

Autofocus on Depth of Interest for 3D Image Coding

Khouloud Samrouth, Olivier Déforges, Yi Liu, Mohamad Khalil, Wassim El Falou

▶ **To cite this version:**

Khouloud Samrouth, Olivier Déforges, Yi Liu, Mohamad Khalil, Wassim El Falou. Autofocus on Depth of Interest for 3D Image Coding. Journal of Electrical and Computer Engineering, 2017, 2017, pp.9689715. 10.1155/2017/9689715 . hal-01502344

HAL Id: hal-01502344

https://univ-rennes.hal.science/hal-01502344

Submitted on 20 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Hindawi
Journal of Electrical and Computer Engineering
Volume 2017, Article ID 9689715, 15 pages
https://doi.org/10.1155/2017/9689715

Research Article

Autofocus on Depth of Interest for 3D Image Coding

Khouloud Samrouth,[1] Olivier Deforges,[1] Yi Liu,[1] Mohamad Khalil,[2] and Wassim EL Falou[2]

[1] *UEB, CNRS UMR 6164, IETR Lab, INSA de Rennes 20, Avenue des Buttes de Coesmes, CS 70839, 35708 Rennes, France*
[2] *Faculty of Engineering I, Lebanese University, Tripoli, Lebanon*

Correspondence should be addressed to Khouloud Samrouth; khouloud.samrout@gmail.com

Received 13 October 2016; Accepted 30 January 2017; Published 22 February 2017

Academic Editor: Panajotis Agathoklis

Copyright © 2017 Khouloud Samrouth et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For some 3D applications, one may want to focus on a specific depth zone representing a region of interest in the scene. In this context, we introduce a new functionality called "autofocus" for 3D image coding, exploiting the depth map as an additional semantic information provided by the 3D sequence. The method is based on a joint "Depth of Interest" (DoI) extraction and coding scheme. First, the DoI extraction scheme consists of a precise extraction of objects located within a DoI zone, given by the viewer or deduced from an analysis process. Then, the DoI coding scheme provides a higher quality for the objects in the DoI at the expense of other depth zones. The local quality enhancement supports both higher SNR and finer resolution. The proposed scheme embeds the Locally Adaptive Resolution (LAR) codec, initially designed for 2D images. The proposed DoI scheme is developed without modifying the global coder framework, and the DoI mask is not transmitted, but it is deduced at the decoder. Results showed that our proposed joint DoI extraction and coding scheme provide a high correlation between texture objects and depth. This consistency avoids the distortion along objects contours in depth maps and those of texture images and synthesized views.

1. Introduction

Recent studies in the 3D technology led to a growing development in 3D applications. Nowadays, next generations of highly advanced multimedia video systems, such as 3D television (3DTV), 3D cinemas, and free viewpoint television (FTV), provide depth perception for the viewers [1, 2] on both large TV screens [3] and mobile phones [4] and allow free navigation within a real-world scene [5, 6]. Such autostereoscopic systems, providing 3D viewing experience without special glasses or other head gear [7, 8], rely on depth maps to give the 3D impression and to synthesize intermediate views at an arbitrary view point [9, 10]. These depth maps can be considered as an additional semantic information on the scene, which can be exploited in a region of interest (RoI) context. Indeed, among advanced functionalities for image coding, such as scalability, lossy/lossless and security, RoI is useful for many applications [11]. In RoI-based image coding, some regions, that are of interest to the viewer, are encoded with a higher fidelity than the rest of the image. There are several 2D applications using the RoI coding such as compression of infrared or digital medical images [12, 13], segmentation [14], and accurate objects location [15]. However, for some 3D applications, the areas of interest can be partially or fully dependent on the depth information [16]. In this context, new research works have been devoted to the application of the RoI feature in the 3D domain [16–20]. In [17], Fan et al. proposed a digital watermarking algorithm for copyright protection of 3D images. The proposed watermarking scheme was based on Depth Perceptual Region of Interest (DP-RoI) [18] from the idea that the Human Visual System (HVS) is more sensitive to the distortion in foreground than background regions in 2D image. The DP-RoI extraction and tracking algorithms were proposed in [18] where Zhang et al. exploited the depth information beside the large contrast of illumination in the texture image to extract the RoI. In [19], Karlsson and Sjöström determined the foreground by using the information in the depth map. The foreground was defined as the RoI. It was used to reallocate bits from the background to the RoI by applying a spatiotemporal filter [21] on the texture image. This method achieved an improvement in the perceived quality of foreground at the expense of a quality reduction in background. In the method proposed by Karlsson, the RoI coding has only been used

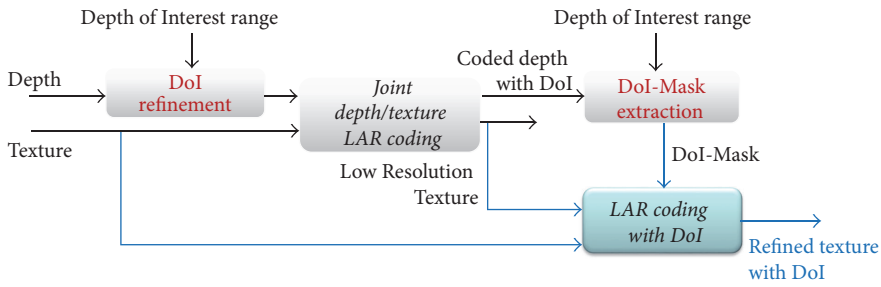


Figure 1: Block diagram of the proposed DoI scheme.

Table 1: Functional comparison for RoI coding.

Codec	RoI resolution unit	Different quality for background/foreground	RoI mask transmission
JPEG2K	Pixel	No	No
JPEGXR	Block	Yes (except lossless)	Yes
H.264	Block	Yes (except lossless)	Yes
HEVC	Block	Yes (except lossless)	Yes
LAR	Pixel	Yes	No

for the ordinary monoscopic texture video, part of the 2D plus depth representation, without the depth map. Zheng et al. [20] proposed a scheme for RoI coding of 3D meshes inspired by JPEG2000 [22]. In [16], Chamaret et al. adaptively adjusted the shift between the two views of a stereoscopic system in order to have a null disparity on the RoI area. This enhanced the visual quality of a given area in the rendered view scene. However, Chamaret et al.'s method was applied on stereoscopic and not on autostereoscopic systems.

Many different contributions have been proposed for the RoI extraction and coding schemes in the 3D domain, but only a few combine both extraction and coding. From a representation and coding point of view, three important criteria must be considered for the RoI coding system: (1) RoI spatial resolution, (2) the RoI-mask cost, and (3) the choice of quality levels between background/foreground. Most of the state-of-the-art codecs apply a block-based RoI coding scheme, where the coding unit is a block of a minimum 16 × 16 pixels, such as JPEGXR [23], H.264 [24], and HEVC [25]. When the RoI spatial resolution is block-based, the RoI-mask cost is reduced. However, it decreases the accuracy along the RoI contours. On the other hand, other state-of-the-art codecs, such as JPEG2000 [26, 27], apply RoI pixel-based scheme but impose the difference of quality between background and foreground. Table 1 gives a functional comparison of these different codecs. Indeed, for 3D applications, the resolution of objects contours in the depth map has to be very fine, in particular when rendering intermediate views. The reason is that 3D rendering software, such as VSRS [28], relies on depth maps to synthesize intermediate views. Several researches [29–32] highlight the influence of depth image compression and its implications on the quality of video plus depth virtual view rendering.

In this paper, we propose a so-called "Depth of Interest" (DoI) scheme for joint representation and coding of the region of interest in the 2D + Z autostereoscopic domain. The related application consists of an automatic focus on a given depth range for a better rendering in this area of interest. It offers a fine extraction of objects with well-defined contours in the depth map, inducing a low distortion among area of interest in the synthesized views. Contrary to the state-of-the-art methods, such a scheme better fits with the three criteria of a RoI coding scheme: (1) it relies on a pixel-based resolution, (2) the RoI-mask transmission is almost free as only the two values limiting the Depth of Interest range are transmitted, and (3) it allows freely choosing different qualities for the different regions in the scene (cf. Table 1). The proposed DoI coding scheme was tested on a large set of MPEG 3D reference sequences (*Balloons*, *Kendo*, *BookArrival*, *Newspaper*, *UndoDancer*, and *GTFly*).

This paper is organised as follows. In Section 2, the global coding scheme is presented. In Section 3, the 2D LAR coder framework and the 2D + Z coding scheme are described. In Section 4, the proposed DoI representation and coding scheme is explained. Section 5 provides the experimental results for DoI extraction, coding performance, and subjective quality tests for texture images and synthesized intermediate views. Finally, we conclude this paper in Section 6.

2. Global Coding Scheme

The proposed scheme embeds the Locally Adaptive Resolution (LAR) codec [33]. LAR is a global coding framework based on a content-based QuadTree partitioning. It provides a lot of functionalities such as *a unique codec for lossy and lossless compression*, resolution and quality scalability, partial/full encryption, 2D region of interest coding, rate control, and rate-distortion optimization [34]. Extensions to scalable 2D + Z coding have been presented in [35].

For 2D + Z applications, it can be assumed that the object of interest is located in a specific depth zone. This depth range which is provided as an input of our global coding scheme defines the Depth of Interest (DoI). First, the DoI representation scheme consists in defining, from the depth map, a binary mask covering the DoI zone. It leads to a fine extraction of objects located within a DoI zone. Then, the DoI coding scheme aims at ensuring a higher quality at DoI zone at the expense of other depth zones. Higher visual quality in the DoI is ensured for both depth and texture images (cf. Figure 1).

