Sparse optimization of non separable vector lifting scheme for stereo image coding

Article *in* Journal of Visual Communication and Image Representation · October 2018

DOI: 10.1016/j.jvcir.2018.10.025

4 authors:

Ismail Bezzine
Université Paris 13 Nord
1 PUBLICATION **14** CITATIONS
SEE PROFILE

Mounir Kaaniche
Université Paris 13 Nord
40 PUBLICATIONS **211** CITATIONS
SEE PROFILE

Saadi Boudjit
Université Sorbonne Paris Nord
63 PUBLICATIONS **431** CITATIONS
SEE PROFILE

Azeddine Beghdadi
Université Sorbonne Paris Nord
329 PUBLICATIONS **3,390** CITATIONS
SEE PROFILE

Sparse optimization of non separable vector lifting scheme for stereo image coding

I. Bezzine[a,*], M. Kaaniche[a], S. Boudjit[a], A. Beghdadi[a]

[a] *L2TI, Institut Galilée, Université Paris 13, Sorbonne Paris Cité, 99 Avenue Jean Baptiste Clément, 93430 Villetaneuse, France*

Abstract

One of the potential 3D imaging techniques relies on the use of stereoscopic systems. The great interest in these systems has resulted in huge amounts of data which need to be compressed for storage and transmission purposes. In this context, vector lifting scheme has been found to be an efficient approach for stereo image coding. For instance, the coding performance depends on the design of the involved lifting operators referred to as prediction and update filters. For this reason, while a non separable vector lifting structure is retained, we investigate in this paper different techniques for optimizing sparse criteria to design the filters used with both views. More precisely, an independent full optimization algorithm as well as a joint algorithm will be developed and studied. Simulations performed on different stereo images demonstrate the effectiveness of the proposed sparse optimization algorithms in terms of quality of reconstruction and bitrate saving.

Keywords: Stereo images, adaptive coding, vector lifting scheme, non separable transform, optimization.

*Corresponding author

Email addresses: ismail.bezzine@univ-paris13.fr (I. Bezzine), mounir.kaaniche@univ-paris13.fr (M. Kaaniche), boudjit@univ-paris13.fr (S. Boudjit), azeddine.beghdadi@univ-paris13.fr (A. Beghdadi)

1. Introduction

Stereoscopic and multiview systems are considered among the recent 3D imaging techniques. In particular, a stereoscopic system consists in generating two images by recording the same scene from two slightly different positions. The obtained images, referred to as left and right images, are then merged by the brain to perceive the scene in three dimensions. For this reason, stereovision has been widely used in various application fields such as 3DTV, computer vision, remote sensing and medicine [1, 2]. Thus, the increasing interest in stereo images has resulted in huge amounts of data which will constitute a problem for its practical use. Therefore, it becomes mandatory to design efficient stereo images coding schemes in order to improve their transmission visual quality while reducing their storage capacity.

A basic approach for stereo images coding may consist in encoding independently the left and right views by employing existing still image encoders. However, such approach may not appear so efficient since it does not exploit the main feature of these images. Indeed, as the stereo images correspond to the same 3D scene, they present similar contents and have high correlation. Thus, efficient stereo image coding schemes could be designed by exploiting the inter-view redundancies. To this respect, the conventional scheme can be described as follows. First, one image, for example the left one, is selected as a reference image, and the other one (i.e the right image) is considered as a target image. Then, the target image is predicted from the reference one thanks to the disparity estimation/compensation (DE/DC) process. The difference between the original target image and the predicted one leads to the generation of the residual image. Finally, the reference and residual images as well as the disparity information are encoded. It is important to note here that this idea is behind most of the existing stereo image coding methods. However, the main differences between them can be classified in two categories. The first category of methods aim to improve the DE/DC process as well as the coding of the disparity (or depth) maps [3, 4, 5, 6, 7, 8]. For instance, while the standard block-matching (BM)

technique is often used to perform the DE/DC step, modified BM [3] and learning dictionary-based techniques have also been developed [4, 5, 6]. Indeed, in [5], a block dependent dictionary is used by linking together disparities yielding similar compensation. In [6], directional prediction model combined with linear predictive scheme is proposed for efficient disparity compensation. For the same purpose, the authors proposed in [7] to use the neighborhood of the homologous pixel in the reference image to predict the pixel of the target image and compute the residual image. This computation step is optimized by minimizing the ℓ_1-norm of the resulting prediction error. Note that the disparity is generally encoded using DPCM techniques followed by an entropy coder while the reference and residual images are often encoded in the transform domain. To this end, the second category of the existing stereo image compression methods have been devoted to the design of efficient decomposition (i.e transform) for coding the reference and residual images. More precisely, some methods have been developed based on the Discrete Cosine Transform (DCT) [9, 10]. However, it has been shown in [11] that residual images contain very narrow vertical edges and DCT yields a moderate energy packing of such images. For this reason, it has been proposed to use a directional DCT to better exploit the characteristics of the residual images [10, 12]. Other encoding methods based on wavelet transforms have also been developed in order to allow high quality scalability and progressive reconstruction of the stereo images [13, 14, 15]. Indeed, a family of wavelet-based coders is investigated in [13]. In [14], a coding method based on adaptive lifting scheme has been developed. In this scheme, an adaptive prediction step is performed according to the local gradient information of the reference image. However, the main drawback of this adaptive coding strategy is that it depends on the reference image which has poor quality at low bitrate and results in a great impact on the image reconstruction process. In [15], a bandelet transform [16] is firstly applied to the left and right images to estimate the disparity map and generate the residual image. Then, the disparity map as well as the bandelet coefficients of the left and residual images are encoded. The main limitation of this method is that it requires to transmit a side information

related to the size of each block transform which will affect its performance at low bitrate. In addition to this kind of methods based on the coding of reference and residual images, an alternative approach that does not directly generate a residual image has been proposed in [17] and [18] for grayscale and color stereo images, respectively. More precisely, the approach consists in using a multi-scale decomposition based on the concept of vector lifting scheme (VLS). Note that, unlike conventional lifting scheme, the VLS is a joint wavelet decomposition that aims at exploiting the inter-view correlations to generate two compact multi-resolution representations of the left and right images. While a separable decomposition has been carried out in [17], its extended non separable version (NS-VLS) has been developed in [19]. Such extension presents two main advantages. First, it allows to better capture the two dimensional characteristics of the edges which are neither horizontal nor vertical. Moreover, it offers more flexibility in the design of an adaptive transform well adapted to the contents of the input images [20, 21].

In this paper, we propose to retain the previous NS-VLS decomposition and focus on the built of content-adaptive decomposition through sparse optimization algorithms. This is achieved using different ℓ_1 based minimization techniques. It is important to note here that sparse optimization algorithms have been recently employed for still image coding [22] whereas this paper deals with their extension to the context of stereo image coding. While *only* a weighted ℓ_1 minimization technique has already been investigated in [19] for stereo image coding and in [23] for hologram compression, this work aims at developing and studying *various* optimization strategies for the design of all the involved filters used with the left and right images. More precisely, in addition to the basic ℓ_2 and ℓ_1 optimization approaches which can be separately applied to each filter of each view, two optimization algorithms based on the weighted ℓ_1 minimization technique are considered. In the first one, we resort to a *full* optimization algorithm where the filters of each view are optimized independently of those used with the other view. However, in the second one, a *joint* optimization algorithm, based on an hybrid weighted ℓ_1 minimization technique, is developed to take

into account the inter-view redundancies.

The remainder of this paper is organized as follows. In Sec. 2, the analysis structure as well as the resulting wavelet subbands of the NS-VLS decomposition are given. The proposed optimization algorithms for the design of the lifting operators of both views are described in Sec. 3. Finally, experimental results are shown in Sec. 4 and some conclusions are drawn in Sec. 5.

2. Non separable vector lifting scheme

2.1. Analysis structure

A conventional separable lifting scheme (LS) [24] consists in splitting the input 1D signal into two sets formed by the even and odd samples, respectively. Then, prediction and update steps are applied to generate the detail and approximation signals. Such structure is referred to as P-U (Predict-Update) LS like the 5/3 transform retained in the JPEG2000 coding standard [25]. As shown in [21], a 1D P-U LS has an equivalent 2D non separable structure that can be obtained by splitting the input image into four polyphase components and applying three prediction steps followed by the update one (P-P-P-U structure) to generate three detail subbands and one approximation subband.

Based on this observation, the 2D NS-VLS decomposition has been derived [19] where intra prediction steps are performed on the reference (i.e left) image and hybrid prediction steps are employed with the target (i.e right) image to exploit the intra and inter-view redundancies. The corresponding analysis structure is illustrated in Fig. 1.

