Weighted Bi-Prediction for Light Field Image Coding

Caroline Conti[*a,b], Paulo Nunes[a,b], Luís Ducla Soares[a,b]
[a]Instituto de Telecomunicações Av. Rovisco Pais 1, 1049-001 Lisbon;
[b]Instituto Universitário de Lisboa (ISCTE-IUL), Av. das Forças Armadas,1649-026 Lisbon, Portugal

ABSTRACT

Light field imaging based on a single-tier camera equipped with a microlens array – also known as integral, holoscopic, and plenoptic imaging – has currently risen up as a practical and prospective approach for future visual applications and services. However, successfully deploying actual light field imaging applications and services will require developing adequate coding solutions to efficiently handle the massive amount of data involved in these systems. In this context, self-similarity compensated prediction is a non-local spatial prediction scheme based on block matching that has been shown to achieve high efficiency for light field image coding based on the High Efficiency Video Coding (HEVC) standard. As previously shown by the authors, this is possible by simply averaging two predictor blocks that are jointly estimated from a causal search window in the current frame itself, referred to as self-similarity bi-prediction. However, theoretical analyses for motion compensated bi-prediction have suggested that it is still possible to achieve further rate-distortion performance improvements by adaptively estimating the weighting coefficients of the two predictor blocks.

Therefore, this paper presents a comprehensive study of the rate-distortion performance for HEVC-based light field image coding when using different sets of weighting coefficients for self-similarity bi-prediction. Experimental results demonstrate that it is possible to extend the previous theoretical conclusions to light field image coding and show that the proposed adaptive weighting coefficient selection leads to up to 5 % of bit savings compared to the previous self-similarity bi-prediction scheme.

Keywords: Light field coding, plenoptic, holoscopic, HEVC, weighted bi-prediction

1 INTRODUCTION

Light Field (LF) imaging based on a single-tier camera equipped with a Microlens Array (MLA) – also known as holoscopic, plenoptic, and integral imaging – has recently become a prospective imaging approach for providing richer content capture, visualization, and manipulation, being applicable in many different areas of research, e.g., 3D television[1,2], biometric recognition[3], and medical imaging[4].

Recognizing the potential of this emerging technology, as well as the new challenges that need to be overcome for successfully introducing LF applications into the consumer market, novel standardization initiatives on LF image and video coding standardization are also emerging. Notably, the Joint Photographic Experts Group (JPEG) committee has recently started the JPEG Pleno standardization initiative[5] that addresses representation and coding of emerging new imaging modalities. In addition, the Moving Picture Experts Group (MPEG) group has also recently started a new work item on coded representations for immersive media (MPEG-I)[6]. The challenge to provide a LF representation with convenient spatial and angular resolutions requires handling a huge amount of data and, thus, efficient coding becomes of utmost importance. In this context, although standardized LF representation and coding solutions are still in an early stage of development, various LF coding solutions have been already proposed in the literature.

Several Light Field Coding (LFC) solutions in the literature try to take advantage of the particular planar intensity distribution of the LF image. Notably, as a result of the used optical system, the raw LF image can be represented as a 2D array of Micro-Images (MIs), and significant cross-correlation exists between these MIs in a neighborhood, which may be exploited for improving compression efficiency. Alternatively, Viewpoint Images (VIs) (a.k.a., subaperture images), or higher resolution views can be extracted/rendered from the LF content to represent the LF data as a set of views and to exploit the inter-view redundancy for LF coding. Additionally, other LF data representations can be also designed by extracting the depth/disparity information, as well as information on the intensity and direction of each light ray (i.e., the

[*] caroline.conti@lx.it.pt; phone +351 218 418 164, www.it.pt/cconti

ray-space) from the captured LF image. Essentially, previous LFC solutions can be then categorized in the following four main approaches, depending on the representation format and coding schemes that are adopted:

i) **LFC based on non-local spatial prediction**[7–12], which relies on a non-local prediction technique to exploit the redundancy between MIs in a spatial neighborhood when encoding the entire raw LF image;

ii) **Transform-based LFC**[13,14], which uses a Discrete Cosine Transform (DCT) or a Discrete Wavelet Transform (DWT) to exploit the redundancy between stacks of MIs or VIs;

iii) **LFC based on inter-view prediction**[15–20], in which a set of MIs, VIs or high resolution views are extracted and coded as a Pseudo-Video Sequence (PVS) or as multiview content; and

iv) **Disparity-assisted coding**[21–23], in which the disparity information is derived from the LF content and encoded along with a sparse set of texture information.

This paper is focused on LFC solutions based on non-local spatial prediction. These solutions are less dependent on a very precise calibration pre-process and are mainly advantageous for applications in which the LF content is consumed by the end user in a format similar to the captured format, or is consumed by using a proprietary LF rendering algorithm that makes use of the same (raw) 2D format. In this context, the authors' previous work has shown that significantly better coding performance can be achieved compared to the state-of-the art High Efficiency Video Coding (HEVC) coding standard by using the concept of Self-Similarity (SS) compensated prediction[7–9]. Similarly to motion estimation, this is possible by simply averaging two predictor blocks that are jointly estimated (based on block matching) from a causal search window in the current frame itself, referred to as self-similarity bi-prediction (Bi-SS)[9]. Motivated by these results, this paper proposes to experimentally analyze if there is still room for improving the Bi-SS prediction by making use of different weighting coefficients for combining the two jointly estimated predictor blocks. While theoretical motivations for weighted motion compensated bi-prediction has been previously presented and analyzed in the literature (e.g., in Girod[24]), its Rate-Distortion (RD) performance has not been analyzed yet for LF image coding. Therefore, this paper presents a comprehensive study of the RD performance for HEVC-based LF image coding when using two different weighted Bi-SS prediction schemes, namely: *i)* using a fixed set of weighting coefficients that are different from the averaging coefficients previously adopted for Bi-SS prediction; *ii)* using an adaptive algorithm that is here proposed for estimating the optimal set of weighting coefficients for each prediction block.

The remainder of this paper is organized as follows: Section 2 reviews the relevant work on LFC solutions based on non-local spatial prediction; Section 3 describes two weighted Bi-SS prediction schemes with fixed and adaptive weighting coefficients; Section 4 presents the test conditions and experimentally analyzes these two weighted Bi-SS prediction schemes; and, finally, Section 5 concludes the paper.

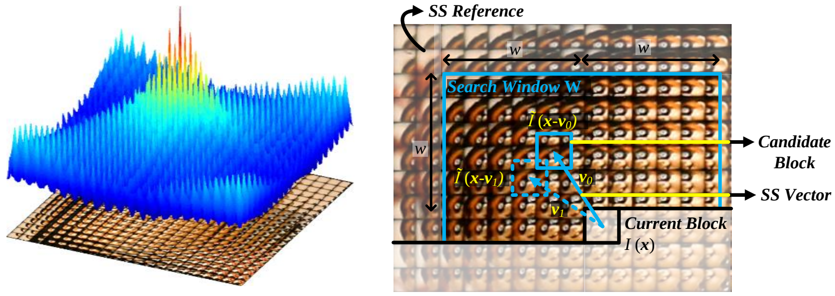
2 RELATED WORK

In the context of LFC based on the non-local spatial predictive approach, previous work of the authors [7–9] showed that further improvements are still possible for LF images with respect to the state-of-the-art for 2D image coding using the HEVC Main Still Picture profile[25,26] by using the concept of SS compensated prediction.