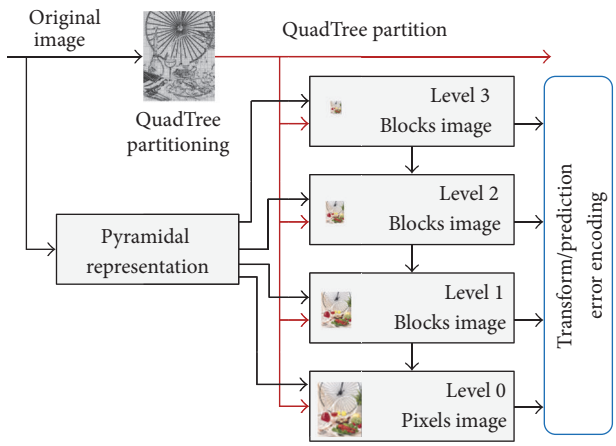


Figure 2: LAR multiresolution coder scheme.

For this purpose, firstly, a DoI refinement is applied on the depth map (Section 4.1). The QuadTree partition of the depth map is adapted to get a higher resolution at the DoI layer, and thus well-defined contours along the objects of interest are obtained. Secondly, a DoI coding is applied on the texture image (Section 4.2). The RoI-based scheme using the defined depth-based binary mask allows setting different qualities inside and outside the DoI. The benefits of the proposed DoI representation scheme are as follows: (1) it is content-based; (2) the DoI coding scheme is inserted as a preprocess, and thus we prevent modifying the coder which reduces the complexity of the coder; (3) the DoI-Mask is not transmitted to the decoder, and thus there is no over cost coding.

3. LAR Coder Framework

3.1. 2D LAR Coder. Locally Adaptive Resolution (LAR) [33] is an efficient multiresolution content-based 2D image coder for both lossless and lossy image compressions (cf. Figure 2).

3.1.1. QuadTree Partitioning. The LAR coder relies on a local analysis of image activity. It starts with the estimation of a QuadTree representation, conditioning a variable-size block partitioning ($1 \times 1, 2 \times 2, \ldots, 64 \times 64, 128 \times 128$ pixels). The block size estimation is performed by comparing the difference of the local maximum and minimum values within the block with a given threshold $\mathrm{Th}_{\mathrm{Quad}}$. Therefore, the local resolution defined by the pixel size depends on the local activity: the smallest blocks are located upon edges, whereas large blocks map homogeneous areas (cf. Figure 3). Then, the main feature of the LAR coder consists in preserving contours while smoothing homogeneous parts of the image. This QuadTree partition is the key structure of the LAR coder.

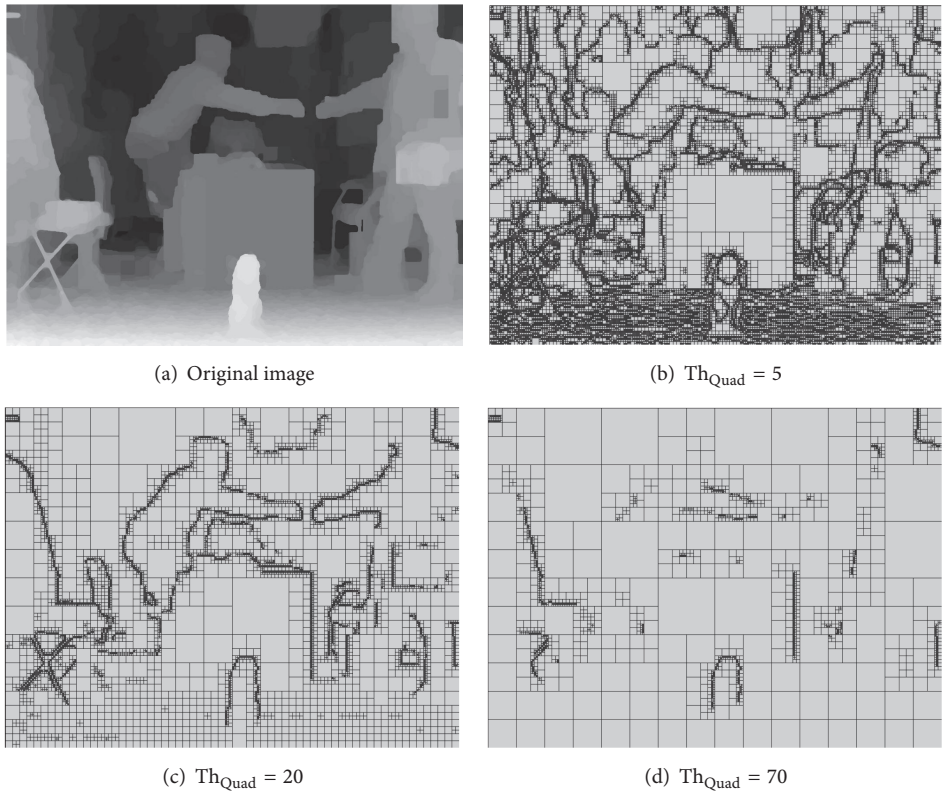
3.1.2. Transform and Prediction. The coder uses a pyramidal representation of the image. Starting from the upper level of the pyramid (lower resolution/larger blocks), a dyadic decomposition conditioned by the QuadTree is performed. Each level of the decomposition integrates a 2×2 block transform (interleaved S + P transform) stage and a prediction one. The resulting prediction error is quantized using an input quantization factor Q_p and coded using an adapted entropy coder. The quantization value for each level is then given by $Q_l = Q_p/2^l$, with l being the pyramidal level (full resolution for $l = 0$). It has been found in [34] that the near optimal choice for the pair $(Q_p, \mathrm{Th}_{\mathrm{Quad}})$ is to set as $\mathrm{Th}_{\mathrm{Quad}} = (2/3)Q_p$. Thus Q_p is the only parameter of the codec, and $Q_p = 1$ induces a lossless compression.

3.2. Global 2D + Z LAR Coding Scheme. The proposed DoI coding scheme integrates a global 2D + Z coding scheme. The global scalable 2D + Z coding scheme has been proposed in [35] (cf. global scheme in Figure 4). In a first step, a QuadTree partition is built considering only the Z image, (Quad_Z). Based on Quad_Z, the 2D image, represented in a (Y, C_b, C_r) colour space, is encoded first at low bit rate and low resolution. Then, the Z image is encoded with an improved prediction stage using the Y component of the previously encoded 2D image. In a second step, the QuadTree partition is refined considering both 2D and Z images (Quad_{ZT}). Finally, based on Quad_{ZT}, the quality of the 2D image is improved by a new coding process at a finer resolution. This joint coding method preserves the spatial coherence between texture and depth images and reduces visual artefacts for synthesized views, especially upon edges. In terms of complexity, this coding scheme has a reduced one compared to JPEG2K and about the same one compared to JPEGXR. For objective efficiency compression point of view, the codec is close to JPEG2K and significantly outperforms JPEGXR (cf. Figure 5).

4. Proposed Depth of Interest (DoI) Representation and Coding Scheme

In this section, we present the proposed joint (Depth of Interest) DoI representation and coding scheme. Firstly, a DoI refinement of depth is presented in Section 4.1. Secondly, a DoI coding of texture is presented in Section 4.2.

4.1. DoI Refinement of Depth. As previously mentioned, quality of compressed depth map will be strongly linked to



(a) Original image (b) $Th_{Quad} = 5$

(c) $Th_{Quad} = 20$ (d) $Th_{Quad} = 70$

Figure 3: Examples of QuadTree partition of BookArrival view 10 frame 33 for different thresholds Th_{Quad}.

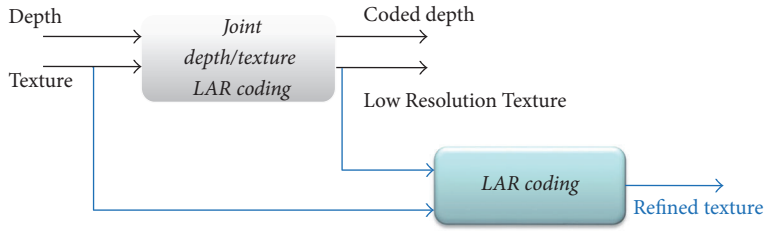


Figure 4: Global 2D + Z coding scheme.

the distortion introduced on contours. Thus, a local quality enhancement of depth inside a DoI should firstly aim at refining the spatial resolution. As the LAR initially considers a fix threshold to build the QuadTree, a possible solution would be to design a new QuadTree algorithm with an adaptive threshold. We opted for another solution preserving the QuadTree estimation stage but introducing a depth map preprocessing. This stage performs a dynamic range adjustment on the depth map so that the Depth of Interest (DoI) gets the largest dynamic range at the expense of the other depth zones (cf. Figure 1).

We introduce the following notations:

(i) In_Depth: input depth map.

(ii) Out_Depth: output depth map after dynamic range adjustment.

(iii) Z_{in_l}, Z_{in_h}: input low and high limits of Depth of Interest range.

(iv) W_{in}: input Depth of Interest window size, with $W_{in} = Z_{in_h} - Z_{in_l}$.

(v) W_{out}: adjusted Depth of Interest window size, with $W_{out} = F \times W_{in}$.

The input depth range $[Z_{in_l}, Z_{in_h}]$ and F coefficient can be considered as approximations of the focal distance and F-number (or relative aperture) in optical system, respectively. In particular, the F parameter will control the degree of image sharpness for the DoI. These parameters will be set by the user.

The proposed piecewise adjustment function is given in Figure 6 and the corresponding algorithm in Procedure 1. We set the constraint that the value of the middle point of the input depth range has to be unchanged.

