# Efficient inter-view bit allocation methods for stereo image coding

Walid Hachicha, Mounir Kaaniche, Azeddine Beghdadi, Faouzi Alaya Cheikh

# Efficient inter-view bit allocation methods for stereo image coding

W. Hachicha, *Student Member IEEE,* M. Kaaniche, *Member IEEE,* A. Beghdadi, *Senior Member IEEE,* and F. A. Cheikh, *Senior Member IEEE*

**Abstract**

In this paper, we present efficient bit allocation methods for stereo image coding purpose. Since the common idea behind most of the existing stereo compression schemes consists of encoding a reference and residual images as well as a disparity map, we mainly focus on the bit allocation issue between the reference and residual images. Generally, this problem is solved in an empirical manner by looking for the optimal rates leading to the minimum distortion value. Thanks to recent approximations of the entropy and distortion functions, we propose accurate and fast bit allocation schemes appropriate for the open-loop- and closed-loop-based stereo coding structures. Experimental results show the benefits which can be drawn from the proposed bit allocation methods.

**Index Terms**

Stereo image, compression, bit allocation, wavelets, quantization, rate-distortion theory.

## I. Introduction

During the last decade, the growing demand for 3D imaging and the recent developments in display and acquisition technologies have increased the usage of stereovision in various application fields such as 3DTV, 3D cinema, video games and videoconferencing. The principle of the stereoscopic system consists of two views, called left and right images, obtained by recording the same scene from two slightly different view angles. As a result, the involved stereo images become prohibitive and constitute

W. Hachicha, M. Kaaniche and A. Beghdadi are with L2TI-Institut Galilée, Université Paris 13, Sorbonne Paris Cité, 99 avenue J. B. Clément, Villetaneuse, France, 93430. E-mail: walid.hachicha@univ-paris13.fr, mounir.kaaniche@univ-paris13.fr, azeddine.beghdadi@univ-paris13.fr. F. A. Cheikh is with the Norwegian Colour and Visual Computing Lab, Gjøvik University College, Norway. E-mail: faouzi.cheikh@hig.no.

a major limitation for developing efficient stereoscopic systems. Therefore, a compression scheme is required to ensure a rapid transmission and compact storage of the stereo images.

To this end, conventional still image compression techniques, such as JPEG [1] and JPEG2000 [2], can be simply used to encode each view. However, this independent coding scheme is not so efficient since it does not exploit the strong correlation between the two views. Indeed, stereo images have similar contents since they correspond to the same 3D scene. Generally, the left and right pixels resulting from the projection of the same 3D point have different positions, and the estimation of the spatial displacement across all the pixels leads to the so-called disparity map. Thus, in order to take into account the inter-image redundancies, more efficient coding schemes have been developed based on the disparity estimation/compensation process [3], [4]. More precisely, the idea behind most of the existing works consists firstly in selecting one image (for example the left one) as a reference image. Then, the other one (target image) is predicted from the reference using the estimated disparity map, and the difference between the original right image and the predicted one, called residual image, is generated. Finally, the reference and residual images as well as the disparity map are encoded. To this respect, the disparity map is often losslessly encoded using a DPCM technique followed by an entropy coder whereas the reference and residual images can be encoded in different transform domains. Indeed, many works have retained the Discrete Cosine Transform (DCT) domain [4], [5], [6], [7]. Moellenhoff and Maier [8] have shown that residual images have two specific properties in the sense that most of their intensity values are near zero and they contain very narrow vertical edges. Thus, they propose to code the stereo residuals with two dimensional (2D)-DCT and a wavelet transform followed by an updated quantization matrix. In [9], a mixed transform is also employed by applying a DCT for the best matching blocks and a Haar wavelet transform for the occluded blocks. Recently, Hachicha *et al.* have proposed a stereo image coder based on the directional structure of the residual image [10]. The latter is transformed using a 1D Directional DCT along the horizontal, vertical, positive diagonal and negative diagonal directions. In addition, due to the well known limitation of a DCT-based coding scheme, recent works have preferred to employ the Discrete Wavelet Transform (DWT). In [11], the authors propose to encode the reference and residual images using different families of DWT followed by an embedded zero-tree coder [12]. Moreover, a new technique based on the concept of Vector Lifting Schemes has been developed in [13]. It consists of designing a joint multiscale representation which encodes simultaneously the left and right images. In a recent work [14], the authors propose to generate the disparity map and the residual one by applying a bandelet transform to the left and the right images.

While most of the existing works are based on the encoding of the reference and residual images

as well as the disparity map, the main contributions of the developed works have been devoted to the three following issues. The first one consists in designing an efficient disparity estimation/compensation technique [15], [16]. The second one is related to the improvement of the transform applied to the images [9], [13]. The third issue focuses on the entropy encoding of the transformed coefficients [17]. However, the bit allocation issue for stereo image coding has rarely been addressed. Indeed, the main question that arises here is: how to distribute an available budget of bits between the two views to be encoded? It is important to note here that the problem of bit allocation has already been investigated in image and video compression standards [18]. Moreover, this problem has been studied in the context of Multi-view image coding in a depth-image-based-rendering framework [19], [20], [21], [22].

Generally, the main idea behind bit allocation is based on Rate-Distortion (R-D) theory and aims at minimizing the average distortion subject to a constraint on the global bitrate. To solve this constrained optimization problem, numerical and analytical methods have been developed. The first ones aim at empirically estimating the R-D plots and resort to some iterative techniques such as the well known Lagrangian approach [23], [24], [25]. The main drawback of this first category of methods is that the generation of the R-D curves presents a high computational cost since a large number of R-D points must be measured. For this reason, analytical methods have been investigated by considering various hypotheses on the distribution and quantizer characteristics in order to provide closed-form expressions of the R-D function [26], [27], [28], [29]. In the context of stereo image coding, there is only few research works devoted to bit allocation between the two views. Among them, Woo and Ortega proposed in [30] a dependent bit allocation scheme between the left and residual images by using dynamic programming technique. However, in most of the developed stereo image coding schemes, the bit allocation problem is solved in an empirical manner. Indeed, for a given budget of bits, the reference and residual images are firstly encoded/decoded at different possible couple of bitrates. Then, the configuration leading to the minimum distortion is selected as the optimal solution. Thus, when the global bitrate increases, the number of configurations becomes very important, and therefore, the optimal coding scheme of the stereo images becomes computationally intensive.

The objective of this paper is to design efficient bit allocation techniques for optimal stereo image coding. To this end, we will consider two standard schemes, based on open-loop and closed-loop structures, which constitute the basic idea behind all the existing stereo compression techniques [11], [13], [15]. More precisely, by resorting to the Lagrangian optimization tool, we develop two approaches. Indeed, in the case of an open-loop structure, and based on recent approximation formulas of the entropy and distortion functions [29], we derive explicit expression of the system of equations that satisfies the

Lagrangian multiplier, which allows us to propose a fast and accurate bit allocation strategy. After that, in the context of a closed-loop-based coding scheme, we propose a two dimensional exponential model for modeling the rate-distortion functions of the reference and residual images.

The remainder of this paper is organized as follows. In Section II, we firstly recall the principle of the retained open and closed loop schemes for stereo image coding. We introduce also the source and quantization models. Then, the two proposed bit allocation methods are described in Section III. Finally, the optimality of these methods in terms of rate-distortion are shown in Section IV and some conclusions are drawn in Section V.

## II. Stereo image coding schemes

As mentioned before, the common idea behind most of the state-of-the-art stereo image compression methods consists of encoding the reference image, the residual one and the disparity field. However, depending on the Disparity Estimation/Disparity Compensation (DE/DC) stage, two main structures, called Open-Loop (OL) and Closed-Loop (CL), are generally employed.

### *A. Retained OL- and CL-based coding schemes*

Fig. 1(a) illustrates the block diagram of an OL-based stereo image coding system.

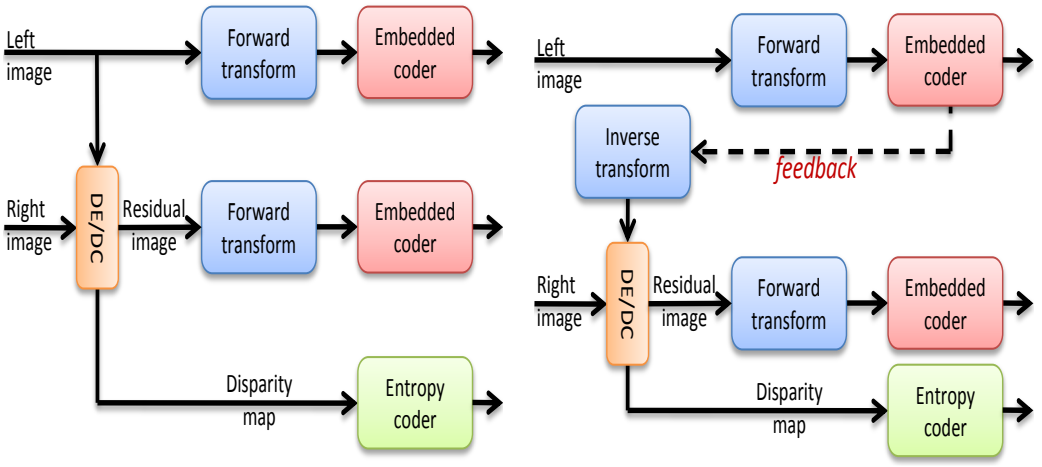



Fig. 1: Stereo image coding scheme based on: (a) Open-Loop structure (b) Closed-Loop structure.

As it can be seen in this figure, the disparity map between the left and right images is firstly estimated.

Then, the left image, selected as a reference, is used to predict the right image through the disparity compensation process. After that, the residual image is generated by computing the difference between the original right image and the predicted one. Finally, the residual image as well as the reference one and the disparity map are encoded.