The SS estimation (depicted in Figure 1) is used to exploit the cross-correlation existing in an MI neighborhood (see Figure 1a) by estimating the prediction block with the highest similarity (according to appropriate criteria) to the current block in the previously coded and reconstructed area of the current picture itself (the SS reference, as seen in Figure 1b). Hence, the relative position between the current and the 'best' candidate block is signaled by an SS vector, $\boldsymbol{v}_0$, (see Figure 1b). Similarly to the conventional HEVC inter P frame prediction, the best SS vector, $\boldsymbol{v}_0^{best}$, for the SS prediction can be found in terms of Rate-Distortion Optimization (RDO) by minimizing the Lagrangian cost function in (1) [27],

$$J_{Uni-SS}^{best} = \min_{\boldsymbol{v}_0} \left\| I(\boldsymbol{x}) - \tilde{I}(\boldsymbol{x} - \boldsymbol{v}_0) \right\|_1 + \lambda\, R(\boldsymbol{v}_0) \quad (1)$$

where $I(\boldsymbol{x})$ is a matrix variable representing the current block at position $\boldsymbol{x} = (x, y)$ in the LF image; $\tilde{I}(\boldsymbol{x} - \boldsymbol{v}_0)$ represents a candidate block in the SS reference, $\tilde{I}$, with $\boldsymbol{x} - \boldsymbol{v}_0 \in \mathbf{W}$ (see Figure 1b); $R(\boldsymbol{v}_0)$ corresponds to an estimated number of bits for encoding the SS vector $\boldsymbol{v}_0$ (i.e., the estimated number of bits necessary to encode the motion vector difference between $\boldsymbol{v}_0$ and its predictor); and λ is the Lagrangian multiplier. In addition, the ℓ_1-norm (or Sum of Absolute Differences (SAD)), $\| \; \|_1$, is used and a limited causal search window $\mathbf{W}$ is adopted. Finally, the SS predictor block, $\hat{I}(\boldsymbol{x})$, is derived



(a) (b)

Figure 1 SS prediction: (a) inherent MI cross-correlation in a light field image neighborhood; (b) Bi-SS estimation process (example of a second candidate block and SS vector for bi-prediction is shown in dashed blue line).

as $\tilde{I}(\boldsymbol{x} - \boldsymbol{v}_0^{best})$. As done in HEVC reference software version 14.0[28], when SAD is used as the distortion measure, λ is given by $\sqrt{\lambda_{Intra}}$, where λ_{Intra} is the Lagrangian multiplier computed for prediction mode selection in intra-coded frames.

Notice that the SS estimation process in (1) only considers a single compensated signal for prediction of the current block, as firstly proposed by the authors[7,8], and for this reason will be hereinafter referred to as Uni-predicted Self-Similarity (Uni-SS) compensated prediction. An MI-based vector prediction scheme has been also proposed by the authors[8] to take advantage of the particular characteristics of the SS prediction data. In this case, three MI-based vector prediction (MIVP) candidate vectors, is included into the HEVC Advanced Motion Vector Prediction (AMVP) and merge candidate lists[29] to further improve the RD performance.

Although not targeting LF image coding, another prediction scheme similar to the Uni-SS compensated prediction[7,8], known as Intra Block Copy (IntraBC)[30], has been recently proposed in the literature in the context of Screen Content Coding (SCC)[30]. In this case, the prediction estimation is performed considering only integer pixel accuracy and the search window is expanded to the entire Coding Block (CB) row or column, or to the entire previously coded area of the picture by using a hash-based search[30].

To further improve the performance of the Uni-SS compensated prediction[8], a jointly estimated Bi-SS estimation and compensation scheme has been also proposed by the authors[9] that is based on the generic concept of superimposed prediction [31], which allows bi-prediction using samples from the same search area. Therefore, these predictor blocks can be located in the same MI and in overlapped pixel positions as illustrated in Figure 1b. Moreover, instead of simply combining two (independent) best uni-predicted candidate blocks for bi-prediction, the locally optimal rate-constrained algorithm[32] is used for jointly estimating these two predictor blocks. More specifically, two possible candidate predictors are derived from the same search area, **W** (Figure 1b), to predict the current block, namely: *i)* the Uni-SS candidate, and *ii)* the Bi-SS candidate.

The Uni-SS candidate predictor corresponds to the previous solution[7,8], in which the predictor block is found by minimizing the Lagrangian cost function in (1). Differently, for the Bi-SS candidate predictor, the iterative algorithm in

Initialization: $k = 0, \quad h_0 = h_1 = 1/2, \quad v_0^{(k)} = v_0^{best}, \quad J_{Bi-SS}^{(k)} = J_{MAX}$

do

$J_{Bi-SS}^{best} = J_{Bi-SS}^{(k)}$

$p = (k)\mathrm{mod}(2),\ q = 1 - p$

$J_{Bi-SS}^{(k+1)}(\boldsymbol{v}_q^{(k+1)} | \boldsymbol{v}_p^{(k)}) = \left\| I(\boldsymbol{x}) - \left[h_q \cdot \tilde{I}(\boldsymbol{x} - \boldsymbol{v}_q^{(k+1)}) + h_p \cdot \tilde{I}(\boldsymbol{x} - \boldsymbol{v}_p^{(k)})\right] \right\|_1 + \lambda \cdot \left[R(\boldsymbol{v}_q^{(k+1)}) + R(\boldsymbol{v}_p^{(k)})\right]$

$\boldsymbol{v}_q^{(k+1)} = \arg\min_{\boldsymbol{v}_q^{(k+1)}} J_{Bi-SS}^{(k+1)}$

$k = k + 1$

while $k \neq K$ **or** $J_{Bi-SS}^{(k)} < J_{Bi-SS}^{best}$

Figure 2 Algorithm for jointly estimating the two predictor blocks $\tilde{I}(\boldsymbol{x} - \boldsymbol{v}_0)$ and $\tilde{I}(\boldsymbol{x} - \boldsymbol{v}_1)$ for the Bi-SS candidate predictor. The index q defines which of the two vectors ($\boldsymbol{v}_0$ or $\boldsymbol{v}_1$) will be optimized in a particular iteration k, while the index p defines the vector that will be kept fixed.

Figure 2 is used to jointly estimate two predictor blocks, which are then linearly combined using the average weighting coefficient $\boldsymbol{h} = \left(\frac{1}{2}, \frac{1}{2}\right)$. This algorithm avoids searching through all possible combinations of two candidate predictor blocks $\tilde{I}(\boldsymbol{x} - \boldsymbol{v}_0)$ and $\tilde{I}(\boldsymbol{x} - \boldsymbol{v}_1)$ inside $\mathbf{W}$. For this, in each algorithm iteration, k, an optimal SS candidate vector $\boldsymbol{v}_q^{(k+1)}$ (with index $q \in \{0,1\}$) is found by minimizing the Lagrangian cost function, $J_{Bi-SS}^{(k+1)}$, conditioned to the optimal SS candidate vector found in the previous iteration $\boldsymbol{v}_p^{(k)}$ (with $p \in \{0,1\}$). Therefore, the algorithm is focused on finding an optimized vector $\boldsymbol{v}_1$ conditioned to a known vector $\boldsymbol{v}_0$ in even iterations, and vice versa in odd iterations. The maximum number of iterations, K, defines a tradeoff between complexity and RD performance and can be adjusted according to the system constraints. Similarly to (1), the Lagrangian cost function shown in Figure 2 is used to find the optimal SS vector in each iteration, where λ is computed as $\sqrt{\lambda_{Intra}}$ [28] and $R\left(\boldsymbol{v}_q^{(k+1)}\right) + R(\boldsymbol{v}_p^{(k)})$ corresponds to the estimated number of bits for encoding the SS vectors $\boldsymbol{v}_0$ and $\boldsymbol{v}_1$ given in each iteration, i.e., the estimated number of bits necessary to encode the motion vector difference between the SS vectors and their predictor vectors and for signaling the vector predictor using AMVP. Finally, the best prediction between Uni-SS and Bi-SS candidates is chosen in terms of conventional RDO[27] by comparing the associated Lagrangian costs J_{Uni-SS}^{best} and J_{Bi-SS}^{best}, respectively found in (1) and Figure 2. Regarding the achieved coding efficiency, it was shown[9] that jointly estimating the two candidate blocks for Bi-SS prediction led to further RD improvements when compared to the case[7,8] in which only one candidate block is estimated.

