Stereoscopic Image Coding Performance using Disparity-Compensated Block Matching Algorithm

Imen Kadri*†, Gabriel Dauphin*, Anissa Mokraoui*
*L2TI, Institut Galilée,
Université Paris 13,
Villetaneuse, France
Email: imen.kadri@enit.rnu.tn,
gabriel.dauphin@univ-paris13.fr,
anissa.mokraoui@univ-paris13.fr

Zied Lachiri†
†LSITI, Ecole Nationale d'Ingénieur de Tunis,
Université de Tunis El Manar,
Tunis,Tunisie
Email: lachiri.z@gmail.com

***Abstract*—This paper focuses on the disparity-compensated stereoscopic image coding. Such approach takes advantage of the existing redundancy between the two views as they are intended to render the visual impression of a 3D-scene, in which inter-view object displacements are understood as depth-related information. The classical approach is based on Block Matching (BM) algorithm, yielding a disparity map with which the predicted image is most similar to its original version. Then, with no modification of the disparity map, the residual image is encoded, yielding a refinement added to the predicted image. The proposed approach, first, improves all the possible predicted images taking into account this refinement, and then, estimates the disparity map as the one with which the predicted image resembles most that same view. Despite the significant increase in the numerical complexity, the substantial improved performance in terms of Peak-Signal to Noise-Ratio (PSNR) of this new approach is evidence of ongoing progress in this field of research.**

***Index Terms*—Stereoscopic Image Coding, Block Matching Algorithm, Disparity, Disparity-Compensation.**

I. Introduction

During the last decades, the use of stereo imaging technology has greatly increased. Applications concern the entertainment industry (3D cinema), video games, medical field (stereoscopic displays) and cartography (aerial stereo-photography).

A stereoscopic image is composed of two views. A specific device is required to make them perceived as two viewpoints of a single 3D-scene, where same points have different positions on each view. This spatial displacement is called disparity. Estimating the true disparity map (i.e. the true depth) remains an extensive research field [1].

In this paper, our concern stems from the increased storage needs when using stereoscopic images, and the existing redundancy between the two views. Improving the performance of stereoscopic image coding technique is also an extensive research field. Hence disparity compensated coding is a classical technique [2], [3]. Additional benefits is obtained when depth-based video coding is combined with view synthesis prediction, and this includes lifting schemes [4]. Higher performances are achieved when different techniques are combined as in H264 or in HEVC, which has been subjectively evaluated in [5]. Two extensions of H264, AVC and MVC are concerned with stereoscopic-image coding, note that they make use of disparity compensated coding [4].

This paper focuses on improving the disparity compensated coding scheme, which consists in coding separately a reference view, losslessly encoding an estimated disparity map and encoding a residual image. The transmitted information enables the decoder to reconstruct the reference view, and using the disparity map to compute a predicted view, to which is added the decoded residual image. This scheme is very similar to motion compensation.

Research within this framework has achieved increased performance when estimating the disparity map, by taking into account its own bit-cost in [6], by using blocks of arbitrary shapes in [7], and by addressing also the illumination compensation in [8]. Reducing the numerical complexity is also a significant research issue. Examples include selecting optimal hyper-parameter values thanks to allocation modeling as opposed to an exhaustive search in [9], and reducing the search area in [10].

This paper is a proof of concept showing the substantial progress to be expected when the disparity map is estimated by taking into account the encoding of the residual. To this end, an algorithm using greedy search has been designed. Section II shows how finding the best performing disparity map can be regarded as solving an optimization problem. The classical approach is derived as a suboptimal solution. Section III derives from a different suboptimal solution a new algorithm. In section IV, experimentations show significant increased performance on some stereoscopic images. Section V concludes the paper.

II. Problem statement

This paper is concerned with coding rectified stereoscopic images using the disparity-compensated coding scheme based on a closed loop. The reference view encoded separately is assumed to be the left view. The right view is then decomposed into K non-overlapping blocks of same size. In traditional stereoscopic image coding, pixel values of each block are first predicted using pixel values of the corresponding block

slightly shifted according to a disparity map. This first predicted view is then improved by processing its lossy decoded residual image. Note that in this paper, the refinement of the predicted view is being referred to as *compensation*. Notations are presented in subsection II-A and used in subsection II-B to set an optimization problem of which the well-known Block Matching algorithm is a suboptimal solution.