While this OL structure is generally used due to its simplicity, it presents the following drawback. Indeed, at the decoder side, and knowing that the disparity is often losslessly encoded/decoded, the reference image is firstly decoded. Then, the residual image is decoded and added to the compensation of the reconstructed left image to recover the right image. Thus, it is clear that the encoder and decoder do not use the same reference image during the prediction stage (i.e compensation process), which makes this first structure sub-optimal.

To overcome this problem, a closed-loop structure has been developed by including a feedback in the encoder as shown in Fig. 1(b). More precisely, the reference image (i.e the left one) is firstly encoded and decoded. Then, in order to generate the residual image, the disparity estimation and compensation processes are performed on the *reconstructed* left image, instead of the *original* left image as used with the OL architecture. In order to reduce the complexity of this structure, some authors have considered that the disparity map is estimated only once from the original reference and target images, while keeping the remaining steps unchanged [30]. Note that this consideration will also be adopted in this work. Therefore, the main difference between the retained OL and CL coding structures concerns only the computation of the residual image. This latter consists in applying the disparity compensation process on the left image followed by a subtraction operation in order to generate the prediction error between the original right image and the predicted one. While the residual image is computed only once from the original right and left images in an OL structure, this computation step is performed many times in the case of a CL structure. More precisely, for each target bitrate, the residual image is generated for different configurations of the decoded left image, which makes the complexity of a CL-based coding scheme higher than that involved in an OL architecture.

It is worth pointing out that a previous study in terms of reconstruction error has shown a main advantage of the CL coding system [11]. Indeed, using such a structure ensures that the distortions of the left and residual images are equal to those evaluated on the left and right images, which constitutes an important feature for the inter-view bit allocation problem. While a CL architecture is well known to be more performant than the OL one, it is important to note that a CL-based coding scheme is more difficult to implement in a real-time encoder. Therefore, in practice, one should consider the trade-off between performance and complexity in the selection of the coding architecture [31].

*B. Source and quantization models*

Once the considered stereo image coding structures have been defined, the reference and residual images, denoted by $I^{(l)}$ and $I^{(e)}$ respectively, are decorrelated by applying an appropriate transformation. To this end, wavelet transforms have been widely used in the most recent image compression algorithms such as the JPEG2000 standard. After that, the transformed coefficients are quantized and encoded using some embedded coders like EBCOT. To this respect, let us define in what follows the source and quantization models.

The sources to be quantized correspond to the $J$ subbands of the left and residual images which will be designated by $I_j^{(l)}$ and $I_j^{(e)}$ with $j \in \{1, \ldots, J\}$. An appropriate distribution for modelling the resulting coefficients is the Generalized Gaussian (GG). The GG probability density function is given by:

$$\forall \xi \in \mathbb{R}, \qquad \tilde{f}_j^{(v)}(\xi) = \frac{\beta_j^{(v)} (\omega_j^{(v)})^{1/\beta_j^{(v)}}}{2\Gamma(1/\beta_j^{(v)})} e^{-\omega_j^{(v)} |\xi|^{\beta_j^{(v)}}} \tag{1}$$

where the superscript $v \in \{l, e\}$ is used to distinguish between the left view and the residual one, $\Gamma$ is the gamma function, $\omega_j^{(v)}$ and $\beta_j^{(v)}$ are respectively the scale and shape parameters which can be estimated using the method of moments or maximum likelihood techniques [32].

During the quantization process, for both images subbands $I_j^{(l)}$ and $I_j^{(e)}$, a uniform scalar quantizer with a quantization step $q_j^{(v)}$, and having a deadzone of size $(2\tau_j^{(v)} - 1)q_j^{(v)}$, where $\tau_j^{(v)} > \frac{1}{2}$ and $v \in \{l, e\}$, is employed. Note that such a quantizer is generally used in wavelet-based image compression schemes [2], [33]. Thus, for each input coefficient, $I_{j,s}^{(v)}$ where $s$ denotes the spatial position, the output of the quantizer $\overline{I}_{j,s}^{(v)}$ is expressed as

$$\overline{I}_{j,s}^{(v)} = 0, \quad \text{if } |\overline{I}_{j,s}^{(v)}| < \left(\tau_j^{(v)} - \tfrac{1}{2}\right) q_j^{(v)}, \quad \text{where } \tau_j^{(v)} > 1/2 \tag{2}$$

and, for all $i \in \mathbb{Z}$, $\overline{I}_{j,s}^{(v)} = r_{i,j}^{(v)}$,

$$\begin{aligned} &(\text{if } (\tau_j^{(v)} + i - \tfrac{3}{2})q_j^{(v)} \leq \overline{I}_{j,s}^{(v)} < (\tau_j^{(v)} + i - \tfrac{1}{2})q_j^{(v)} \text{ and } i \geq 1) \\ \text{or } &(\text{if } (-\tau_j^{(v)} + i + \tfrac{1}{2})q_j^{(v)} < \overline{I}_{j,s}^{(v)} \leq (-\tau_j^{(v)} + i + \tfrac{3}{2})q_j^{(v)} \text{ and } i \leq -1), \end{aligned}$$

where the reconstruction levels are given by

$$\forall i \geq 1, \qquad r_{i,j}^{(v)} = -r_{-i,j}^{(v)} = (\tau_j^{(v)} + i - 1)q_j^{(v)} \tag{3}$$

Based on these source and quantization models, we focus now on the bit allocation problem for encoding the quantized coefficients of the left and residual images.

## III. Proposed bit allocation methods

The objective of this section is to develop fast and efficient bit allocation methods for stereo image coding systems based on the standard OL and CL structures.

### A. Motivation

As mentioned in Section I, a crucial step of the bit allocation algorithms consists in estimating the R-D functions. More precisely, in a subband coding context (i.e. wavelet-based coders), the R-D curves must be generated for each wavelet subband. To this end, a large number of R-D operating points, that range from low to high bitrate, must be measured in order to obtain accurate R-D curves. So, such approach is extremely complex. Moreover, the generated R-D curves are found to be much irregular, especially at low bit-rates, and may have a nonconvex behavior, which would affect the robustness of the bit allocation algorithms. To overcome these drawbacks, we prefer in this paper to resort to an analytical approach by exploiting recent approximations of the entropy and distortion functions. While analytical-based R-D algorithms use generally high rate-distortion models, it is important to note that our approximations are very precise at low and high bitrate.

Before describing the proposed bit allocation methods, we firstly introduce the rate and distortion approximation functions.

### B. Rate and distortion approximation

As frequently considered in the design of R-D algorithms, we approximate the bitrate of the sources by a zero-order entropy of the quantized coefficients [34]. Thus, by assuming that the wavelet coefficients are modeled by a GG distribution, and based on recent approximation of the entropy of a quantized GG random variable [29], the entropy of the quantized subbands $\overline{I}_j^{(l)}$ and $\overline{I}_j^{(e)}$, denoted respectively by $H_j^{(l)}$ and $H_j^{(e)}$, can be approximated as follows:

$$
\begin{aligned}
&\text{for } v \in \{l, e\},\\
&H_j^{(v)}(q_j^{(v)}) = -\mathsf{p}_{0,j}^{(v)} \ln \mathsf{p}_{0,j}^{(v)} - 2\mathsf{p}_{1,j}^{(v)} \ln \mathsf{p}_{1,j}^{(v)} + \left(h_{\beta_j^{(v)}}(1) - \ln((\omega_j^{(v)})^{1/\beta_j^{(v)}} q_j^{(v)})\right)\\
&\qquad \times \left(1 - Q_{1/\beta_j^{(v)}}\left(\omega_j^{(v)}\left(\tau_j^{(v)} + \frac{1}{2}\right)^{\beta_j^{(v)}} (q_j^{(v)})^{\beta_j^{(v)}}\right)\right)\\
&\qquad + \frac{{\omega_j^{(v)}}^{1/\beta_j^{(v)}} (\tau_j^{(v)} + \frac{1}{2}) q_j^{(v)}}{\Gamma(1/\beta_j^{(v)})}\, e^{-\omega_j^{(v)}(\tau_j^{(v)}+\frac{1}{2})^{\beta_j^{(v)}} (q_j^{(v)})^{\beta_j^{(v)}}},
\end{aligned}
\tag{4}
$$

where

- $\mathsf{p}_{0,j}^{(v)}$ is the probability of the zero level:

$$\mathsf{p}_{0,j}^{(v)} = 2\int_0^{q_j{}^{(v)}(\tau_j^{(v)}-\frac{1}{2})} \widetilde{f}_j^{(v)}(\xi)d\xi = Q_{1/\beta_j^{(v)}}\Big(\omega_j^{(v)}(\tau_j^{(v)}-\frac{1}{2})^{\beta_j^{(v)}}(q_j^{(v)})^{\beta_j^{(v)}}\Big), \tag{5}$$

- $\mathsf{p}_{1,j}^{(v)}$ is the probability of the $r_{1,j}^{(v)}$ reconstruction level

$$\begin{aligned}\mathsf{p}_{1,j}^{(v)} = \int_{(\tau_j^{(v)}-\frac{1}{2})q_j^{(v)}}^{(\tau_j^{(v)}+\frac{1}{2})q_j^{(v)}} \widetilde{f}_j^{(v)}(\xi)d\xi =& \frac{1}{2}Q_{1/\beta_j^{(v)}}\Big(\omega_j^{(v)}\big((\tau_j^{(v)}+\frac{1}{2})q_j^{(v)}\big)^{\beta_j^{(v)}}\Big) \\ &- \frac{1}{2}Q_{1/\beta_j^{(v)}}\Big(\omega_j^{(v)}\big((\tau_j^{(v)}-\frac{1}{2})q_j^{(v)}\big)^{\beta_j^{(v)}}\Big),\end{aligned} \tag{6}$$

- $h_{\beta_j^{(v)}}$ is the differential entropy of a GG distribution [35]:

$$\begin{aligned}h_{\beta_j^{(v)}}(\omega_j^{(v)}) &= -\int_{-\infty}^{\infty} \widetilde{f}_j^{(v)}(\xi)\ln \widetilde{f}_j^{(v)}(\xi)\,d\xi \\ &= \ln\Big(\frac{2\Gamma(1/\beta_j^{(v)})}{\beta_j^{(v)}{\omega_j^{(v)}}^{1/\beta_j^{(v)}}}\Big) + \frac{1}{\beta_j^{(v)}},\end{aligned} \tag{7}$$

- $Q_a$ with $a \in \mathbb{R}_+^*$ is the normalized incomplete Gamma function:

$$\forall \xi \in \mathbb{R}, \qquad Q_a(\xi) = \frac{1}{\Gamma(a)}\int_0^{\xi} \theta^{a-1}e^{-\theta}d\theta. \tag{8}$$

Fig. 2(a) illustrates the approximation of the entropy of a given GG source. Note that the entropy plotted in solid line is obtained by firstly quantizing the source at different quantization levels and then evaluating the resulting entropy values, whereas the entropy plotted with star symbol is obtained by using directly the approximation model (4). Thus, it can be noticed that (4) yields an accurate approximation of the entropy for any set of quantization parameters.