While 2D non separable lifting operators are used, it is worth pointing out that the main feature of this VLS-based decomposition concerns the prediction stages. For instance, a conventional P-P-P-U lifting structure is first applied to the left and right images. Since the left image is selected as a reference image and encoded in intra-mode, an hybrid prediction stage is added to the first lifting steps used with the right view to exploit simultaneously the intra and inter-view redundancies based on the information coming from the left view

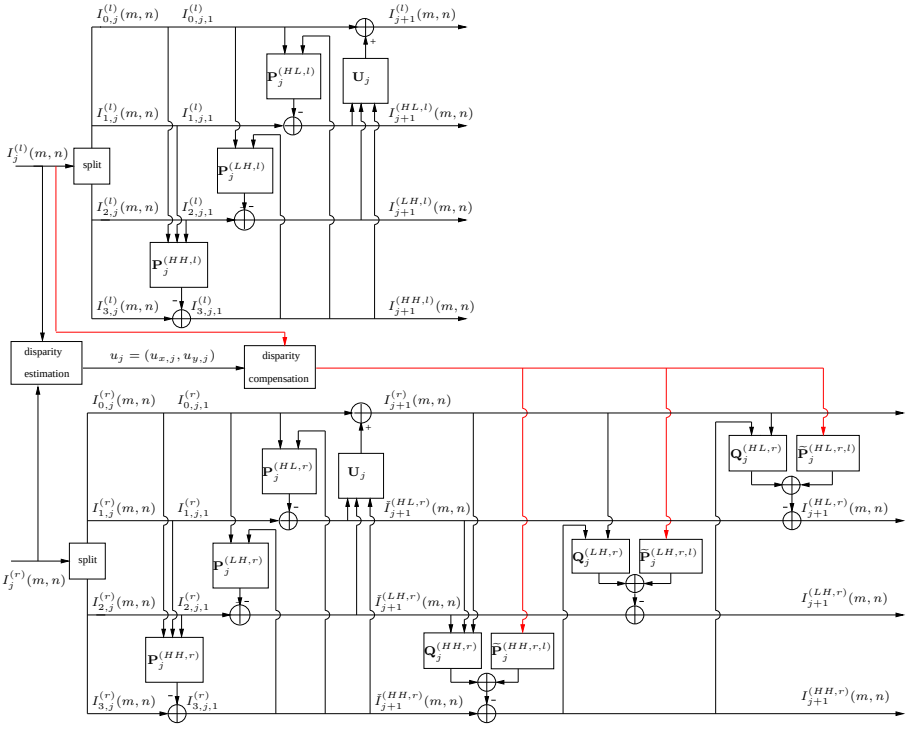

Figure 1: NS-VLS decomposition structure.

(highlighted with the red color in Fig. 1). In addition to the illustration of the basic concept of a NS-VLS, we should note that the main notations have been also included in the above figure to better understand the mathematical equations used in the next sections.

In what follows, the different lifting operators used in this structure as well as the computation of the resulting wavelet coefficients will be described.

2.2. Wavelet representations of the stereo pairs

As a multiscale transform, our decomposition will be described for a given resolution level $j \in \mathbb{N}^*$. Let us denote by $I_j^{(l)}$ and $I_j^{(r)}$ the approximation subbands of the left and right images. Note that $j = 0$ corresponds to the original stereo images $I^{(l)}$ and $I^{(r)}$. Moreover, for each view $v \in \{l, r\}$, the image $I_j^{(v)}$ has four polyphase components $I_{0,j}^{(v)}(m,n) = I_j^{(v)}(2m, 2n)$, $I_{1,j}^{(v)}(m,n) =$

$I_j^{(v)}(2m, 2n+1)$, $I_{2,j}^{(v)}(m,n) = I_j^{(v)}(2m+1, 2n)$, and $I_{3,j}^{(v)}(m,n) = I_j^{(v)}(2m+1, 2n+1)$.

As it can be shown in Fig. 1, a non separable lifting stage, composed of three prediction steps and an update one, is applied to the left image to produce three detail subband coefficients oriented diagonally $I_{j+1}^{(HH,l)}$, vertically $I_{j+1}^{(LH,l)}$ and horizontally $I_{j+1}^{(HL,l)}$ as well as the approximation coefficients $I_{j+1}^{(l)}$. These signals can be computed as follows:

$$I_{j+1}^{(HH,l)}(m,n) = I_{3,j}^{(l)}(m,n) - \Big((\mathbf{P}_{0,j}^{(HH,l)})^\top \mathbf{I}_{0,j}^{(HH,l)} + (\mathbf{P}_{1,j}^{(HH,l)})^\top \mathbf{I}_{1,j}^{(HH,l)} + (\mathbf{P}_{2,j}^{(HH,l)})^\top \mathbf{I}_{2,j}^{(HH,l)}\Big), \quad (1)$$

$$I_{j+1}^{(LH,l)}(m,n) = I_{2,j}^{(l)}(m,n) - \Big((\mathbf{P}_{0,j}^{(LH,l)})^\top \mathbf{I}_{0,j}^{(LH,l)} + (\mathbf{P}_{1,j}^{(LH,l)})^\top \underline{\mathbf{I}}_{j+1}^{(HH,l)}\Big), \quad (2)$$

$$I_{j+1}^{(HL,l)}(m,n) = I_{1,j}^{(l)}(m,n) - \Big((\mathbf{P}_{0,j}^{(HL,l)})^\top \mathbf{I}_{0,j}^{(HL,l)} + (\mathbf{P}_{1,j}^{(HL,l)})^\top \bar{\mathbf{I}}_{j+1}^{(HH,l)}\Big), \quad (3)$$

$$I_{j+1}^{(l)}(m,n) = I_{0,j}^{(l)}(m,n) + \Big((\mathbf{U}_{0,j}^{(HL,l)})^\top \mathbf{I}_{j+1}^{(HL,l)} + (\mathbf{U}_{1,j}^{(LH,l)})^\top \mathbf{I}_{j+1}^{(LH,l)} + (\mathbf{U}_{2,j}^{(HH,l)})^\top \mathbf{I}_{j+1}^{(HH,l)}\Big), \quad (4)$$

where for each $i \in \{0,1,2\}$ and $o \in \{HL, LH, HH\}$,

- $\mathbf{P}_{i,j}^{(o,l)} = (p_{i,j}^{(o,l)}(s,t))_{(s,t)\in\mathcal{P}_{i,j}^{(o,l)}}$ is the vector of prediction filter coefficients and $\mathcal{P}_{i,j}^{(o,l)}$ denotes its support,
- $\mathbf{I}_{i,j}^{(o,l)} = (I_{i,j}^{(l)}(m+s, n+t))_{(s,t)\in\mathcal{P}_{i,j}^{(o,l)}}$ is a reference vector that allows to compute $I_{j+1}^{(o,l)}(m,n)$,
- $\underline{\mathbf{I}}_{j+1}^{(HH,l)} = (I_{j+1}^{(HH,l)}(m+s, n+t))_{(s,t)\in\mathcal{P}_{1,j}^{(LH,l)}}$ and $\bar{\mathbf{I}}_{j+1}^{(HH,l)} = (I_{j+1}^{(HH,l)}(m+s, n+t))_{(s,t)\in\mathcal{P}_{1,j}^{(HL,l)}}$ are used in the second and third prediction steps,
- $\mathbf{U}_{i,j}^{(o,l)} = (u_{i,j}^{(o,l)}(s,t))_{(s,t)\in\mathcal{U}_{i,j}^{(o,l)}}$ is an update weighting vector with support $\mathcal{U}_{i,j}^{(o,l)}$,
- $\mathbf{I}_{j+1}^{(o,l)} = (I_{j+1}^{(o,l)}(m+s, n+t))_{(s,t)\in\mathcal{U}_{i,j}^{(o,l)}}$ is the reference vector containing the samples used in the update step.

Unlike the conventional lifting scheme applied to the reference image, an improved one is applied to the target one (i.e the right image). Indeed, let us recall that the key idea behind vector lifting scheme [17] consists in using hybrid (intra and inter) prediction steps. For instance, as it can be seen from Fig. 1, the prediction steps used with the target image use some samples from

the current view as well as their matching ones in the reference image. To this respect, for the right image, Eqs. (1)-(4) are firstly applied to produce three intermediate detail subbands and an approximation one denoted respectively by $\check{I}_{j+1}^{(HH,r)}$, $\check{I}_{j+1}^{(LH,r)}$, $\check{I}_{j+1}^{(HL,r)}$ and $I_{j+1}^{(r)}$. Then, a second hybrid prediction stage, composed of three steps, is added to exploit at the same time the intra and inter-view redundancies in the stereo images. This is achieved by using the estimated disparity field denoted by $u_j = (u_{x,j}, u_{y,j})$. For notation concision, the disparity compensated left image on a given matching sample (m, n), given by $I_j^{(l)}(m+u_{x,j}(m,n), n+u_{y,j}(m,n))$, will be simply replaced by $I_j^{(c)}(m,n)$. Let us denote its corresponding four polyphase components by $I_{0,j}^{(c)}(m,n)$, $I_{1,j}^{(c)}(m,n)$, $I_{2,j}^{(c)}(m,n)$ and $I_{3,j}^{(c)}(m,n)$. Therefore, the final detail subbands of the right image are given by:

$$\begin{aligned} I_{j+1}^{(HH,r)}(m,n) = \check{I}_{j+1}^{(HH,r)}(m,n) - \Big(& (\mathbf{Q}_{0,j}^{(HH,r)})^\top \check{\mathbf{I}}_{0,j+1}^{(HH,r)} + (\mathbf{Q}_{1,j}^{(HH,r)})^\top \check{\mathbf{I}}_{1,j+1}^{(HH,r)} \\ & + (\mathbf{Q}_{2,j}^{(HH,r)})^\top \check{\mathbf{I}}_{2,j+1}^{(HH,r)} + (\widetilde{\mathbf{P}}_{0,j}^{(HH,r,l)})^\top \mathbf{I}_{0,j}^{(HH,c)} + (\widetilde{\mathbf{P}}_{1,j}^{(HH,r,l)})^\top \mathbf{I}_{1,j}^{(HH,c)} \\ & + (\widetilde{\mathbf{P}}_{2,j}^{(HH,r,l)})^\top \mathbf{I}_{2,j}^{(HH,c)} + (\widetilde{\mathbf{P}}_{3,j}^{(HH,r,l)})^\top \mathbf{I}_{3,j}^{(HH,c)} \Big), \end{aligned} \tag{5}$$

$$\begin{aligned} I_{j+1}^{(LH,r)}(m,n) = \check{I}_{j+1}^{(LH,r)}(m,n) - \Big(& (\mathbf{Q}_{0,j}^{(LH,r)})^\top \check{\mathbf{I}}_{0,j+1}^{(LH,r)} + (\mathbf{Q}_{1,j}^{(LH,r)})^\top \underline{\mathbf{I}}_{j+1}^{(HH,r)} \\ & + (\widetilde{\mathbf{P}}_{0,j}^{(LH,r,l)})^\top \mathbf{I}_{0,j}^{(LH,c)} + (\widetilde{\mathbf{P}}_{1,j}^{(LH,r,l)})^\top \mathbf{I}_{1,j}^{(LH,c)} + (\widetilde{\mathbf{P}}_{2,j}^{(LH,r,l)})^\top \mathbf{I}_{2,j}^{(LH,c)} \\ & + (\widetilde{\mathbf{P}}_{3,j}^{(LH,r,l)})^\top \mathbf{I}_{3,j}^{(LH,c)} \Big), \end{aligned} \tag{6}$$

$$\begin{aligned} I_{j+1}^{(HL,r)}(m,n) = \check{I}_{j+1}^{(HL,r)}(m,n) - \Big(& (\mathbf{Q}_{0,j}^{(HL,r)})^\top \check{\mathbf{I}}_{0,j+1}^{(HL,r)} + (\mathbf{Q}_{1,j}^{(HL,r)})^\top \underline{\mathbf{I}}_{j+1}^{(HH,r)} \\ & + (\widetilde{\mathbf{P}}_{0,j}^{(HL,r,l)})^\top \mathbf{I}_{0,j}^{(HL,c)} + (\widetilde{\mathbf{P}}_{1,j}^{(HL,r,l)})^\top \mathbf{I}_{1,j}^{(HL,c)} + (\widetilde{\mathbf{P}}_{2,j}^{(HL,r,l)})^\top \mathbf{I}_{2,j}^{(HL,c)} \\ & + (\widetilde{\mathbf{P}}_{3,j}^{(HL,r,l)})^\top \mathbf{I}_{3,j}^{(HL,c)} \Big), \end{aligned} \tag{7}$$

where for every $i \in \{0,1,2,3\}$ and $o \in \{HL, LH, HH\}$,

- $\mathbf{Q}_{i,j}^{(o,r)} = (q_{i,j}^{(o,r)}(s,t))_{(s,t)\in\mathcal{Q}_{i,j}^{(o,r)}}$ is an intra prediction weighting vector whose support is denoted by $\mathcal{Q}_{i,j}^{(o,r)}$,
- $\widetilde{\mathbf{P}}_{i,j}^{(o,r,l)} = (p_{i,j}^{(o,r,l)}(s,t))_{(s,t)\in\widetilde{\mathcal{P}}_{i,j}^{(o,r,l)}}$ is an inter prediction weighting vector whose support is denoted by $\widetilde{\mathcal{P}}_{i,j}^{(o,r,l)}$,
- $\check{\mathbf{I}}_{0,j+1}^{(o,r)} = (I_{j+1}^{(r)}(m+s,n+t))_{(s,t)\in\mathcal{Q}_{0,j}^{(o,r)}}$ is a reference vector used to compute $I_{j+1}^{(o,r)}(m,n)$,

• $\check{\mathbf{I}}_{1,j+1}^{(HH,r)} = (\check{I}_{j+1}^{(HL,r)}(m+s,n+t))_{(s,t)\in\mathcal{Q}_{1,j}^{(HH,r)}}$ and $\check{\mathbf{I}}_{2,j+1}^{(HH,r)} = (\check{I}_{j+1}^{(LH,r)}(m+s,n+t))_{(s,t)\in\mathcal{Q}_{2,j}^{(HH,r)}}$ are two reference vectors used to compute $I_{j+1}^{(HH,r)}(m,n)$,
• $\underline{\mathbf{I}}_{j+1}^{(HH,r)} = (I_{j+1}^{(HH,r)}(m+s,n+t))_{(s,t)\in\mathcal{Q}_{1,j}^{(LH,r)}}$ and $\bar{\mathbf{I}}_{j+1}^{(HH,r)} = (I_{j+1}^{(HH,r)}(m+s,n+t))_{(s,t)\in\mathcal{Q}_{1,j}^{(HL,r)}}$ are two intra prediction vectors used to compute $I_{j+1}^{(LH,r)}(m,n)$ and $I_{j+1}^{(HL,r)}(m,n)$,
• $\mathbf{I}_{i,j}^{(o,c)} = (I_{i,j}^{(c)}(m+s,n+t))_{(s,t)\in\widetilde{\mathcal{P}}_{i,j}^{(o,r,l)}}$ is a reference vector containing the matching samples used to compute $I_{j+1}^{(o,r)}(m,n)$.

Finally, at the last resolution level $j = J$, instead of coding the approximation subband $I_J^{(r)}$, a residual subband $e_J^{(r)}$ is generated by computing the difference between the right approximation subband and the disparity compensated left one:

$$e_J^{(r)}(m,n) = I_J^{(r)}(m,n) - I_J^{(c)}(m,n). \tag{8}$$

Once the considered NS-VLS has been defined, we investigate in the next section techniques for optimizing sparse criteria to design the lifting operators used with the left and right images.

3. Proposed sparse optimization algorithms

Since the coding performance of wavelet-based coding scheme depends on the choice of the lifting operators, a great attention should be paid to the design of the prediction and update filters of both views. To this respect, two kinds of optimization strategies could be adopted and will be described in what follows.

3.1. Independent full optimization of the stereo pairs

A straightforward solution consists in applying classical optimization methods used in the context of lifting-based still image coding to each view separately. Thus, the two lifting structures used with the left and right images can be firstly optimized in an independent way. To this end, we will resort to ℓ_2, ℓ_1 and weighted ℓ_1 based minimization techniques.

3.1.1. ℓ_1-based minimization technique

While the update filter of each view $\mathbf{U}_j^{(v)}$, with $v \in \{r, l\}$, will be optimized by minimizing the error between the approximation subband $I_{j+1}^{(v)}$ and the decimated subband obtained after an ideal low-pass filtering of $I_j^{(v)}$ [21], we will focus on the optimization methods for designing the different prediction filters. To this respect, the standard approach consists in minimizing the variance (i.e the ℓ_2-norm) of the output detail subband since the latter can be seen as a prediction error [26, 27]. In a recent work [22], ℓ_1-based minimization techniques have been proposed in the context of *one* stage lifting structure for still image coding purpose. It is important to note here that the use of ℓ_1 criterion presents two main advantages. First, minimizing an ℓ_1 criterion allows to generate sparse representation which could achieve good coding performance [28]. Moreover, from the information theory point of view, it has been shown that, at high bitrate, the minimization of the entropy of the detail subbands is closely related to the minimization of their ℓ_β-norm where β is the shape parameter of a generalized gaussian distribution (GGD) used for modeling the detail coefficients [29]. Indeed, knowing that the wavelet detail subbands $I_{j+1}^{(o,v)}$ are generally multiplied by the weights $\sqrt{w_{j+1}^{(o,v)}}$ before the entropy encoding, and if we consider $\beta = 1$, the resulting differential entropy can be obtained as follows:

$$\frac{1}{M_j N_j \alpha_{j+1}^{(o,v)} \ln(2)} \sum_{m=1}^{M_j} \sum_{n=1}^{N_j} \left| I_{j+1}^{(o,v)}(m,n) \right| + \log_2 \left(2\alpha_{j+1}^{(o,v)} \sqrt{w_{j+1}^{(o,v)}} \right) \tag{9}$$

where (M_j, N_j) represent the dimensions of the subband $I_{j+1}^{(o,v)}$, $\alpha_{j+1}^{(o,v)}$ is the scale parameter of the GGD which can be estimated using a classical maximum likelihood estimate, and the weights $w_{j+1}^{(o,v)}$ are computed based on the wavelet filters used for the reconstruction process as shown in [30, 31].