The Out_Depth image is then used for the QuadTree estimation, whereas the input depth is encoded with the same quantization parameter for the whole image.

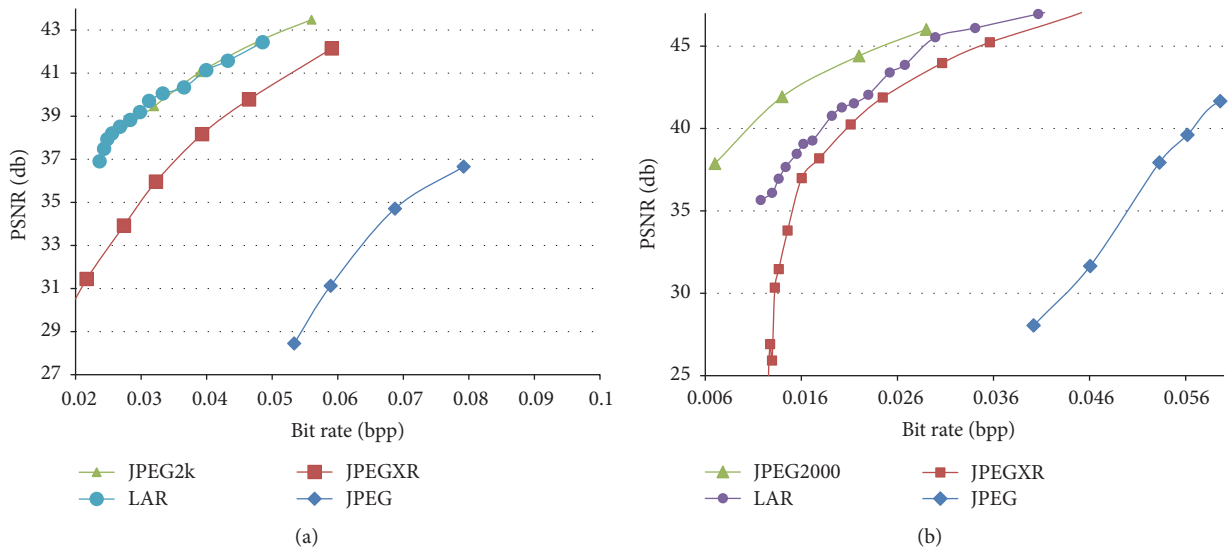


Figure 5: Rate-distortion curves of depth images for (a) GTFly frame 157 view 1 and (b) UndoDancer frame 250 view 1.

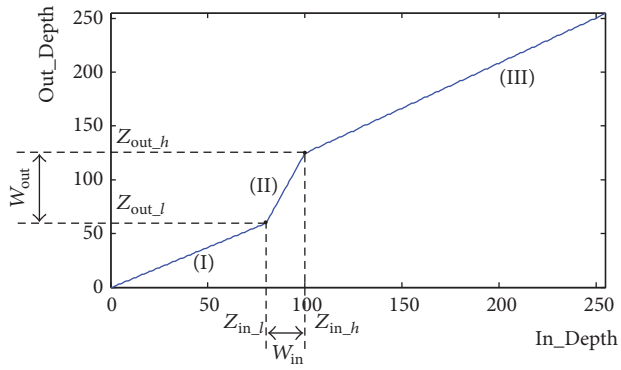


Figure 6: Output dynamic range adjustment of depth values as a function of input values.

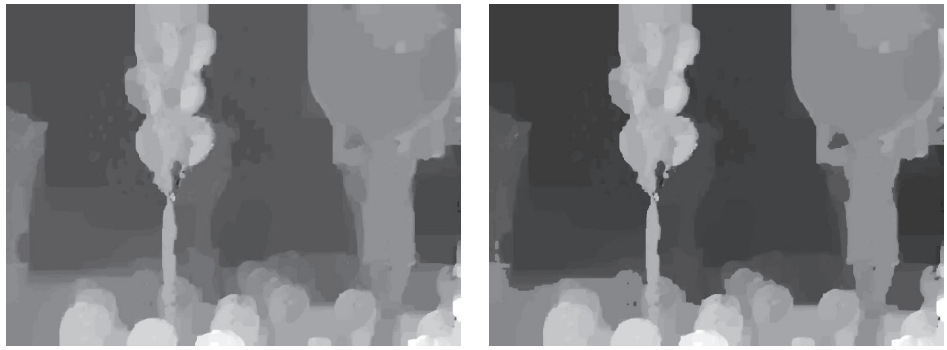


Figure 7: Original input binary depth mask of Balloons view 5 frame 1 (1024 × 768 pixels) for foreground zone with $Z_{in_l} = 128$, $Z_{in_h} = 255$.

4.1.1. QuadTree Results after Depth Map Preprocessing. Figures 7 and 10 represent the DoI zone within Z_{in_l} and Z_{in_h} in the original input depth map (original input binary depth mask). Figures 8(a) and 8(b) present the original and the adjusted depth maps, respectively. Indeed, after the dynamic range adjustment of the depth map, the DoI is mapped by smaller blocks. Thus, it leads to an increase of the local resolution inside the DoI area in the QuadTree (cf. Figures 9 and 11).

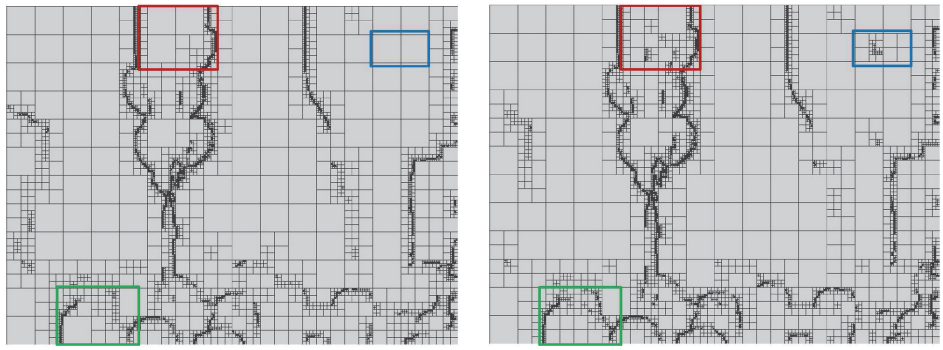
4.1.2. Objective and Visual Results of Decoded Depth Maps in DoI Context. In this section, we compare the depth map coding between the original 2D + Z LAR solution and the proposed one including the adjustment stage. For the same bit rate, the previous preprocessing step decreases the global objective quality of the depth image in comparison with the classic coding. However, it increases the local quality within the DoI with a gain up to 11 dB for the given configuration as the local resolution in the DoI is increased, (cf. Figures 12 and 14). Furthermore, we can notice a significant visual quality improvement in the DoI (cf. Figures 13 and 15). More precisely, it is obviously visible that edges of objects within DoI are more accurately encoded in comparison with the classic coding scheme.

4.2. DoI Coding of Texture. The "DoI" coding of texture means that the texture image has to be compressed with unequal quality according to the given DoI (cf. Figure 1).



(a) (b)

Figure 8: Depth map of Balloons view 5 frame 1: (a) original and (b) after dynamic range adjustment of foreground zone with $Z_{\text{in_}l} = 128$, $Z_{\text{in_}h} = 255$, $F = 1.3$.



(a) (b)

Figure 9: QuadTree results of depth map Balloons view 5 frame 1: (a) from original map with $\text{Th}_{\text{Quad}} = 28$ and (b) after dynamic range adjustment of foreground zone with $Z_{\text{in_}l} = 128$, $Z_{\text{in_}h} = 255$, $F = 1.3$, and $\text{Th}_{\text{Quad}} = 29$.

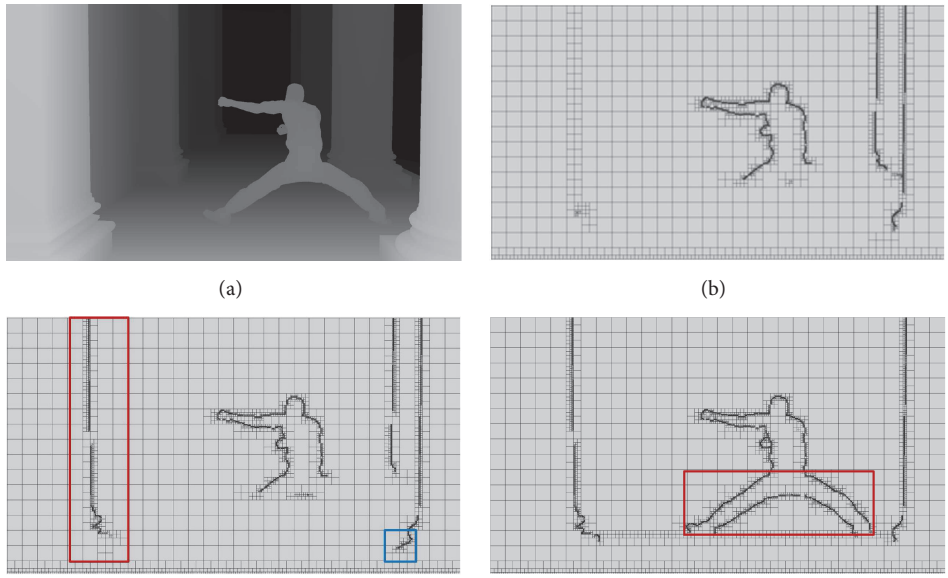


(a) (b)

Figure 10: Original input binary depth mask of UndoDancer view 1 frame 250 (1920 × 1080 pixels) (a) for foreground zone with $Z_{\text{in_}l} = 128$, $Z_{\text{in_}h} = 255$ and (b) for depth zone within $Z_{\text{in_}l} = 100$, $Z_{\text{in_}h} = 121$.