A. Notations

The following notations are illustrated on Fig. 1 showing both the classical disparity coding and decoding schemes separated by a dashed line.

$\mathbf{I}_l$ (upper left corner of Fig. 1) denotes the reference original left view. $\widehat{\mathbf{I}}_l$ is the left view reconstructed version yielded by the decoder (lower left corner of Fig. 1), the coder computes it too (center of the upper part). $\mathbf{I}_r$ (upper right corner and center of the upper part of Fig. 1) represents the original right view. $\mathbf{I}_p$ is the right view predicted version without compensation and $\widehat{\mathbf{I}}_r$ its compensated version yielded by the decoder (lower right corner of Fig. 1). $\mathbf{R}$ represents the residual image, that is the difference between the original right view and its prediction.

C_{q_l} (upper left corner of Fig. 1) denotes a lossy encoding operation which is controlled by parameter $q_l \in \mathbf{Q}_l$. $\mathbf{Q}_l$ is a set containing all allowed parameter values. It is used to specifically encode $\mathbf{I}_l$. D_l is the decoding operation applied on $C_{q_l}(\mathbf{I}_l)$. C_{q_r} (upper right corner of Fig. 1) is a lossy encoding operation which is also controlled by the parameter $q_r \in \mathbf{Q}_r$ to compress $\mathbf{R}$. D_r is the decoding operation applied on $C_{q_r}(\mathbf{R})$. C (center upper part of Fig. 1) is a lossless encoding operation of the disparity map $\mathbf{d}$. D refers to the inverse operation of C.

With these notations, a first description of the disparity compensated coding scheme is given by:

$$\widehat{\mathbf{I}}_l = D_l(C_{q_l}(\mathbf{I}_l)), \quad (1)$$

$$\mathbf{d} = D(C(\mathbf{d})), \quad (2)$$

$$\mathbf{R} = \mathbf{I}_r - \mathbf{I}_p, \quad (3)$$

$$\widehat{\mathbf{I}}_r = \mathbf{I}_p + D_r(C_{q_r}(\mathbf{R})). \quad (4)$$

There remains the core of the disparity compensated coding scheme which relies on two operators located at the upper part of Fig. 1.

- IP: the Image Predictor transforms $\widehat{\mathbf{I}}_l$ into $\mathbf{I}_p$ using $\mathbf{d} = (d_1, \ldots, d_K)$, a disparity map:

$$\mathbf{I}_p\left(\begin{bmatrix} i_k + \Delta i, \\ j_k + \Delta j \end{bmatrix}\right) = \widehat{\mathbf{I}}_l\left(\begin{bmatrix} i_k + \Delta i, \\ j_k + \Delta j + d_k \end{bmatrix}\right), \quad (5)$$

  where k ranges from 1 to K. $\mathbf{I}_r$ is decomposed into non-overlapping blocks that are k-labeled. $\mathcal{B}$ is the set of all internal-block positive displacements, each displacement is denoted as $[\Delta i, \Delta j]^T \in \mathcal{B}$. $[i_k + \Delta i, j_k + \Delta j]^T_{(\Delta i, \Delta j) \in \mathcal{B}}$ is the set of all pixel positions in the k-block.
- DE: the Disparity Estimator yields a disparity map from $\widehat{\mathbf{I}}_l$ and $\mathbf{I}_r$ using the classical Block Matching algorithm. This algorithm, here denoted as BM, assigns to d_k the disparity value for which $\mathbf{I}_p$ matches most to $\mathbf{I}_r$ on the k-block, in Eq. (6):

$$d_k = \underset{d \in \mathbf{S}}{\arg\min} \sum_{\begin{bmatrix} \Delta i \\ \Delta j \end{bmatrix} \in \mathcal{B}} \left(\begin{array}{c} \widehat{\mathbf{I}}_l \begin{bmatrix} i_k + \Delta i \\ j_k + \Delta j + d \end{bmatrix} \\ -\mathbf{I}_r \begin{bmatrix} i_k + \Delta i \\ j_k + \Delta j \end{bmatrix} \end{array} \right)^2, \quad (6)$$

  where $\mathbf{S}$ is the searching window containing all allowed disparity values.

