EFFICIENT JOINT MULTISCALE DECOMPOSITION FOR COLOR STEREO IMAGE CODING

O. Dhifallah[1], *M. Kaaniche*[2], *A. Benazza-Benyahia*[1]

[1] COSIM Lab.,
Sup'Com, Carthage Univ., Tunisia
dhifallahoussama@gmail.com
benazza.amel@supcom.rnu.tn

[2] Institut Galilée, L2TI,
Université Paris 13,
Sorbonne Paris Cité, France
mounir.kaaniche@univ-paris13.fr

ABSTRACT

With the recent advances in stereoscopic display technologies, the demand for developing efficient stereo image coding techniques has increased. While most of the existing approaches have been proposed and studied in the case of monochrome stereo images, we are interested in this paper in compressing color stereo data. More precisely, we design a multiscale decomposition, based on the concept of vector lifting scheme, that jointly exploits the cross-view and inter-color channels redundancies. Moreover, our decomposition is well adapted to the contents of these data. Experimental results performed on natural stereo images show the benefits which can be drawn from the proposed coding method.

Index Terms— Color stereoscopic image, disparity estimation, compression, joint coding schemes, lifting scheme.

1. INTRODUCTION

Stereo Images (SI) correspond to the acquisition for a given scene of two views from two slightly different angles. Thus, a pair of a left and right view is obtained. Such images present the advantage of containing the 3D information of the underlying scene. This ability to provide depth perception has increased the interest of SI in various applications such as 3DTV, videoconferences and safe navigation. As a result, a huge amount of data has been generated. Moreover, due to the benefits of using the color information in different stereovision problems, the involved data size becomes prohibitive and constitutes a major limitation for developing stereoscopic applications. Consequently, it is necessary to develop efficient compression techniques for storage and transmission purposes.

To this end, the intuitive approach would be to apply conventional still image compression methods independently on each view. However, such approach is not so efficient since it does not exploit the cross-view dependencies resulting from the observation of a common scene from 2 viewpoints [1]. Indeed, the projection of a given 3D point onto the left and right view planes results in two pixels at different locations. The difference in location between these two homologous pixels represents the disparity information. Hence, most of SI coding methods aim at capturing the cross-view similarities by firstly estimating the disparity map. Then, one view (say the right one), referred to as target view, is predicted from the left view. After that, the difference between the original right view and the predicted one, called residual image, is generated. Finally, the disparity map, the reference image (i.e. the left one) and the residual one are encoded. While the disparity map is often encoded using a DPCM technique, the reference and residual images have been encoded in different transform domains. The most widely used ones are the discrete cosine transform [1, 2] and the discrete wavelet transform [3, 4]. Recently, a novel approach based on a generalized version of lifting scheme, called Vector Lifting Scheme (VLS), has been proposed in [5]. Indeed, instead of encoding a reference and a residual image, a joint multiscale decomposition is developed to produce a pair of multiresolution representations of the left and the right images.

It is important to note that these works as well as most of the reported ones have neglected the color information by simply considering the monochrome (grayscale) versions of the stereo data. To this respect, we focus in this paper on the compression issue of color stereoscopic images. More specifically, we propose to extend the VLS-based decomposition [5] to the context of *color* SI. Our main contribution consists in *jointly* exploiting the spatial, cross-view and inter-color channel redundancies. Moreover, our decomposition is adapted to the contents of the color components.

This paper is organized as follows. Sec. 2 is dedicated to a brief review of disparity estimation methods and the VLS-based decomposition. In Sec. 3, the proposed extended coding scheme is presented. Finally, in Sec. 4, experimental results are given and some conclusions are drawn in Sec. 5.

