between the camera images in order to reconstruct the views of the 3D scenes. The decoding scheme is illustrated in the Fig. 5. The quantized coefficients from partitions A are easily recovered by arithmetic decoding. The coefficients from the partition B however have to be reconstructed by Slepian-Wolf decoding.

These missing coefficients are first predicted in each subband, with the help of the coefficients from the partition A. The unknown coefficients are simply predicted by interpolation from the four nearest neighbours in the partition A. This simple spatial prediction exploits the correlation among successive coefficients. It forms the predicted subband $\hat{U}_j$. The side information is then built by refining the value of the predicted coefficients by motion compensation between the approximations of the different images. The decoder implements motion estimation on the sphere to exploit the redundancy between the omnidirectional images from different cameras. Next, the coset index that corresponds to the coefficients in partition B are then recovered by using the syndrome bits of the LDPC code, as well as the side information created by prediction and motion estimation. Finally, the SLP coefficients are recovered by coset decoding with help of the side information. The main steps of the joint decoder algorithm are detailed in the rest of this section

B. Side information generation

The subbands built on the spatial prediction of the missing coefficients are refined by motion estimation on the sphere. The inverse SLP transform is first applied on the predicted subbands to generate the predicted images $\hat{X}$ and $\hat{Y}$. The correlation between these images is captured by implementing multi-resolution motion estimation on the sphere [23]. The resulting motion vectors are then used for constructing an estimate $\tilde{X}$ of the image X that serves as side information for decoding the image X. In particular, the side information image is first constructed by applying motion compensation from the image $\hat{Y}$. It then undergoes a SLP decomposition, similar to the transform implemented at the encoder. The coefficients corresponding to the partition A are then substituted by the coefficients that have been correctly received from the encoder, in order to reduce the estimation error. The same process is implemented to generate the side information image $\tilde{Y}$ with motion compensation based on the predicted image $\hat{X}$.

7

The exploitation of the correlation between images by motion estimation permits to refine the value of the predicted coefficients from the partition B. The motion estimation process can be repeated on the side information images in order to further improve the image approximation. We have observed empirically that it is advantageous to repeat the motion compensation a second time. This step is represented with a dashed line on the block scheme in the Fig. 5 . Further iterations however do not improve significantly the side information. The resulting side information images are transformed with SLP to build the side information subbands that are used for decoding the coset indexes.

C. Coefficient decoding

The coefficients from partition B are finally recovered by Slepian-Wolf decoding. The side information described above is used by the LDPC decoder together with the syndromes bits (S_x, S_y) to decode the coset indexes in each subband U_j. An LDPC decoder bank uses n LDPC decoders to decode each bit plane successively, starting from MSB to LSB bit planes. While decoding the bit plane b_n, the previously decoded $n-1$ bit planes $b_0, b_1, ..., b_{n-1}$ are used as the additional side information by the LDPC decoder. LDPC decoding is implemented with a Belief propagation algorithm, where the confidence level is initialized at the variable node using the following log likelihood ratio (LLR)

$$LLR = \log\left(\frac{P(b_n = 0|\tilde{U}_j, b_0, b_1, ...b_{n-1})}{P(b_n = 1|\tilde{U}_j, b_0, b_1, ...b_{n-1})}\right). \tag{6}$$

The coset indexes of the subband U_j are reconstructed when all the bit planes are decoded. The coefficients in each subband are finally computed by decoding the coset indexes. The decoded coefficient corresponds to the coefficient in the coset that is the closest to the side information $\tilde{U}_j$. Once all the subband coefficients are decoded, the image is reconstructed by inverse SLP.

VI. EXPERIMENTAL RESULTS

A. Setup

We evaluate the performance of our system on both synthetic and natural spherical images. Synthetic spherical image set Room is shown in Figure 6 respectively and the natural spherical image set Lab is shown in Fig. 7.

Image X   **Image Y**

Fig. 6. Original Room images

The Spherical Laplacian Pyramid is implemented with an 7-tap axisymmetric filter $H(m) = \{-0.0625\ 0\ 0.5625\ 1\ 0.5625\ 0\ -0.0625\}$. These values are chosen so that the filter has a bandwidth close to the one of a half band filter. The SLP decomposition is further carried out in the Fourier domain in order to speed up the computations. The SLP implements four levels of decompositions in the results presented below. The multi-resolution motion estimation at the decoder is carried out on blocks of size 4. Finally, the performance are measured in terms of PSNR, where the mean square error is evaluated using the inner product on the sphere.