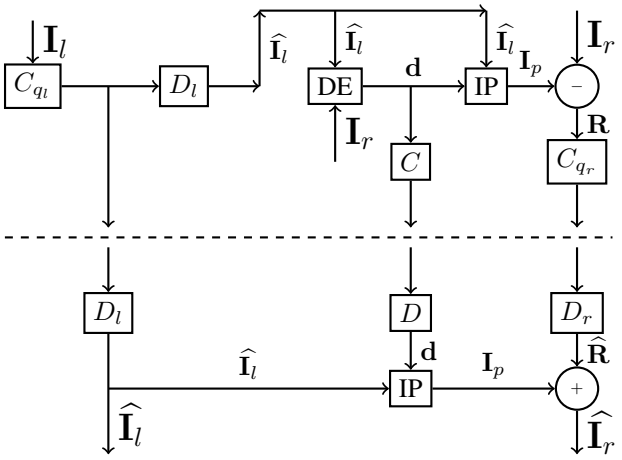

Fig. 1: Disparity compensated coding scheme (the dashed line separates the coder above from the decoder below).

The bitrate, denoted b, is deduced from the bit streams $C_{q_l}(\mathbf{I}_l)$, $C(\mathbf{d})$ and $C_{q_r}(\mathbf{R})$:

$$b(\mathbf{I}_l, \mathbf{d}, \mathbf{I}_r) = \frac{|C_{q_l}(\mathbf{I}_l)| + |C(\mathbf{d})| + |C_{q_r}(\mathbf{R})|}{|\mathbf{I}_l| + |\mathbf{I}_r|}, \quad (7)$$

where $|.|$ is the set cardinal number, here it helps counting the number of bits above and the number of pixels below.

B. Optimization problem statement

The aim of a coding/decoding scheme is a trade-off between getting the highest quality (e.g. visual rendering) while using the least amount of bits accounted for by Eq. (7). Here this trade-off is rephrased into finding the best quality (i.e. visual rendering) within a constrained bit-budget. As the end user observing the reconstructed stereoscopic image is generally a human being, our focus should be the extent to which the visual experience is being preserved. Designing objective quality metric modelling the quality of this visual experience is an existing research field as exemplified in [11]. As of now, no objective quality metric has proven to be completely reliable when applied to stereoscopic images. We are considering a quality metric, denoted as J to be used as a cost function in an optimization problem. It is for the sake of simplifying the description of the proposed algorithm in section III. As a consequence, at the coder level the choice of $q_l, q_r, \mathbf{d}$ should within a constrained bit-budget minimize J. However, numerical complexity being an important issue, the two following

assumptions seem to be required to make this optimization problem tractable:

(A1) *Up to a non-decreasing or non-increasing mapping, J is the sum of a metric computed independently on each blocks of $\widehat{\mathbf{I}}_r$ and $\mathbf{I}_r$ and a metric computed on $\widehat{\mathbf{I}}_l$ and $\mathbf{I}_l$.*

(A2) *Bit-rate control is addressed solely by selecting q_l and q_r.*

With these assumptions and denoting b_a the allowed bit-budget, the optimization problem addressing the optimal choice of $q_l, q_r, \mathbf{d}$ at the coder level is:

$$\mathbf{d}(q_l, q_r) = \quad \arg\min_{\mathbf{d}\in\mathbf{S}^K} \sum_k \Delta J_r\left(\widehat{\mathbf{I}}_r\,,\, \mathbf{I}_r\right) \tag{8}$$

$$J(q_l, q_r) = \quad J_l(\widehat{\mathbf{I}}_l, \mathbf{I}_l) + \sum_k \Delta J_r(\widehat{\mathbf{I}}_r, \mathbf{I}_r) \tag{9}$$

$$(q_l, q_r) = \quad \arg\min_{b(\mathbf{I}_l,\mathbf{d},\mathbf{I}_r)\le b_a} J(q_l, q_r), \tag{10}$$

where Eq. (8) is the modified DE-operator replacing Eq. (6), Eq. (9) accounts for assumption (A1) and Eq. (10) for assumption (A2).