2. STEREO IMAGE CODING BACKGROUND

2.1. Disparity estimation

Most of the reported SI coding techniques are based on the disparity estimation/compensation step. To this end, several methods have already been developed for the monochrome and color SI [6–8]. In the context of coding application, block-based ones are often retained since a single disparity vector is assigned to all the pixels within the block. To this respect, block-matching technique is often employed [9, 10]. More precisely, the right image is firstly partitioned into nonoverlapping blocks which are then compared to the candidate blocks of the left image located in a given search area. Finally, the disparity vector $\mathbf{v} = (v_x, v_y)$ is obtained by minimizing some dissimilarity criterion $\mathcal{D}$. In the case of color stereo images, the disparity estimation problem can be simply formulated as follows:

$$(v_x, v_y)(x, y) = \arg \min_{(v_x,v_y)\in\Omega} \sum_{c=1}^{C} \mathcal{D}\Big(I^{(\mathrm{r},c)}(x, y)\,; I^{(\mathrm{l},c)}(x + v_x(x, y), y + v_y(x, y))\Big) \quad (1)$$

where Ω is the range of candidate disparity values, $I^{(\mathrm{l},c)}$ and $I^{(\mathrm{r},c)}$ denote respectively the c^{th} color channel for the left and right images for $c \in \{1, \ldots, C\}$, C representing the number of color channels.

Generally, the sum of square differences or the sum of absolute differences is retained as a dissimilarity measure. Moreover, when the stereo images are rectified, the disparity is purely horizontal (i.e. $v_y = 0$). For the sake of simplicity, this assumption will be considered in the following.

2.2. VLS-based decomposition for monochrome SI

As aforementioned, most of the reported works are devoted to the coding of monochrome stereo images ($C = 1$). Thus, we describe in this part the VLS for a given channel c of the left image $I^{(\mathrm{l},c)}$ and the right one $I^{(\mathrm{r},c)}$. More precisely, a conventional lifting structure, composed of a prediction and update steps, is applied to the reference image $I^{(\mathrm{l},c)}$. At each resolution level $(j+1)$, the associated detail coefficients $\widetilde{d}^{(\mathrm{l},c)}_{j+1}$ and the approximation ones $\widetilde{I}^{(\mathrm{l},c)}_{j+1}$ are computed as follows:

$$\widetilde{d}^{(\mathrm{l},c)}_{j+1}(x, y) = I^{(\mathrm{l},c)}_{j}(x, 2y+1) - \lfloor \sum_{k\in\mathcal{P}^{(\mathrm{l},c)}_j} p^{(\mathrm{l},c)}_{j,k} I^{(\mathrm{l},c)}_{j}(x, 2y+2k) \rfloor \quad (2)$$

$$\widetilde{I}^{(\mathrm{l},c)}_{j+1}(x, y) = I^{(\mathrm{l},c)}_{j}(x, 2y) + \lfloor \sum_{k\in\mathcal{U}^{(\mathrm{l},c)}_j} u^{(\mathrm{l},c)}_{j,k} \widetilde{d}^{(\mathrm{l},c)}_{j+1}(x, y-k) \rfloor \quad (3)$$

where the set $\mathcal{P}^{(\mathrm{l},c)}_j$ (resp. $\mathcal{U}^{(\mathrm{l},c)}_j$) and the coefficients $p^{(\mathrm{l},c)}_{j,k}$ (resp. $u^{(\mathrm{l},c)}_{j,k}$) denote respectively the support and the weights of the predictor (resp. update) operator.

Then, for the right image, a similar structure is also applied to generate an intermediate detail signal and the approximation one. After that, a second prediction step has been added to exploit at the same time the intra and inter-view redundancies thanks to the estimated disparity map. Thus, the resulting decomposition is given by:

$$\widetilde{d}^{(\mathrm{r},c)}_{j+1}(x, y) = I^{(\mathrm{r},c)}_{j}(x, 2y+1) - \lfloor \sum_{k\in\mathcal{P}^{(\mathrm{r},c)}_j} p^{(\mathrm{r},c)}_{j,k} I^{(\mathrm{r},c)}_{j}(x, 2y+2k) \rfloor \quad (4)$$

$$\widetilde{I}^{(\mathrm{r},c)}_{j+1}(x, y) = I^{(\mathrm{r},c)}_{j}(x, 2y) + \lfloor \sum_{k\in\mathcal{U}^{(\mathrm{r},c)}_j} u^{(\mathrm{r},c)}_{j,k} \widetilde{d}^{(\mathrm{r},c)}_{j+1}(x, y-k) \rfloor, \quad (5)$$