An LFC solution similar to the Bi-SS solution[9] has been also proposed in the literature[10] to extend the SS prediction concept by using HEVC inter B frame bi-prediction LF image coding. However, in this case, to guarantee that the two prediction signals came from two different MIs, the search area was proposed to be separated into two non-overlapping parts[10] to perform the prediction estimation as in conventional HEVC bi-prediction. For this reason, this solution is referred to here as LFC Restricted-SS[10] solution.

3 WEIGHTED BI-SS PREDICTION

Motivated by the authors' previous results[7–9], two different weighted Bi-SS prediction schemes are proposed in this section aiming at improving the Bi-SS prediction RD performance by considering different weighting factors for bi-prediction. These are: *i)* Weighted Bi-SS prediction with fixed weighting coefficients; and *ii)* Adaptive weighted Bi-SS prediction.

3.1 Weighted Bi-SS Prediction with Fixed Weighting Coefficients

This represents the simplest solution – referred to as LFC Weighted Bi-SS (Fixed) – in which a fixed set of weighting coefficients, $\mathbf{h} = (h_0, h_1 = 1 - h_0)$ (see Figure 2), are used for combining the two predictor blocks for Bi-SS prediction. The goal of this Weighted Bi-SS (Fixed) solution is to analyze if there is a better balance between the two predictor blocks for jointly estimated bi-prediction that leads to a better RD performance compared to the averaging coefficients that has been adopted in the previous LFC Bi-SS solution[9].

For this, the HEVC weighted prediction signaling [33] is used. Basically, the usage of explicit weighted prediction in HEVC is activated by a flag in the Picture Parameter Set (PPS), and different integer weighting factors, w_p, and offset values, o_p, can be assigned for prediction in each slice [33].

The resulting predictor block $\hat{I}(\boldsymbol{x})$ for weighted bi-prediction can be then derived by [33]:

$$\hat{I}(\boldsymbol{x}) = \left\lfloor \frac{w_0 \cdot \tilde{I}(\boldsymbol{x} - \boldsymbol{v}_0) + w_1 \cdot \tilde{I}(\boldsymbol{x} - \boldsymbol{v}_1) + (o_0 + o_1 + 1) \cdot 2^{LWD}}{2 \cdot 2^{LWD}} \right\rfloor \tag{2}$$

where LWD is a log weight denominator rounding factor [33] used to normalize the integer weighting factors and the sub-sample interpolation filtering process [34]. These weighting parameters (i.e., LWD, w_0, w_1, o_0, and o_1) are then coded in the slice header using variable length codes (notably, using zero-order Exponential-Golomb coding) and sent to the decoder.

Therefore, the following four LFC Weighted Bi-SS (Fixed) solutions are considered, which are analyzed in terms of their RD performance in Section 4.2:

- *LFC Weighted Bi-SS* $\left(\mathbf{h} = \left(\frac{7}{8}, \frac{1}{8}\right)\right)$ – In this case, weighted Bi-SS prediction is used with weighting coefficients $h_0 = \frac{7}{8}$ and $h_1 = \frac{1}{8}$. In (2), this corresponds to having $w_0 = 7$, $w_1 = 1$, $LWD = 2$, and $o_0 = o_1 = 0$.

- *LFC Weighted Bi-SS* $\left(\mathbf{h} = \left(\frac{3}{4}, \frac{1}{4}\right)\right)$ – In this case, weighted Bi-SS prediction is used, where the average weighting coefficients in Figure 2 are replaced by $h_0 = \frac{3}{4}$ and $h_1 = \frac{1}{4}$. In (2), this corresponds to having $w_0 = 3$, $w_1 = 1$, $LWD = 1$, and $o_0 = o_1 = 0$.
- *LFC Weighted Bi-SS* $\left(\mathbf{h} = \left(\frac{1}{4}, \frac{3}{4}\right)\right)$ – In this case, weighted Bi-SS prediction is used, where the average weighting coefficients in Figure 2 are replaced by $h_0 = \frac{1}{4}$ and $h_1 = \frac{3}{4}$. In (2), this corresponds to having $w_0 = 1$, $w_1 = 3$, $LWD = 1$, and $o_0 = o_1 = 0$.
- *LFC Weighted Bi-SS* $\left(\mathbf{h} = \left(\frac{1}{8}, \frac{7}{8}\right)\right)$ – In this case, weighted Bi-SS prediction is used with weighting coefficients $h_0 = \frac{1}{8}$ and $h_1 = \frac{7}{8}$. In (2), this corresponds to having $w_0 = 1$, $w_1 = 7$, $LWD = 2$, and $o_0 = o_1 = 0$.

3.2 Adaptive Weighted Bi-SS Prediction

The adaptive weighted Bi-SS solution proposed in this section – referred to as LFC Weighted Bi-SS (Adaptive) – is motivated by the theoretical analysis proposed for motion compensated prediction by Girod[24], which suggests that further RD performance improvements for bi-prediction can be achieved by adaptively estimating the weighting coefficients for each block being coded. Therefore, this Weighted Bi-SS (Adaptive) solution is used to experimentally analyze if it is possible to generalize the theoretical assumptions from Girod[24] for HEVC-based LF image coding using jointly estimated Bi-SS prediction.