A suboptimal solution is to assume, as in Fig. 1, that the modified DE-operator is fed with $\mathbf{I}_r$ and $\widehat{\mathbf{I}}_l$ and more specifically that equation Eq. (8) can be replaced by:

$$\mathbf{d}(q_l) = \arg\min_{\mathbf{d}\in S^K} \sum_k \Delta J_r\left(\mathbf{I}_p, \mathbf{I}_r\right), \tag{11}$$

where S^K is the set of all possible disparity maps and $\mathbf{I}_p$ is defined in Eq. (5). Since both $\mathbf{I}_p$ and ΔJ_r are computed independently on each block (see Eq. (5) and (A1)), this suboptimal solution is actually equivalent to:

$$d_k(q_l) = \arg\min_{d\in S} \Delta J_r\left(\mathbf{I}_p, \mathbf{I}_r\right). \tag{12}$$

Let us now suppose that the objective quality metric chosen is the PSNR (Peak Signal to Noise Ratio) whose specific definition for stereoscopic images is:

$$\begin{aligned}&\mathrm{PSNR}(\mathbf{I}_l, \widehat{\mathbf{I}}_l, \mathbf{I}_r, \widehat{\mathbf{I}}_r) = \\ &\quad 10\log_{10}\left(\frac{255^2(|\mathbf{I}_l|+|\mathbf{I}_r|)}{\sum_{i,j}(\widehat{\mathbf{I}}_l(i,j)-\mathbf{I}_l(i,j))^2+\sum_{i,j}(\widehat{\mathbf{I}}_r(i,j)-\mathbf{I}_r(i,j))^2}\right),\end{aligned} \tag{13}$$

where pixel values are ranging from 0 to 255, i, j span all pixel positions and $|\mathbf{I}_l|, |\mathbf{I}_r|$ denote the number of pixels. As $x \mapsto 10\log_{10}\left(\frac{255^2(|\mathbf{I}_l|+|\mathbf{I}_r|)}{x}\right)$ is a non-increasing mapping, the PSNR fullfills (A1) with:

$$\Delta J_r\left(\mathbf{I}_p, \mathbf{I}_r\right) = \sum_{\left[\begin{smallmatrix}\Delta i\\ \Delta j\end{smallmatrix}\right]\in\mathcal{B}} \left(\begin{array}{c}\mathbf{I}_p\left[\begin{smallmatrix} i_k+\Delta i\\ j_k+\Delta j\end{smallmatrix}\right] \\ \\ -\mathbf{I}_r\left[\begin{smallmatrix} i_k+\Delta i\\ j_k+\Delta j\end{smallmatrix}\right]\end{array}\right)^2. \tag{14}$$

Using equation Eq. (5), we can see, that in the PSNR-context, Eq. (12) is actually equivalent to Eq. (6), that is to the well-known BM algorithm.

Table I summarizes the disparity-compensated coding algorithm when equation Eq. (12) is used as a suboptimal solution, it is referred to here as the BM algorithm even when the quality metric is not the PSNR.

III. Proposed Disparity-Compensated Block Matching algorithm

The proposed Disparity-Compensated Block Matching algorithm, denoted as DCBM algorithm, is different from the BM algorithm in that Eq. (8) is no longer simplified into Eq. (11). DCBM algorithm is derived from a different suboptimal solution involving much greater numerical complexity.

The DCBM algorithm is computed in $K+1$ steps. In the first step, the disparity map is computed using the classical block matching algorithm. The initial disparity map is $d_k(0, q_l, q_r) = \arg\min_{d\in S} \Delta J_r\left(\mathbf{I}_p, \mathbf{I}_r\right)$ yielding no dependency to q_r (k ranges from 1 to K).

Assume that at each step $t \in \{1, \ldots, K\}$, a disparity map $\mathbf{d}(t-1, q_l, q_r)$ has already been computed. For each $s \in \mathbf{S}$, a predicted image $\mathbf{I}_p(t, q_l, q_r, s)$ is computed taking into account s on the t^{th} block and $d_k(t-1, q_l, q_r)$ for all other blocks:

$$\begin{aligned}&\mathbf{I}_p(t, q_l, q_r, s)\left[\begin{smallmatrix} i_k+\Delta i\\ j_k+\Delta j\end{smallmatrix}\right] = \\ &\quad\begin{cases} D_l C_{q_l}\mathbf{I}_l\left[\begin{smallmatrix} i_k+\Delta i\\ j_k+\Delta j+d_k(t-1,q_l,q_r)\end{smallmatrix}\right] & \text{if } k\neq t\\ \\ D_l C_{q_l}\mathbf{I}_l\left[\begin{smallmatrix} i_k+\Delta i\\ j_k+\Delta j+s\end{smallmatrix}\right] & \text{if } k=t\end{cases}\end{aligned} \tag{15}$$

with $(\Delta i, \Delta j)$ spanning $\mathcal{B}$ and k ranging from 1 to K.