$$\widehat{d}^{(\mathrm{r},c)}_{j+1}(x, y) = \widetilde{d}^{(\mathrm{r},c)}_{j+1}(x, y) - \lfloor \sum_{k\in\mathcal{Q}^{(\mathrm{r},c)}_j} q_{j,k} \widetilde{I}^{(\mathrm{r},c)}_{j+1}(x, y+k) + \sum_{k\in\mathcal{P}^{(\mathrm{r,l},c)}_j} p^{(\mathrm{r,l},c)}_{j,k} I^{(\mathrm{l},c)}_{j}(x + v_{x,j}(x, 2y+1), 2y+1-k) \rfloor, \quad (6)$$

where the notations used in Eqs. (4) and (5) are similar to those used with the left view and, $\mathcal{Q}^{(\mathrm{r},c)}_j$ and $q_{j,k}$ (resp. $\mathcal{P}^{(\mathrm{r,l},c)}_j$ and $p^{(\mathrm{r,l},c)}_{j,k}$) are the support and the weights of the second intra (resp. inter)-image predictor. We should note that during the disparity compensation process that aims at exploiting the inter-view redundancies, the disparity map $v_{x,j}$ is obtained by scaling by 2^j and subsampling the initial full resolution map v_x, since the dimensions of the wavelet subbands are divided by 2 from a given resolution level to the next one. In case of non-integer values of the disparity map, a bilinear interpolation could be used to compute the disparity-compensated intensity value $I^{(\mathrm{l},c)}_{j}(x + v_{x,j}(x, 2y + 1), 2y + 1 - k)$.

It is clear that this VLS-based is described for a given column x of the left and right images. Thus, once performed on all the columns, it is applied along the lines yielding an approximation subband and three detail subbands for each image. The same process is applied on the approximation subbands, over J levels, to produce a multiresolution representations of both images. Finally, at the last resolution level, instead on encoding the approximation subband $I^{(\mathrm{r},c)}_J$ of the right image, we predict it from that of the reference view $I^{(\mathrm{l},c)}_J$ and encode the resulting residual one:

$$e^{(\mathrm{r},c)}_J(x, y) = I^{(\mathrm{r},c)}_J(x, y) - \lfloor I^{(\mathrm{l},c)}_J(x + v_{x,J}(x, y), y) \rfloor. \quad (7)$$

3. EXTENDED VLS FOR COLOR STEREO IMAGE

To compress a color SI, a straightforward solution consists in applying the previous decomposition *separately* to each color channel of left and right images. However, such technique exploits only the intra and inter-view redundancies. Thus, in order to design an efficient coding scheme well adapted to the content of the color stereo images, we propose to extend the VLS to a more general framework by further taking into account the correlation existing between the channel components.

3.1. A generic scheme

Without loss of generality, we will assume that the number C of color components amounts 3 for the sake of simplicity. Let us denote by $\mathrm{P} = (c_1, c_2, c_3)$ denote a permutation of the index channels of the color SI.

First, one channel (for example c_1) is a selected as a reference channel, and their associated views $(I^{(\mathrm{l},c_1)}, I^{(\mathrm{r},c_1)})$ are encoded by applying the VLS decomposition presented in Sec. 2.

Then, for the subsequent channel c_2, the detail signal associated to the reference view $I^{(\mathrm{l},c_2)}$ is computed by using an hybrid prediction step that exploits the correlation existing between the current channel $I^{(\mathrm{l},c_2)}$ and the previous encoded one $I^{(\mathrm{l},c_1)}$. Thus, Eq. (2) becomes:

$$
\begin{aligned}
\widetilde{d}_{j+1}^{(\mathrm{l},c_2)}(x,y) = I_j^{(\mathrm{l},c_2)}(x,2y+1) - \\
\lfloor \sum_{k \in \mathcal{P}_j^{(\mathrm{l},c_2)}} p_{j,k}^{(\mathrm{l},c_2)} I_j^{(l,c_2)}(x, 2y+2k) \\
+ \sum_{k \in \mathcal{P}_j^{(\mathrm{l},c_1,c_2)}} p_{j,k}^{(\mathrm{l},c_1,c_2)} I_j^{(\mathrm{l},c_1)}(x, 2y+1-k) \rfloor
\end{aligned}
\quad (8)
$$