The first step is the extraction of a binary mask image (DoI-Mask) that will be further used to define the region of interest, considering that only the depth range $(Z_{\text{in_}l}, Z_{\text{in_}h})$ is transmitted to the decoder. This extraction is simply performed by a binarization step of the coded depth image, with the depth range as input parameter. The considered depth image is the coded one for two main reasons. The first one is that the process has to be duplicated at the decoder side. The second reason is that the binary mask has to be seen as a subset of the QuadTree partition.

The second step is the quality enhancement for the DoI for the texture. As mentioned in Section 3.2, the texture image is encoded in two passes: first, the 2D + Z images are compressed at low bit rate based on Quad_Z only, and then the 2D image is enhanced considering Quad_{ZT}. In this coding scheme, the same quantization factor Q_p is applied at both passes.



(a) (b) (c) (d)

Figure 11: (a) Original Depth map of UndoDancer view 1 frame 250, (b) QuadTree from original depth map with $\mathrm{Th}_{\mathrm{Quad}} = 47$, (c) QuadTree from depth map after dynamic range adjustment of foreground zone with $Z_{\mathrm{in_}l} = 128$, $Z_{\mathrm{in_}h} = 255$, $F = 1.3$, and $\mathrm{Th}_{\mathrm{Quad}} = 60$, and (d) QuadTree from depth map after dynamic range adjustment of foreground zone with $Z_{in_l} = 128$, $Z_{\mathrm{in_}h} = 255$, $F = 7.0$, and $\mathrm{Th}_{\mathrm{Quad}} = 71$.

Input: In_Depth, $Z_{\mathrm{in_}l}, Z_{\mathrm{in_}h}, F$
Output: Out_Depth
$\delta W = W_{\mathrm{out}} - W_{\mathrm{in}} = (F-1) \times W_{\mathrm{in}}$
$Z_{\mathrm{out_}l} = Z_{\mathrm{in_}l} - \dfrac{\delta W}{2}$
$Z_{\mathrm{out_}h} = Z_{\mathrm{in_}h} + \dfrac{\delta W}{2}$
for all i **do**
if In_Depth(i) $\le Z_{\mathrm{in_}l}$ **then**
$b = \dfrac{Z_{\mathrm{out_}l}}{Z_{\mathrm{in_}l}}$
(I) Out_Depth(i) = $b \times$ In_Depth(i).
end if
if $Z_{\mathrm{in_}l} <$ In_Depth(i) $< Z_{\mathrm{in_}h}$ **then**
$b = \dfrac{Z_{\mathrm{out_}h} - Z_{\mathrm{out_}l}}{Z_{\mathrm{in_}h} - Z_{\mathrm{in_}l}}$
$c = Z_{\mathrm{out_}l} - b \times Z_{\mathrm{in_}l}$
(II) Out_Depth(i) = $b \times$ In_Depth(i) + c.
end if
if In_Depth(i) $> Z_{\mathrm{in_}h}$ **then**
$b = \dfrac{Z_{\mathrm{out_}h} - 255}{Z_{\mathrm{in_}h} - 255}$
$c = 255 \times (1 - b)$
(III) Out_Depth(i) = $b \times$ In_Depth(i) + c.
end if
end for
{return the adjusted depth map}
return Out_Depth

Procedure 1: Dynamic range adjustment of input depth map.

We introduce two different ways of texture quality enhancement. The first one consists in a global SNR quality enhancement using the concept of region-level coding introduced in [33] for RoI coding: the image is represented by regions mapping the QuadTree, and each region is independently encoded at a SNR quality level. The original solution allowed only two regions (RoI and non-RoI); we extended it until eight regions with their own quantization parameter, starting from label 0 for the RoI. Thus, the DoI coding system now defines the DoI-Mask as the RoI-mask and implements at least three quantization parameters: $Q_{p_Z_LR_T}$ for the depth and Low Resolution Texture images, $Q_{p_\mathrm{Ref}_T_\mathrm{DoI}}$ for refined texture image inside the DoI, and $Q_{p_\mathrm{Ref}_T_\mathrm{NDoI}}$ for the refined texture image outside it (cf. Figure 16). In this paper, we present results for this configuration only, but other scenarios are possible as a total of seven quantization levels are available for the non-DoI zones. For instance, a simple one consists of first dividing the non-DoI zones into N depth ranges ($N < 8$), defining N regions, and then applying a progressive quantization in such a way that the quantization factor is weighted by the distance of the region to the DoI.

The second enhancement way is to obtain a local resolution refinement for the texture image inside the DoI only. It simply consists of refining the Quad_Z only inside the DoI by masking the input texture image for the Quad_{ZT} estimation ($\mathrm{Quad}_{ZT_\mathrm{DoI}}$). For this mode, no information has to be transmitted to the decoder, and a unique quantization factor is applied (cf. Figure 17).

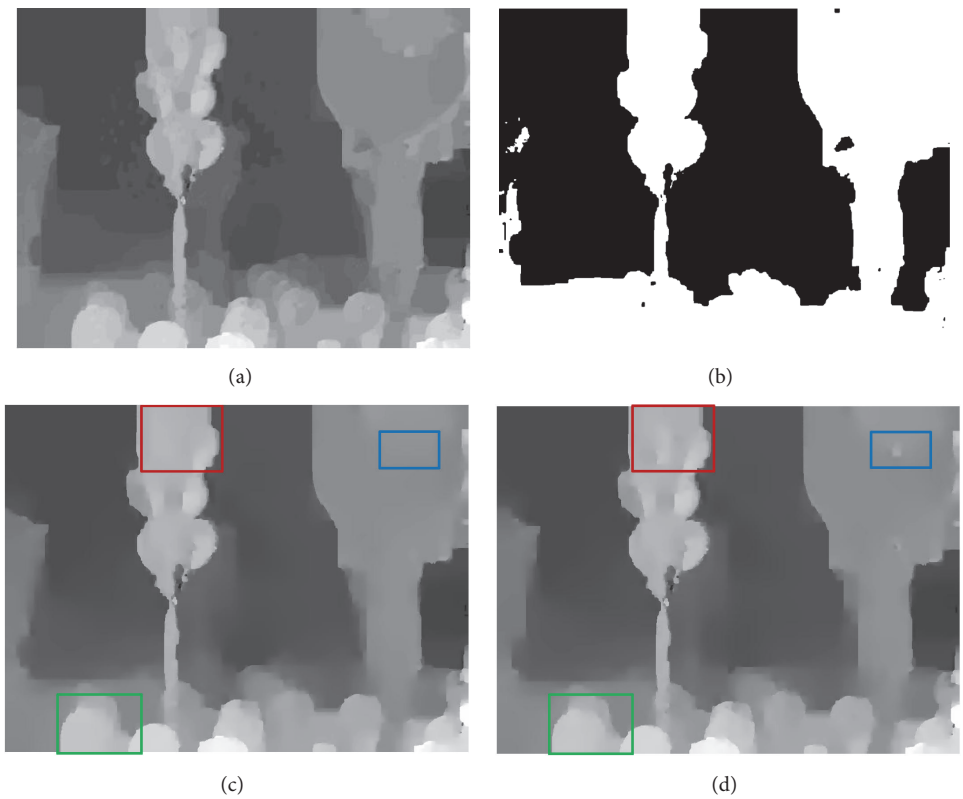


Figure 12: Visual quality comparison over depth map of Balloons view 5 frame 1 at 0.048 bpp: (a) original depth map, (b) input binary depth mask, (c) depth coded by classic LAR (Global PSNR = 36.75 dB), and (d) depth coded with DoI refinement of foreground zone with $Z_{\text{in}_l} = 128$, $Z_{\text{in}_h} = 255$, and $F = 1.3$ (Global PSNR = 36.53 dB).

A joint SNR enhancement and local resolution refinement solution also is feasible. Examples of the different solutions are presented in the next section.

The two proposed texture quality enhancement methods are both efficient, yet simple. Both methods are based on the LAR coder framework which has a low complexity as mentioned in Section 1. Thus, a relatively small additional computational cost is associated with both methods.

5. Experimental Results

5.1. Evaluation Methodology. After illustrating the preprocessing and encoding depth maps in Section 4.1, in this section, we focus on the texture coding aspects, with SNR quality and local resolution enhancement. Finally, we explore the results of the proposed coding scheme in synthesized views context. To the best of our knowledge, the proposed global representation and coding scheme is unique in terms of combined functionalities, so comparisons with state of the art are not feasible. However, we provide in the following comparative results for RoI with block instead of pixel accuracy. More details about the compression efficiency of the 2D + Z compared to the state of the art can be found in [35].

5.2. Objective and Visual Results of Decoded Texture Images in DoI Context. Some examples of DoI coding for texture images are provided in the following. More particularly, Figures 18 and 21 show the original texture, the original input depth mask (defined from the original depth map of the DoI zone within Z_{in_l} and Z_{in_h}), and the DoI-Mask (defined from the decoded refined depth map) with different resolutions: full resolution as available for the proposed scheme and subsampled one by 8 × 8 or 16 × 16 block resolution, generally provided by state-of-the-art coders.

First, examples of SNR quality enhancement are provided. Figures 19 and 22 show a zoom on the visual quality of decoded texture image coded with classic LAR and with LAR-RoI using the full resolution DoI-Mask and a subsampled one. Results show that pixel accuracy on RoI contours gives a better visual quality (cf. Figures 20 and 23).