Compensation transforms $\mathbf{I}_p(t, q_l, q_r, s)$ into $\widehat{\mathbf{I}}_r(t, q_l, q_r, s)$:

$$\begin{aligned}&\widehat{\mathbf{I}}_r(t, q_l, q_r, s) = \\ &\quad\mathbf{I}_p(t, q_l, q_r, s) + D_r(C_{q_r}\left(\mathbf{I}_r - \mathbf{I}_p(t, q_l, q_r, s)\right))\end{aligned} \tag{16}$$

Finally $\Delta J_r(\widehat{\mathbf{I}}_r, \mathbf{I}_r)$ is computed and the best disparity is selected as in Eq. (17):

$$d_k(t, q_l, q_r) = \begin{cases} d_k(t-1, q_l, q_r) & \text{if } k\neq t\\ \arg\min\limits_{s\in\mathbf{S}} \Delta J_r\left(\widehat{\mathbf{I}}_r(t, q_l, q_r, s), \mathbf{I}_r\right) & \text{if } k=t\end{cases} \tag{17}$$

The DCBM algorithm is summarized in Table II.

Note that the increased numerical complexity when using DCBM stems from the necessity to compress at each block a new image each time a new disparity value is considered. It is thought that, research focusing into modelling the ΔJ_r-distortions of JPEG coding and their relationship with the choice of the disparity, may reduce sharply this increased numerical complexity as only this ΔJ_r-distortion model would have to be computed each time a new disparity value is considered.

IV. Experiments and results

The performance of the proposed DCBM algorithm is compared to the BM algorithm on some stereoscopic images extracted from the Middleburry database [12] and the Deimos database [13].

In these experiments, the PSNR (defined in Eq. (13)) is used both as a performance measure and as the cost function used in the algorithms, resulting in ΔJ_r being defined as in Eq. (14).

TABLE I: Disparity compensated image coding in the J-context and using BM algorithm

Input: $\mathbf{I}_l, \mathbf{I}_r, b_a$

Output: $q_l, q_r, C_{q_l}(\mathbf{I}_l), C(\mathbf{d}), C_{q_r}(\mathbf{R}), b, J$

For each $q_l \in \mathbf{Q}_l$
- Compute $C_{q_l}(\mathbf{I}_l)$, $\widehat{\mathbf{I}}_l$ with Eq. (4) and $J_l(\widehat{\mathbf{I}}_l, \mathbf{I}_l)$
- For each $k \in \{1 \dots K\}$
  - For each $d \in \mathbf{S}$
    - Compute on the k-block, $\mathbf{I}_p$ with Eq. (5) and $\Delta J_r(\mathbf{I}_p, \mathbf{I}_r)$
  - End
  - Select d_k with Eq. (12) using $\Delta J_r(\mathbf{I}_p, \mathbf{I}_r)$
- End
- Compute $C(\mathbf{d})$ using $\mathbf{d} = (d_1, \dots, d_K)$
- Compute $\mathbf{I}_p$ with Eq. (5) and $\mathbf{R} = \mathbf{I}_r - \mathbf{I}_p$ with Eq. (3)
- For each $q_r \in \mathbf{Q}_r$
  - Compute $C_{q_r}(\mathbf{R})$
  - Compute $\sum_k \Delta J_r(\widehat{\mathbf{I}}_r, \mathbf{I}_r)$ on all blocks
  - Compute $J(q_l, q_r)$ with Eq. (9) using $J_l(\widehat{\mathbf{I}}_l, \mathbf{I}_l)$ and $\sum_k \Delta J_r(\widehat{\mathbf{I}}_r, \mathbf{I}_r)$
  - Compute $b(\mathbf{I}_l, \mathbf{d}, \mathbf{I}_r)$ with Eq. (7) using $C_{q_l}(\mathbf{I}_l), C(\mathbf{d}), C_{q_r}(\mathbf{R})$
- End

End

Select q_l, q_r with Eq. (10) using $J(q_l, q_r)$ and $b(\mathbf{I}_l, \mathbf{d}, \mathbf{I}_r)$

Compute $C_{q_l}(\mathbf{I}_l)$, $\widehat{\mathbf{I}}_l$ with Eq. (4) and $J_l(\widehat{\mathbf{I}}_l, \mathbf{I}_l)$