where $\mathcal{P}_j^{(\mathrm{l},c_1,c_2)}$ and $p_{j,k}^{(\mathrm{l},c_1,c_2)}$ represent the support and the weights of the cross-spectral predictor (between c_1 and c_2). After that, an update step similar to that used in Eq. (3) is applied to generate the approximation coefficients $\widetilde{I}_{j+1}^{(\mathrm{l},c_2)}$. It is worth pointing out that the update step is kept unchanged since the dependencies are often exploited at the prediction stage. Similarly, for the right view, after generating $\widetilde{d}_{j+1}^{(\mathrm{r},c_2)}$ and $\widetilde{I}_{j+1}^{(\mathrm{l},c_2)}$ based on Eqs. (4) and (5), the hybrid prediction step will be applied to exploit further the cross-color channel correlation between c_2 and c_1, leading to the following detail coefficients:

$$
\begin{aligned}
\widehat{d}_{j+1}^{(\mathrm{r},c_2)}(x,y) = \widetilde{d}_{j+1}^{(\mathrm{r},c_2)}(x,y) - \lfloor \sum_{k \in \mathcal{Q}_j^{(\mathrm{r},c_2)}} q_{j,k} \widetilde{I}_{j+1}^{(\mathrm{r},c_2)}(x, y+k) \\
+ \sum_{k \in \mathcal{P}_j^{(\mathrm{r},\mathrm{l},c_2)}} p_{j,k}^{(\mathrm{r},\mathrm{l},c_2)} I_j^{(\mathrm{l},c_2)}(x + v_{x,j}(x,2y+1), 2y+1-k) \\
+ \sum_{k \in \mathcal{P}_j^{(\mathrm{r},c_1,c_2)}} p_{j,k}^{(\mathrm{r},c_1,c_2)} I_j^{(\mathrm{r},c_1)}(x, 2y+1-k) \rfloor.
\end{aligned}
\quad (9)
$$

Moreover, as shown in Eq. (7), at the last resolution level J, the residual subband $e_J^{(\mathrm{r},c_2)}$ is rewritten as:

$$
\begin{aligned}
e_J^{(\mathrm{r},c_2)}(x,y) = I_J^{(\mathrm{r},c_2)}(x,y) - \lfloor p_{J,k}^{(\mathrm{r},\mathrm{l},c_2)} I_J^{(l,c_2)}(x + v_{x,J}(x,y), y) \\
+ p_{J,k}^{(\mathrm{r},c_1,c_2)} I_J^{(\mathrm{r},c_1)}(x,y) \rfloor.
\end{aligned}
\quad (10)
$$

Finally, for the last color channel c_3 of the left and right views, the extension consists now in exploiting its correlation with simultaneously the previous encoded ones c_1 and c_2. Therefore, Eqs. (8), (9) and (10) becomes as follows:

$$
\begin{aligned}
\widetilde{d}_{j+1}^{(\mathrm{l},c_3)}(x,y) = I_j^{(\mathrm{l},c_3)}(x,2y+1) - \\
\lfloor \sum_{k \in \mathcal{P}_j^{(\mathrm{l},c_3)}} p_{j,k}^{(\mathrm{l},c_3)} I_j^{(\mathrm{l},c_3)}(x, 2y+2k) \\
+ \sum_{k \in \mathcal{P}_j^{(\mathrm{l},c_1,c_3)}} p_{j,k}^{(\mathrm{l},c_1,c_3)} I_j^{(\mathrm{l},c_1)}(x, 2y+1-k) \\
+ \sum_{k \in \mathcal{P}_j^{(\mathrm{l},c_2,c_3)}} p_{j,k}^{(\mathrm{l},c_2,c_3)} I_j^{(\mathrm{l},c_2)}(x, 2y+1-k) \rfloor,
\end{aligned}
\quad (11)
$$