For this, the Weighted Bi-SS (Adaptive) prediction is evaluated for all CB sizes (i.e., from 64×64 down to 8×8). Moreover, to perform the weighted Bi-SS estimation, each CB can be further split into smaller Prediction Blocks (PBs) considering all partition patterns defined for HEVC inter coding[33]. Similarly to the LFC solutions discussed in Section 2, the Lagrangian formulation in (3) can be used for finding, for each PB partition, the set of optimal predictor parameters $\{\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{h}\}$ that minimizes the prediction error in the weighted Bi-SS prediction subjected to the number of bits, $R(\boldsymbol{v}_0) + R(\boldsymbol{v}_1) + R(\boldsymbol{h})$, necessary for signaling each of these predictor parameters to the decoder.

$$J_{Weighted\ Bi-SS}(\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{h}) = \left\| I(\boldsymbol{x}) - [h_0 \cdot \tilde{I}(\boldsymbol{x} - \boldsymbol{v}_0) + h_1 \cdot \tilde{I}(\boldsymbol{x} - \boldsymbol{v}_1)] \right\|_1 + \lambda \cdot [R(\boldsymbol{v}_0) + R(\boldsymbol{v}_1) + R(\boldsymbol{h})] \quad (3)$$

Given the very large number of coding possibilities of HEVC (e.g., due to the very flexible PB partition patterns and sizes, estimation considering quarter-pixel precision, and using an optimized set of vector predictors) and to avoid significantly increasing the complexity for solving the optimization problem in (3), the iterative algorithm in Figure 3 is proposed to derive a Weighted Bi-SS candidate predictor, which considers a limited set of five possible weighting coefficients, $\mathcal{H} = \left\{\left(\frac{7}{8}, \frac{1}{8}\right), \left(\frac{3}{4}, \frac{1}{4}\right), \left(\frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{4}, \frac{3}{4}\right), \left(\frac{1}{8}, \frac{7}{8}\right)\right\}$. Therefore, the set of coefficients, $\boldsymbol{h}_i \in \mathcal{H}$, that minimizes the locally optimal constrained Bi-SS solution in Figure 2 conditioned to $\boldsymbol{h}_i$ – i.e., $J_{Bi-SS}\,(\boldsymbol{v}_0, \boldsymbol{v}_1 \mid \boldsymbol{h}_i)$ indicated in Figure 3 – is chosen as the best weighting coefficients, $\boldsymbol{h}^{best}$, and the Weighted Bi-SS candidate predictor is then found by using the weighted linear combination of the corresponding jointly estimated predictor blocks $\tilde{I}(\boldsymbol{x} - \boldsymbol{v}_0^{best})$ and $\tilde{I}(\boldsymbol{x} - \boldsymbol{v}_1^{best})$ (Figure 3).

Initialization: $J^{best}_{Weighted\ Bi-SS} = J_{MAX}$

for each $\boldsymbol{h}_i \in \mathcal{H}$ **do**

Find the optimal jointly estimated predictor blocks according to Figure 2:

$$J_{Weighted\ Bi-SS}(\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{h}_i) = \underset{\boldsymbol{v}_0, \boldsymbol{v}_1}{\operatorname{argmin}}\ J_{Bi-SS}\,(\boldsymbol{v}_0, \boldsymbol{v}_1 | \boldsymbol{h}_i)$$

If $J_{Weighted\ Bi-SS} < J^{best}_{Weighted\ Bi-SS}$ **then**

$J^{best}_{Weighted\ Bi-SS} = J_{Weighted\ Bi-SS}, \quad \boldsymbol{v}_0^{best} = \boldsymbol{v}_0, \quad \boldsymbol{v}_1^{best} = \boldsymbol{v}_1, \quad \boldsymbol{h}^{best} = \boldsymbol{h}_i$

end

end for

Figure 3 Iterative algorithm used for solving the problem in (3) and determining an optimized weighted Bi-SS predictor. Essentially, the set of weights, $\boldsymbol{h}_i$, that minimizes the locally optimal constrained Bi-SS solution conditioned to $\boldsymbol{h}_i$, $J_{Bi-SS}(\boldsymbol{v}_0, \boldsymbol{v}_1 | \boldsymbol{h}_i)$, is chosen as the best weighting coefficients, $\boldsymbol{h}^{best}$, and the Weighted Bi-SS candidate predictor is found by using the weighted linear combination of the corresponding jointly estimated predictor blocks $\tilde{I}(\boldsymbol{x} - \boldsymbol{v}_0^{best})$ and $\tilde{I}(\boldsymbol{x} - \boldsymbol{v}_1^{best})$.

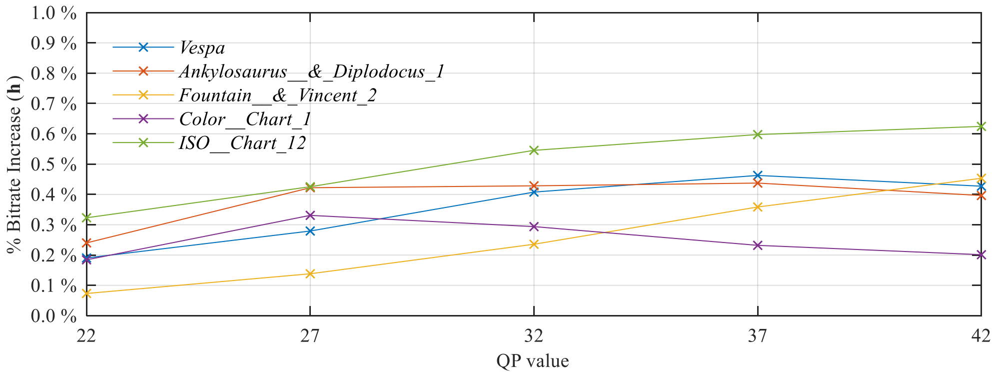


Figure 4 Estimated bounds for entropy coding the chosen set of weighting coefficients when coding each LF image considered in Section 4 (see Figure 5) using the proposed Weighted Bi-SS (Adaptive) prediction for five different QP values: 22, 27, 32, 37, and 42. The percentage of bitrate increase is calculated as $-n_{PB}\sum p(\boldsymbol{h}_i)\cdot\log_2 p(\boldsymbol{h}_i)/T_{bits}$, where: $-\sum p(\boldsymbol{h}_i)\cdot\log_2 p(\boldsymbol{h}_i)$ is the minimum average number of bits needed for signaling the chosen set of weighting coefficients, $\boldsymbol{h}_i$, for each PB partition; $p(\boldsymbol{h}_i)$ is the probability of usage of $\boldsymbol{h}_i$; n_{PB} is the number of PB partitions that make use of the proposed Weighted Bi-SS (Adaptive) prediction; T_{bits} is the total number of bits for encoding the entire LF image (without weighting coefficients signaling).

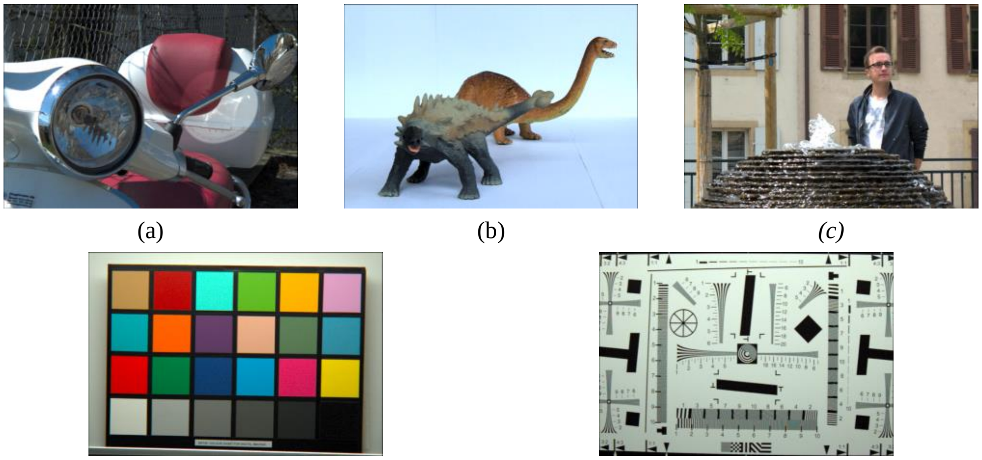
It should be noticed that the Weighted Bi-SS (Adaptive) prediction proposed in this paper aims at experimentally analyzing if there is room for improving the jointly estimated Bi-SS prediction efficiency for LF image coding by using an optimal set of weighting coefficients $\boldsymbol{h}^{best}$ as theoretically suggested for motion compensated prediction[24]. For this reason, and since the choice of $\boldsymbol{h}^{best}$ is highly dependent on how efficiently these sets of weighting coefficients are entropy encoded, the number of bits $R(\boldsymbol{h})$ in (3) are not considered when iteratively solving the problem in Figure 3 and are not accounted for in the RD performance of the proposed LFC Weighted Bi-SS (Adaptive) solution.