Then, Figure 24 shows an example of local resolution enhancement for DoI in the texture image. A higher resolution is accorded to the texture image in the DoI only which leads to an increase in the objective quality as well as in the visual quality in the DoI.

5.3. Visual Results of Synthesized Views in DoI Context. The final and most important issue in 2D + Z image coding is the visual quality of the resulting synthesized views. With the depth and texture information, intermediate views at an arbitrary view point can be synthesized with the View Synthesis Reference Software (VSRS 3.0) [28]. In this set of

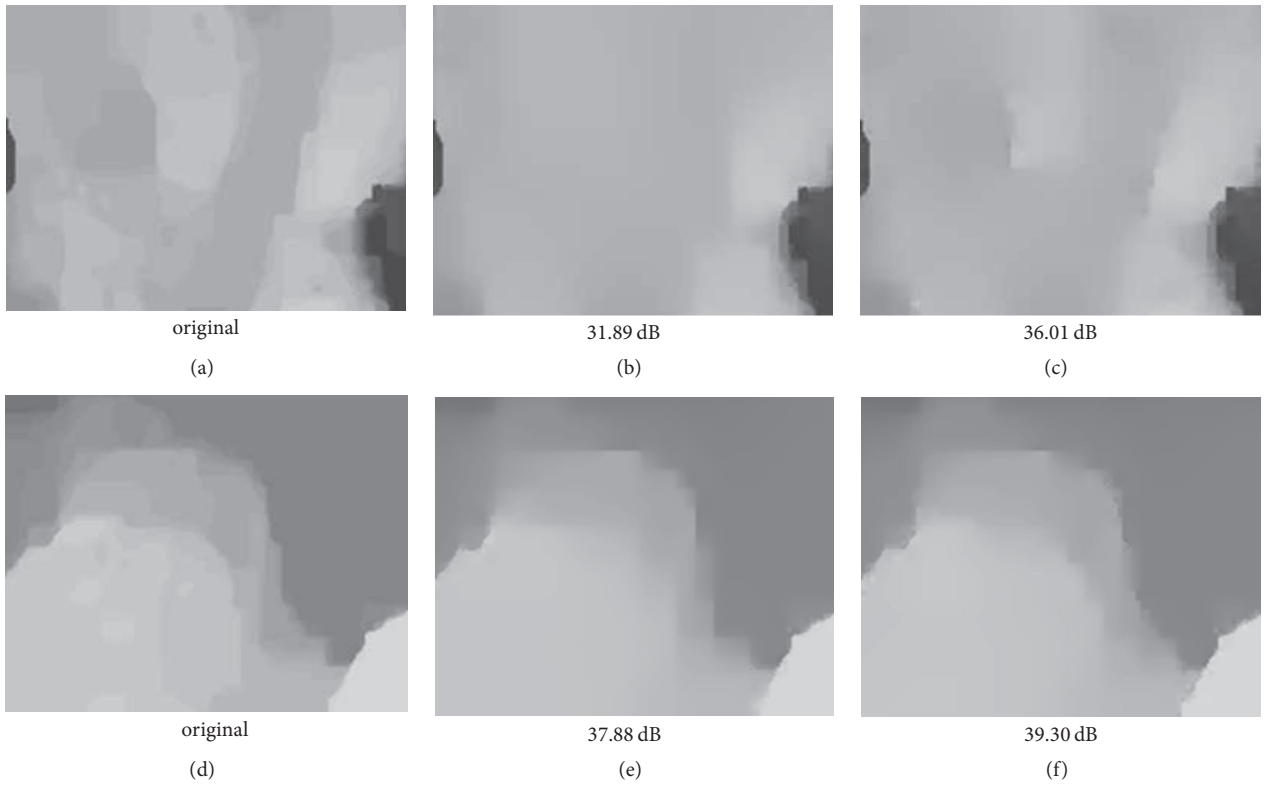


FIGURE 13: Zoom on DoI zones in depth map of Balloons view 5 frame 1 at 0.048 bpp: (a, d) original, (b, e) coded by classic LAR, and (c, f) coded with DoI refinement of foreground zone with $Z_{\text{in_}l} = 128$, $Z_{\text{in_}h} = 255$, $F = 1.3$.

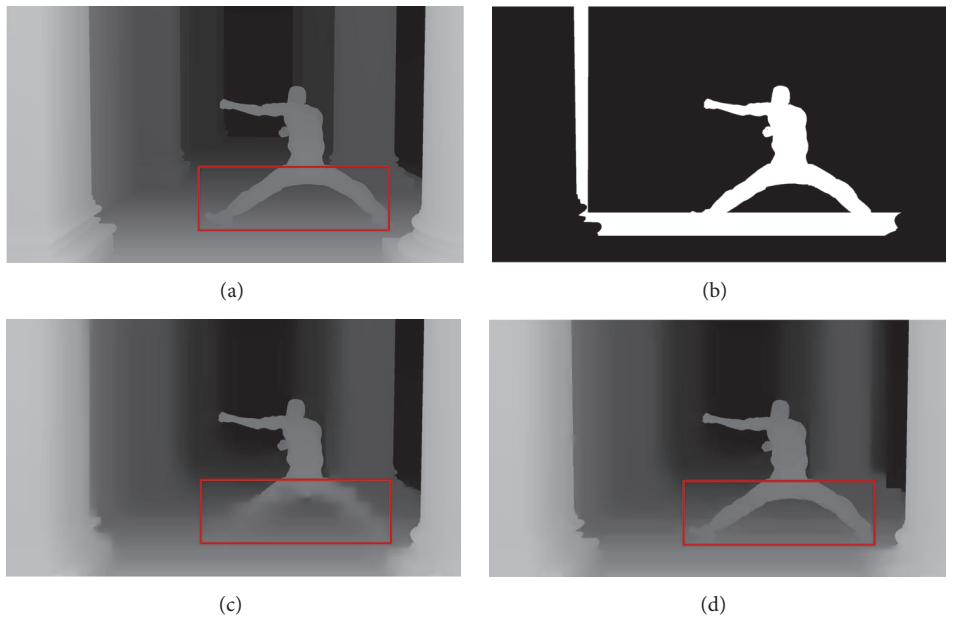


FIGURE 14: Visual quality comparison over depth map of UndoDancer view 1 frame 250 at 0.014 bpp: (a) original depth map, (b) input binary depth mask, (c) depth coded by classic LAR (Global PSNR = 35.66 dB), and (d) depth coded with DoI refinement of depth zone within $Z_{\text{in_}l} = 100$ and $Z_{\text{in_}h} = 121$ with $F = 7.0$ (Global PSNR = 32.88 dB).

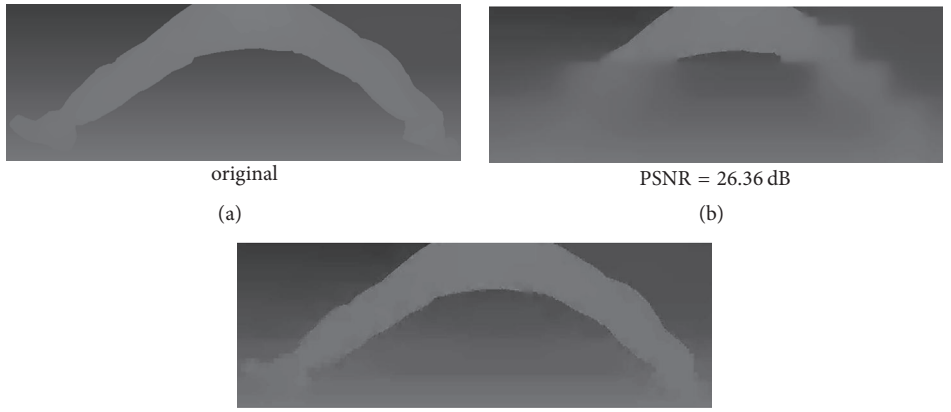


FIGURE 15: Zoom on DoI zones in depth map of UndoDancer view 1 frame 250 at 0.014 bpp: (a) original, (b) coded by classic LAR, and (c) coded with DoI refinement of depth zone within $Z_{\text{in}_l} = 100$ and $Z_{\text{in}_h} = 121$ with $F = 7.0$.

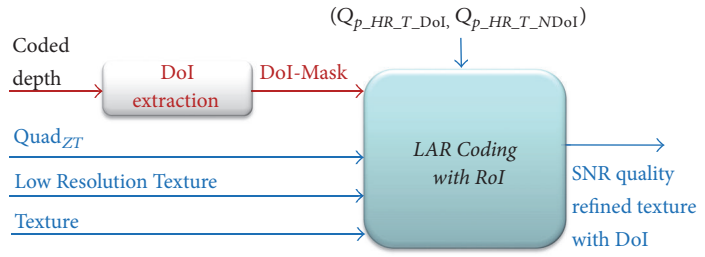


FIGURE 16: RoI coding based on DoI-Mask.