$$
\begin{aligned}
\widehat{d}_{j+1}^{(\mathrm{r},c_3)}(x,y) = \widetilde{d}_{j+1}^{(\mathrm{r},c_3)}(x,y) - \lfloor \sum_{k \in \mathcal{Q}_j^{(\mathrm{r},c_3)}} q_{j,k} \widetilde{I}_{j+1}^{(\mathrm{r},c_3)}(x, y+k) \\
+ \sum_{k \in \mathcal{P}_j^{(\mathrm{r},\mathrm{l},c_3)}} p_{j,k}^{(\mathrm{r},\mathrm{l},c_3)} I_j^{(\mathrm{l},c_3)}(x + v_{x,j}(x,2y+1), 2y+1-k) \\
+ \sum_{k \in \mathcal{P}_j^{(\mathrm{r},c_1,c_3)}} p_{j,k}^{(\mathrm{r},c_1,c_3)} I_j^{(\mathrm{r},c_1)}(x, 2y+1-k) \\
+ \sum_{k \in \mathcal{P}_j^{(\mathrm{r},c_2,c_3)}} p_{j,k}^{(\mathrm{r},c_2,c_3)} I_j^{(\mathrm{r},c_2)}(x, 2y+1-k) \rfloor,
\end{aligned}
\quad (12)
$$

$$
\begin{aligned}
e_J^{(\mathrm{r},c_3)}(x,y) = I_J^{(\mathrm{r},c_3)}(x,y) \text{-} \lfloor p_{J,k}^{(\mathrm{r},\mathrm{l},c_3)} I_J^{(\mathrm{l},c_3)}(x + v_{x,J}(x,y), y) \\
+ p_{J,k}^{(\mathrm{r},c_1,c_3)} I_J^{(\mathrm{r},c_1)}(x,y) + p_{J,k}^{(\mathrm{r},c_2,c_3)} I_J^{(\mathrm{r},c_2)}(x,y) \rfloor,
\end{aligned}
\quad (13)
$$

where notations similar to those used previously are employed to represent the inter-view redundancies as well as the correlation of the current channel c_3 with c_1 and c_2. It is worth pointing out that the choice of appropriate weights is a key issue for the compactness of the multiscale representation. While the update coefficients are simply set to 1/4 as performed in the 5/3 lifting scheme of JPEG2000 [11], the prediction coefficients are optimized by minimizing the variance of the detail coefficients which allows us to build a coding scheme well adapted to the contents of the color stereo data.

3.2. Permutation of color components

In the previous section, the extended VLS for color stereo image is described for a given permutation $\mathrm{P} = (c_1, c_2, c_3)$ by considering c_1 as a reference channel which is encoded independently of the other color channels. Then, the second channel c_2 is encoded by exploiting its correlation with c_1. Finally, c_3 is encoded based on c_1 and c_2. Therefore, an important issue in color SI coding schemes concerns the choice of the permutation of the color components. For instance, if the RGB space is considered, it can be noticed that the possible permutations correspond to the following six modes:

• Mode 1: $\mathrm{P} = (R, G, B)$; • Mode 2: $\mathrm{P} = (R, B, G)$
• Mode 3: $\mathrm{P} = (G, R, B)$; • Mode 4: $\mathrm{P} = (G, B, R)$
• Mode 5: $\mathrm{P} = (B, R, G)$; • Mode 6: $\mathrm{P} = (B, G, R)$

Once the proposed decomposition is performed for these six permutations, an optimization criterion should be defined to select the optimal mode. In this respect, we propose to use a criterion reflecting the compactness of the multiscale representation which is defined directly on the resulting wavelet coefficients in order to avoid the encoding and decoding processes. More precisely, this appropriate criterion is expressed as follows:

$$\min \sum_{c \in \{R,G,B\}} \sum_{j=1}^{3J+1} \frac{1}{4^j} \Big(H_j^{(\mathrm{l},c)} + H_j^{(\mathrm{r},c)} \Big) \tag{14}$$

where $H_j^{(\mathrm{l},c)}$ and $H_j^{(\mathrm{r},c)}$ represent respectively the entropy of the j^{th}-wavelet subband of the c^{th} color channel for the left and right images. Finally, it is worth pointing out that, since the left and right images present similar contents, we propose to employ the same permutation P for the two views. This strategy presents the advantage of simplifying the optimization procedure.