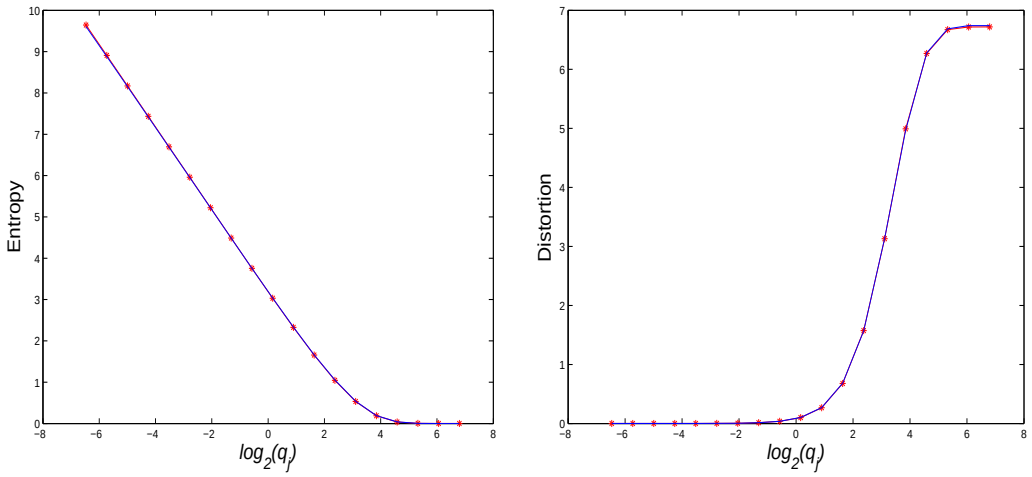
Concerning the distortion function, it could be estimated using the second order moment of the quantization error which represents the standard mean square error. While the distortion of a quantized GG random variable is given in [29] for a general order moment greater than or equal to 1, the distortion of

the quantized left and residual subbands, $D_j^{(l)}$ and $D_j^{(e)}$, are given by:

$\forall v \in \{l, e\}$,

$$D_j^{(v)}(q_j^{(v)}) = {\omega_j^{(v)}}^{-2/\beta_j^{(v)}} \frac{\Gamma(3/\beta_j^{(v)})}{\Gamma(1/\beta_j^{(v)})} Q_{3/\beta_j^{(v)}}\left(\omega_j^{(v)}\left((\tau_j^{(v)} + \frac{1}{2})q_j^{(v)}\right)^{\beta_j^{(v)}}\right)$$
$$- 2{(\omega_j^{(v)})}^{-1/\beta_j^{(v)}} \frac{\Gamma(2/\beta_j^{(v)})}{\Gamma(1/\beta_j^{(v)})} r_{1,j}^{(v)} \left(Q_{2/\beta_j^{(v)}}\left(\omega_j^{(v)}\left((\tau_j^{(v)} + \frac{1}{2})q_j^{(v)}\right)^{\beta_j^{(v)}}\right) - Q_{2/\beta_j^{(v)}}\left(\omega_j^{(v)}\left((\tau_j^{(v)} - \frac{1}{2})q_j^{(v)}\right)^{\beta_j^{(v)}}\right)\right)$$
$$+ (r_{1,j}^{(v)})^2 \left(Q_{1/\beta_j^{(v)}}\left(\omega_j^{(v)}\left((\tau_j^{(v)} + \frac{1}{2})q_j^{(v)}\right)^{\beta_j^{(v)}}\right) - Q_{1/\beta_j^{(v)}}\left(\omega_j^{(v)}\left((\tau_j^{(v)} - \frac{1}{2})q_j^{(v)}\right)^{\beta_j^{(v)}}\right)\right)$$
$$+ \frac{1}{12}(q_j^{(v)})^2 \left(1 - Q_{1/\beta_j^{(v)}}\left(\omega_j^{(v)}\left((\tau_j^{(v)} + \frac{1}{2})q_j^{(v)}\right)^{\beta_j^{(v)}}\right)\right). \quad (9)$$

Similarly to the entropy function, we illustrate in Fig. 2(b) the approximation of the distortion function. Indeed, the curve plotted in solid line is obtained by quantizing the source at different quantization steps and computing the resulting quantization error, whereas the star symbols correspond to the values provided by the approximation model (9). Again, it can be observed that (9) leads to good approximation of the distortion at low and high bitrate.




(a) (b)

Fig. 2: In (a) (resp. (b)), the entropy (resp. distortion) function, plotted in solid line, and the approximated one, plotted in star symbol, using Eq. (4) (resp. Eq. (9)). The parameters of the GG source are: $\beta_j = 0.75$ and $\omega_j = 1$.

### C. Bit allocation approach for OL-based coding scheme

Once the distortion and entropy functions have been defined, the bit allocation problem can be formulated. More specifically, our objective consists of finding the entropy values of the different subbands $\mathbf{H} = \left(H_1^{(l)}, \ldots, H_J^{(l)}, H_1^{(e)}, \ldots, H_J^{(e)}\right) \in [0, +\infty[^{2J}$ minimizing the average distortion subject to the constraint that the total bitrate is smaller than or equal to a given bitrate $R_{\max}$:

$$\begin{aligned} &\min \sum_{j=1}^{J} \left(\rho_j^{(l)} D_j^{(l)}(H_j^{(l)}) + \rho_j^{(e)} D_j^{(e)}(H_j^{(e)})\right) \\ &\text{subject to } \sum_{j=1}^{J} s_j \left(H_j^{(l)} + H_j^{(e)}\right) + R^{(d)} \leq R_{\max}, \end{aligned} \tag{10}$$

where $R^{(d)}$ represents the bitrate required to encode losslessly the disparity map, $s_j$ corresponds to the fraction of total coefficients in the $j^{th}$ subband, and $(\rho_j^{(l)}, \rho_j^{(e)}) \in [0, +\infty[^2$ are two weights which take into account for the nonorthogonality of the wavelet transform. Indeed, in the case of biorthogonal transforms or redundant frames, the weighting procedure of the distortion in the wavelet domain ensures a good approximation of the distortion in the spatial domain [36]. Moreover, the introduced weights $(\rho_j^{(l)}, \rho_j^{(e)})_{1 \leq j \leq J}$ can be exploited in order to optimize perceptual criteria better fitting the Human Visual System (HVS) characteristics. For example, in [37], the wavelet subbands are multiplied by a single invariant weighting factor derived from the Contrast Sensitivity Function in order to account for the Contrast Sensitivity of the HVS.

The constrained minimization problem (10) is solved using the standard Lagrangian optimization technique. For any Lagrangian multiplier $\lambda < 0$, the Lagrangian functional $\widetilde{\mathcal{J}}(\mathbf{H}, \lambda)$ is expressed as:

$$\begin{aligned} \widetilde{\mathcal{J}}(\mathbf{H}, \lambda) = &\sum_{j=1}^{J} \left(\rho_j^{(l)} D_j^{(l)}(H_j^{(l)}) + \rho_j^{(e)} D_j^{(e)}(H_j^{(e)})\right) \\ &-\lambda \Big( \sum_{j=1}^{J} s_j (H_j^{(l)} + H_j^{(e)}) + R^{(d)} - R_{\max} \Big). \end{aligned} \tag{11}$$

Under the assumption of the differentiability, and after imposing the zero gradient condition, we find that the optimal entropy values are solutions of the following system:

$$\forall\ j \in \{1, \ldots, J\}, \quad \begin{cases} \frac{\rho_j^{(l)}}{s_j} \frac{\partial D_j^{(l)}}{\partial H_j^{(l)}} = \lambda, \\ \frac{\rho_j^{(e)}}{s_j} \frac{\partial D_j^{(e)}}{\partial H_j^{(e)}} = \lambda. \end{cases} \tag{12}$$

This equation shows that the optimal entropies correspond to the points having the same slope $\lambda$ on the R-D curves, $(H_j^{(l)}, \frac{\rho_j^{(l)}}{s_j} D_j^{(l)})$ and $(H_j^{(e)}, \frac{\rho_j^{(e)}}{s_j} D_j^{(e)})$, of both the left and residual image subbands. Thus, it becomes necessary to compute the R-D curves of the different subbands. This step may be computationally

heavy since a large number of R-D operating points must be measured. To deal with this problem, it has been recently proposed to perform the computation for a small number of R-D points, followed by a smooth spline interpolation, in order to fit the R-D curve [25]. Indeed, once the R-D operating points are obtained, a fitting procedure is generally performed to guarantee the convexity of the R-D functions. However, this procedure may not result in a good fit which could affect the robustness of the bit allocation algorithm.