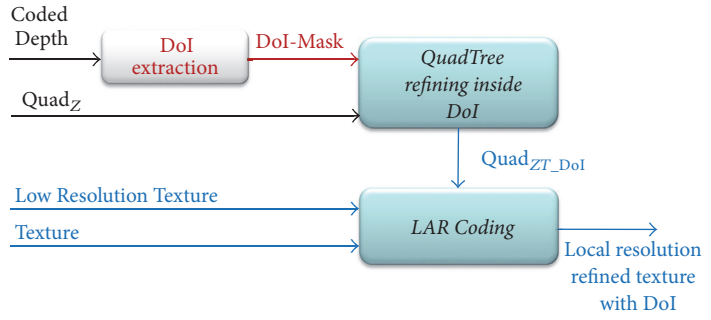
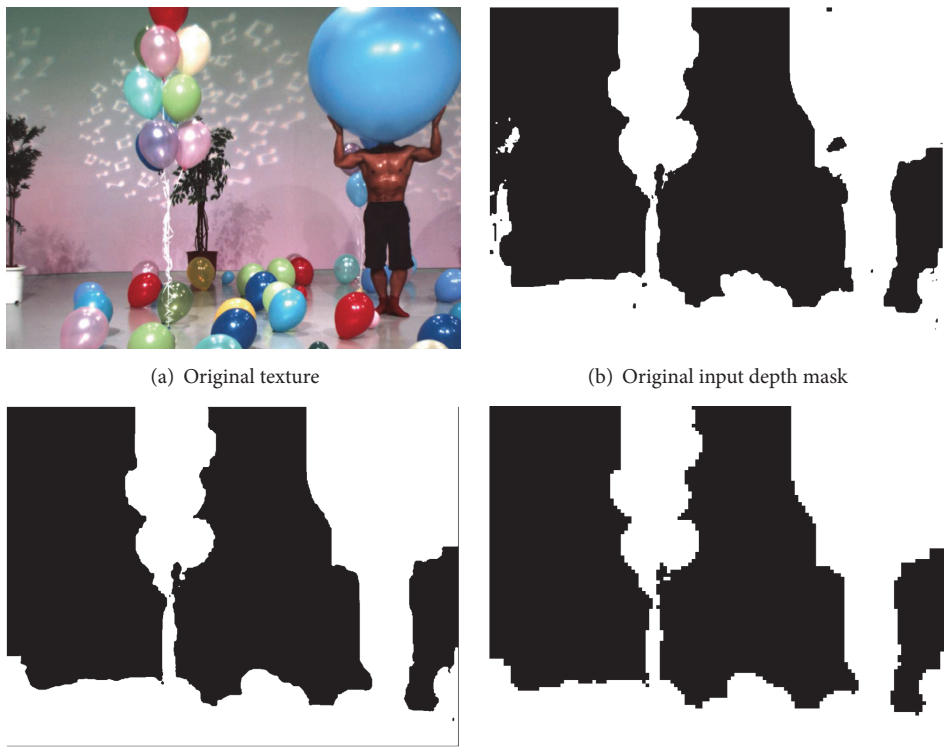


FIGURE 17: Local resolution refinement inside the DoI only.

experiments, we consider the texture images and the depth maps coded at low bit rate with and without the proposed DoI scheme, in order to evaluate the compression effect of the proposed technique on the synthesized views (cf. Figure 25). In order to compare the proposed DoI scheme with the RoI block-based approach (such as H264 and HEVC), we study the effect of RoI resolution on synthesized views by decoding texture images using DoI-Mask at different resolutions: full resolution and 8×8 and 16×16 block resolution. It is clearly noticeable that the quality within the DoI in the intermediate views synthesized from depth images decoded by the proposed DoI scheme is much better than the one synthesized from depth and texture images decoded by the classic LAR at the same bit rate, (cf. Figures 26(c) and 26(d)). Moreover, the decoded texture images with block-based DoI scheme lead to a low quality synthesized views especially upon DoI contour, while the decoded texture images with our proposed pixel-based DoI scheme lead to a fine and accurate quality upon DoI contour in the synthesized views (cf. Figure 26).

6. Conclusion

In this paper a joint content-based scheme, called "Depth of Interest" (DoI) scheme, for representation and coding of the region of interest in the 3D autostereoscopic domain is



(a) Original texture (b) Original input depth mask

(c) DoI-mask (d) 8 × 8 subsampled DoI-mask

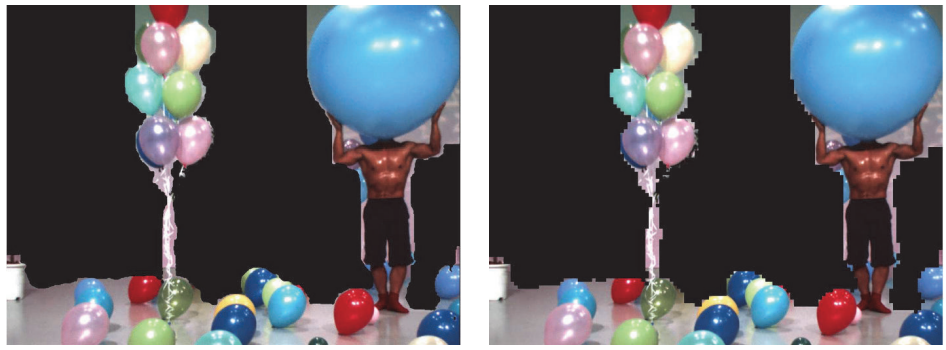
FIGURE 18: Mask of the Depth of Interest for foreground zone with $Z_{\text{in}_l} = 128$ and $Z_{\text{in}_h} = 255$ of (a) Balloons view 5 frame 1, extracted (b) from original depth, (c) from depth coded with DoI refinement scheme, and (d) from depth coded with DoI refinement scheme for 8 × 8 block resolution.



(a) (b) (c) (d)

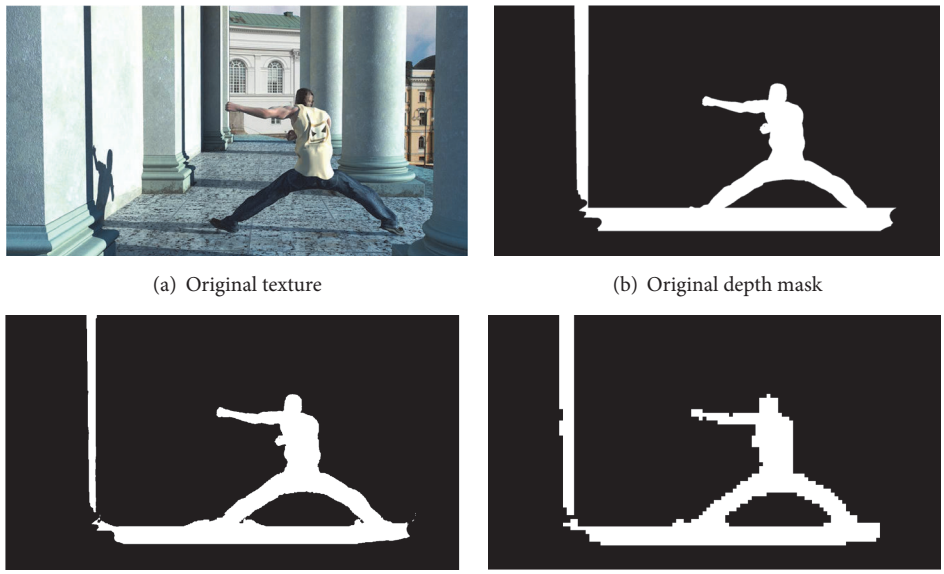
FIGURE 19: Zoom on visual quality of texture view for Balloons view 5 frame 1 (a) original, coded at 0.18 bpp, (b) by classic LAR (Global PSNR = 32.70 dB), (c) by applying the SNR quality enhancement ($Q_{p_\text{Ref}_T_\text{DoI}} = 25$, $Q_{p_\text{Ref}_T_\text{NDoI}} = 120$) using the full-resolution DoI-Mask of foreground zone with $Z_{\text{in}_l} = 128$ and $Z_{\text{in}_h} = 255$ (Global PSNR = 33.05 dB, DoI coded at 0.54 bpp, $\text{PSNR}_{\text{DoI}} = 36.87$ dB; non-DoI coded at 0.12 bpp, $\text{PSNR}_{\text{Non-DoI}} = 30.74$ dB), and (d) by applying the SNR quality enhancement for a 8 × 8 block resolution DoI-Mask.

presented. It ensures a higher quality in depth zones of the sequence that are of interest to the viewer. The proposed scheme embeds the LAR (Locally Adaptive Resolution) codec. The DoI representation scheme consists of defining, from the depth map, a binary mask covering the DoI zone. Then, the DoI coding scheme is applied on both depth and texture images. For this purpose, a preprocess, consisting of a dynamic range adjustment, is applied on the depth map in order to increase the resolution in the QuadTree partition at the DoI zone. Then, for the texture image, we use a RoI-based



(a) (b)

Figure 20: Extracted region of interest from texture coded with RoI using (a) the full-resolution DoI-Mask and (b) 8 × 8 block resolution DoI-Mask.



(a) Original texture (b) Original depth mask

(c) DoI-Mask (d) 16 × 16 subsampled DoI-Mask

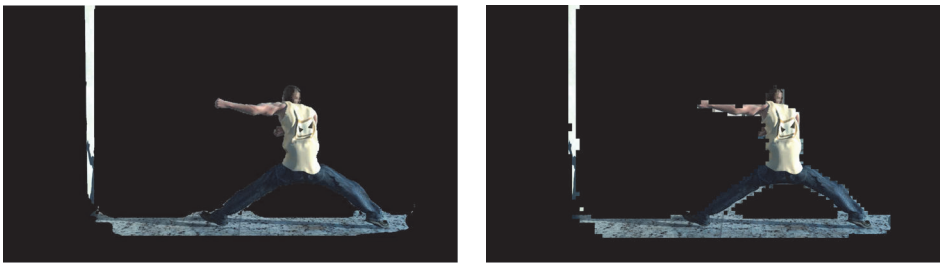
Figure 21: Mask of the Depth of Interest zone within $Z_{\text{in}_l} = 100$ and $Z_{\text{in}_h} = 121$ of (a) UndoDancer view 1 frame 255, extracted (b) from original depth, (c) from depth coded with DoI refinement scheme, and (d) from depth coded with DoI refinement scheme for 16 × 16 block resolution.