4. EXPERIMENTAL RESULTS

Simulations are performed on six rectified color stereo images downloaded from the Middlebury stereovision website [1]. As described in Sec. 3.2, we have used the most common color system which is the RGB space, as its three coordinates are the reference colors in almost all the image acquisition processes. The disparity map is estimated by using the block-matching technique with a 8 × 8 block size. In order to evaluate the performance of the proposed color stereo image coding approach, we consider the following decompositions carried out over $J = 3$ resolution levels. The first one consists in applying the 9/7 wavelet transform, retained in the lossy compression mode of JPEG2000 [11], separately to each color channel of each view. This scheme is designated by "Independent scheme". The second one resorts to the standard stereo coding scheme performed *individually* on each pair $(I^{(\mathrm{l},c)}, I^{(\mathrm{r},c)})$. As aforementioned, this standard scheme encodes a reference image and a residual one by using the 9/7 transform. This scheme is denoted by "Standard residual scheme". The third one corresponds to the recent VLS-based decomposition described in Subsec. 2.2 applied also *individually* to each pair $(I^{(\mathrm{l},c)}, I^{(\mathrm{r},c)})$. This scheme is designated by "Classical VLS". While the two previous schemes exploit *only* the inter-view redundancies, we finally propose to improve them by further taking into account the inter-color channel correlations according to the strategy presented in Sec. 3. The extended VLS decomposition scheme is called "Improved VLS". Moreover, a straightforward extension of the standard residual scheme consists in generating residual images for a given channel c, by performing an hybrid prediction step that uses *simultaneously* the information coming from the other channels of the same view as well as that coming from the same channel of the other view. The improved approach is denoted by "Improved residual scheme". We should note here that the VLS-based decompositions have been carried out by using the following supports of the predictor and update operators:

$$\begin{aligned} &\forall c_i \in \{c_1, c_2, c_3\}, \quad \mathcal{P}_j^{(\mathrm{r,l},c_i)} = \{-1, 0, 1\}, \text{and} \\ &\qquad \mathcal{P}_j^{(\mathrm{l},c_i)} = \mathcal{P}_j^{(\mathrm{r},c_i)} = \mathcal{U}_j^{(\mathrm{l},c_i)} = \mathcal{U}_j^{(\mathrm{r},c_i)} = \mathcal{Q}_j^{(\mathrm{r},c_i)} = \{0, 1\}, \\ &\forall c_i \neq c_j, \quad \mathcal{P}_j^{(\mathrm{l},c_i,c_j)} = \mathcal{P}_j^{(\mathrm{r},c_i,c_j)} = \{0\}. \end{aligned} \tag{15}$$

The coding performance of the considered methods is measured in terms of the average PSNR given by:

$$\mathrm{PSNR} = 10 \log_{10} \left(\frac{225^2}{(MSE^{(\mathrm{l})} + MSE^{(\mathrm{r})})/2} \right) \tag{16}$$

where $MSE^{(\mathrm{l})}$ (resp. $MSE^{(\mathrm{r})}$) is the average of the mean square error of the three color components for the left (resp. right) view.

Fig. 1 shows the variations of the PSNR versus the bitrate for the SI "Laundry". It can be noticed that exploiting *jointly* the inter-view and inter-color channel correlations (i.e "Improved residual scheme" and "Improved VLS") clearly outperform the coding schemes where only the inter-view redundancies are exploited. Moreover, our proposed VLS-based color SI coding method achieves an improvement of about 0.3-2 dB compared to the standard residual-based coding approach. Fig. 2 displays a zoom applied on the reconstructed target image of "Art" for the improved versions of the standard residual and VLS-based coding approaches. The decoded images as well as the used quality metrics PSNR, SSIM and the structural distortion (StSD) [12] show that the proposed VLS decomposition significantly improves the visual quality of reconstruction. Finally, we have also used the Bjontegaard metric [13] to evaluate the performance of the proposed approach for different color SIs. The results are illustrated in Table 1 for low and high bitrates corresponding respectively to the four bitrate points $\{0.15, 0.2, 0.25, 0.3\}$ and $\{0.85, 0.9, 0.95, 1\}$ bpp. This table indicates that the proposed VLS approach outperforms the classical VLS (resp. Improved residual scheme) by about 1-2 dB (resp. 0.5-2 dB).