For each $k \in \{1 \dots K\}$
- For each $d \in \mathbf{S}$
  - Compute on the k-block, $\mathbf{I}_p$ with Eq. (5) and $\Delta J_r(\mathbf{I}_p, \mathbf{I}_r)$
- End
- Select d_k with Eq. (12) using $\Delta J_r(\mathbf{I}_p, \mathbf{I}_r)$

End

Compute $C(\mathbf{d})$ using $\mathbf{d} = (d_1, \dots, d_K)$

Compute $\mathbf{I}_p$ with Eq. (5) and $\mathbf{R} = \mathbf{I}_r - \mathbf{I}_p$ with Eq. (3)

Compute $C_{q_r}(\mathbf{R})$

Compute $\sum_k \Delta J_r(\widehat{\mathbf{I}}_r, \mathbf{I}_r)$ on all blocks

Compute J with Eq. (9) using $J_l(\widehat{\mathbf{I}}_l, \mathbf{I}_l)$ and $\sum_k \Delta J_r(\widehat{\mathbf{I}}_r, \mathbf{I}_r)$

Compute b with Eq. (7) using $C_{q_l}(\mathbf{I}_l), C(\mathbf{d}), C_{q_r}(\mathbf{R})$

TABLE II: Proposed Disparity-Compensated Block Matching algorithm.

Input: $\mathbf{I}_l, \mathbf{I}_r, b_a$

Output: $q_l, q_r, C_{q_l}(\mathbf{I}_l), C(\mathbf{d}), C_{q_r}(\mathbf{R}), b, J$

For each $q_l \in \mathbf{Q}_l$
- Compute $C_{q_l}(\mathbf{I}_l)$, $\widehat{\mathbf{I}}_l$ with Eq. (4) and $J_l(\widehat{\mathbf{I}}_l, \mathbf{I}_l)$
- Compute $\mathbf{d}(0, q_l)$ with Eq. (12)
- For each $q_r \in \mathbf{Q}_r$
  - For each $t \in \{1, \dots, K\}$
    - For each $s \in \mathbf{S}$
      - Compute $\mathbf{I}_p(t, q_l, q_r, s)$ with Eq. (15)
      - Compute $\widehat{\mathbf{I}}_r(t, q_l, q_r, s)$ with Eq. (16)
      - Compute $\Delta J_r\left(\widehat{\mathbf{I}}_r(t, q_l, q_r, s), \mathbf{I}_r\right)$
    - End
    - Compute $\mathbf{d}(t, q_l, q_r)$ with Eq. (17) using $\Delta J_r\left(\widehat{\mathbf{I}}_r, \mathbf{I}_r\right)$
  - End
- End

End

Select q_l, q_r with Eq. (10) using $J(q_l, q_r)$ and $b(\mathbf{I}_l, \mathbf{d}, \mathbf{I}_r)$

Compute $C_{q_l}(\mathbf{I}_l)$, $\widehat{\mathbf{I}}_l$ with Eq. (4) and $J_l(\widehat{\mathbf{I}}_l, \mathbf{I}_l)$

For each $t \in \{1 \dots K\}$
- For each $s \in \mathbf{S}$
  - Compute $\mathbf{I}_p(t, q_l, q_r, s)$ with Eq. (15)
  - Compute $\widehat{\mathbf{I}}_r(t, q_l, q_r, s)$ with Eq. (16)
  - Compute $\Delta J_r\left(\widehat{\mathbf{I}}_r(t, q_l, q_r, s), \mathbf{I}_r\right)$
- End
- Compute $\mathbf{d}(t, q_l, q_r)$ with Eq. (17) using $\Delta J_r\left(\widehat{\mathbf{I}}_r, \mathbf{I}_r\right)$

End

Compute $C(\mathbf{d})$ using $\mathbf{d} = (d_1, \dots, d_K)$

Compute $\mathbf{I}_p$ with Eq. (5) and $\mathbf{R} = \mathbf{I}_r - \mathbf{I}_p$ with Eq. (3)

Compute $C_{q_r}(\mathbf{R})$

Compute $\sum_k \Delta J_r(\widehat{\mathbf{I}}_r, \mathbf{I}_r)$ on all blocks

Compute J with Eq. (9) using $J_l(\widehat{\mathbf{I}}_l, \mathbf{I}_l)$ and $\sum_k \Delta J_r(\widehat{\mathbf{I}}_r, \mathbf{I}_r)$