Therefore, in order to overcome these drawbacks, we propose to employ the recent approximation functions addressed in Section III-B. Since the rate and distortion functions, given by Eqs. (4) and (9), are expressed with respect to the quantization step $q_j^{(v)}$, Eq. (12) could be rewritten as:

$$\forall\ j \in \{1,\ldots,J\}, \quad \begin{cases} \frac{\rho_j^{(l)}}{s_j} \frac{\partial D_j^{(l)}}{\partial q_j^{(l)}} \left( \frac{\partial H_j^{(l)}}{\partial q_j^{(l)}} \right)^{-1} (q_j^{(l)}(\lambda)) = \lambda \\ \frac{\rho_j^{(l)}}{s_j} \frac{\partial D_j^{(e)}}{\partial q_j^{(e)}} \left( \frac{\partial H_j^{(e)}}{\partial q_j^{(e)}} \right)^{-1} (q_j^{(e)}(\lambda)) = \lambda \end{cases} \tag{13}$$

where, for each $v \in \{l, e\}$, the explicit expressions of the derivatives of $D_j^{(v)}$ and $H_j^{(v)}$ with respect to $q_j^{(v)}$ can be easily derived from the developed approximations (4) and (9).

Thus, for a given value $\lambda$, the quantization steps $q_j^{(l)}$ and $q_j^{(e)}$ for the left and residual images can be obtained from (13). Finally, the rate allocation problem is solved by finding the slope value $\lambda^*$ that satisfies:

$$\sum_{j=1}^{J} s_j \Big( H_j^{(l)}((q_j^{(l)})^*) + H_j^{(e)}((q_j^{(e)})^*) \Big) = R_{\max} - R^{(d)}. \tag{14}$$

where $(q_j^{(l)})^*$ and $(q_j^{(e)})^*$ represent the quantization parameters obtained from (13) for the slope $\lambda^*$. To find the optimal value $\lambda^*$, simple algorithms such as Newton and bisection methods can be used.

Note that these algorithms converge in few iterations to the optimal solution if and only if the R-D curves are both convex and differentiable. It should be noted also that the convexity of the R-D function is an important property since it allows to reduce the duality gap between the original constrained optimization problem (10) and the Lagrangian relaxation (11). Therefore, it becomes necessary to study the convexity of the R-D curves resulting from our theoretical approximations of the rate and distortion functions given by (4) and (9). To this end, many experiments have been carried out on a large dataset of stereo images. It is worth pointing that the convexity property is often satisfied for the most wavelet subbands, as it can be seen in Figures 3(a) and 3(c). However, we have observed that it is not the case for some subbands. For example, Figures 3(b) and 3(d) show a particular case where $\frac{\partial D_j}{\partial H_j}$ is a non monotonic function.

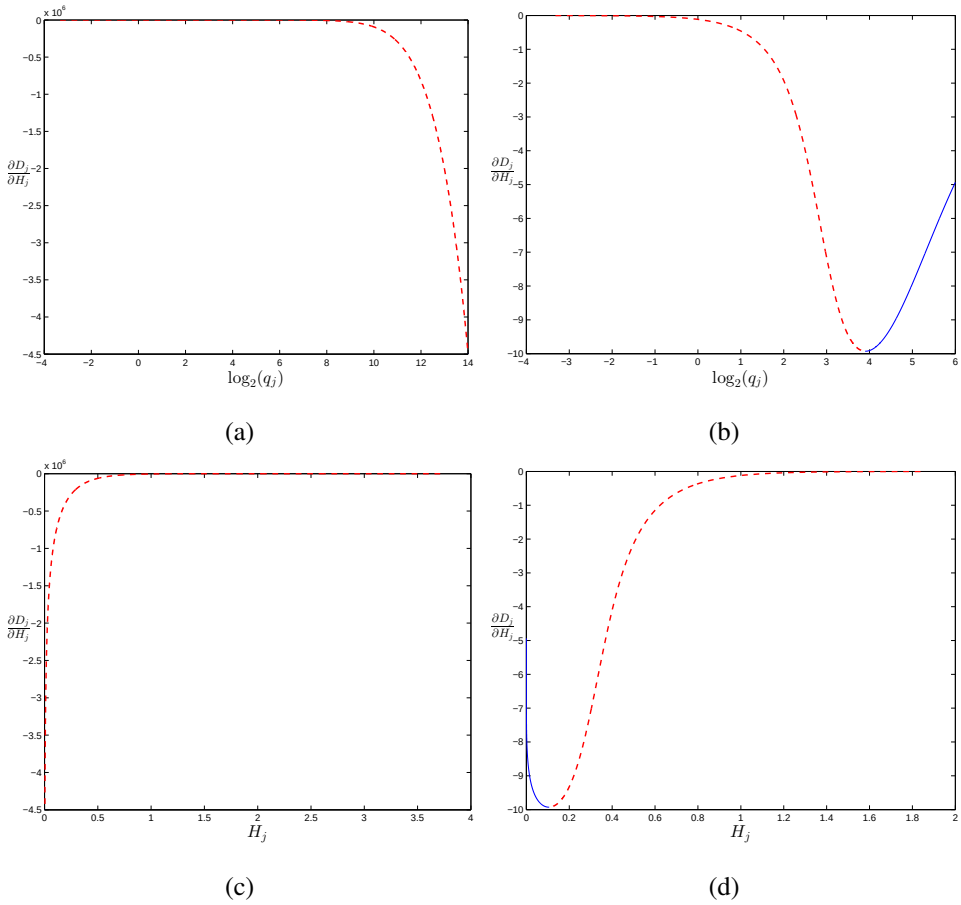



Fig. 3: In the top side (resp. bottom side), the first derivative function $\frac{\partial D_j}{\partial H_j}$ versus $\log_2(q_j)$ (resp. $H_j$) obtained with: (left side) $\beta_j = 0.5$ and $\omega_j = 0.1$, (right side) $\beta_j = 1.5$ and $\omega_j = 0.1$.

To solve this problem, we have determined the inflexion point $\bar{q}_j^{(v)}$ of the curve by solving this equation:

$$\forall\ v \in \{l, e\},\quad \frac{\partial^2 D_j^{(v)}}{\partial^2 H_j^{(v)}} = 0 \quad \Longleftrightarrow \quad \frac{\partial^2 D_j^{(v)}}{\partial^2 q_j^{(v)}} \frac{\partial H_j^{(v)}}{\partial q_j^{(v)}} - \frac{\partial^2 H_j^{(v)}}{\partial^2 q_j^{(v)}} \frac{\partial D_j^{(v)}}{\partial q_j^{(v)}} = 0, \tag{15}$$

In this case, during the determination of $q_j^{(l)}$ and $q_j^{(e)}$ that satisfy Eq. (13), we are restricted to the search interval $]0, \bar{q}_j^{(v)}]$ with $v \in \{l, e\}$. It is important to note that, in this interval, we have observed that the possible entropy values range from low to high bitrate as shown in Fig. 3(d), which could not affect the optimality of the bit allocation scheme.

### *D. Bit allocation approach for CL-based coding scheme*

Contrary to an OL-based coding scheme where the reference and target images are quantized independently, the bit allocation problem for a CL-based coding scheme is complicated by the dependencies

arising from using the disparity compensation (i.e prediction stage) based on the quantized reference image. In this case, this problem is referred to as dependent bit allocation for stereo image coding. As mentioned in Section II-A, we assume in this part that the disparity map is estimated only once from the original left and right images. Therefore, we only focus on the bit allocation between the two views. Since the disparity compensation process, used to generate the residual image from the quantized reference one, and the quantization stage of both images are coupled with each other, the bit allocation problem can be reformulated as follows:

$$\min \sum_{j=1}^{J} \left( \rho_j^{(l)} D_j^{(l)}(q_j^{(l)}) + \rho_j^{(e)} D_j^{(e)}(q_j^{(e)} \mid \mathbf{q}^{(l)}) \right)$$

$$\text{subject to} \sum_{j=1}^{J} s_j \left( H_j^{(l)}(q_j^{(l)}) + H_j^{(e)}(q_j^{(e)} \mid \mathbf{q}^{(l)}) \right) + R^{(d)} \leq R_{\max}, \quad (16)$$

where $\mathbf{q}^{(l)} = (q_j^{(l)})_{1 \leq j \leq J}$.

While the entropy $H_j^{(l)}(q_j^{(l)})$ and distortion $D_j^{(l)}(q_j^{(l)})$ of the reference subbands have a closed form expressions for a GG distribution, deriving the conditional ones of the residual image $H_j^{(e)}(q_j^{(e)} \mid \mathbf{q}^{(l)})$ and $D_j^{(e)}(q_j^{(e)} \mid \mathbf{q}^{(l)})$ is a difficult task. For this reason, we propose simply to perform the disparity compensation process to generate the residual image, and then deduce their resulting entropy and distortion values. More precisely, and after applying the wavelet transform, each wavelet subband of the reference image is firstly quantized at $N$ quantization steps $(q_j^{(l,i)})_{1 \leq i \leq N}$. The resulting entropy and distortion values can be computed from (4) and (9). Then, for a given $i$-th quantized reference image, we proceed as follows:

① First, the inverse wavelet transform is applied to the quantized reference image in order to obtain the reconstructed version. Note that this step corresponds to the decoding process that appears in the closed loop structure, as it can be seen in Fig. 1.

② Secondly, the disparity compensation stage is performed on the reconstructed reference image. The residual image is then generated and decomposed using the wavelet transform.

③ Finally, the residual subbands are quantized at $N$ quantization steps $(q_j^{(e,i)})_{1 \leq i \leq N}$, and their resulting entropy and distortion values are also computed using (4) and (9).