Therefore, instead of minimizing the ℓ_2-norm, each prediction filter $\mathbf{P}_j^{(o,l)}$, $\mathbf{P}_j^{(o,r)}$ and $\mathbf{P}_j^{(o,r,l)} = \left(\mathbf{Q}_j^{(o,r)}, \widetilde{\mathbf{P}}_j^{(o,r,l)} \right)^\top$ will be optimized by minimizing the ℓ_1-norm of their output detail subbands $I_{j+1}^{(o,l)}$ and $I_{j+1}^{(o,r)}$. Thus, for the intra prediction

filters $\mathbf{P}_j^{(o,l)}$ and $\mathbf{P}_j^{(o,r)}$, the criterion is expressed as:

$$\begin{aligned}
&\forall\ v \in \{r,l\}, \forall\ o \in \{HL, LH, HH\}, \forall\ i \in \{1,2,3\},\\
\mathcal{J}_{\ell_1}^{(v)}(\mathbf{P}_j^{(o,v)}) &= \sum_{m=1}^{M_j}\sum_{n=1}^{N_j}\left|I_{j+1}^{(o,v)}(m,n)\right|\\
&= \sum_{m=1}^{M_j}\sum_{n=1}^{N_j}\left|I_{i,j}^{(v)}(m,n) - (\mathbf{P}_j^{(o,v)})^{\top}\tilde{\mathbf{I}}_j^{(o,v)}(m,n)\right| \qquad (10)
\end{aligned}$$

with $I_{i,j}^{(v)}(m,n)$ is the $(i+1)^{th}$ polyphase component of the view $I_j^{(v)}$ to be predicted, $\mathbf{P}_j^{(o,v)}$ is the prediction operator vector to be optimized, and $\tilde{\mathbf{I}}_j^{(o,v)}(m,n)$ is the reference vector containing the samples used in the prediction step. Similarly, for the hybrid prediction filters $\mathbf{P}_j^{(o,r,l)}$ used in the second lifting stage with the right image, the ℓ_1 criterion will be rewritten as:

$$\begin{aligned}
&\forall\ o \in \{HL, LH, HH\}, \forall\ i \in \{1,2,3\},\\
\mathcal{J}_{\ell_1}^{(r)}(\mathbf{P}_j^{(o,r,l)}) &= \sum_{m=1}^{M_j}\sum_{n=1}^{N_j}\left|\check{I}_{i,j}^{(r)}(m,n) - (\mathbf{P}_j^{(o,r,l)})^{\top}\tilde{\mathbf{I}}_j^{(o,r,l)}(m,n)\right| \qquad (11)
\end{aligned}$$

where $\tilde{\mathbf{I}}_j^{(o,r,l)}(m,n)$ is a reference vector containing the samples from right and disparity compensated left images used in the prediction step, and $\check{I}_{i,j}^{(r)}$ is the polyphase component to be predicted in the second lifting stage. According to Fig. 1, the four polyphase components of the second lifting stage are defined as:

$$\begin{cases}
\check{I}_{0,j}^{(r)}(m,n) = I_{j+1}^{(r)}(m,n)\\
\check{I}_{1,j}^{(r)}(m,n) = \check{I}_{j+1}^{(HL,r)}(m,n)\\
\check{I}_{2,j}^{(r)}(m,n) = \check{I}_{j+1}^{(LH,r)}(m,n)\\
\check{I}_{3,j}^{(r)}(m,n) = \check{I}_{j+1}^{(HH,r)}(m,n)
\end{cases} \qquad (12)$$

To minimize this criterion, we propose to use the proximity operators tool [32] which has been found to be efficient for solving nonsmooth optimization problem [33, 34]. Based on this tool, the minimization of the above ℓ_1 criterion (for example the criterion given by Eq. (10)) is equivalent to the following

minimization problem:

$$\forall\, o \in \{HL, LH, HH\}, \forall\, i \in \{1,2,3\},$$
$$\min_{\mathbf{z}_j^{(o,v)} \in V} \sum_{m=1}^{M_j} \sum_{n=1}^{N_j} \left| I_{i,j}^{(v)}(m,n) - z_j^{(o,v)}(m,n) \right| + \imath_V(\mathbf{z}_j^{(o,v)}), \tag{13}$$

where $\imath_V$ is the indicator function and V is the vector space given by:

$$V = \{\mathbf{z}_j^{(o,v)} = \left(z_j^{(o,v)}(m,n)\right)_{\substack{1 \le m \le M_j \\ 1 \le n \le N_j}} \in \mathbb{R}^{M_j \times N_j} \mid \exists\, \mathbf{P}_j^{(o,v)},$$
$$\forall\, (m,n) \in \{1,\ldots,M_j\} \times \{1,\ldots,N_j\}, z_j^{(o,v)}(m,n) = (\mathbf{P}_j^{(o,v)})^\top \tilde{\mathbf{I}}_j^{(o,v)}(m,n)\}.$$

After that, the Douglas Rachford (DR) algorithm will be applied to solve our minimization problem and obtain the optimized prediction filter $\mathbf{P}_j^{(o,v)}$. For more details on the proximity operators as well as the DR algorithm, the reader is referred to [22].

3.1.2. Weighted ℓ_1-based minimization technique

In the previous part, each prediction filter $\mathbf{P}_j^{(o,v)}$ has been separately optimized by minimizing the ℓ_1-norm of its associated detail subband $I_{j+1}^{(o,v)}$. However, if we focus on the lifting structure applied to the left image, it can be seen in Fig. 1 that the diagonal detail coefficients resulting from the first prediction step are used in the second and third prediction steps to generate the left detail coefficients oriented vertically $I_{j+1}^{(LH,l)}$ and horizontally $I_{j+1}^{(HL,l)}$, respectively. Therefore, instead of minimizing the ℓ_1-norm of the diagonal detail coefficients, it becomes more interesting to optimize the first prediction filter $\mathbf{P}_j^{(HH,l)}$ by minimizing a weighted sum of the ℓ_1-norms of the three detail subbands of the left image. Concerning the filters $\mathbf{P}_j^{(LH,l)}$ and $\mathbf{P}_j^{(HL,l)}$, they will be simply optimized by minimizing the ℓ_1-norm of their corresponding detail coefficients $I_{j+1}^{(LH,l)}$ and $I_{j+1}^{(HL,l)}$ since the second and third predictions are two independent steps.

Regarding the prediction filters used with the lifting structure applied to the right image, it can be also seen from Fig. 1 that the first three intermediate detail coefficients $\check{I}_{j+1}^{(o,r)}$ as well as the final diagonal detail ones $I_{j+1}^{(HH,r)}$ are

involved in the last two prediction steps to generate the final horizontal and vertical detail subbands. Therefore, and similarly to the left image, the first four prediction filters, $\mathbf{P}_j^{(o,r)}$ with $o \in \{HH, LH, HL\}$ and $\mathbf{P}_j^{(HH,r,l)}$, used with the right image should be optimized by minimizing the weighted sum of the ℓ_1-norms of the three detail subbands of the right image. Finally, the last prediction filters $\mathbf{P}_j^{(LH,r,l)}$ and $\mathbf{P}_j^{(HL,r,l)}$ are optimized by minimizing the ℓ_1-norm of their corresponding detail coefficients $I_{j+1}^{(LH,r)}$ and $I_{j+1}^{(HL,r)}$.

Therefore, the weighted ℓ_1 criterion used with the left and right images can be expressed for each view as follows:

$$\begin{aligned} &\forall\ v \in \{r,l\}, \forall\ o \in \{HL, LH, HH\}, \\ &\mathcal{J}_{w\ell_1}^{(v)}(\mathbf{P}_j^{(o,v)}) = \sum_{o\in\{HL,LH,HH\}} \sum_{m=1}^{M_j} \sum_{n=1}^{N_j} \frac{1}{\alpha_{j+1}^{(o,v)}} \left| I_{j+1}^{(o,v)}(m,n) \right|. \end{aligned} \tag{14}$$

To minimize this weighted ℓ_1 criterion, we need first to rewrite the weighted criterion (i.e $I_{j+1}^{(o,v)}$) as a function of the filter to be optimized $\mathbf{P}_j^{(o,v)}$. To this respect, and since the weighted ℓ_1 minimization concerns only the first prediction filter for the left image and the first four prediction filters for the right image, let us introduce the notation $\left(I_{i,j,q}^{(v)}\right)_{i\in\{0,1,2,3\}}$ the four polyphase components obtained from the inputs $\left(\hat{I}_{i,j,q-1}^{(v)}\right)_{i\in\{0,1,2,3\}}$ after the q-th prediction step while $q \in \{1\}$ (resp. $q \in \{1,2,3,4\}$) in the case of the left (resp. right) view. To illustrate these components, $\left(I_{i,j,1}^{(v)}\right)_{i\in\{0,1,2,3\}}$ (i.e for $q=1$) have been inserted in Fig. 1 after the first prediction step used with both views. Thus, for each $i \in \{0,1,2,3\}$, we have:

$$\begin{cases} \hat{I}_{i,j,q-1}^{(v)}(m,n) = I_{i,j}^{(v)}(m,n) & \text{for} \quad q=1, \quad \forall\ v \in \{r,l\} \\ \hat{I}_{i,j,q-1}^{(r)}(m,n) = I_{i,j,q-1}^{(r)}(m,n) & \text{for} \quad q \in \{2,3\} \\ \hat{I}_{i,j,q-1}^{(r)}(m,n) = \check{I}_{i,j}^{(r)}(m,n) & \text{for} \quad q=4 \end{cases} \tag{15}$$