B. Channel model evaluation

Before analyzing the performance of the distributed coding scheme, we propose to evaluate the channel model that is used for designing the Slepian-Wolf encoder. We first show in Figure 8 the distribution of the error between the subband Y_0 and the corresponding side information subband $\tilde{Y}_0$ in the Room image dataset. The error is computed only on the coefficients of the partition B. We can see that the error follows a Laplacian distribution, as expected.
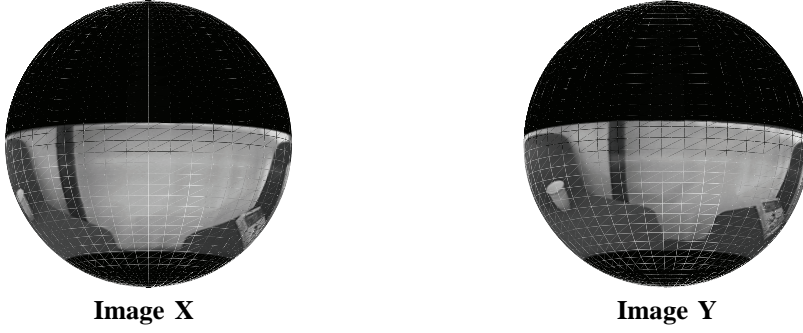
Image X   **Image Y**
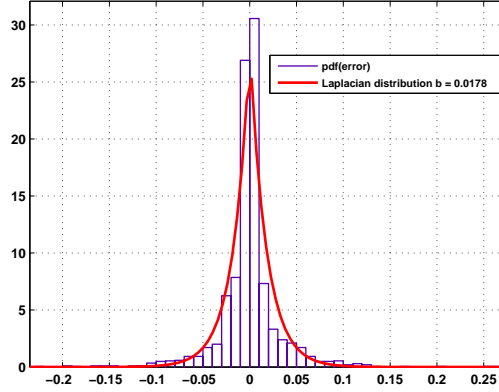
Fig. 7. Original Lab images

Fig. 8. Distribution of the approximation for a sample subband of the Room image. The fitting curve shows a zero-mean Laplacian distribution with $\lambda = 0.0178$.

Then, we estimate the constants α_j and β_j that are used to compute the conditional entropy in Eq. (4) for both Room and Lab image sets. These constants are the same for all image sets and differ only in the respective subbands. We have obtained $\alpha_3 = 1.04$ and $\beta_3 = 2.44$ for the LL subband and $\alpha_j = 0.54$ and $\beta_j = 1.92$ for the detail subband.

Finally, we evaluate the constant γ that captures the benefit of the motion estimation. We have obtained a value of $\gamma_3 = 1.25$ for the LL subband in both image sets. For the detail subbands the parameter γ_j is found to be 1.6, 1.4, 1.2 and 1.1 starting from the lowest to the highest resolution subbands, respectively. The value of γ_j is decreasing when the resolution increases, since the ME is mostly efficient in capturing the correlation in the low frequency subbands.

C. Coding performance

We first compare the performance of proposed DSC solution (using estimated correlation model) with an independent coding scheme, and a joint encoding scheme. In the independent coding scheme, the images X and Y are encoded independently using SLP-based strategy. The images X and Y are transformed using four SLP decomposition levels. Compression is achieved by first quantizing the coefficients [8] and further the quantized indexes are entropy coded (e.g., arithmetic coding), similarly as the coefficients of the partition A in the distributed coding scheme. The joint encoding scheme is based on a SLP decomposition, followed by quantization and arithmetic coding. The image Y is selected as the reference image and X is the predicted from Y. The reference image is encoded using four SLP decomposition levels. Multiresolution motion estimation with a 3-level SLP decomposition and blocks of size 8 is used to predict the image Y. The residual error after motion estimation is also encoded using SLP based strategy. The motion vectors of the successive resolution levels are differentially encoded. Finally, the rate allocation between the reference and the predicted images is chosen such that the rate-distortion performance is maximized. The corresponding rate distributions are given in Tables I and II, where the bits used for the motion vectors are included in the budget of the predicted frames.