Nevertheless, Figure 4 gives some insight into the RD performance bounds for LF image coding using the proposed Weighted Bi-SS (Adaptive) prediction. For this, an estimation of the minimum bitrate increase for entropy coding the chosen set of weighting coefficients, $\boldsymbol{h}$, for each PB partition is presented in Figure 4 for all tested LF images (see Section 4) considering five different Quantization Parameter (QP) values. This rough estimation is considered when experimentally assessing the RD performance of the proposed Weighted Bi-SS (Adaptive) prediction in Section 4.

As an alternative to the Weighted Bi-SS (Adaptive) prediction, an SS-skip mode can be also used for coding a CB that contains only a single PB, and in this case, all predictor parameters $\{\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{h}\}$ are directly derived using the HEVC merge technique[33]. Therefore, only an index is signaled which identifies the chosen predictor parameters from a list of merge candidates. In this work, the merge candidate list is derived as follows:

- **Spatial merge vector candidates** – Up to four spatial candidates from the set of five[29] neighboring blocks that were coded with Weighted Bi-SS mode are included. In this case, the set of weighting coefficients, $\boldsymbol{h}$, is directly derived from the set of weighting coefficients adopted by these spatial neighboring blocks, where $\boldsymbol{h} \in \mathcal{H}$.
- **MI-based vector candidates** – The maximum size of the merge candidate list is signaled in the slice header syntax (being equal to five as defined by default in HEVC standard[29]). After selecting the spatial candidates, up to three MIVP merge candidates, as previously proposed by the authors[8], are included into the merge candidate list until the maximum number of candidates is reached. In this case, the set weighting coefficients is set to $\boldsymbol{h} = \left(\frac{1}{2}, \frac{1}{2}\right)$.
- **Additional merge candidates** – Furthermore, if the merge candidate list is still not fully populated, bi-predicted candidates can also be derived by combining two existing candidates from different reference picture lists. When the list is still not completely filled, zero motion candidates are included to complete the list. In both cases, the set weighting coefficients is set to $\boldsymbol{h} = \left(\frac{1}{2}, \frac{1}{2}\right)$.

It should be noticed that the HEVC merge technique can be also used to extend the concept of the direct modes from H.264/AVC standard[35]. Thus, this merge mode is evaluated for almost all PB partition patterns as an alternative to the Weighted Bi-SS prediction estimation. In this case, the index of the chosen merge candidate is encoded and transmitted along with the residual block.



(a) (b) (c)

(d) (e)

Figure 5 Example of a central view rendered from each LF test image: (a) *Vespa*, (b) *Ankylosaurus_&_Diplodocus_1*, (c) *Fountain_&_Vincent_2*, (d) *Color_Chart_1*, and (e) *ISO_Chart_12*.

Finally, the encoder will choose the best, among Weighted Bi-SS, SS-skip and conventional HEVC intra prediction, in an RDO manner.

4 PERFORMANCE ASSESSMENT

This section assesses the performance of the two weighted Bi-SS prediction schemes proposed in this paper, i.e., LFC Weighted Bi-SS (Fixed), and LFC Weighted Bi-SS (Adaptive) solutions. For this purpose, the test conditions are firstly introduced in Section 4.1 and, then, the obtained results for each proposed solution are presented, respectively, in Sections 4.2 and 4.3.

4.1 Test Conditions

The test conditions to experimentally evaluate the performance of the proposed LFC Weighted Bi-SS solutions can be summarized as follows:

- **Test Images** – Five light field images are considered to obtain representative RD results. These are[36] (see Figure 5): *Vespa*, *Ankylosaurus_&_Diplodocus_1*, *Fountain_&_Vincent_2*, *Color_Chart_1*, and *ISO_Chart_12*. These images were acquired using a Lytro Illum camera[36] and have resolution of 7728×5368 (with 15×15 MIs). The (raw) LF test images were converted to the Y'CbCr 4:2:0 color format before being encoded.
- **Codec Software Implementation** – The reference software of HEVC version 14.0[28] was used as the benchmark, as well as the base software for implementing the proposed LFC solutions with weighted Bi-SS prediction.
- **Search Range** – A search window with size w=128 as depicted in Figure 1b was adopted for all tested LF images.
- **Search Strategy** – The full search algorithm with the HEVC quarter-pixel accuracy was used.
- **Coding Configuration** – The results are presented using the Main Still Picture profile[29] and five QP values are considered, i.e., 22, 27, 32, 37, and 42.
- **RD Evaluation** – The objective quality evaluation was performed in terms of the luma PSNR of the raw LF image. The rate is presented in bits per pixel (bpp), which is calculated as the total number of bits needed for encoding all scalable layers divided by the number of pixels in the LF raw image. The results are shown using Bjøntegaard Delta (BD)[37] metric in terms of rate (referred to here as BD-Rate).

4.2 RD Performance for Fixed Weighting Coefficients

This section experimentally analyzes the RD performance of the LFC Weighted Bi-SS (Fixed) solution when different sets of weighting coefficients, h_0 and h_1 (see Section 3.1), are used for the jointly estimated Bi-SS prediction. For this, the RD performance achieved by each of the four LFC Weighted Bi-SS (Fixed) solutions presented in Section 3.1 are compared against the LFC Uni-SS[8] and the LFC Bi-SS[9] solutions (see Section 2).

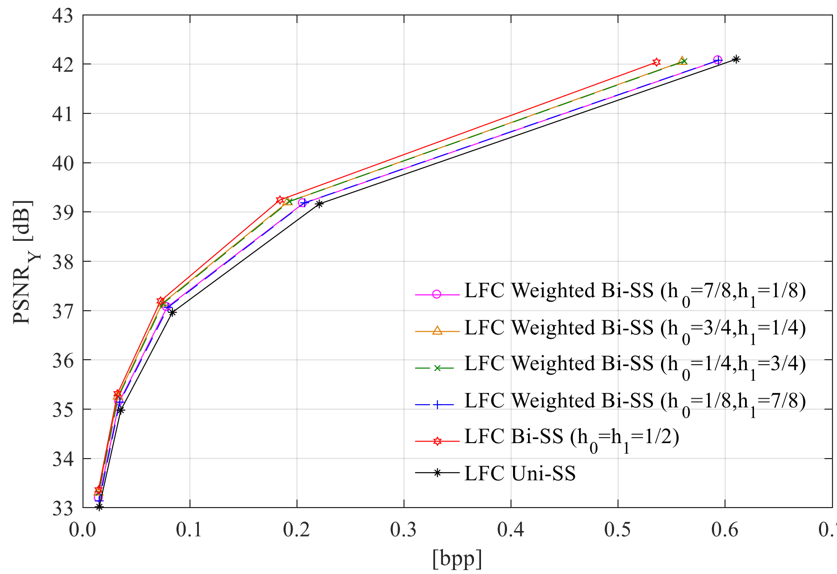


Figure 6 RD performance of the proposed LFC Weighted Bi-SS solution considering different sets of weighting coefficients that are fixed for coding the entire LF image. The results are illustrated for coding the LF test image *Vespa* using five QP values (i.e., 22, 27, 32, 37, and 42) and are compared to the author's previously proposed LFC Uni-SS[8] and LFC Bi-SS[9].