By repeating these three steps for the $N$ possible quantized reference images, we deduce the average distortion of the left and residual images for the different configurations of the quantization steps $(q_j^{(l,i)}, q_j^{(e,i')})_{1 \leq i,i' \leq N}$ (i.e the entropy values $(H_j^{(l,i)}, H_j^{(e,i')})_{1 \leq i,i' \leq N}$). In this way, we obtain a joint R-D function, denoted by $D_j(H_j^{(l)}, H_j^{(e)})$, where the used $N \times N$ distortion points are viewed as the results

of a two dimensional function. For this reason, we refer the associated operating R-D points to as R-D surfaces.

Once the computation is performed for some operating points, the R-D surfaces are generated by fitting these points. Inspired by the Mallat's analysis [38], we assume that the distortion can be parameterized by a decreasing exponential function when the rate increases. As the global distortion is obtained by averaging the distortion of the reference and residual images, we propose to approximate these R-D surfaces by a sum of two exponential functions. Therefore, the joint R-D function can be expressed as:

$$\forall j \in \{1, \ldots, J\}, \quad D_j(H_j^{(l)}, H_j^{(e)}) = a_j^{(l)} e^{-b_j^{(l)} H_j^{(l)}} + a_j^{(e)} e^{-b_j^{(e)} H_j^{(e)}}, \tag{17}$$

where, for $v \in \{l, e\}$, the parameters $a_j^{(v)}$ and $b_j^{(b)}$ are nonnegative reals, which are determined by minimizing the error between the generated R-D operating points and the analytical expression of the distortion (17).

As an example, we illustrate in Fig. 4 the R-D points for $N = 15$, presented with circle symbol, as well as the R-D surfaces resulting from the fitting exponential step. This figure shows the four subbands obtained over one-level wavelet decomposition of the left and residual images. It can be noticed that the data are well fitted by the joint R-D model. This is confirmed by performing a goodness-of-fit test such as the Normalized Root Mean Square Error (NRMSE).

It is clear that the proposed joint distortion function is convex and differentiable since Eq. (17) is the sum of two exponential functions. Thus, we can apply now the Lagrangian optimization method. Indeed, the Lagrangian functional $\mathcal{J}(\mathbf{H}, \lambda)$, in the case of the CL-based coding scheme, can be rewritten as:

$$\mathcal{J}(\mathbf{H}, \lambda) = \sum_{j=1}^{J} D_j(H_j^{(l)}, H_j^{(e)}) - \lambda\Big(\sum_{j=1}^{J} s_j(H_j^{(l)} + H_j^{(e)}) + R^{(d)} - R_{\max}\Big). \tag{18}$$

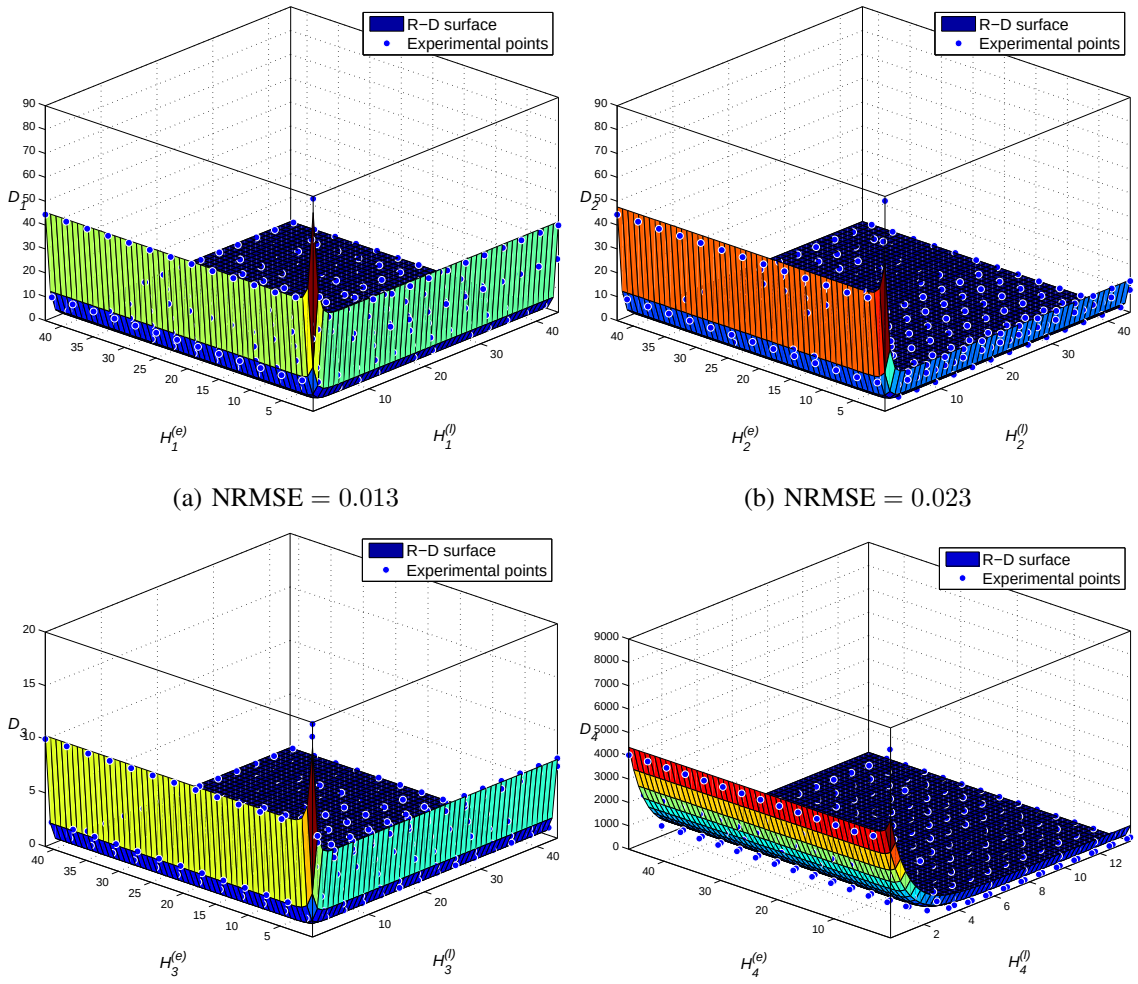
Thanks to the analytical joint distortion function, it can be checked that the optimal entropy values having the same slope $\lambda$ are simply obtained as follows:

$$\forall\ j \in \{1, \ldots, J\}, \quad \begin{cases} H_j^{(l)}(\lambda) = \frac{1}{b_j^{(l)}} \ln\left(-\frac{a_j^{(l)} b_j^{(l)}}{\lambda s_j}\right) \\ H_j^{(e)}(\lambda) = \frac{1}{b_j^{(e)}} \ln\left(-\frac{a_j^{(e)} b_j^{(e)}}{\lambda s_j}\right) \end{cases} \tag{19}$$

Finally, as performed in the previous subsection, the optimal solution is reached by looking for the best value $\lambda^*$ that satisfies the rate constraint condition:

$$\sum_{j=1}^{J} s_j(H_j^{(l)}(\lambda^*) + H_j^{(e)}(\lambda^*)) = R_{\max} - R^{(d)}. \tag{20}$$

We should note that the fitting operation performed to estimate the RD surface parameters involves some additional complexity to the bit allocation method for a CL-based coding scheme, compared to the OL




(a) NRMSE $= 0.013$ (b) NRMSE $= 0.023$

(c) NRMSE $= 0.009$ (d) NRMSE $= 0.038$

Fig. 4: R-D surfaces for the wavelet subbands of the "Drumsticks" stereo image with $N = 15$ (a) horizontal detail subband, (b) vertical detail subband, (c) diagonal detail subband, (d) approximation subband.

case where close approximation of the entropy and distortion functions are available for any given set of quantization steps. However, in the case of a CL-based coding scheme, an explicit expression of the optimal entropy values having the same slope $\lambda$ are provided in (19), contrary to the case of an OL coding scheme where a numerical algorithm is required to solve the system (13) and deduce the optimal entropy.

## IV. EXPERIMENTAL RESULTS

In this part, we study the performance of the proposed bit allocation methods for the OL- and CL-based stereo coding schemes. Simulations have been carried out on different standard stereo images downloaded from the stereovision website [1, 2, 3]. Concerning the Middlebury datasets, the gray scale versions of views 1 and 2 have been used in our simulation. The disparity map is estimated based on the block-matching technique with a $8 \times 8$ block size, and then losslessly encoded by using an arithmetic coder. We should note that the average coding cost $R^{(d)}$ of the losslessly encoded disparity map is about 0.03-0.05 bits per pixel (bpp). For the OL- and CL-based coding schemes, the reference and residual images are encoded by employing, over 3 resolution levels, the 9/7 transform retained for the lossy compression mode of JPEG2000 [2]. The entropy coder EBCOT has been used to encode the wavelet coefficients. Note that the weights $(\rho_j^{(l)}, \rho_j^{(e)})_{1 \leq j \leq J}$ for the wavelet subbands of the reference and residual images are computed by using the procedure described in [36]. Moreover, the method of moments is adopted to estimate the GG distribution parameters of the different wavelet subbands.

In order to study the accuracy and the optimality of the proposed bit allocation methods, we compare them with the optimal allocation strategy which is often used in the reported stereo image compression works. This strategy, designated in what follows by "Exhaustive search", considers several configurations of rate allocation $(R^{(l)}, R^{(e)})$, where $R^{(l)} + R^{(e)} = R_{\text{max}} - R^{(d)}$, and $R^{(l)}$, $R^{(e)}$ and $R^{(d)}$ denote the bitrate of the left image, the residual one and the disparity map, respectively. The optimal solution corresponds to the couple $(R^{(l)}, R^{(e)})$ leading to the minimum distortion evaluated through the Mean Square Error of the left image $\text{MSE}^{(l)}$ and the right one $\text{MSE}^{(r)}$ obtained after the decoding process. Furthermore, we compare these allocation methods with a prior given rate allocation which is commonly used in practice [21]. More precisely, a fraction $\gamma$ (resp. $1 - \gamma$) of the available bitrate $R_{\text{max}} - R^{(d)}$ is assigned to the reference view (resp. the residual one). Such allocation method will be denoted by "Fixed". To this end, we consider here different such allocation by taking a medium and high $\gamma$ values, given respectively by $50\%$ and $75\%$. We have also tested for comparison the state-of-the-art bit allocation method developed in [30]. This method will designated by "Woo and Ortega [30]".