The comparison between the distributed, independent and joint coding scheme is given in the Fig. 9 in terms of rate-distortion performance. We observe that the distributed coding scheme performs close to joint encoding algorithm that is based on the same representation and coding strategy. We see also that our proposed DSC scheme clearly outperforms independent coding scheme. In particular, the gain is reaches 1.5 dB for the Room images and 1.3 dB for the Lab images. We further compute
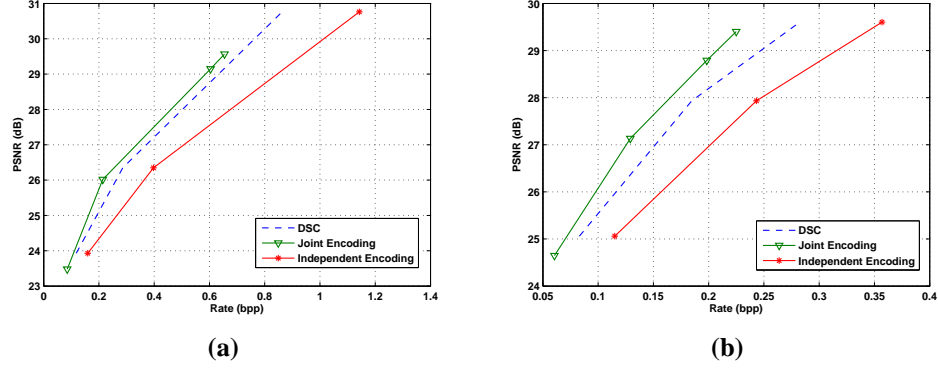
9

Fig. 9. Rate-distortion performance comparison between the proposed DSC scheme, and joint and independent coding strategies for (a) Room (b) Lab image sets.

TABLE I
ROOM IMAGES: DISTRIBUTION OF BITS BETWEEN REFERENCE IMAGE AND THE PREDICTED IMAGE IN JOINT ENCODING

Reference Image (Y)		Predicted Image (X)		Total Rate (bpp)	Mean PSNR (dB)
Bits	PSNR (dB)	Bits	PSNR (dB)		
5293	23.8	297	23.19	0.0853	23.47
13161	26.11	864	25.93	0.214	26
37694	30.5	1887	28.12	0.604	29.15
37694	30.5	5197	28.8	0.6545	29.57

TABLE II
LAB IMAGES: DISTRIBUTION OF BITS BETWEEN REFERENCE IMAGE AND THE PREDICTED IMAGE IN JOINT ENCODING

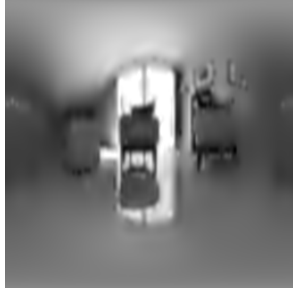| Reference Image (Y)		Predicted Image (X)		Total Rate (bpp)	Mean PSNR (dB)
Bits	PSNR (dB)	Bits	PSNR (dB)		
3742	25.26	219	24.11	0.0604	24.64
7759	28.06	707	26.37	0.1292	27.13
11469	29.73	1528	28.02	0.1983	28.79
11469	29.73	3274	29.1	0.2250	29.40

the rate savings between DSC and independent coding schemes for the same reconstruction quality. Tables III and IV tabulate the percentage of rate saving at different reconstructed qualities for the Room and Lab images respectively. We could see that bit saving is approximately 25%, for both images. The reconstructed Room image is finally shown in the Figure 10 for two sample bit rates. The reconstructed images are shown as the planar image in (θ, ϕ) coordinates to show the particular features of the spherical images.

TABLE III
RATE SAVINGS FOR ROOM IMAGE Y

PSNR (dB).	Bits		Bits saved	% Bit saving
	DSC	Independent		
23.8	3928	5293	1365	25.8
26.1	9526	13161	3635	27.6
30.5	28166	37694	9528	25.3

TABLE IV
RATE SAVINGS FOR LAB IMAGE X

PSNR (dB).	Bits		Bits saved	% Bit saving
	DSC	Independent		
25.26	2852	3797	945	24.9
28.06	6214	8179	1965	24
29.73	9542	11920	2378	20

Rate:0.14 bpp, PSNR: 26.1 dB   *Rate:0.43 bpp, PSNR: 30.5 dB*

Fig. 10. Reconstructed image X in the Room scene.