Table 1 Prediction statistics for coding the LF image *Vespa* using the LFC weights Bi-SS with fixed weighting coefficients compared to the previous LFC Bi-SS solution[9] (for QP value 22)

LFC Solution	**Prediction Mode Statistics**			**SS Prediction Statistics**		
	Intra	**SS**	**SS-skip**	**Uni-SS**	**Bi-SS**	**Merge**
Bi-SS ($h_0 = h_1 = 1/2$)	29.9%	65.6%	4.5%	14.0%	9.9%	76.1%
Weighted Bi-SS ($h_0 = 7/8, h_1 = 1/8$)	39.0%	57.8%	3.2%	28.2%	8.8%	63.1%
Weighted Bi-SS ($h_0 = 3/4, h_1 = 1/4$)	33.4%	62.5%	4.1%	14.7%	13.1%	72.1%
Weighted Bi-SS ($h_0 = 1/4, h_1 = 3/4$)	32.7%	63.2%	4.1%	16.2%	11.9%	71.9%
Weighted Bi-SS ($h_0 = 1/8, h_1 = 7/8$)	38.2%	58.5%	3.2%	28.0%	9.4%	62.6%

Since the results and the conclusions are consistent for all tested LF test images in Figure 5, and to avoid significantly increasing the size of this paper, Figure 6 illustrates RD performance only for the LF image *Vespa* (Figure 5a). From these results, it can be seen that the LFC Bi-SS solution using the average weighting coefficients always outperforms all the LFC Weighted Bi-SS (Fixed) solutions. Moreover, the more equally distributed the weighting coefficients are, the better the RD efficiency of the Bi-SS prediction is shown to be. In addition, in all cases, the LFC Weighted Bi-SS (Fixed) solutions always outperform the LFC Uni-SS solution.

Additionally, comparing the achieved RD coding performance with the prediction statistics observed when coding the LF image *Vespa*, as presented in Table 1, it can be seen that the worse RD performance achieved when a more unbalanced set of weighting coefficients is adopted (e.g., when $h_0 = \frac{7}{8}$ and $h_1 = \frac{1}{8}$.) is always associated to a further increase in the percentage of usage of the HEVC Intra prediction and the Uni-SS candidate predictor.

4.3 RD Performance for Adaptive Weighting Coefficients

This section experimentally analyzes the RD performance of the LFC Weighted Bi-SS (Adaptive) solution compared to the following benchmark solutions (as reviewed in Section 2): *i)* HEVC[28]; *ii)* HEVC SCC[30]; *iii)* LFC Uni-SS[8]; *iv)* LFC Restricted-SS[10]; and *v)* LFC Bi-SS[9]. The results are then shown in Table 2 in terms of the BD-Rate metric [37] with respect to these benchmark solutions.

Table 2 RD Performance for the proposed LFC Weighted Bi-SS (Adaptive) for each LF image in Figure 5 in terms of BD-Rate metric[37] with respect to each benchmark solution

LF Image	**BD-Rate w.r.t. HEVC[28]**	**BD-Rate w.r.t. HEVC SCC[30]**	**BD-Rate w.r.t. LFC Uni-SS[8]**	**BD-Rate w.r.t. LFC Restricted-SS[10]**	**BD-Rate w.r.t. LFC Bi-SS[9]**
(a)	-51.6 %	-41.2 %	-22.7 %	-14.9 %	-2.5 %
(b)	-83.4 %	-68.0 %	-37.5 %	-14.5 %	**-5.5 %**
(c)	-52.2 %	-31.1 %	-18.4 %	-12.1 %	-2.4 %
(d)	-67.6 %	-39.0 %	-18.4 %	-20.2 %	-4.0 %
(e)	-61.1 %	-44.1 %	-23.0 %	-18.0 %	-2.7 %
Average	**-63.2 %**	**-44.7 %**	**-24.0 %**	**-16.0 %**	**-3.4 %**

Table 3 RD Performance for each LF image in Figure 5, considering the performance bounds illustrated in Figure 4 for entropy coding the set of weighting coefficients, $\boldsymbol{h}$, chosen for each PB partition. Results are shown in terms of BD-Rate metric[37] with respect to each benchmark solution

LF Image	**BD-Rate w.r.t. HEVC[28]**	**BD-Rate w.r.t. HEVC SCC[30]**	**BD-Rate w.r.t. LFC Uni-SS[8]**	**BD-Rate w.r.t. LFC Restricted-SS[10]**	**BD-Rate w.r.t. LFC Bi-SS[9]**
(a)	-51.4 %	-40.9 %	-22.4 %	-14.6 %	-2.1 %
(b)	**-83.4 %**	**-67.8 %**	**-37.2 %**	-14.1 %	**-5.1 %**
(c)	-52.0 %	-30.9 %	-18.2 %	-11.9 %	-2.1 %
(d)	-67.5 %	-38.8 %	-18.2 %	-20.0 %	-3.8 %
(e)	-60.8 %	-43.8 %	-22.5 %	**-17.6 %**	-2.2 %
Average	**-63.0 %**	**-44.5 %**	**-23.7 %**	**-15.6 %**	**-3.0 %**

From these results, it can be seen that the proposed LFC Weighted Bi-SS (Adaptive) solution is able to further improve the RD performance for jointly estimated Bi-SS prediction leading to up to 5.5 % of bit savings when compared to the previously proposed LFC Bi-SS solution[9], without considering the extra bits for signaling the set of weighting coefficients, $\boldsymbol{h}$, for each PB partition.

To infer the impact of signaling these weighting coefficients, Table 3 presents the same kind of RD results of Table 2 adjusted by the entropy of these weighting coefficients as depicted in Figure 4. In this case, the proposed LFC Weighted Bi-SS (Adaptive) still presents significant bit savings of up to 5.1 % compared to the LFC Bi-SS solution. This shows that it is possible to extend the theoretical conclusions presented in the literature for motion compensated bi-prediction to LF coding with the jointly estimated Bi-SS prediction.

In addition, it is important to notice that, for all tested solutions shown in Tables 2 and 3, improved RD performance comes with the price of additional computational load. Generally, the coding solutions in Tables 2 and 3 can be ordered in terms of encoding complexity as: LFC Weighted Bi-SS (Adaptive) > LFC Bi-SS > LFC Restricted-SS > LFC Uni-SS > HEVC SCC > HEVC (intra coding). Therefore, for use cases in which the computation complexity is not an issue, the proposed LFC Weighted Bi-SS (Adaptive) is able to achieve significant bit savings of up to: 83.4 % with respect to HEVC; 67.8 % with respect to HEVC SCC; 37.2 % with respect to LFC Uni-SS; and 17.6 % with respect to LFC Restricted-SS.