Compute b with Eq. (7) using $C_{q_l}(\mathbf{I}_l), C(\mathbf{d}), C_{q_r}(\mathbf{R})$

To reduce the important amount of computations, the left view is not encoded and the PSNR is computed using only $\widehat{\mathbf{I}}_r$ and $\mathbf{I}_r$ with a simplified definition. The bitrate in bit-per-pixel (bpp) takes into account the amounts of bits to encode $\mathbf{d}$ and $\mathbf{R}$ and Eq. (7) is modified into: $b(\mathbf{d}, \mathbf{I}_r) = \frac{|C(\mathbf{d})| + |C_{q_r}(\mathbf{R})|}{|\mathbf{I}_r|}$. Arithmetic coding [14] has been chosen to encode $\mathbf{d}$ and JPEG to encode the $\mathbf{R}$. To reduce the numerical complexity, the available range of JPEG-hyperparameter values has been reduced in $\mathbf{Q}_r$ to $\{10, 20, \dots, 90\}$. Of great importance is the choice of the block size, set to 8×8 as it matches the JPEG-block size. Pixel values are ranging from 0 to 255. The searching window $\mathbf{S}$ is, unless otherwise specified, $\{-14, \dots, 15\}$.

Table III compares the performance of the DCBM algorithm to the BM algorithm using Bjontegaard metric [15]. On 10 stereoscopic images, the table shows on the second column the average increased PSNR-value at equal bitrate and on the third column the average bitrate decreasing ratio at equal distortion level. It appears that when using DCBM algorithm the increase in performance is substantial, above 1 dB on most tested images.

The rate-distortion curves of the DCBM algorithm (red curve connecting circles) and the BM algorithm (green curve connecting plus signs) are provided on Fig. 2. These results

TABLE III: Average increased performance of the DCBM algorithm compared to BM algorithm using the Bjontegaard metric.

Image	ΔPSNR (dB)	bpp (%)
House	+5.69	-44.6
Art	+2.39	-37.9
Bowling1	+1.85	-20.8
Books	+1.65	-27.5
Rubik	+1.63	-60.1
Aloe	+1.19	-30.1
Wood	+1.14	-18.9
Teddy	+1.1	-21
Barn	+0.44	-13
Deimos_573-12	+0.1	-4.4

have been carried out on the stereoscopic image "Bowling1", of size is 417×370, where the original right view is shown on Fig. 3. The increase in the PSNR-value is important and fairly constant over a wide range of bitrates (0.4 bpp and 1 bpp). Simulations have been done using Matlab in a Windows environment on a computer having the following characteristics: 3.7GHz, one processor, quad cores. Each point on Fig. 2 needs 0.8 seconds using the BM algorithm and about 4 hours using the DCBM algorithm.

Fig. 4 shows the reconstructed right views using the BM algorithm at 0.39 bpp on the left and DCBM algorithm at 0.38

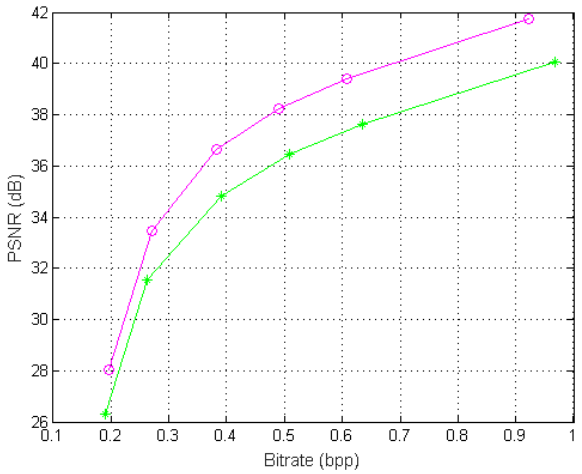

Fig. 2: Rate-distortion of the DCBM (red curve connecting circles) and BM (green curve connecting plus signs) algorithms with $\mathbf{S} = \{-14, \ldots, 15\}$ and "Bowling1" stereoscopic image.

bpp on the right. Differences between the two images can be observed when looking closely at the shadow, cast by the white left pin on the grey big ball: white block-coding JPEG-artifacts that can be seen on the BM-image are no-longer visible on the DCBM-image, even though JPEG is also used in the DCBM algorithm.

Fig. 3: Right view of the "Bowling1" stereoscopic image.