Based on these notations, each detail subband $I_{j+1}^{(o,v)}$ can be written as a function of the filter to be optimized $\mathbf{P}_j^{(o,v)}$ as follows:

$$\begin{aligned} &\forall\ o \in \{HH, LH, HL\}, \\ &I_{j+1}^{(o,v)}(m,n) = y_{j,q}^{(o,v)}(m,n) - (\mathbf{P}_j^{(o,v)})^\top \mathbf{I}_{j,q}^{(o,v)}(m,n) \end{aligned} \tag{16}$$

where the signal to be predicted $y_{j,q}^{(o,v)}(m,n)$ as well as the reference vector $\mathbf{I}_{j,q}^{(o,v)}(m,n)$ are given by:

$$y_{j,q}^{(o,v)}(m,n) = \sum_{i' \in \mathbb{I}_i} \sum_{k,l} h_{i',j,q}^{(o,v)}(k,l) I_{i',j,q}^{(v)}(m-k,n-l) + \sum_{k,l} h_{i,j,q}^{(o,v)}(k,l) \hat{I}_{i,j,q-1}^{(v)}(m-k,n-l), \quad (17)$$

$$\mathbf{I}_{j,q}^{(o,v)}(m,n) = \Big(\sum_{k,l} h_{i,j,q}^{(o,v)}(k,l) \hat{I}_{i',j,q-1}^{(v)}(m-k-r,n-l-s) \Big)_{\substack{(r,s) \in \mathcal{P}_j^{(o,v)} \\ i' \in \mathbb{I}_i}} \quad (18)$$

with $\mathbb{I}_i = \{0,1,2,3\} \backslash \{i\}$ and i is the index number of the polyphase component to be predicted by the current prediction filter under optimization.
Once the different terms involved in the weighted ℓ_1 criterion are defined, the DR algorithm in a three-fold product space could be applied to solve our minimization problem. More details about DR algorithm through a formulation in three fold product space can be found in [22] and references therein.

3.1.3. Full optimization algorithm

As mentioned before, for the left image, the optimization of the filter $\mathbf{P}_j^{(HH,l)}$ depends on the optimization of the filters $\mathbf{P}_j^{(LH,l)}$ and $\mathbf{P}_j^{(HL,l)}$ since the weighted sum of the ℓ_1-norms of all the detail subband coefficients is minimized. On the other hand, the optimization of the filters $\mathbf{P}_j^{(LH,l)}$ and $\mathbf{P}_j^{(HL,l)}$ depends on the optimization of $\mathbf{P}_j^{(HH,l)}$ since the latter allows to compute the diagonal detail subband which is used in the second and third prediction steps. Similarly, for the right image, the optimization of the first four prediction filters depends on the optimization of the last two ones, and vice-versa. Therefore it becomes interesting to resort to an iterative algorithm which jointly optimizes the different prediction filters. To this respect, we start by optimizing all the filters used with the left image independently of the right one. Then, all the filters of the right image will be optimized. Thus, our first independent full optimization algorithm can be described as follows:

First algorithm

① Optimization of the left image filters:

(a) Initialize the iteration number it to 0

• Optimize separately the three prediction filters $\mathbf{P}^{(o,l)}$ by minimizing the ℓ_1 criterion $\mathcal{J}_{\ell_1}^{(l)}(\mathbf{P}_j^{(o,l)})$. The new filters will be designated respectively by $\mathbf{P}_j^{(HH,l,0)}$, $\mathbf{P}_j^{(LH,l,0)}$, and $\mathbf{P}_j^{(HL,l,0)}$.

• Optimize the update filter [21].

• Compute the constant values $\frac{1}{\alpha_{j+1}^{(o,l,0)}}$, the weights $w_{j+1}^{(o,l,0)}$ and the differential entropy of the three resulting detail subbands.

(b) for $it = 1, 2, 3, \ldots$

• Set $\mathbf{P}_j^{(LH,l)} = \mathbf{P}_j^{(LH,l,it-1)}$, $\mathbf{P}_j^{(HL,l)} = \mathbf{P}_j^{(HL,l,it-1)}$, and optimize $\mathbf{P}_j^{(HH,l)}$ by minimizing the weighted ℓ_1 criterion $\mathcal{J}_{w\ell_1}^{(l)}(\mathbf{P}_j^{(HH,l)})$. Let $\mathbf{P}_j^{(HH,l,it)}$ be the new optimal filter.

• Set $\mathbf{P}_j^{(HH,l)} = \mathbf{P}_j^{(HH,l,it)}$, and optimize $\mathbf{P}_j^{(LH,l)}$ as well as $\mathbf{P}_j^{(HL,l)}$ by minimizing $\mathcal{J}_{\ell_1}^{(l)}(\mathbf{P}_j^{(LH,l)})$ and $\mathcal{J}_{\ell_1}^{(l)}(\mathbf{P}_j^{(HL,l)})$, respectively. Let $\mathbf{P}_j^{(LH,l,it)}$ and $\mathbf{P}_j^{(HL,l,it)}$ be the new optimal filters.

• Optimize the update filter.

• Compute the new constant values $\frac{1}{\alpha_{j+1}^{(o,l,it)}}$, the weights $w_{j+1}^{(o,l,it)}$ and the differential entropy of the three resulting detail subbands.

② Optimization of the right image filters:

(a) Initialize the iteration number it to 0

• Optimize separately the three intra prediction filters $\mathbf{P}_j^{(o,r)}$ by minimizing the ℓ_1 criterion $\mathcal{J}_{\ell_1}^{(r)}(\mathbf{P}_j^{(o,r)})$.

• Optimize the update filter.

• Optimize separately the three inter prediction filters $\mathbf{P}_j^{(o,r,l)}$ by minimizing the ℓ_1 criterion $\mathcal{J}_{\ell_1}^{(r)}(\mathbf{P}_j^{(o,r,l)})$.

• Compute the constant values $\frac{1}{\alpha_{j+1}^{(o,r,0)}}$, the weights $w_{j+1}^{(o,r,0)}$ and the differential entropy of the three resulting detail subbands.

(b) for $it = 1, 2, 3, \ldots$

• Optimize $\mathbf{P}_j^{(HH,r)}$, while setting all the other filters equal to those obtained in the previous iteration $(it-1)$, by minimizing the weighted criterion $\mathcal{J}_{w\ell_1}^{(r)}(\mathbf{P}_j^{(HH,r)})$. Let $\mathbf{P}_j^{(HH,r,it)}$ be the new optimal filter.

• Set $\mathbf{P}_j^{(HH,r)} = \mathbf{P}_j^{(HH,r,it)}$, and optimize $\mathbf{P}_j^{(LH,r)}$ by minimizing the weighted criterion $\mathcal{J}_{w\ell_1}^{(r)}(\mathbf{P}_j^{(LH,r)})$. Let $\mathbf{P}_j^{(LH,r,it)}$ be the new optimal filter.

• Set $\mathbf{P}_j^{(LH,r)} = \mathbf{P}_j^{(LH,r,it)}$, and optimize $\mathbf{P}_j^{(HL,r)}$ by minimizing the weighted criterion $\mathcal{J}_{w\ell_1}^{(r)}(\mathbf{P}_j^{(HL,r)})$. Let $\mathbf{P}_j^{(HL,r,it)}$ be the new optimal filter.

• Optimize the update filter.

• Set $\mathbf{P}_j^{(HL,r)} = \mathbf{P}_j^{(HL,r,it)}$, and optimize $\mathbf{P}_j^{(HH,r,l)}$ by minimizing the weighted criterion $\mathcal{J}_{w\ell_1}^{(r)}(\mathbf{P}_j^{(HH,r,l)})$. Let $\mathbf{P}_j^{(HH,r,l,it)}$ be the new optimal filter.

• Set $\mathbf{P}_j^{(HH,r,l)} = \mathbf{P}_j^{(HH,r,l,it)}$, and optimize $\mathbf{P}_j^{(LH,r,l)}$ as well as $\mathbf{P}_j^{(HL,r,l)}$ by minimizing $\mathcal{J}_{\ell_1}^{(r)}(\mathbf{P}_j^{(LH,r,l)})$ and $\mathcal{J}_{\ell_1}^{(r)}(\mathbf{P}_j^{(HL,r,l)})$, respectively. Let $\mathbf{P}_j^{(LH,r,l,it)}$ and $\mathbf{P}_j^{(HL,r,l,it)}$ be the new optimal filters.

• Compute the new constant values $\frac{1}{\alpha_{j+1}^{(o,r,it)}}$, the weights $w_{j+1}^{(o,r,it)}$ and the differential entropy of the three resulting detail subbands.

3.2. Joint optimization method of the stereo pairs

3.2.1. Motivation

According to Fig. 1, the lifting stage applied to the left image is similar to the first stage applied to the right image used to generate the approximation subband $I_{j+1}^{(r)}$ and three intermediate detail subbands $\check{I}_{j+1}^{(o,r)}$ with $o \in \{HH, LH, HL\}$. Moreover, one of the main feature of stereo images is that they present high inter-view correlations since they correspond to the same 3D scene. Therefore, instead of optimizing each view independently of the other one, it becomes interesting to design a joint optimization approach to take into account the previous observations.

3.2.2. Hybrid weighted ℓ_1 minimization and optimization algorithm

To exploit the correlation existing between the left and right, we propose first to assume that the filters of the first lifting stage employed with the right image $(\mathbf{P}_j^{(o,r)}, \mathbf{U}_j^{(r)})$ are similar to those used with the left image $(\mathbf{P}_j^{(o,r)}, \mathbf{U}_j^{(r)})$.