Finally, we compare in Figure 11 the average performance of the distributed coding scheme with independent coding implemented by the JPEG compression standard. The equiangular grid of the spherical image is represented as a 2D planar image. A baseline JPEG scheme is use to encode both unwarped X and Y independently. We can see that both independent and the distributed coding schemes based on the SLP decomposition outperform JPEG at low coding rates, thanks to efficient data processing on the sphere. At higher rates, the mode of representation of the information becomes less important, and JPEG provides improved performance.
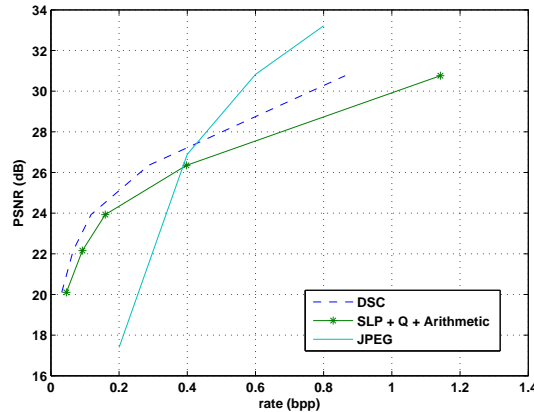

Fig. 11. Average rate-distortion performance for encoding both images X and Y using the distributed coding scheme, and independent coding with JPEG.

D. DSC scheme analysis

We analyze in more details the behavior of the distributed coding scheme. In particular, we examine the rate balance between the two encoders, by comparing the RD performance of the images X and Y. Fig. 12 shows the RD curves for the test images. As expected, the DSC scheme balances the encoding rates, since the encoding rates between the images X and Y are quite similar at a given reconstruction quality.

Lastly, we study the effect of imprecise estimation of the coding rate in the Slepian-Wolf encoder. For both image sets, we compare the rate-distortion characteristics between the estimated correlation model described in the paper, and an exact oracle model where the Laplacian distribution parameter b is known a priori at the encoder. The comparison is presented in Fig. 13 for the image X of the Room and Lab sets. We can see that the methodology proposed in this paper for estimating the channel rate performs very similar to the exact model. The performance degradation due to inexact rate estimation stays smaller than 0.2 dB.

VII. CONCLUSIONS

In this paper we have presented a transform-based distributed source coding scheme for the balanced representation of 3D scenes with stereo omnidirectional cameras. The images are decomposed by Spherical Laplacian Pyramid, and coefficients
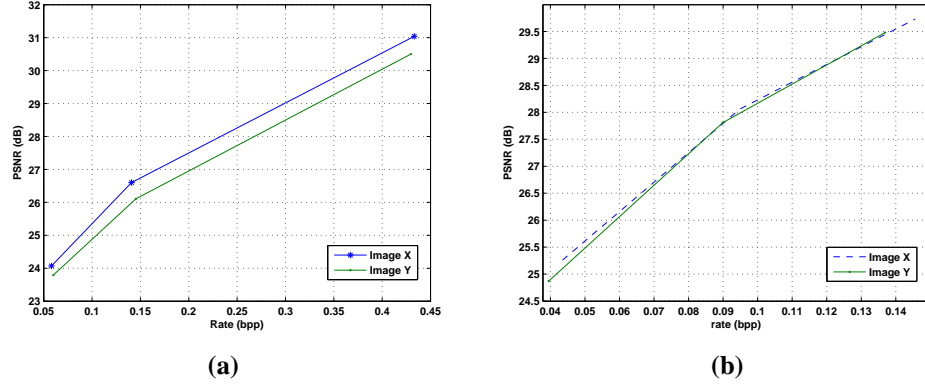
Fig. 12. Rate-distortion comparison between the images X and Y, to examine the rate balance among the encoders (a) Room (b) Lab
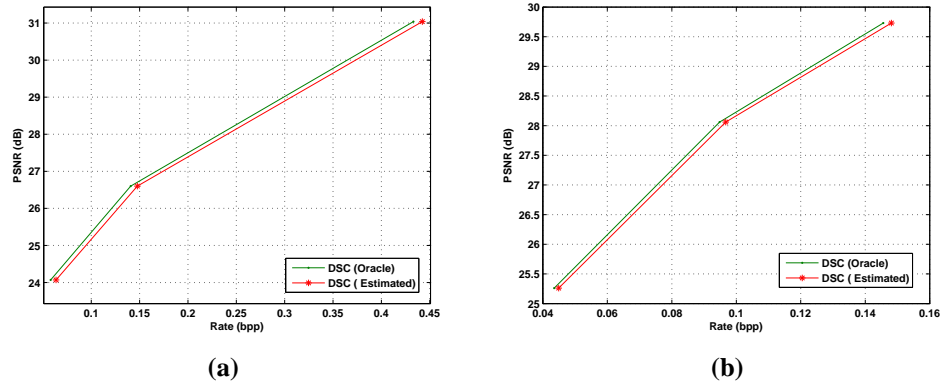