[1] http://vision.middlebury.edu/stereo/

[2] http://vasc.ri.cmu.edu/idb/html/stereo/index.html

[3] http://vasc.ri.cmu.edu/idb/html/jisct/

The reported results are given in terms of the average bitrate $R_{\text{av}}$ and its corresponding PSNR measure:

$$R_{\text{av}} = \frac{R^{(l)} + R^{(e)} + R^{(d)}}{2}, \tag{21}$$

$$\text{PSNR} = 10 \log_{10} \frac{255^2}{(\text{MSE}^{(l)} + \text{MSE}^{(r)})/2}. \tag{22}$$

Figures 5 and 6 illustrate the performance of the proposed bit allocation methods in the case of OL- and CL-based stereo image coding schemes. It can be noticed that the best a priori rate allocation depends generally on the stereo data, the coding architecture and the range of the target bitrate. Moreover, compared to such fixed ratio allocation strategy, the state-of-the-art bit allocation method [30] improves the R-D results. However, its performance remains close or inferior to that obtained with our approaches. Overall, we observe that the proposed allocation methods and the exhaustive search schemes are often better. For this reason, we have evaluated in what follows the behavior of the proposed allocation methods with respect to the exhaustive search strategy in terms of Bjontegaard metric, often used to describe the distance between two R-D curves [39]. Table I gives the difference in terms of PSNR and bitrate saving. The results are provided for low, middle and and high bitrates corresponding to the four target bitrate points $\{0.15, 0.2, 0.3, 0.4\}$, $\{0.5, 0.6, 0.7, 0.8\}$ and $\{0.9, 1, 1.1, 1.2\}$ bpp, respectively.

For the case of OL-based coding scheme, it can be noticed from Fig. 5 and Table I that the difference between the R-D results of the proposed method and the optimal strategy ranges from 0.01 to 0.3 dB. It is important to note that this degradation is not only due to the developed theoretical approach, and can be mainly explained by the non optimality of the OL structure. Indeed, while the exhaustive search procedure looks for the best couple $(R^{(l)}, R^{(r)})$ leading to the minimum distortion of the left and right images reconstructed after the decoding process, it is important to emphasize that the theoretical allocation method is performed on the reference and residual images by minimizing their corresponding distortions $D^{(l)}$ and $D^{(e)}$. On the other hand, as mentioned at the end of Section II, the main limitation of the OL structure is that these resulting distortions do not correspond to those computed on the left and right images, which may lead to a suboptimal solution.

However, since this drawback does not exist in a CL structure, we remark that, in this case, the R-D curves obtained by our proposed approache are very close to those resulting from the optimal allocation procedure based on the exhaustive search method. Indeed, a very negligeable degradation always inferior to 0.1 dB is observed.

Although a PSNR degradation is observed for some stereo images in the case of an OL structure as depicted in Table I, Fig. 7-9 show that the visual quality of the reconstructed stereo images obtained by the proposed allocation methods are similar to those resulting from the exhaustive search method.

While achieving R-D performance close to that obtained with the optimal allocation strategy, our proposed methods reduce significantly the required execution time, as shown in Table II. Simulations are carried out by using an Intel Core i7 (3.2 GHz) computer with a Matlab implementation. It can be observed that the optimal algorithm is too computationally expensive since an exhaustive search method is performed whereas our proposed bit allocation methods take only very few seconds. The execution time indicates that our method is about twenty (resp. three) times faster than the optimal strategy (resp. the state-of-the-art method [30]). Note that we report here the computational time relative to the encoding process since the bit allocation step is only performed on the encoder side. At the decoder side, the execution time is the same for the all considered methods.

All these results confirm the effectiveness of the proposed bit allocation methods for stereo image compression purpose in terms of computational time and coding performance trade-offs.

## V. Conclusion

In this work, we have focused on the bit allocation problem for stereo image compression. More precisely, the standard open-loop and closed-loop-based coding structures have been considered. To this end, two bit allocation methods based on recent approximation of R-D functions have been developed. Compared to the optimal allocation procedure where an exhaustive search strategy is performed in order to find the optimal rates assigned to each view, the proposed bit allocation methods lead to R-D results very close to the optimal ones, especially in the case of the popular closed-loop structure. In future work, the development of analytical expressions of the conditional rate and distortion functions could be investigated. Moreover, the proposed inter-view bit allocation methods can be extended to a multiview/video coding framework.

## References

[1] G. Wallace, "The JPEG still picture compression standard," *Communications of the ACM*, pp. 30–44, 1991.

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

[3] M. E. Lukacs, "Predictive coding of multi-viewpoint image sets," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 521–524, 1986.

[4] M. G. Perkins, "Data compression of stereopairs," *IEEE Transactions on Communications*, vol. 40, no. 4, pp. 684–696, 1992.

[5] H. Aydinoglu and M. Hayes III, "Stereo image coding," in *IEEE International Symposium on Circuits and Systems*, vol. 1, 1995, pp. 247–250.

[6] W. Woo and A. Ortega, "Stereo image compression based on disparity field segmentation," in *SPIE Conference on Visual Communications and Image Processing*, vol. 3024, 1997, pp. 391–402.

[7] U. Ahlvers, U. Zoelzer, and S. Rechmeier, "FFT-based disparity estimation for stereo image coding," in *International Conference on Image Processing*, vol. 1, 2003, pp. 761–764.

[8] M. Moellenhoff and M. Maier, "Transform coding of stereo image residuals," *IEEE Transactions on Image Processing*, vol. 7, no. 6, pp. 804–812, June 1998.

[9] T. Frajka and K. Zeger, "Residual image coding for stereo image compression," *Optical Engineering*, vol. 42, no. 1, pp. 182–189, 2003.

[10] W. Hachicha, A. Beghdadi, and F. Cheikh, "1D directional DCT-based stereo residual compression," in *European signal Processing Conference*, Marrakech, Morocco, 2013.

[11] N. Boulgouris and M. Strintzis, "A family of wavelet-based stereo image coders," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 10, pp. 898–903, 2002.

[12] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Image Processing*, vol. 41, no. 12, pp. 3445–3462, 1993.

[13] M. Kaaniche, A. Benazza-Benyahia, B. Pesquet-Popescu, and J.-C. Pesquet, "Vector lifting schemes for stereo image coding," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2463–2475, 2009.

[14] A. Maalouf and M.-C. Larabi, "Bandelet-based stereo image coding," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2010, pp. 698–701.

[15] W. Woo and A. Ortega, "Overlapped block disparity compensation with adaptive windows for stereo image coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 2, pp. 194 –200, 2000.

[16] L. Lucas, N. Rodrigues, E. da Silva, and S. de Faria, "Stereo image coding using dynamic template-matching prediction," in *IEEE International Conference on Computer as a Tool*, 2011, pp. 1–4.

[17] J. Ellinas and M. Sangriotis, "Stereo image compression using wavelet coefficients morphology," *Image and Vision Computing*, vol. 22, no. 4, pp. 281–290, 2004.

[18] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H. 264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, 2003.

[19] F. Shao, G. Jiang, W. Lin, M. Yu, and Q. Dai, "Joint bit allocation and rate control for coding multi-view video plus depth based 3D video," *IEEE Transactions on MultiMedia*, vol. 15, no. 8, pp. 1843–1454, December 2013.

[20] G. Cheung, V. Velisavljevic, and A. Ortega, "On dependent bit allocation for multiview image coding with depth-image-based rendering," *IEEE Transactions on Image Processing*, vol. 20, no. 11, pp. 3179–3194, 2011.

[21] B. Rajaei, T. Maugey, H.-R. Pourreza, and P. Frossard, "Rate-distortion analysis of multiview coding in a DIBR framework," *Annals of telecommunications*, pp. 1–14, 2012.

[22] A. Gelman, P. L. Dragotti, and V. Velisavljevic, "Multiview image coding using depth layers and an optimized bit allocation." *IEEE Transactions on Image Processing*, vol. 21, no. 9, pp. 4092–4105, 2012.

[23] S.-W. Wu and A. Gersho, "Rate-constrained optimal block-adaptive coding for digital tape recording of HDTV," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 1, no. 1, pp. 100–112, 1991.

[24] K. Ramchandran, A. Ortega, and M. Vetterli, "Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders," *IEEE Transactions on Image Processing*, vol. 3, no. 5, pp. 533–545, 1994.

[25] T. André, M. Cagnazzo, M. Antonini, and M. Barlaud, "JPEG2000-compatible scalable scheme for wavelet-based video coding," *EURASIP Journal on Image and Video Processing*, vol. 2007, 2007.

[26] G. Sullivan, "Efficient scalar quantization of exponential and Laplacian random variables," *IEEE Transactions on Information Theory*, vol. 42, no. 5, pp. 1365–1374, 1996.

[27] A. György, T. Linder, and K. Zeger, "On the rate-distortion function of random vectors and stationary sources with mixed distributions," *IEEE Transactions on Information Theory*, vol. 45, no. 6, pp. 2110–2115, 1999.

[28] A. Fraysse, B. Pesquet-Popescu, and J.-C. Pesquet, "On the uniform quantization of a class of sparse sources," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3243–3263, 2009.

[29] M. Kaaniche, A. Fraysse, B. Pesquet-Popescu, and J.-C. Pesquet, "A bit allocation method for sparse source coding," *IEEE Transactions on Image processing*, vol. 23, no. 1, pp. 137–152, January 2014.