5. CONCLUSION

In this paper, we have presented a new technique based on VLS for color stereo image coding that exploits simultaneously the correlations between the color channels and the two views. Experimental results have shown the benefits of the proposed joint coding approach. In a future work, other optimization criteria for the selection of the optimal color permutation could be investigated. Moreover, in addition to the retained RGB color space, it would be interesting to consider other popular color systems such as the YUV space.

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

(a) Original target image (b) PSNR=29.82 dB, SSIM=0.80, StSD=0.70 (c) PSNR=30.60 dB, SSIM=0.83, StSD=0.66

Fig. 2. Zoom applied on the reconstructed right image of the "Art" S.I at 0.25 bpp using (b) Improved residual scheme (c) Improved VLS.

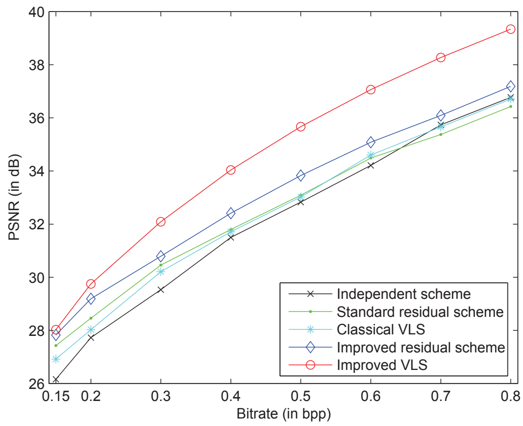

Fig. 1. PSNR (in dB) versus the bitrate (in bpp) after JPEG2000 progressive encoding of the stereo pair "Laundry".

Table 1. The average PSNR differences (in dB) at low and high bitrates.

Stereo Image	Improved VLS w.r.t Classical VLS		Improved VLS w.r.t Improved residual scheme	
	low	**high**	**low**	**high**
Laundry	1.69	2.61	0.69	2.15
Art	1.06	1.65	0.67	1.40
Dwarves	1.10	1.34	0.44	0.95
Drumsticks	1.12	1.74	0.35	1.50

REFERENCES

[1] M. G. Perkins, "Data compression of stereo pairs," *IEEE Trans. on Communic.*, vol. 40, no. 4, pp. 684–696, April 1992.

[2] W. Hachicha, A. Beghdadi, and F. A Cheikh, "1D directional DCT-based stereo residual compression," in *EUSIPCO*, Marrakech, Morocco, 2013.

[3] N. V. Boulgouris and M. G. Strintzis, "A family of wavelet-based stereo image coders," *IEEE Trans. on CSVT*, vol. 12, no. 10, pp. 898–903, October 2002.

[4] R. Darazi, A. Gouze, and B. Macq, "Adaptive lifting scheme-based method for joint coding 3D-stereo images with luminance correction and optimized prediction," in *IEEE ICASSP*, Taipei, April 2009, pp. 917–920.

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

[6] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frames stereo correspondance algorithms," *Int. J. of Computer Vision*, vol. 47, no. 1, pp. 7–42, 2002.

[7] K. Mühlmann, D. Maier, J. Hesser, and R. Männer, "Calculating dense disparity maps from color stereo images, an efficient implementation," *Int. J. of Computer Vision*, vol. 47, no. 1-3, pp. 79–88, April 2002.

[8] M. Bleyer and S. Chambon, "Evaluation of different methods for using colour information in global stereo matching approaches (regular paper)," in *Congress of the Internat. Soc. for Photogrammetry and Remote Sensing*, Beijing, China, 2008.

[9] A. Koschan, V. Rodehorst, and K. Spiller, "Color stereo vision using hierarchical block matching and active color illumination," in *Internat. Conf. on Pattern Recognition*, Vienna, Austria, Aug. 1996, vol. 1, pp. 835–839.

[10] T. Belli, M. Cord, and S. Philipp-Foliguet, "Colour contribution for stereo image matching," in *Internat. Conf. on Color in Graphics and Image Processing*, Saint-Etienne, France, October 2000, pp. 317–322.

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

[12] V. De Silva, H. K. Arachchi, E. Ekmekcioglu, and A. Kondoz, "Toward an impairment metric for stereoscopic video: A full-reference video quality metric to assess compressed stereoscopic video," *IEEE Trans. on Image Processing*, vol. 22, no. 9, pp. 3392–3404, September 2013.

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