Fig. 13. RD comparison between estimated correlation model Vs exact correlation model (a) Room Image X (b) Lab Image X

are partitioned and Slepian-Wolf encoded in independent encoders. A joint decoder efficiently exploits both the intra- and interview correlation by coefficient prediction and motion estimation. Our scheme outperforms the independent coding and performs close to a joint encoding solution based on similar coding principles. The proposed scheme is shown to be quite robust to inexact estimation of the correlation and Slepian-Wolf coding rate. It provides an interesting and low complexity solution for simple networks of omnidirectional cameras, since it does not require any calibration nor hierarchy between sensors.

REFERENCES

[1] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources", IEEE Trans. on Inform. Theory, vol. 19(4), pp. 471-480, July 1973.
[2] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side-information at the decoder", IEEE Trans. on Inform. Theory, vol. 22(1), pp. 1-10, Jan. 1976.
[3] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS)", IEEE Trans. on Inform. Theory, vol. 49(3), pp. 626-643, Mar. 2003.
[4] J. Garcia-Frias and Y. Zhao, "Compression of correlated binary sources using turbo codes," IEEE Commun. Lett., vol. 5(10), pp. 417-419, Oct. 2001.
[5] A. Aaron and B. Girod, "Compression with side information using turbo codes," Proc. of IEEE DCC, pp. 252-261, Apr. 2002.
[6] A. D. Liveris, Z. Xiong and C. N. Georghiades, "Compression of binary sources with side-information at the decoder using LDPC codes", IEEE Commun. Lett., vol. 6(10), pp. 440-442, Oct. 2002.
[7] M. Grangetto, E. Magli and G. Olmo, "Distributed Arithmetic Coding", IEEE Commun. Lett., vol. 11(11), pp. 883-885, Nov. 2007.
[8] J.L. Salinas and R.L. Baker, "Laplacian pyramid encoding: optimum rate and distortion allocations,", Proc. of the IEEE ICASSP, pp. 1957-1960, May 1989.
[9] R. Zamir, S. Shamai and U. Erez, "Nested Linear/Lattice Codes for Structured Multiterminal Binning, " IEEE Trans. on Infom. Theory, vol.48(6), pp. 1250 - 1276, June 2002.
[10] D. MacKay, "Good error-correcting codes based on very sparse matrices," IEEE Trans. on Infom. Theory, vol. 45(3), pp. 399431, Mar. 1999.
[11] D. Mackay and R.M. Neal, "Near Shannon limit performance of low density parity check codes," Electronics Lett. , vol. 33(6), pp. 457 -458, Mar. 1997.
[12] A. Aaron, S. Rane, E. Setton, and B. Girod, "Transform-domain Wyner-Ziv codec for video," Proc. of the SPIE Visual Commun. and Image Proc., pp. 520-528, Jan. 2004.
[13] R. Puri, A. Majumdar and K. Ramchandran, "PRISM: A Video Coding Paradigm With Motion Estimation at the Decoder," IEEE Trans. on Image Proc. , vol. 16(10), pp. 2436-2448, Oct. 2007.
[14] B. Girod, A.Aaron, S.Rane, and D.Rebollo-monedero, " Distributed Video Coding," Proc. of the IEEE, vol. 93, pp.71-83, Jan. 2005.
[15] X. Zhu, A. Aaron and B. Girod, "Distributed compression for large camera arrays", Proc. IEEE SSP 2003, pp. 30-33, Sept. 2003.
[16] F. Yang, Q. Dai and G. Ding, "Multi-view Images Coding based on Multiterminal Source Coding," Proc. of IEEE ICASSP, pp. 1037-1040, Apr.2007.

[17] N. Gehrig and P.L.Dragotti, "Distributed Compression of Multi-View Images using a Geometric Approach", Proc. of IEEE ICIP, Sept. 2007.

[18] T. Tillo, B. Penna, P. Frossard and P. Vandergheynst, "Distributed Coding of Spherical Images with Jointly Refined Decoding," Proc. of IEEE MMSP,pp. 1-4, Nov. 2005.

[19] V. Thirumalai, I. Tosic and P. Frossard, "Distributed Coding of Multiresolution Omnidirectional Images," Proc. of IEEE ICIP, pp. 345-348, Sept. 2007.