Thus, for notations concision, the three intra prediction filters as well as the update one will be simply denoted by $\mathbf{P}_j^{(o)}$ and $\mathbf{U}_j$:

$$\begin{cases} \mathbf{P}_j^{(HH,r)} = \mathbf{P}_j^{(HH,l)} = \mathbf{P}_j^{(HH)}, \\ \mathbf{P}_j^{(LH,r)} = \mathbf{P}_j^{(LH,l)} = \mathbf{P}_j^{(LH)}, \\ \mathbf{P}_j^{(HL,r)} = \mathbf{P}_j^{(HL,l)} = \mathbf{P}_j^{(HL)}, \\ \mathbf{U}_j^{(r)} = \mathbf{U}_j^{(l)} = \mathbf{U}_j. \end{cases} \tag{19}$$

Moreover, since these filters are applied both to the left and right images, we also propose to design an hybrid weighted ℓ_1 criterion defined simultaneously on the stereo pairs. More precisely, this criterion is the weighted sum of the ℓ_1-norm of the three detail subbands of the left image as well as the three intermediate detail subbands of the right one. Therefore, the new hybrid weighted ℓ_1 criterion can be expressed as follows:

$$\forall\ o \in \{HL, LH, HH\},$$

$$\mathcal{J}_{w\ell_1}^{(r,l)}(\mathbf{P}_j^{(o)}) = \sum_{o\in\{HL,LH,HH\}} \sum_{m=1}^{M_j} \sum_{n=1}^{N_j} \Big(\frac{1}{\alpha_{j+1}^{(o,l)}} \Big| I_{j+1}^{(o,l)}(m,n) \Big| + \frac{1}{\alpha_{j+1}^{(o,r)}} \Big| \check{I}_{j+1}^{(o,r)}(m,n) \Big| \Big) \tag{20}$$

Similarly to the weighted ℓ_1 criterion given by Eq. (14), the new hybrid one will also be minimized using the DR in a product space.

Once the prediction filters used in the first lifting stage of both views have been jointly optimized, the last three inter prediction filters used with the right image will be optimized as performed in the previous optimization algorithm (i.e by minimizing the weighted sum of the ℓ_1-norms of the three detail subbands of the right image).

Therefore, the second optimization algorithm can be summarized as follows.

Second algorithm

① Optimization of the intra prediction filters used with the left and right images:

(a) Initialize the iteration number it to 0

- Optimize separately the three prediction filters $\mathbf{P}_j^{(o)}$ by minimizing

the ℓ_1 criterion $\mathcal{J}_{\ell_1}^{(l)}(\mathbf{P}_j^{(o)})$. The new filters will be designated respectively by $\mathbf{P}_j^{(HH,0)}$, $\mathbf{P}_j^{(LH,0)}$, and $\mathbf{P}_j^{(HL,0)}$.

• Optimize the update filter of the left image.

• Set the intra prediction filters and update one equal to those obtained with the left image (Eq. 19).

• Compute the constant values $\frac{1}{\alpha_{j+1}^{(o,l,0)}}$ and $\frac{1}{\alpha_{j+1}^{(o,r,0)}}$, the weights $w_{j+1}^{(o,l,0)}$ and $w_{j+1}^{(o,r,0)}$, as well as the differential entropy of the six resulting detail subbands.

(b) for $it = 1, 2, 3, \ldots$

• Set $\mathbf{P}_j^{(LH)} = \mathbf{P}_j^{(LH,it-1)}$, $\mathbf{P}_j^{(HL)} = \mathbf{P}_j^{(HL,it-1)}$, and optimize $\mathbf{P}_j^{(HH)}$ by minimizing the hybrid weighted ℓ_1 criterion $\mathcal{J}_{w\ell_1}^{(r,l)}(\mathbf{P}_j^{(HH)})$. Let $\mathbf{P}_j^{(HH,it)}$ be the new optimal filter.

• Set $\mathbf{P}_j^{(HH)} = \mathbf{P}_j^{(HH,it)}$, and optimize $\mathbf{P}_j^{(LH)}$ as well as $\mathbf{P}_j^{(HL)}$ by minimizing $\mathcal{J}_{w\ell_1}^{(r,l)}(\mathbf{P}_j^{(LH)})$ and $\mathcal{J}_{w\ell_1}^{(r,l)}(\mathbf{P}_j^{(HL)})$, respectively. Let $\mathbf{P}_j^{(LH,it)}$ and $\mathbf{P}_j^{(HL,it)}$ be the new optimal filters.

• Optimize the update filter.

• Compute the new constant values $\frac{1}{\alpha_{j+1}^{(o,l,it)}}$ and $\frac{1}{\alpha_{j+1}^{(o,r,it)}}$, the weights $w_{j+1}^{(o,l,it)}$ and $w_{j+1}^{(o,r,it)}$, as well as the differential entropy of the six resulting detail subbands.

② Optimization of the remaining inter prediction filters used with the right image:

(a) Initialize the iteration number it to 0

• Apply the first intra lifting stage to the right image using the optimal filters obtained with the left image, and optimize separately the three inter prediction filters $\mathbf{P}_j^{(o,r,l)}$ by minimizing the ℓ_1 criterion $\mathcal{J}_{\ell_1}^{(r)}(\mathbf{P}_j^{(o,r,l)})$. The new filters will be denoted respectively by $\mathbf{P}_j^{(HH,r,l,0)}$, $\mathbf{P}_j^{(LH,r,l,0)}$, and $\mathbf{P}_j^{(HL,r,l,0)}$.

• Compute the constant values $\frac{1}{\alpha_{j+1}^{(o,r,0)}}$, the weights $w_{j+1}^{(o,r,0)}$ and the differential entropy of the three resulting detail subbands.

(b) for $it = 1, 2, 3, \ldots$

• Set $\mathbf{P}_j^{(LH,r,l)} = \mathbf{P}_j^{(LH,r,l,it-1)}$, $\mathbf{P}_j^{(HL,r,l)} = \mathbf{P}_j^{(HL,r,l,it-1)}$, and optimize $\mathbf{P}_j^{(HH,r,l)}$ by minimizing the weighted criterion $\mathcal{J}_{w\ell_1}^{(r)}(\mathbf{P}_j^{(HH,r,l)})$. Let $\mathbf{P}_j^{(HH,r,l,it)}$ be the new optimal filter.

• Set $\mathbf{P}_j^{(HH,r,l)} = \mathbf{P}_j^{(HH,r,l,it)}$, and optimize $\mathbf{P}_j^{(LH,r,l)}$ as well as $\mathbf{P}_j^{(HL,r,l)}$ by minimizing $\mathcal{J}_{\ell_1}^{(r)}(\mathbf{P}_j^{(LH,r,l)})$ and $\mathcal{J}_{\ell_1}^{(r)}(\mathbf{P}_j^{(HL,r,l)})$, respectively. Let $\mathbf{P}_j^{(LH,r,l,it)}$ and $\mathbf{P}_j^{(HL,r,l,it)}$ be the new optimal filters.

• Compute the new constant values $\frac{1}{\alpha_{j+1}^{(o,r,it)}}$, the weights $w_{j+1}^{(o,r,it)}$ and the differential entropy of the three resulting detail subbands.

It is important to note here that the convergence of the two proposed optimization algorithms is achieved in few iterations (after about 5 or 6 iterations) where the weighted ℓ_1 minimization technique performed on each prediction filter takes about 4-5 seconds for an image of size 512×512 using a Matlab implementation. For instance, compared to the optimization strategy developed in [19], the proposed joint optimization algorithm increases slightly the execution time (2 seconds per filter) since an hybrid weighted ℓ_1 minimization technique (given by Eq. (20)) is employed. Moreover, this joint algorithm presents two main advantages compared to the first independent full one. Indeed, in addition to an efficient exploitation of the characteristics of the stereo images through the design of an hybrid criterion, it simplifies the optimization process and reduces the bitrate of the filter coefficients that should be transmitted to the decoder (thanks to the assumption given by Eq. (19)).

4. Experimental results

Simulations were conducted on different stereo images taken from various datasets such as VASC CMU and middlebury ones [35, 36]. In order to illustrate the proposed sparse optimization of NS-VLS in the context of stereo image coding, and since our non separable lifting structure is the 2D extension of 1D P-U LS (as explained at the beginning of Sec. 2), we will consider the P-U 5/3 LS, known also as (2,2) wavelet transform, which has been selected for the JPEG2000 coding standard. Note that the spatial prediction and update

supports of its extended 2D non separable structure (used with the left image and the right one in the first lifting stage) can be found in [21]. For the second lifting stage used with the right image, the intra prediction filters $\mathcal{Q}_{i,j}^{(o,r)}$ have the same spatial supports of the other prediction filters used with the left image as well as the first intra lifting stage employed with the right image, while the spatial supports of the inter prediction filters $\widetilde{\mathbf{P}}_{i,j}^{(o,r,l)}$ are defined by the set $\widetilde{\mathcal{P}}_{i,j}^{(o,r,l)} = \{(s,t),\ \text{with } s \in \{-1,0,1\} \text{ and } t \in \{-1,0,1\}\}$ (where, $s = t = 0$ corresponds to the matching pixel in the disparity compensated left image $I_j^{(c)}(m,n)$ of the current sample to be predicted $I_{i,j}^{(r)}(m,n)$).
Thus, to show the performance of the proposed optimization methods, we will consider the following ones carried out over three resolution levels:

- The first one consists in coding independently the left and right images by applying the 5/3 transform to each view. In the following, this method will be denoted by "Independent".