V. Conclusion

This paper proposed a new stereoscopic image coding algorithm where the block-based disparity map is selected taking into account also the compensation effect (i.e. a refinement of the predicted image using the encoded residual image). As compared to the classical approach, a substantial increase in performance is observed on most tested stereoscopic images. However this comes at the cost of a huge increase in numerical complexity, and this study is to be considered as a proof of concept. It is thought that adequate modelling of JPEG distortion may hinder the increase in numerical complexity.

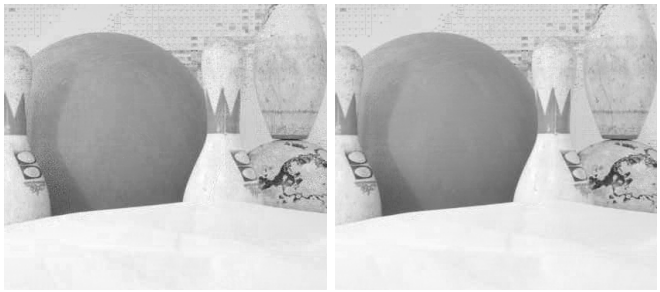

Fig. 4: Reconstructed "Bowling1" right view using: (i) algorithm BM at $b = 0.39$ bpp (left figure), (ii) algorithm DCBM at $b = 0.38$ bpp (right figure).

References

[1] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondance algorithms," *International Journal of Computer Vision, IJCV*, vol. 47, no. 1, pp. 7–42, April 2002.

[2] M. Flierl, A. Mavlankar, and B. Girod, "Motion and disparity compensated coding for multi-view video," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1474–1484, 2007.

[3] W. Woo and A. Ortega, "Stereo image compression with disparity compensation using the MRF model," in *Visual Communications and Image Processing*, vol. 2727, 1996, pp. 1–14.

[4] F. Dufaux, B. Pesquet-Popescu, and M. Cagnazzo, *Emerging Technologies for 3D Video: Creation, Coding, Transmission and Rendering*, 1st ed. Wiley Publishing, 2013.

[5] P. Hanhart, M. Rerabek, P. Korshunov, and T. Ebrahimi, "Subjective evaluation of HEVC intra coding for still image compression," [JCT-VC contribution] AhG4, Tech. Rep., January 2013.

[6] A. Kadaikar, G. Dauphin, and A. Mokraoui, "Modified block matching algorithm improving rate-distortion performance for stereoscopic image coding," in *2015 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, Dec 2015, pp. 478–483.

[7] G. Shen, W. Kim, A. Ortega, J. Lee, and H. Wey, "Edge-aware intra prediction for depth-map coding," in *2010 IEEE International Conference on Image Processing*, Sept 2010, pp. 3393–3396.

[8] Y. Chen, M. M. Hannuksela, L. Zhu, A. Hallapuro, M. Gabbouj, and H. Li, "Coding techniques in multiview video coding and joint multiview video model," in *2009 Picture Coding Symposium*, May 2009, pp. 1–4.

[9] W. Hachicha, M. Kaaniche, A. Beghdadi, and F. A. Cheikh, "Efficient inter-view bit allocation methods for stereo image coding," *IEEE Transactions on Multimedia*, vol. 17, no. 6, pp. 765–777, June 2015.

[10] R. Pan, Z.-X. Hou, and Y. Liu, "Fast algorithms for inter-view prediction of multiview video coding," *Journal of Multimedia*, vol. 6, pp. 191–201, April 2011.

[11] A. Benoit, P. L. Callet, P. Campisi, and R. Cousseau, "Using disparity for quality assessment of stereoscopic images," in *15th IEEE International Conference on Image Processing*, November 2008, pp. 389 – 392.

[12] D. Scharstein and C. Pal, "Learning conditional random fields for stereo," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2007, p. 8, http://vision.middlebury.edu/stereo/data/.

[13] K. Fliegel, P. Páta, M. Klíma, M. Blažek, and J. Havlín, "Open source database of images DEIMOS: high dynamic range images," *Proc. SPIE*, vol. 7798, pp. 1016–1023, 2010.

[14] P. G. Howard and J. Vitter, "Arithmetic coding for data compression," *Proceedings of the IEEE*, vol. 82, pp. 857 – 865, July 1994.

[15] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," ITU-T VCEG Meeting, Austin, Texas, USA, Document VCEG-M33, 2001.