[20] D. Varodayan, Y.C- Lin, A. Mavlankar, M. Flierl and B. Girod, " Wyner-Ziv Coding of Stereo Images with Unsupervised Learning of Disparity, " Proc. of PCS, Nov. 2007.

[21] Methods for constructing LDPC codes: Available in URL http://www.cs.utoronto.ca/pub/radford/LDPC-2001-05-04/pchk.html.

[22] Q. Xu and Z. Xiong, "Layered Wyner-Ziv Video Coding," IEEE Trans. Image Proc. , vol. 15(12), pp. 3791-3802, Dec. 2006.

[23] I. Tosic, I. Bogdanova, P. Frossard and P. Vandergheynst, "Multiresolution Motion Estimation for Omnidirectional Images", Proc. of EUSIPCO 2005, Sept. 2005.

[24] S. Mallat, "A Wavelet Tour of Signal Processing," Academic Press, 1998.

[25] P. J. Burt and E. H. Adelson, "The laplacian pyramid as a compact image code," IEEE Trans. on Commun., vol. 31(4), pp. 532540, 1983.

[26] I. Bogdanova, P. Vandergheynst, J.-P. Antoine, L. Jacques and M. Morvidone, "Discrete wavelet frames on the sphere," Proc. of EUSIPCO, Sept. 2004.

[27] F.M.J. Willems, "Totally asynchronous Slepian-Wolf data compression," IEEE Trans. Inform. Theory, vol.34, pp. 35-44, Jan. 1988.

[28] S. Cheng and Z. Xiong, "Successive Refinement for the WynerZiv Problem and Layered Code Design," IEEE Trans. on Sig. Proc., vol.53(8), pp. 3269-3281, Aug. 2005.

[29] S.S. Pradhan and K. Ramchandran, "Distributed source coding: Symmetric rates and applications to sensor networks," Proc. of IEEE DCC, pp. 363-372, Mar. 2000.

[30] V. Stanković, A.D. Liveris, Z. Xiong, and C.N. Georghiades, "Design of Slepian-Wolf codes by channel code partitioning," Proc. of IEEE DCC, pp. 302-311, Mar. 2004.

[31] M. Sartipi and F. Fekri, "Distributed Source Coding in the Wireless Sensor Networks using LDPC codes: The entire Slepian- Wolf Rate Region," Proc. of IEEE WCNC, pp. 1939-1944, Mar. 2005.

[32] N. Gehrig and P.L. Dragotti, "Distributed Compression in Camera Sensor Networks", Proc. of IEEE MMSP, Sept. 2004.

[33] S. Li and K. Fukumori, "Spherical Stereo for the Construction of Immersive VR Environment, " Proc. of IEEE VR, pp. 217-222, Mar. 2005.

[34] C. Geyer and K. Daniilidis, "Catadioptric projective geometry," International Journal of Computer Vision, vol. 45(3), pp. 223-243, Dec. 2001.

[35] P. Tan and J. Li, "A Practical and Optimal Symmetric Slepian-Wolf Compression strategy using Syndrome Formers and Inverse Syndrome Formers," Proc. Allerton Conf. on Communication, Control and Computing, Sept. 2005.

[36] M. Grangetto, E. Magli and G. Olmo, "Symmetric Distributed Arithmetic Coding of Correlated Sources," Proc. of IEEE MMSP, pp. 111-115, Oct. 2007.

[37] V. Thirumalai, I.Tosic and P. Frossard, "Balanced Distributed Coding of Omnidirectionl Images," Accepted for publication in VCIP-2008. (Available in the URL: http://lts4www.epfl.ch/ vijay/publications.php)

[38] C. Geyer and K. Daniilidis, "Catadioptric Projective Geometry," International Journal of Computer Vision, vol. 45(3), pp. 223 - 243, 2001.

[39] D. Healy Jr. and D. Rockmore and P. Kostelec and S. Moore, "Ffts for the 2-sphere - improvements and variations," *Journal of Fourier Analysis and Applications*, vol. 9, no. 3, pp. 341 – 385, 2003.

[40] E.H. Adelson and J.R. Bergen, *Computational Models of Visual Processing*.   M. Landy and J.A. Movshon, eds., MIT Press, Cambridge, 2001, pp. 3 – 20.

[41] A. Jagmohan, A. Sehgal and N. Ahuja," Two-channel predictive multiple description coding", Proc. of IEEE ICIP, pp. 670 -673, Sept.2005.