- The second method represents the state-of-the-art method which consists in coding the left image and the residual one by using the 5/3 wavelet transform. Let us recall that this approach, which will be designated by "Residual", is behind most of the developed stereo image coding schemes.

- While the residual image is generated in the previous method by computing the prediction error between each pixel of the target image and its corresponding one in the reference one, the third method proposed in [7] aims to use the neighborhood of the homologous pixel to predict the pixel of the target image. This computation step is optimized by minimizing the ℓ_1-norm of the resulting prediction error. This method will be designated by "Residual-OPT-L1 [7]".

- The fourth one corresponds to the NS-VLS where the prediction filters are optimized separately by minimizing the variance (i.e ℓ_2-norm) of the detail coefficients. This method will be denoted by "NS-VLS-OPT-L2".

- The fifth method corresponds to the NS-VLS where the prediction filters

are optimized separately by minimizing the ℓ_1-norm of the detail coefficients. This method will be denoted by "NS-VLS-OPT-L1".

- The sixth and seventh methods represent the proposed independent full and joint algorithms used to optimize the NS-VLS. These methods will be designated by "NS-VLS-OPT-WL1-Full" and "NS-VLS-OPT-WL1-Joint", respectively.

All these approaches are firstly compared in terms of rate-distortion performance. Figs. 2 and 3 illustrate the variations of the PSNR versus the bitrate for the "houseof" and "ball" stereo images. Note that the average bitrate as well as the average reconstruction error have been used to evaluate the performance of all these methods. It can be observed that residual-based coding scheme leads to better results compared to the independent coding scheme especially at low bitrate. The optimized residual scheme [7] outperforms the previous one by about 0.15-0.5 dB. Moreover, using sparse ℓ_1 minimization technique improves the ℓ_2 one by achieving a gain of about 0.2-0.4 dB in terms of PSNR. Further improvements are achieved by resorting to the proposed fully and joint weighted ℓ_1 minimization techniques. The gain is about 0.1-0.2 dB compared to the standard ℓ_1 minimization technique. It should be also noted that the results obtained by the joint optimization approach are close to those obtained with the independent fully approach. For instance, one can observe that a very small improvement is achieved by the independent full optimization strategy. Such behavior is expected since, in the full optimization approach, all filters of both views are optimized, whereas in the joint optimization strategy, the intra prediction filters used in the first lifting stage with the right view are assumed to be equal to those used with the left image. Thus, due to the aforementioned advantages of the joint optimization approach (simplifying the optimization process and reducing the transmission cost of the filter coefficients), the latter optimization algorithm is more appropriate from a practical point of view.

After that, and since the joint and full independent optimization methods have similar performances, we will evaluate the relative gain of the proposed joint op-

timization algorithm "NS-VLS-OPT-WL1-Joint" by using the Bjontegaard metric [37]. For instance, the gains of the joint optimization algorithm with respect to the standard ℓ_2 one "NS-VLS-OPT-L2" as well as the ℓ_1 one "NS-VLS-OPT-L1" are provided in Tables 1 and 2 for low, middle and high bitrates, which are obtained by considering the following four bitrate points $\{0.15, 0.2, 0.25, 0.3\}$, $\{0.6, 0.65, 0.7, 0.75\}$ and $\{1.25, 1.3, 1.35, 1.4\}$ bpp, respectively. Note that a bitrate saving with respect to a given reference method corresponds to a negative value. Compared to the sparse ℓ_1 optimization, the joint optimization algorithm leads to a gain of about 1-3% and 0.1-0.2 dB in terms of bitrate saving and PSNR, respectively. The gain becomes much more important compared to the standard ℓ_2 optimization algorithm and it reaches 11.6% and 0.85 dB in terms of bitrate saving and PSNR, respectively.

Finally, we have evaluated the proposed joint optimization algorithm in terms of visual quality of reconstruction. Figs. 4 and 5 display the reconstructed (i.e decoded) target images, for the "art" and "dolls" stereo pairs, using joint optimization algorithm, the standard independent ℓ_2 one as well as the state-of-the-art residual based coding method. Their corresponding PSNR and SSIM [38] values are also provided. Note that a zoom is applied to the reconstructed images to better illustrate the differences between them. It can be noticed that the residual-based coding method may lead to some blocking artifacts. This problem is reduced by resorting to a NS-VLS decomposition, as it can be seen from the results obtained with "NS-VLS-OPT-L2". Moreover, compared to the latter, the joint optimization method improves again the visual quality while better preserving the object edges.

5. Conclusion

In this paper, we have investigated different sparse optimization techniques to design the prediction filters of a non separable vector lifting scheme for stereo image coding purpose. In this context, an independent full optimization of the left and right image filters as well as a joint optimization method have been

developed. Experimental results carried out on different real stereo images have shown the efficiency of these methods. In a future work, we plan to extend this work to the context of multiview coding. Moreover, other perceptual criteria could be investigated for the design of the different lifting operators.

Acknowledgements

This work was made possible by NPRP grant number NPRP8-140-2-065 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

References

[1] Y. Munz, K. Moorthy, A. Dosis, J. D. Hernandez, S. Bann, F. Bello, S. Martin, A.Darzi, T. Rockall, The benefits of stereoscopic vision in robotic-assisted performance on bench models, Surgical endoscopy 18 (4) (2004) 611–616.

[2] I. Feldmann, W. Waizenegger, N. Atzpadin, O. Schreer, Real-time depth estimation for immersive 3D videoconferencing, in: 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video, Tampere, 2010, pp. 1–4.

[3] A. Kadaikar, G. Dauphin, A. Mokraoui, Sequential block-based disparity map estimation algorithm for stereoscopic image coding, Elsevier Signal Processing: Image Communication 39 (PA) (2015) 159–172.

[4] D. Palaz, I. Tosic, P. Frossard, Sparse stereo image coding with learned dictionaries, in: IEEE International Conference on Image Processing, Quebec, Canada, 2011, p. 4 pages.

[5] G. Dauphin, M. Kaaniche, A. Mokraoui, Block dependent dictionary based disparity compensation for stereo image coding, in: IEEE International Conference on Image Processing, Quebec, Canada, 2015, p. 5 pages.

[6] L. F. Lucas, N. M. Rodrigues, C. L. Pagliari, E. A. Silva, S. M. Faria, Recurrent pattern matching based stereo image coding using linear predictors, Multidimensional Systems and Signal Processing 28 (4) (2017) 1393–1416.

[7] W. Hachicha, M. Kaaniche, A. Beghdadi, F. A. Cheikh, Optimized residual image for stereo image coding, in: European Workshop on Visual Information Processing, Paris, France, 2014, p. 6 pages.

[8] I. Tabus, Patch-based conditional context coding of stereo disparity images, IEEE Signal Processing Letters 21 (10) (2014) 1220–1224.

[9] O. Woo, A. Ortega, Stereo image compression based on disparity field segmentation, in: SPIE Conference on Visual Communications and Image Processing, Vol. 3024, San Jose, California, 1997, pp. 391–402.

[10] W. Hachicha, A. Beghdadi, F. A. Cheikh, 1D directional DCT-based stereo residual compression, in: European Signal Processing Conference, Marrakech, Morocco, 2013, p. 5 pages.

[11] M. S. Moellenhoff, M. W. Maier, Characteristics of disparity-compensated stereo image pair residuals, Signal Processing: Image Communications 14 (1998) 49–55.

[12] M. S. Moellenhoff, M. W. Maier, Transform coding of stereo image residuals, IEEE Transactions on Image Processing 7 (6) (1998) 804–812.

[13] N. V. Boulgouris, M. G. Strintzis, A family of wavelet-based stereo image coders, IEEE Transactions on Circuits and Systems for Video Technology 12 (10) (2002) 898–903.

[14] R. Darazi, A. Gouze, B. Macq, Adaptive lifting scheme-based method for joint coding 3D-stereo images with luminance correction and optimized prediction, in: IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, 2009, pp. 917–920.

[15] A. Maalouf, M.-C. Larabi, Bandelet-based stereo image coding, in: IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, Texas, United States, 2010, pp. 698–701.

[16] E. L. Pennec, S. Mallat, Sparse geometric image representations with bandelets, IEEE Transactions on Image Processing 14 (4) (2005) 423–438.

[17] M. Kaaniche, A. Benazza-Benyahia, B. Pesquet-Popescu, J.-C. Pesquet, Vector lifting schemes for stereo image coding, IEEE Transactions on Image Processing 18 (11) (2009) 2463–2475.

[18] O. Dhifallah, M. Kaaniche, A. Benazza-Benyahia, Efficient joint multiscale decomposition for color stereo image coding, in: European Signal Processing Conference, Lisbon, Portugal, 2014, p. 5 pages.

[19] M. Kaaniche, B. Pesquet-Popescu, J.-C. Pesquet, ℓ_1-adapted non separable vector lifting schemes for stereo image, in: European Signal Processing Conference, Bucharest, Romania, 2012, p. 5 pages.