To complement this analysis, Table 4 summarizes some statistics of relevant coding results when using the proposed LFC Weighted Bi-SS (Adaptive), such as: percentages of prediction mode usage, SS bi-prediction usage, and usage of each set of weighting coefficients $\boldsymbol{h} \in \mathcal{H}$. From these results, it can be seen that for all tested LF images, the average weighting coefficients are still the most used for Bi-SS prediction. However, comparing the results of the first row of this table (for the LF image *Vespa*) with the results presented in the first row of Table 1 for coding using the LFC Bi-SS solution, it can be observed that the better RD performance of the LFC Weighted Bi-SS (Adaptive) solution is directly related to a significant increase in the percentage of usage of the Bi-SS prediction candidate predictor – i.e., from 9.9 % when using the LFC Bi-SS in Table 1, to 20.6 % when using the LFC Weighted Bi-SS (Adaptive) – and a consequent decrease in the Uni-SS candidate predictor usage.

Table 4 Prediction statistics for coding each LF image in Figure 5 using the proposed LFC Weighted Bi-SS (Adaptive) solution (for QP value 22)

LF Image	Prediction Mode Statistics			SS Prediction Statistics			Weighted Bi-SS Prediction Statistics				
	Intra	**SS**	**SS-skip**	**Uni-SS**	**Bi-SS**	**Merge**	$\boldsymbol{h}=\left(\frac{7}{8},\frac{1}{8}\right)$	$\boldsymbol{h}=\left(\frac{3}{4},\frac{1}{4}\right)$	$\boldsymbol{h}=\left(\frac{1}{2},\frac{1}{2}\right)$	$\boldsymbol{h}=\left(\frac{1}{4},\frac{3}{4}\right)$	$\boldsymbol{h}=\left(\frac{1}{8},\frac{7}{8}\right)$
(a)	25.5%	69.9%	4.6%	7.8%	20.6%	71.5%	3.2%	21.2%	69.9%	4.8%	0.9%
(b)	12.0%	87.3%	0.7%	3.2%	23.6%	73.3%	1.3%	15.4%	76.8%	6.1%	0.5%
(c)	36.8%	61.4%	1.8%	5.5%	17.1%	77.5%	4.6%	21.2%	69.6%	4.2%	0.4%
(d)	17.0%	78.1%	4.9%	4.3%	21.7%	74.1%	2.3%	17.2%	74.0%	5.8%	0.7%
(e)	16.5%	80.9%	2.7%	6.0%	29.9%	64.2%	2.8%	20.6%	69.3%	6.5%	0.8%

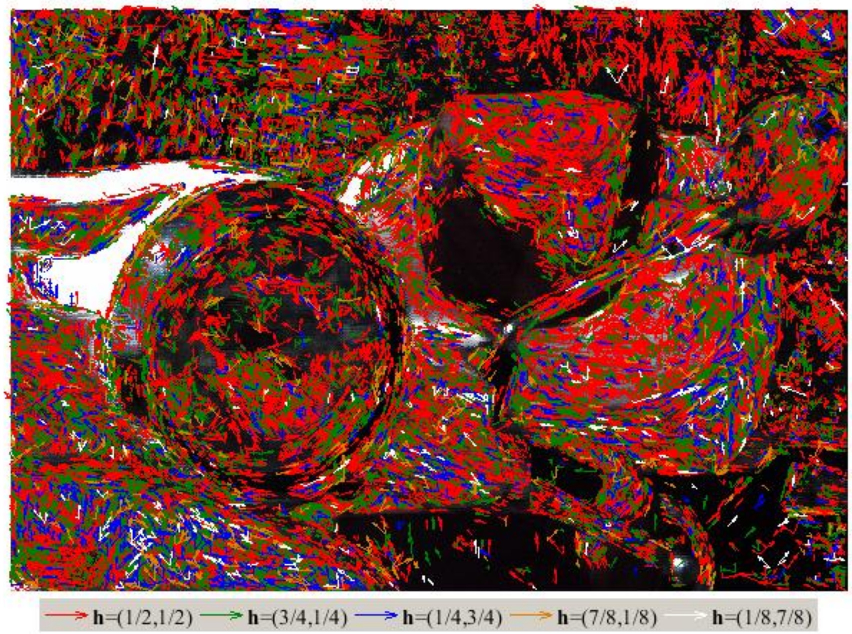


Figure 7 Example of the weighted Bi-SS vectors distribution along the (raw) LF image when coding the LF test image *Vespa*.

Furthermore, comparing the percentage of usage of the set of weighting coefficients $\boldsymbol{h} \in \mathcal{H}$ in Table 4 with their distribution along the raw LF image, as illustrated in Figure 7 for the LF image *Vespa*, it can be observed that although the percentage of usage of each $\boldsymbol{h} \in \mathcal{H}$ is not uniformly distributed, these different weights seem to be equally distributed along the encoded raw LF image.

5 FINAL REMARKS

This paper presented a comprehensive study of the RD performance for HEVC-based LF image coding when using different sets of weighting coefficients for Bi-SS prediction. For this, two Weighted Bi-SS prediction schemes were proposed and analyzed, namely: *i)* using a fixed set of weighting coefficients that are different from the averaging coefficients previously adopted for Bi-SS prediction; *ii)* using an adaptive algorithm for estimating the optimal set of weighting coefficients for each PB partition. Regarding the first solution, experimental results suggests that the more equally distributed the weighting coefficients are, the better the RD efficiency for Bi-SS solution is shown to be, showing that the averaging coefficients is the one that leads to the better RD coding performance. Regarding the second Weighted Bi-SS solution, it was observed that further RD performance improvements can be achieved when adaptively choosing the set of weighting coefficients for each PB partition, leading to up to 5% of bit savings compared to the previous Bi-SS solution. Moreover, statistical analysis of the coding process show that this improvement is mainly due to an increase in Bi-SS prediction efficiency, significantly increasing its percentage of usage for LF image coding. Therefore, this paper has extended the previous theoretical assumptions showing that the adaptive weighting coefficient selection is advantageous for improving the performance for jointly estimated Bi-SS prediction.

Finally, future work will include to propose a better iterative algorithm for estimating the best weighting coefficients and to design an efficient entropy coding scheme for signaling the chosen weighting coefficients for each PB partition.

6 ACKNOWLEDGEMENT

The authors acknowledge the support of FCT (*Fundação para a Ciência e a Tecnologia*, Portugal), under the project UID/EEA/50008/2013.

REFERENCES

[1] Aggoun, A., Tsekleves, E., Swash, M. R., Zarpalas, D., Dimou, A., Daras, P., Nunes, P., Soares, L. D., "Immersive 3D Holoscopic Video System," IEEE Multimed. **20**(1), 28–37 (2013).

[2] Arai, J., "Integral Three-Dimensional Television (FTV Seminar)," ISO/IEC JTC1/SC29/WG11 M34199, Sapporo, Japan (2014).

[3] Raghavendra, R., Raja, K. B.., Busch, C., "Presentation Attack Detection for Face Recognition Using Light Field Camera," IEEE Trans. Image Process. **24**(3), 1060–1075 (2015).

[4] Xiao, X., Javidi, B., Martinez-Corral, M., Stern, A., "Advances in Three-Dimensional Integral Imaging: Sensing, Display, and Applications [Invited]," Appl. Opt. **52**(4), 546–560 (2013).

[5] "JPEG Pleno Call for Proposals on Light Field Coding.", ISO/IEC JTC 1/ SC29/WG1 N74014, Geneva, Switzerland (2017).