(a) (b)

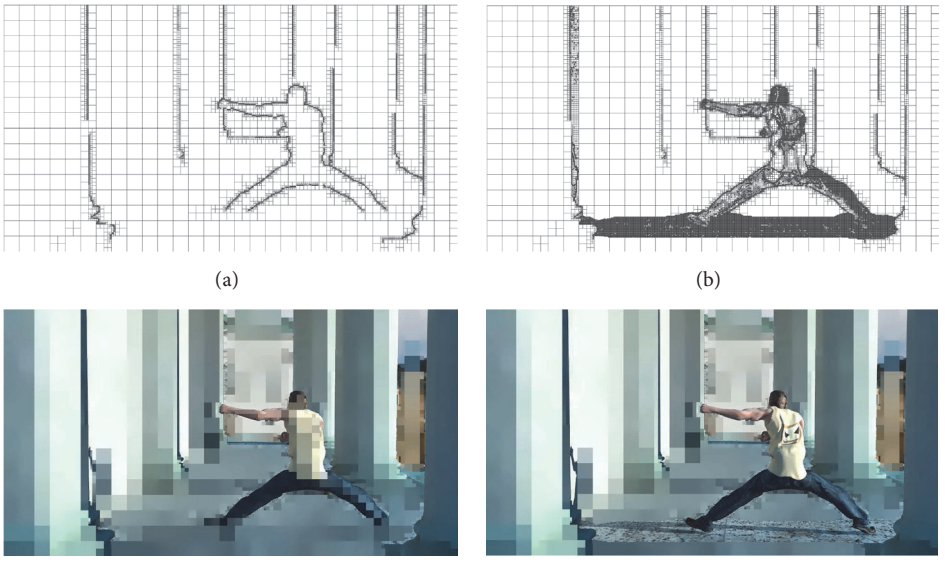
(c) (d)

Figure 22: Zoom on visual quality of texture view for UndoDancer view 1 frame 250 (a) original, coded at 0.2 bpp (b) by classic LAR (Global PSNR = 24 dB), (c) by applying the SNR quality enhancement ($Q_{p_\text{Ref}_T_\text{DoI}} = 25$, $Q_{p_\text{Ref}_T_\text{NDoI}} = 120$) using the full-resolution DoI-Mask of depth zone within $Z_{\text{in}_l} = 100$ and $Z_{\text{in}_h} = 121$ (Global PSNR = 28.82 dB, DoI coded at 0.98 bpp, $\text{PSNR}_{\text{DoI}} = 36.06$ dB, non-DoI coded at 0.17 bpp, $\text{PSNR}_{\text{Non-DoI}} = 28.1$ dB), and (d) by applying the SNR quality enhancement for 16 × 16 block resolution DoI-Mask.



(a) (b)

Figure 23: Extracted region of interest from texture coded with RoI using (a) the full-resolution DoI-Mask; (b) 16 × 16 block resolution DoI-Mask.



(a) (b)

(c) (d)

Figure 24: Visual quality comparison of texture view for UndoDancer view 1 frame 250 with $\mathrm{Th}_{\mathrm{Quad}} = 20, Q_p = 30$. (a) Quad_Z, (b) $\mathrm{Quad}_{ZT_\mathrm{DoI}}$, (c) Low Resolution Texture at 0.06 bpp (Global PSNR = 17.2 dB), and (d) local resolution refined texture at 0.23 bpp ($\mathrm{PSNR}_{\mathrm{DoI}} = 32.24$ dB).

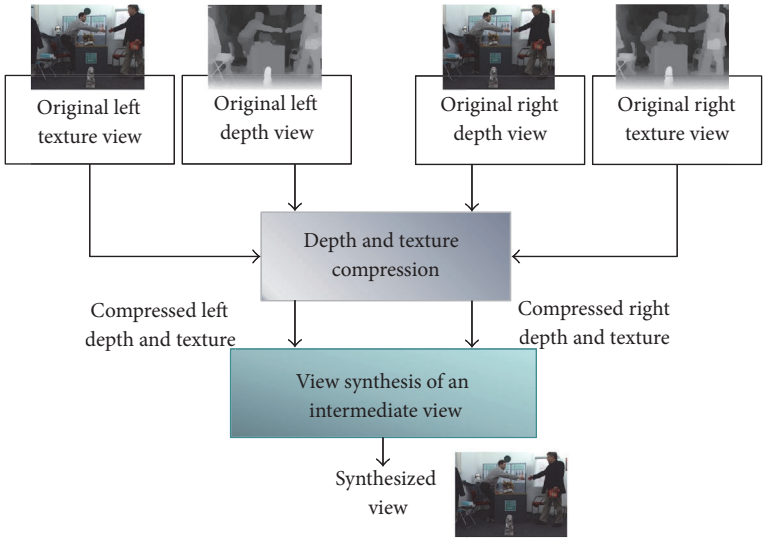


Figure 25: Synthesized view rendering framework.

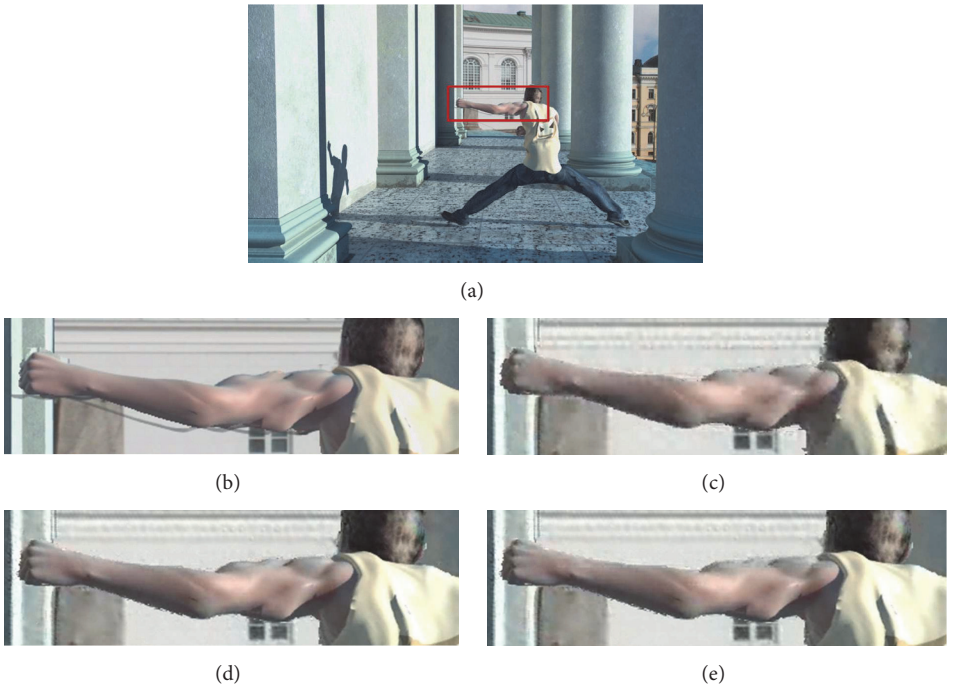


FIGURE 26: Visual quality comparison of synthesized view for (a) UndoDancer view 3 frame 250 using (b) uncompressed depth and texture; depth coded at 0.014 bpp and texture coded at 0.2 bpp (c) by classic LAR; (d) by the proposed DoI refinement of depth scheme and SNR quality enhancement of texture with a full-resolution DoI-Mask; (e) by the proposed DoI scheme with the 16 × 16 block resolution DoI-Mask.

scheme using the defined depth-based binary mask. The main strength of this scheme is that it provides a high correlation between texture objects and depth and allows high quality along objects contours in depth and texture images as well as in synthesized views. In future work, we will focus on a nonlinear adjustment of the dynamic range and we will also study multiple DoI zones.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] A. Smolic and P. Kauff, "Interactive 3-D video representation and coding technologies," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 98–110, 2005.

[2] C. Fehn, R. de la Barré, and S. Pastoor, "Interactive 3-DTV-concepts and key technologies," *Proceedings of the IEEE*, vol. 94, no. 3, pp. 524–538, 2006.

[3] Phillips Research Press Information, "Philips 3d information display solution adds extra dimension to in-store messaging," 2005, http://www.research.philips.com/newscenter/archive/.

[4] J. Harrold and G. J. Woodgate, "Autostereoscopic display technology for mobile 3DTV applications," in *Proceedings of the Stereoscopic Displays and Virtual Reality Systems XIV*, San Jose, Calif, USA, January 2007.

[5] A. Smolic, K. Mueller, P. Merkle et al., "3D video and free viewpoint video—technologies, applications and MPEG standards," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '06)*, pp. 2161–2164, July 2006.

[6] M. Tanimoto, "Overview of FTV (free-viewpoint television)," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '09)*, pp. 1552–1553, New York, NY, USA, July 2009.

[7] N. A. Dodgson, "Autostereo displays: 3d without glasses," in *Proceedings of the Electronic Information Displays Conference (EID '97)*, vol. 38, pp. 31–36, November 1997.

[8] A. Vetro, S. Yea, and A. Smolic, "Toward a 3D video format for auto-stereoscopic displays," in *Applications of Digital Image Processing XXXI*, vol. 7073 of *Proceedings of SPIE*, pp. 1–12, September 2008.

[9] A. Smolic, K. Mueller, N. Stefanoski et al., "Coding algorithms for 3DTV—a survey," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 11, pp. 1606–1621, 2007.