[30] W. Woo and A. Ortega, "Optimal blockwise dependent quantization for stereo image coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 6, pp. 861 –867, 1999.

[31] S. V. Leuven, J. D. Cock, G. V. Wallendael, R. V. de Walle, R. Garrido-Cantos, J. L. Martnez, and P. Cuenca, "Combining open- and closed-loop architectures for H.264/AVC-to-SVC transcoding," in *International Conference on Image Processing*, 2011, pp. 1661–1664.

[32] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Transactions on Image Processing*, vol. 11, no. 2, pp. 146–158, February 2002.

[33] Y. Wang, J. T. Rucker, and J. E. Fowler, "3D tarp coding for the compression of hyperspectral images," *IEEE Geoscience and Remote Sensing Letters*, pp. 136–140, 2004.

[34] H. Gish and J. Pierce, "Asymptotically efficient quantizing," *IEEE Transactions on Information Theory*, vol. 14, no. 5, pp. 676–683, 1968.

[35] W. Szepanski, "$\Delta$-entropy and rate distortion bounds for generalized Gaussian information sources and their application to image signals," *Electronics Letters*, vol. 16, no. 3, pp. 109–111, 1980.

[36] S. Parrilli, M. Cagnazzo, and B. Pesquet-Popescu, "Distortion evaluation in transform domain for adaptive lifting schemes," in *International Workshop on Multimedia Signal Processing*, Cairns, Queensland, Australia, October 2008.

[37] M. J. Nadenau, J. Reichel, and M. Kunt, "Wavelet-based color image compression: exploiting the contrast sensitivity function," *IEEE Transactions on Image Processing*, vol. 12, no. 1, pp. 58–70, 2003.

[38] S. Mallat, *A wavelet tour of signal processing*. Access Online via Elsevier, 1999.

[39] G. Bjontegaard, "Calculation of average PSNR differences between RD curves," ITU SG16 VCEG-M33, Austin, TX, USA, Tech. Rep., April 2001.

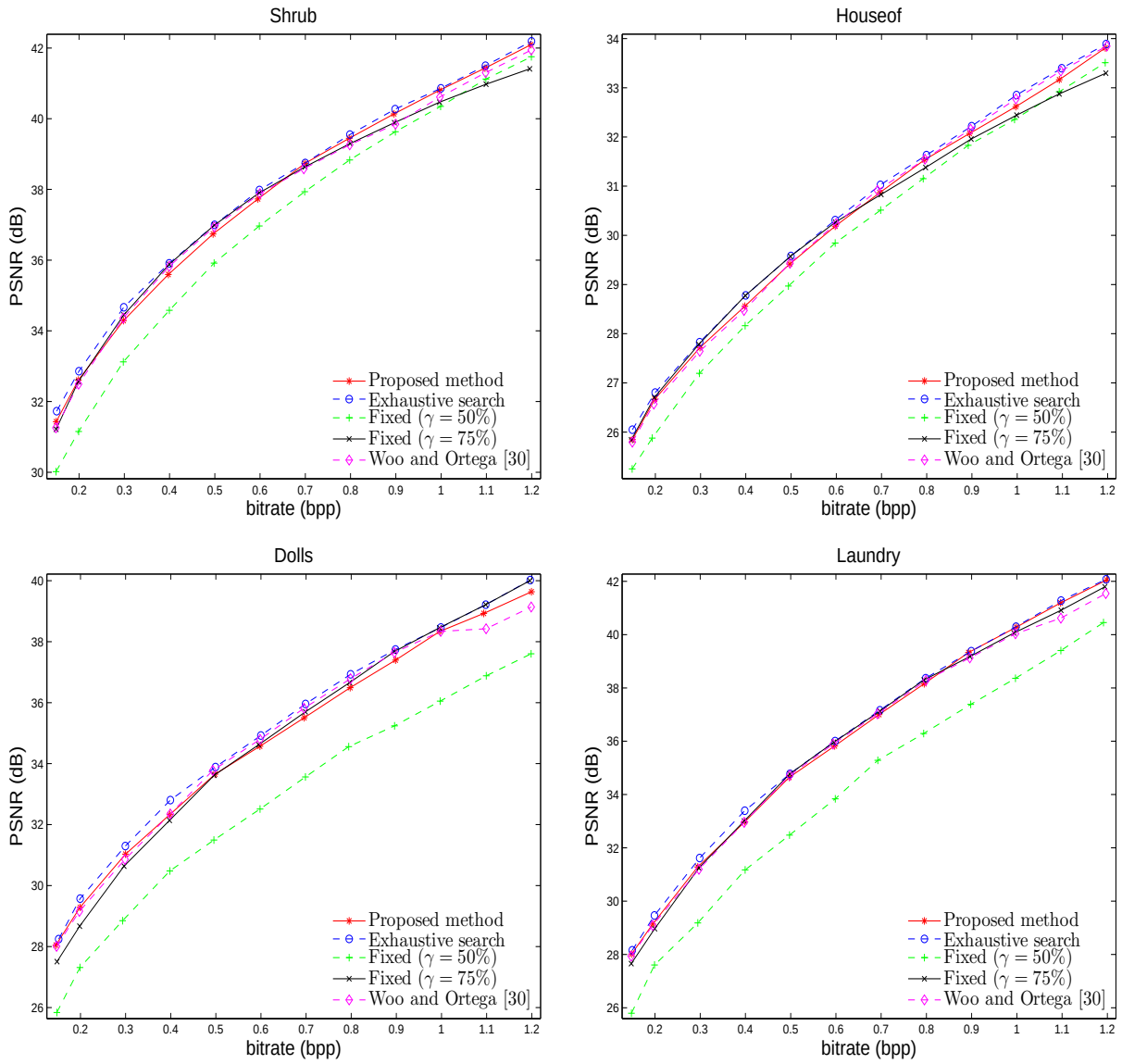



Fig. 5: Rate-distortion performance of the different bit allocation methods for the open-loop-based coding scheme of the images “Shrub”, “Houseof”, “Dolls” and “Laundry”.

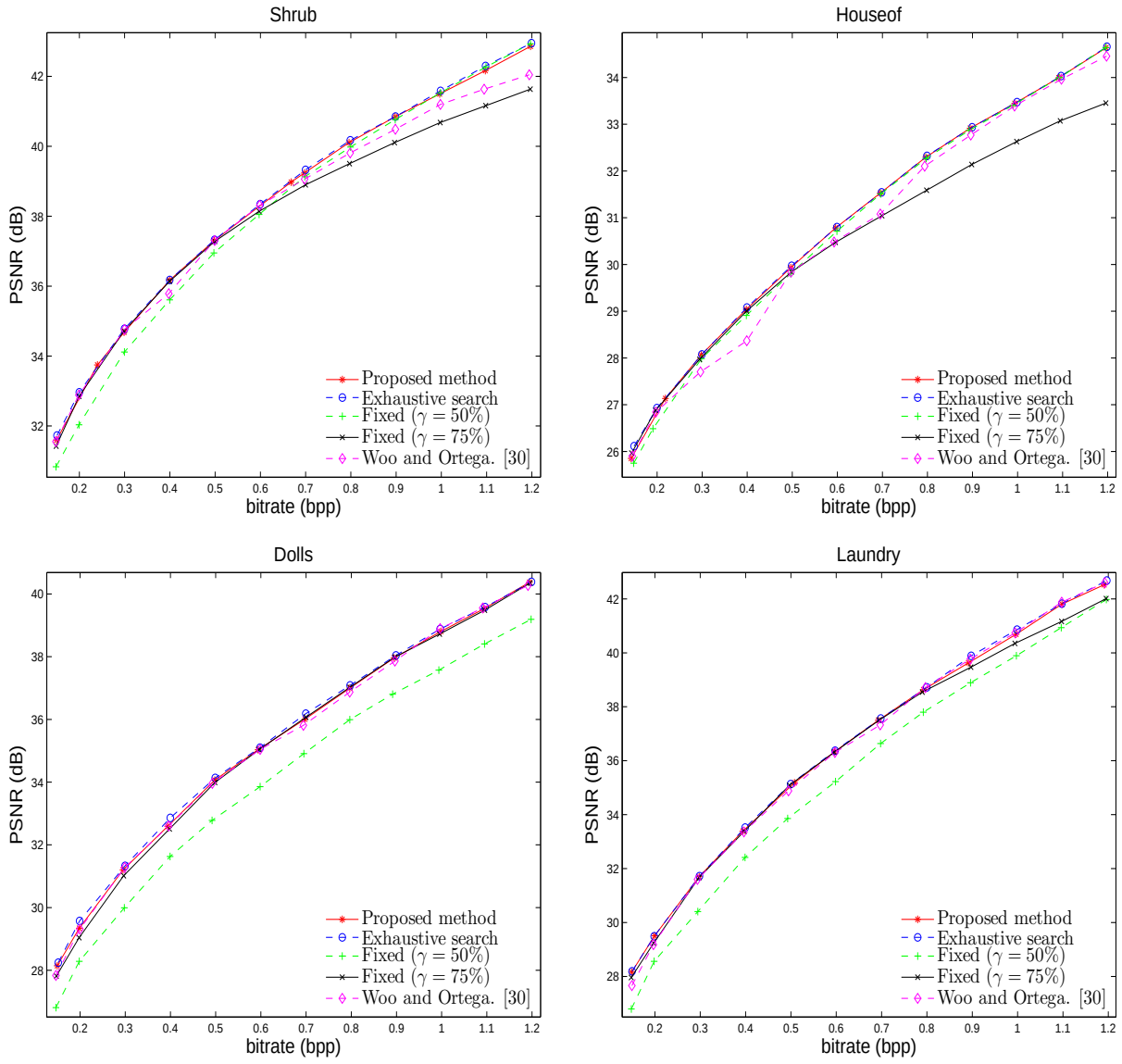



Fig. 6: Rate-distortion performance of the different bit allocation methods for the closed-loop-based coding scheme of the images “Shrub”, “Houseof”, “Dolls” and “Laundry”.