[6] Wegner, K., Lafruit, G., eds., "Call for Immersive Visual Test Material," ISO/IEC JTC1/SC29/WG11 N16766, Hobart, Australia (2017).

[7] Conti, C., Nunes, P., Soares, L. D., "New HEVC Prediction Modes for 3D Holoscopic Video Coding," 2012 19th IEEE Int. Conf. Image Process., 1325–1328, Orlando, FL, US (2012).

[8] Conti, C., Soares, L. D., Nunes, P., "HEVC-Based 3D Holoscopic Video Coding using Self-Similarity Compensated Prediction," Signal Process. Image Commun. **42**, 59–78 (2016).

[9] Conti, C., Nunes, P., Soares, L. D., "HEVC-Based Light Field Image Coding with Bi-Predicted Self-Similarity Compensation," 2016 IEEE Int. Conf. Multimed. Expo Work., 1–4, Seattle, WA, US (2016).

[10] Li, Y., Sjostrom, M., Olsson, R., Jennehag, U., "Coding of Focused Plenoptic Contents by Displacement Intra Prediction," IEEE Trans. Circuits Syst. Video Technol. **26**(7), 1308–1319 (2016).

[11] Lucas, L. F. R., Conti, C., Nunes, P., Soares, L. D., Rodrigues, N. M. M., Pagliari, C. L., da Silva, E. A. B., de Faria, S. M. M., "Locally Linear Embedding-Based Prediction for 3D Holoscopic Image Coding using HEVC," 2014 Proc. 22nd Eur. Signal Process. Conf., 11–15, Lisbon, Portugal (2014).

[12] Liu, D., An, P., Ma, R., Yang, C., Shen, L., "3D Holoscopic Image Coding Scheme Using HEVC with Gaussian Process Regression," Signal Process. Image Commun. **47**, 438–451 (2016).

[13] Zaharia, R., Aggoun, A., McCormick, M., "Adaptive 3D-DCT Compression Algorithm for Continuous Parallax 3D Integral Imaging," Signal Process. Image Commun. **17**(3), 231–242 (2002).

[14] Aggoun, A., "Compression of 3D Integral Images Using 3D Wavelet Transform," J. Disp. Technol. **7**(11), 586–592 (2011).

[15] Shi, S., Gioia, P., Madec, G., "Efficient Compression Method for Integral Images using Multi-View Video Coding," 2011 18th IEEE Int. Conf. Image Process., 137–140, Brussels, Belgium (2011).

[16] Dick, J., Almeida, H., Soares, L. D., Nunes, P., "3D Holoscopic Video Coding Using MVC," 2011 IEEE EUROCON - Int. Conf. Comput. as a Tool, 1–4, Lisbon, Portugal (2011).

[17] Olsson, R., "Empirical Rate-Distortion Analysis of JPEG 2000 3D and H. 264/AVC Coded Integral Imaging Based 3D-Images," 2008 3DTV Conf. True Vis. - Capture, Transm. Disp. 3D Video, 113–116, Istanbul, Turkey (2008).

[18] Vieira, A., Duarte, H., Perra, C., Tavora, L., Assuncao, P., "Data Formats for High Efficiency Coding of Lytro-Illum Light Fields," 2015 Int. Conf. Image Process. Theory, Tools Appl., 494–497, Orleans, France (2015).

[19] Perra, C., Assuncao, P., "High Efficiency Coding of Light Field Images based on Tiling and Pseudo-Temporal Data Arrangement," 2016 IEEE Int. Conf. Multimed. Expo Work., 1–4, Seattle, WA, US (2016).

[20] Liu, D., Wang, L., Li, L., Zhiwei Xiong., Feng Wu., Wenjun Zeng., "Pseudo-Sequence-Based Light Field Image Compression," 2016 IEEE Int. Conf. Multimed. Expo Work., 1–4, Seattle, WA, US (2016).

[21] Choudhury, C., Chaudhuri, S., "Disparity Based Compression Technique for Focused Plenoptic Images," Proc. 2014 Indian Conf. Comput. Vis. Graph. Image Process. - ICVGIP '14, 1–6, Bangalore, India (2014).
[22] Graziosi, D. B., Alpaslan, Z. Y., El-Ghoroury, H. S., "Depth Assisted Compression of Full Parallax Light Fields," Proc. SPIE 9391, Stereosc. Displays Appl. XXVI, 93910Y, San Francisco, CA, US (2015).
[23] Li, Y., Sjöström, M., Olsson, R., Jennehag, U., "Scalable Coding of Plenoptic Images by Using a Sparse Set and Disparities.," IEEE Trans. Image Process. **25**(1), 80–91 (2016).
[24] Girod, B., "Efficiency Analysis of Multihypothesis Motion-Compensated Prediction for Video Coding.," IEEE Trans. Image Process. **9**(2), 173–183 (2000).
[25] Ugur, K., Lainema, J., "Updated Results on HEVC Still Picture Coding Performance," JCTVC-M0041, Incheon, South Korea (2013).
[26] Hanhart, P., Řeřábek, M., Korshunov, P., Ebrahimi, T., "Subjective Evaluation of HEVC Intra Coding for Still Image Compression," JCTVC-L0380, Geneva, Switzerland (2013).
[27] Sullivan, G. J., Wiegand, T., "Rate-Distortion Optimization for Video Compression," IEEE Signal Process. Mag. **15**(6), 74–90 (1998).
[28] Il-Koo Kim., McCann, K., Sugimoto, K., Bross, B., Han, W.-J.., Sullivan, G., "High Efficiency Video Coding (HEVC) Test Model 14 (HM14) Encoder Description," JCTVC-P1002, San José, CA, US (2014).
[29] Sullivan, G. J., Ohm, J.-R., Han, W.-J., Wiegand, T., "Overview of the High Efficiency Video Coding (HEVC) Standard," IEEE Trans. Circuits Syst. Video Technol. **22**(12), 1649–1668 (2012).
[30] Kwon, D.-K., Budagavi, M., "Fast Intra Block Copy (IntraBC) Search for HEVC Screen Content Coding," 2014 IEEE Int. Symp. Circuits Syst., 9–12, Melbourne, Australia (2014).
[31] Flierl, M., Girod, B., Video Coding with Superimposed Motion-Compensated Signals - Applications to H.264 and Beyond, 1sted., Springer US (2004).
[32] Flierl, M., Wiegand, T., Girod, B., "A Locally Optimal Design Algorithm for Block-Based Multi-Hypothesis Motion-Compensated Prediction," Proc. DCC '98 Data Compression Conf., 239–248, Snowbird, UT, US (1998).
[33] Sze, V., Budagavi, M., Sullivan, G. J., eds., High Efficiency Video Coding (HEVC): Algorithms and Architectures, Springer International Publishing, Cham (2014).
[34] Wien, M., High Efficiency Video Coding, Springer Berlin Heidelberg, Berlin, Heidelberg (2015).
[35] Wiegand, T., Sullivan, G. J., Bjøntegaard, G., Luthra, A., "Overview of the H.264/AVC Video Coding Standard," IEEE Trans. Circuits Syst. Video Technol. **13**(7), 560–576 (2003).
[36] Řeřábek, M., Ebrahimi, T., "New Light Field Image Dataset," 8th Int. Conf. Qual. Multimed. Exp., Lisbon, Portugal (2016).
[37] Bjøntegaard, G., "Calculation of Average PSNR Differences between RD Curves," VCEG-M33, Austin, TX, US (2001).