[10] P. Merkle, A. Smolic, K. Muller, and T. Wiegand, "Multiview video plus depth representation and coding," in *Proceedings of the 14th IEEE International Conference on Image Processing (ICIP '07)*, vol. 1, pp. 201–204, San Antonio, Tex, USA, 2007.

[11] I. Himawan, W. Song, and D. Tjondronegoro, "Automatic region-of-interest detection and prioritisation for visually optimised coding of low bit rate videos," in *Proceedings of the IEEE Workshop on Applications of Computer Vision (WACV '13)*, pp. 76–82, January 2013.

[12] G. Hu, M. Xiao, and S. Yuan, "Detecting automatically and compression algorithm for infrared image based on region of interest," in *Proceedings of the International Forum on Computer Science-Technology and Applications (IFCSTA '09)*, pp. 50–53, December 2009.

[13] M. Firoozbakht, J. Dehmeshki, M. Martini et al., "Compression of digital medical images based on multiple regions of interest," in *Proceedings of the 4th International Conference on Digital Society (ICDS '10)*, pp. 260–263, St. Maarten, Netherlands, February 2010.

[14] M. Nawaz, J. Cosmas, A. Adnan, and M. Ali, "Inter-intra frame segmentation using colour and motion for region of interest coding of video," in *Proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB '11)*, Nuremberg, Germany, June 2011.

[15] Y. Li, W. Li, and Y. Ma, "Accurate iris location based on region of interest," in *Proceedings of the International Conference on Biomedical Engineering and Biotechnology (iCBEB '12)*, pp. 704–707, Macau, Macao, May 2012.

[16] C. Chamaret, S. Goddefroy, P. Lopez, and O. Le Meur, "Adaptive 3d rendering based on region-of-interest," in *Stereoscopic Displays and Applications XXI*, vol. 7524 of *Proceedings of SPIE*, pp. 1–12, February 2010.

[17] S.-L. Fan, M. Yu, G.-Y. Jiang, F. Shao, and Z.-J. Peng, "A digital watermarking algorithm based on region of interest for 3D image," in *Proceedings of the 8th International Conference on Computational Intelligence and Security (CIS '12)*, pp. 549–552, Guangzhou, China, November 2012.

[18] Y. Zhang, G. Jiang, M. Yu, Y. Yang, Z. Peng, and K. Chen, "Depth perceptual region-of-interest based multiview video coding," *Journal of Visual Communication and Image Representation*, vol. 21, no. 5-6, pp. 498–512, 2010.

[19] L. S. Karlsson and M. Sjöström, "Region-of-interest 3D video coding based on depth images," in *Proceedings of the 3DTV-Conference*, pp. 141–144, Istanbul, Turkey, May 2008.

[20] H. Zheng, B. Liu, and H. Zhang, "Region-of-interest coding of 3D mesh based on wavelet transform," in *Proceedings of the 3rd International Conference on Image and Graphics (ICIG '04)*, pp. 438–441, IEEE, Hong Kong, December 2004.

[21] L. S. Karlsson, *Spatio-temporal pre-processing methods for region-of-interest video coding [Licenciate Thesis 21]*, Department of Information Technology and Media, Mid Sweden University, Sundsvall, Sweden, 2007, http://urn.kb.se/resolve?urn=urn:nbn:se:miun:diva-51.

[22] JPEG 2000 Part I Final Committee Draft Version 1.0, ISO/IEC FCD15444-1: 2000, Annex H, March 2000.

[23] J. A. Raja, G. Raja, and A. K. Khan, "Selective compression of medical images using multiple regions of interest," *Life Science Journal*, vol. 10, no. 9, pp. 394–397, 2013.

[24] I. A. Fernandez, P. R. Alface, T. Gan, R. Lauwereins, and C. De Vleeschouwer, "Integrated H.264 region-of-interest detection, tracking and compression for surveillance scenes," in *Proceedings of the 18th International Packet Video Workshop (PV '10)*, pp. 17–24, December 2010.

[25] K. Misra, A. Segall, M. Horowitz, S. Xu, A. Fuldseth, and M. Zhou, "An overview of tiles in HEVC," *IEEE Journal on Selected Topics in Signal Processing*, vol. 7, no. 6, pp. 969–977, 2013.

[26] V. Sanchez, A. Basu, and M. K. Mandal, "Prioritized region of interest coding in JPEG2000," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 9, pp. 1149–1155, 2004.

[27] C. Christopoulos, J. Askelöf, and M. Larsson, "Efficient methods for encoding regions of interest in the upcoming JPEG2000 still image coding standard," *IEEE Signal Processing Letters*, vol. 7, no. 9, pp. 247–249, 2000.

[28] M. Tanimoto, T. Fuji, K. Suzuki, N. Fukushima, and Y. Mori, "Reference softwares for depth estimation and view synthesis," Tech. Rep. M15377, ISO/IECJTC1/SC29/WG11, MPEG, Archamps, France, 2008.

[29] P. Merkle, Y. Morvan, A. Smolic et al., "The effect of depth compression on multiview rendering quality," in *Proceedings of the 2008 3DTV-Conference: The True Vision—Capture, Transmission and Display of 3D Video (3DTV-CON '08)*, pp. 245–248, May 2008.

[30] E. Bosc, M. Pressigout, and L. Morin, "Focus on visual rendering quality through content-based depth map coding," in *Proceedings of the 28th Picture Coding Symposium (PCS '10)*, pp. 158–161, Nagoya, Japan, December 2010.

[31] Y. Liu, S. Ma, Q. Huang, D. Zhao, W. Gao, and N. Zhang, "Compression-induced rendering distortion analysis for texture/depth rate allocation in 3D video compression," in *Proceedings of the Data Compression Conference (DCC '09)*, vol. 9, pp. 352–361, IEEE, Snowbird, Utah, USA, March 2009.

[32] A. Tikanmäki, A. Gotchev, A. Smolic, and K. Müller, "Quality assessment of 3D video in rate allocation experiments," in *Proceedings of the 12th IEEE International Symposium on Consumer Electronics (ISCE '08)*, April 2008.

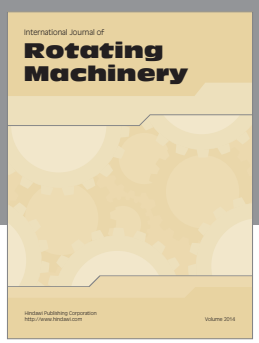
[33] O. Deforges, M. Babel, L. Bedat, and J. Ronsin, "Color LAR codec: a color image representation and compression scheme based on local resolution adjustment and self-extracting region representation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 8, pp. 974–987, 2007.

[34] K. Samrouth, F. Pasteaua, and O. Deforgesa, "Quality constraint and rate-distortion optimization for predictive image coders," in *Image Processing: Algorithms and Systems XI*, vol. 8655 of *Proceedings of SPIE*, pp. 1–9, Burlingame, Calif, USA, February 2013.

[35] K. Samrouth, O. Deforges, Y. Liu, F. Pasteau, M. Khalil, and W. Falou, "Efficient depth map compression exploiting correlation with texture data in multiresolution predictive image coders," in *Proceedings of the IEEE International Conference on Multimedia and Expo Workshops (ICMEW '13)*, pp. 1–6, San Jose, Calif, USA, July 2013.

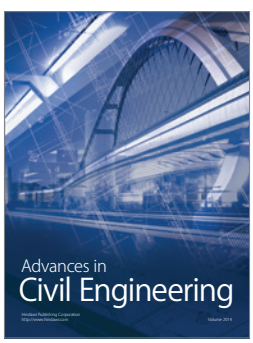


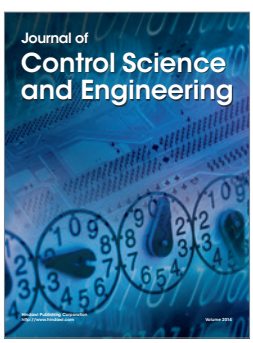






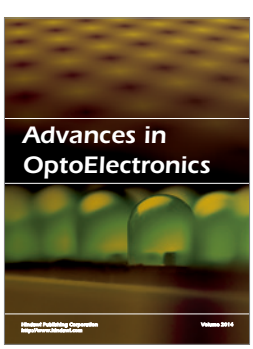






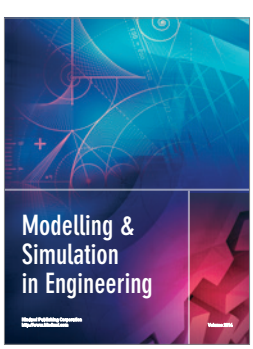






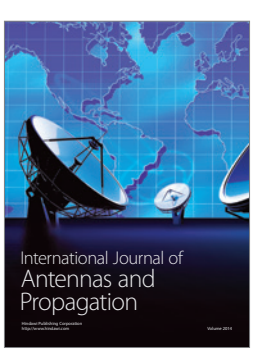




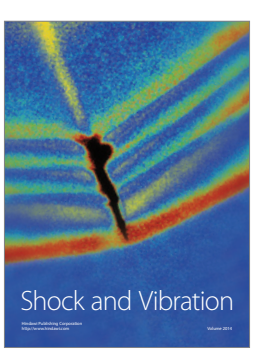


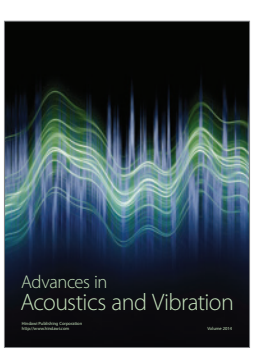












Hindawi

Submit your manuscripts at
https://www.hindawi.com