TABLE I: The average PSNR differences and the bitrate saving at low, middle and high bitrates between proposed methods and the optimal strategy.

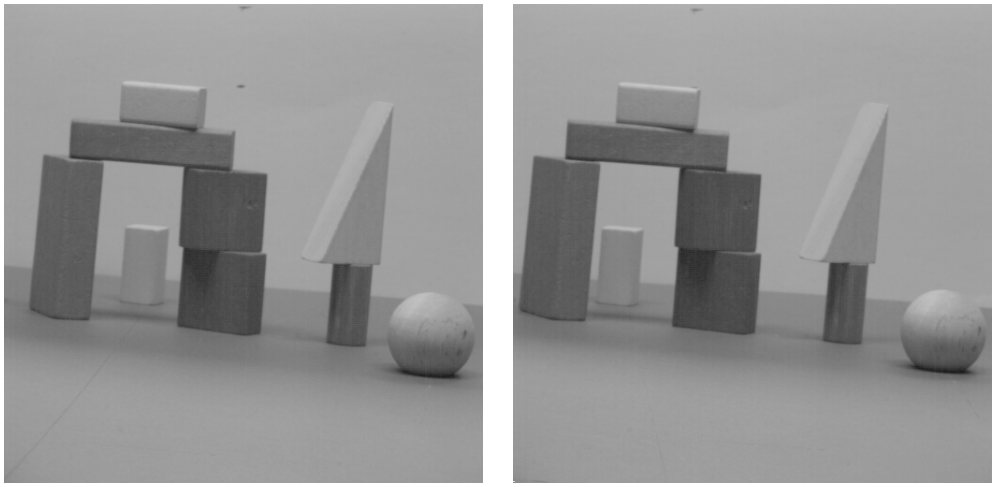
| Image | Scheme | PSNR gain (dB) | | | Bitrate saving (%) | | |
|---|---|---|---|---|---|---|---|
| | | low | middle | high | low | middle | high |
| Shrub | OL | -0.295 | -0.145 | -0.062 | 6.908 | 2.474 | 0.918 |
| | CL | -0.039 | -0.028 | -0.083 | 0.920 | 0.503 | 1.176 |
| Houseof | OL | -0.137 | -0.128 | -0.186 | 4.800 | 2.934 | 3.231 |
| | CL | -0.042 | -0.001 | -0.004 | 1.186 | 0.011 | 0.061 |
| Pentagon | OL | -0.109 | -0.079 | -0.093 | 3.714 | 1.731 | 1.623 |
| | CL | -0.013 | -0.005 | -0.022 | 0.416 | 0.096 | 0.356 |
| Arch | OL | -0.017 | -0.010 | -0.061 | 0.769 | 0.388 | 1.906 |
| | CL | -0.018 | -0.049 | -0.157 | 0.432 | 1.625 | 4.223 |
| Ball | OL | -0.028 | -0.006 | -0.184 | 1.643 | 0.192 | 6.237 |
| | CL | -0.044 | -0.012 | -0.013 | 1.912 | 0.416 | 0.235 |
| Art | OL | -0.280 | -0.345 | -0.264 | 5.868 | 5.152 | 3.313 |
| | CL | -0.028 | -0.055 | -0.024 | 0.507 | 0.775 | 0.297 |
| Dolls | OL | -0.261 | -0.354 | -0.227 | 5.969 | 5.695 | 2.669 |
| | CL | -0.131 | -0.074 | -0.051 | 2.791 | 1.277 | 0.645 |
| Drumsticks | OL | -0.155 | -0.324 | -0.257 | 3.451 | 4.509 | 2.597 |
| | CL | -0.019 | -0.055 | -0.043 | 0.409 | 0.764 | 0.365 |
| Laundry | OL | -0.246 | -0.155 | -0.038 | 4.688 | 2.031 | 0.420 |
| | CL | -0.012 | -0.023 | -0.116 | 0.199 | 0.283 | 1.075 |
| Reindeer | OL | -0.249 | -0.209 | -0.254 | 4.626 | 3.360 | 3.571 |
| | CL | -0.032 | -0.042 | -0.029 | 0.567 | 0.654 | 0.310 |
| Average | OL | -0.178 | -0.175 | -0.163 | 4.243 | 2.847 | 2.648 |
| | CL | -0.038 | -0.034 | -0.054 | 0.934 | 0.640 | 0.874 |


(a) Original stereo image

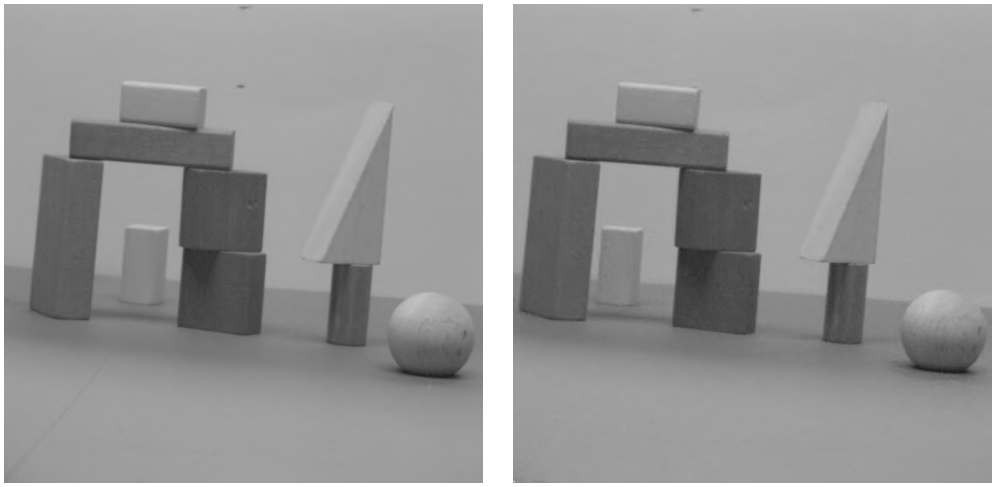

(b) $R_{\text{av}} = 0.1986$ bpp, PSNR $= 32.61$ dB


(c) $R_{\text{av}} = 0.1994$ bpp, PSNR $= 32.84$ dB

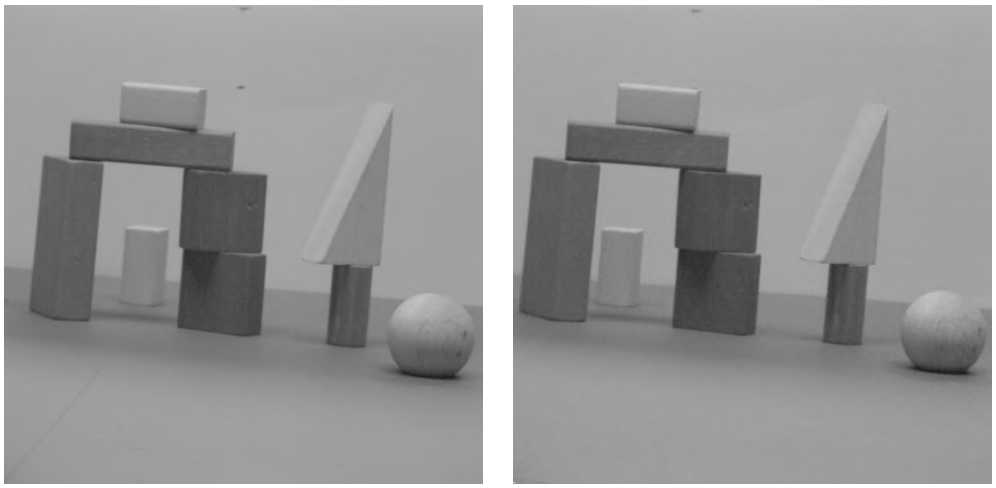
Fig. 7: Original "Shrub" stereo image and the reconstructed ones with an OL-based coding scheme using: (b) proposed method (c) exhaustive search strategy.



(a) Original stereo image



(b) $R_{\text{av}} = 0.1993$ bpp, PSNR $= 43.71$ dB



(c) $R_{\text{av}} = 0.1985$ bpp, PSNR $= 43.72$ dB

Fig. 8: Original "Arch" stereo image and the reconstructed ones with an OL-based coding scheme using: (b) proposed method (c) exhaustive search strategy.

(a) Original ‘Art’ stereo image

(b) $R_{\text{av}} = 0.1975$ bpp, PSNR $= 30.40$ dB

(c) $R_{\text{av}} = 0.1993$ bpp, PSNR $= 30.47$ dB

Fig. 9: Original “Art” stereo image and the reconstructed ones with a CL-based coding scheme using: (b) proposed method (c) exhaustive search strategy.

TABLE II: Encoding time of the different approaches.

| SI | Algorithm | | Encoding Time (s) |
|---|---|---|---|
| Shrub | OL | Proposed | 2.84 |
| | | Exhaustive search | 70.59 |
| | | Woo and Ortega | 6.80 |
| | CL | Proposed | 6.01 |
| | | Exhaustive search | 78.90 |
| | | Woo and Ortega | 12.47 |
| Pentagon | OL | Proposed | 3.02 |
| | | Exhaustive search | 75.18 |
| | | Woo and Ortega | 11.99 |
| | CL | Proposed | 4.72 |
| | | Exhaustive search | 83.61 |
| | | Woo and Ortega | 14.07 |
| Art | OL | Proposed | 3.03 |
| | | Exhaustive search | 54.45 |
| | | Woo and Ortega | 8.52 |
| | CL | Proposed | 3.56 |
| | | Exhaustive search | 60.73 |
| | | Woo and Ortega | 8.99 |