[20] V. Chappelier, C. Guillemot, Oriented wavelet transform for image compression and denoising, IEEE Transactions on Image Processing 15 (10) (2006) 2892–2903.

[21] M. Kaaniche, A. Benazza-Benyahia, B. Pesquet-Popescu, J.-C. Pesquet, Non separable lifting scheme with adaptive update step for still and stereo image coding, Elsevier Signal Processing: Special issue on Advances in Multirate Filter Bank Structures and Multiscale Representations 91 (12) (2011) 2767–2782.

[22] M. Kaaniche, B. Pesquet-Popescu, A. Benazza-Benyahia, J.-C. Pesquet, Adaptive lifting scheme with sparse criteria for image coding, EURASIP Journal on Advances in Signal Processing: Special Issue on New Image and Video Representations Based on Sparsity 2012, 22 pages.

[23] Y. Xing, M. Kaaniche, B. Pesquet-Popescu, F. Dufaux, Sparse based adaptive non separable vector lifting scheme for holograms compression, in: International Conference on 3D Imaging, Liège, Belgium, 2015, p. 8 pages.

[24] W. Sweldens, The lifting scheme: a new philosophy in biorthogonal wavelet construction, in: Wavelet Applications in Signal and Image Processing III, SPIE, San-Diego, CA, USA, 1995, pp. 68–79.

[25] D. Taubman, M. Marcellin, JPEG2000: Image Compression Fundamentals, Standards and Practice, Kluwer Academic Publishers, Norwell, MA, USA, 2001.

[26] O. N. Gerek, A. E. Çetin, Adaptive polyphase subband decomposition structures for image compression, IEEE Transactions on Image Processing 9 (10) (2000) 1649–1660.

[27] A. Gouze, M. Antonini, M. Barlaud, B. Macq, Design of signal-adapted multidimensional lifting schemes for lossy coding, IEEE Transactions on Image Processing 13 (12) (2004) 1589–1603.

[28] J. Chen, J. Hou, L.-P. Chau, Light field compression with disparity-guided sparse coding based on structural key views, IEEE Transactions on Image Processing 27 (1) (2018) 314–324.

[29] H. Gish, J. N. Pierce, Asymptotically efficient quantizing, IEEE Transactions on Information Theory 14 (5) (1969) 676–683.

[30] B. Usevitch, Optimal bit allocation for biorthogonal wavelet coding, in: Data Compression Conference, Snowbird, USA, 1996, pp. 387–395.

[31] S. Parrilli, M. Cagnazzo, B. Pesquet-Popescu, Distortion evaluation in transform domain for adaptive lifting schemes, in: International Workshop on Multimedia Signal Processing, Cairns, Queensland, Australia, 2008, p. 6 pages.

[32] J.-J. Moreau, Proximité et dualité dans un espace hilbertien, Bulletin de la Societé Mathématique de France 93 (1965) 273–288.

[33] C. Chaux, P. Combettes, J.-C. Pesquet, V. Wajs, A variational formulation for framebased inverse problems, Inverse Problems 23 (4) (2007) 1495–1518.

[34] P. L. Combettes, V. R. Wajs, Signal recovery by proximal forward-backward splitting, Multiscale Modeling and Simulation 4 (4) (2005) 1168–1200.

[35] H. Hirschmller, D. Scharstein, Evaluation of cost functions for stereo matching, in: International Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 2007, p. 8 pages.

[36] D. Scharstein, H. Hirschmüller, Y. Kitajima, G. Krathwohl, N. Nesic, X. Wang, P. Westling, High-resolution stereo datasets with subpixel-accurate ground truth, in: German Conference on Pattern Recognition, Münster, Germany, 2014, p. 12 pages.

[37] G. Bjontegaard, Calculation of average PSNR differences between RD curves, Tech. rep., ITU SG16 VCEG-M33, Austin, TX, USA (April 2001).

[38] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Transactions on Image Processing 13 (4) (2004) 600–612.

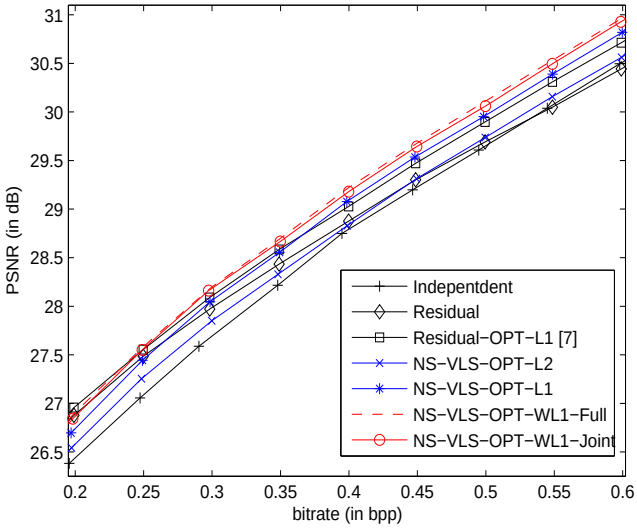

Figure 2: **PSNR (in dB) versus the bitrate (bpp) after JPEG2000 encoding for the "houseof" stereo pair.**

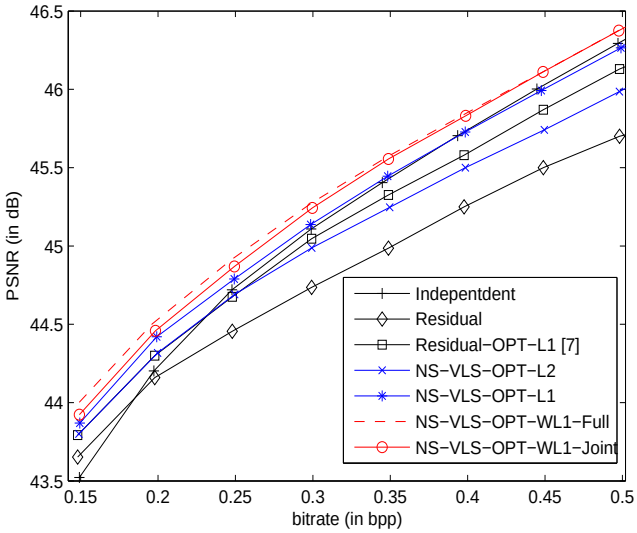

Figure 3: **PSNR (in dB) versus the bitrate (bpp) after JPEG2000 encoding for the "ball" stereo pair.**

Table 1: The average PSNR differences and the bitrate saving at low, medium and high bitrates. The gain of "NS-VLS-OPT-WL1-Joint" w.r.t NS-VLS-OPT-L1.

	bitrate saving (in %)			**PSNR gain (in dB)**		
Images	**low**	**middle**	**high**	**low**	**middle**	**high**
Houseof	-2.95	-2.30	-1.38	0.10	0.12	0.11
Ball	-1.83	-2.81	-1.98	0.05	0.10	0.11
Moebius	-0.93	-2.28	-1.14	0.05	0.14	0.10
Art	-0.92	-0.66	-0.56	0.05	0.06	0.06
Dolls	-2.14	-1.33	-1.03	0.10	0.10	0.11
Playtable	-2.15	-3.05	-1.93	0.10	0.18	0.19
Piano	-1.90	-1.17	-1.24	0.10	0.09	0.15
Teddy	-1.30	-1.11	-0.95	0.07	0.08	0.08
Jadeplant	-2.34	-1.14	-1.08	0.11	0.10	0.11

Table 2: The average PSNR differences and the bitrate saving at low, medium and high bitrates. The gain of "NS-VLS-OPT-WL1-Joint" w.r.t NS-VLS-OPT-L2.

	bitrate saving (in %)			**PSNR gain (in dB)**		
Images	**low**	**middle**	**high**	**low**	**middle**	**high**
Houseof	-8.50	-7.44	-5.54	0.25	0.38	0.39
Ball	-7.74	-11.67	-6.67	0.14	0.39	0.36
Moebius	-7.35	-8.96	-7.20	0.33	0.55	0.61
Art	-7.78	-9.17	-8.26	0.41	0.64	0.86
Dolls	-7.78	-6.12	-6.45	0.32	0.41	0.67
Playtable	-3.70	-8.29	-7.48	0.17	0.49	0.71
Piano	-7.32	-6.92	-6.13	0.32	0.50	0.71
Teddy	-4.41	-4.81	-4.85	0.18	0.32	0.41
Jadeplant	-6.61	-5.69	-5.06	0.28	0.40	0.50

(a)

(b): PSNR=31.14 dB, SSIM=0.818

(c): PSNR=31.04 dB, SSIM=0.829

(d): PSNR=31.45 dB, SSIM=0.848

Figure 4: (a) Original "art" right image. Zoom applied on the reconstructed image at 0.3 bpp using: (b) Residual scheme, (c) NS-VLS-OPT-L2, (d) NS-VLS-OPT2-WL1-Joint.

(a)

(b): PSNR=27.62 dB, SSIM=0.720

(c): PSNR=28.17 dB, SSIM=0.766

(d): PSNR=28.43 dB, SSIM=0.780

Figure 5: (a) Original "dolls" right image. Zoom applied on the reconstructed image at 0.3 bpp using: (b) Residual scheme, (c) NS-VLS-OPT-L2, (d) NS-VLS-OPT2-WL1-Joint.