



HAL
open science

A content based method for perceptually driven joint color/depth compression

Emilie Bosc, Luce Morin, Muriel Pressigout

► **To cite this version:**

Emilie Bosc, Luce Morin, Muriel Pressigout. A content based method for perceptually driven joint color/depth compression. IS&T/SPIE Electronic Imaging, Jan 2012, San Francisco, United States. pp.8288-82. hal-00670373

HAL Id: hal-00670373

<https://hal.science/hal-00670373>

Submitted on 15 Feb 2012

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.

A content based method for perceptually driven joint color/depth compression

E. Bosc^{*a}, L. Morin^a, M. Pressigout^a

^aIETR UMR CNRS 6164, INSA de Rennes, 20, avenue des Buttes de Coesmes, 35708 Rennes Cedex7, France

ABSTRACT

Multi-view Video plus Depth (MVD) data refer to a set of conventional color video sequences and an associated set of depth video sequences, all acquired at slightly different viewpoints. This huge amount of data necessitates a reliable compression method. However, there is no standardized compression method for MVD sequences. H.264/MVC compression method, which was standardized for Multi-View-Video representation (MVV), has been the subject of many adaptations to MVD. However, it has been shown that MVC is not well adapted to encode multi-view depth data.

We propose a novel option as for compression of MVD data. Its main purpose is to preserve joint color/depth consistency. The originality of the proposed method relies on the use of the decoded color data as a prior for the associated depth compression. This is meant to ensure consistency in both types of data after decoding. Our strategy is motivated by previous studies of artifacts occurring in synthesized views: most annoying distortions are located around strong depth discontinuities and these distortions are due to misalignment of depth and color edges in decoded images. Thus the method is meant to preserve edges and to ensure consistent localization of color edges and depth edges.

To ensure compatibility, colored sequences are encoded with H.264. Depth maps compression is based on a 2D still image codec, namely LAR (Locally adapted Resolution). It consists in a quad-tree representation of the images. The quad-tree representation contributes in the preservation of edges in both color and depth data.

The adopted strategy is meant to be more perceptually driven than state-of-the-art methods. The proposed approach is compared to H.264 encoding of depth images. Objective metrics scores are similar with H.264 and with the proposed method, and visual quality of synthesized views is improved with the proposed approach.

Keywords: 3DTV, compression, multi-view video plus depth

1. INTRODUCTION

3D Video applications¹, such as 3D Television (3DTV) or Free Viewpoint Television (FTV), require the use of several conventional video sequences to ensure depth sensation, or to offer novel views of a scene. For these purposes, the use of color and geometry information of the scene is the key. MVD data refer to a specific representation of an observed scene and meet this need. They consist in a set of conventional color video sequences and an associated set of depth video sequences, all acquired at slightly different viewpoints.

A first issue refers to the need for an efficient MVD compression method, considering the huge amount of data to be processed. Up to now, there is no standardized compression method for MVD sequences. Most of the proposed compression methods rely on the extension of state-of-the-art 2D codecs. The most popular is H264/AVC³ whose 3D extension (standardized for Multi-View-Video representation, MVV), namely H.264/MVC for Multi-view Video Coding⁴, has been the subject of many adaptations for MVD compression⁵. However, the exploitations of the spatial inter-view redundancies in both types of data turn out to be insufficient in particular cases. For instance, Merkle *et al.*⁵ observed that in case of large disparity between the different views of multi-view sequences, the predictions structures did not result in an improved coding efficiency.

A second issue refers to the synthesis of novel views from decoded data. New intermediate viewpoints can be generated from depth and color data through Depth-Image-Based-Rendering⁶ (DIBR) methods. Previous studies already pointed

out the impact of depth encoding on the synthesized frames. Compression-related artifacts that may be imperceptible in depth maps cause important distortion during the synthesis process⁷.

Many methods have been proposed recently in order to address the aforementioned issues. Various encoding strategies are possible to achieve depth map compression. Several studies have proposed bit-rate-control methods^{8,9}, relying on the objective quality of the resulting synthesized views, or on a distortion model¹⁰. A popular and efficient strategy is the post-processing of depth maps after decoding¹¹. Depth-adapted encoding methods^{12,13,14} have also been proposed. Section 0 gives a review of these methods. Our work is in line with the depth-adapted encoding strategy since the method proposed in this paper relies on a content-based representation of the depth map.

The main purpose of this novel framework is to preserve the consistency between color and depth data. Our strategy is motivated by previous studies⁷ of artifacts occurring in synthesized views: most annoying distortions are located around strong depth discontinuities¹⁵ and these distortions are due to misalignment of depth and color edges in decoded images. Thus the method is meant to preserve edges and to ensure consistent localization of color edges and depth edges. It is based on a 2D still image codec, namely LAR¹⁶ (Locally adapted Resolution). The LAR codec is based on a quad-tree representation of the images. In this quad-tree, the smaller the blocks, the higher the probability of the presence of a depth discontinuity. Analogously, big blocks correspond to smooth areas. The quad-tree representation contributes in the preservation of depth transitions when target bit-rate decreases. Another original contribution of the proposed method relies on the use of the decoded color data as an anchor for the enhancement of the associated decoded depth, together with information provided by the quad-tree structure. This is meant to ensure consistency in both types of data after decoding.

This paper is organized as follows: Section 2 introduces the compression issues for MVD data. Section 3 presents the proposed method. Section 4 defines the experimental protocol and gives the results. Finally Section 5 concludes the paper.

2. COMPRESSION OF MVD SEQUENCES AND QUALITY OF SYNTHESIZED VIEWS

This section presents the main issues related to MVD compression and a review of the proposed methods addressing these problems in the literature.

Most of the proposed compression methods for MVD data rely on the extension of state-of-the-art 2D codecs. Sets of color and depth sequences can be separately encoded through existing 2D methods. This is an evident encoding strategy because depth maps, being monochromatic signals, are considered as conventional sequences. However, depth maps are not natural images. They provide structural information of the scene: large and smooth regions often belong to the same depth plane. The closer the depth plane is from the acquiring camera, the lighter the region. This leads to smooth areas with sharp edges. The edges correspond to depth transitions.

Previous studies^{17,18,19} have shown that coding artifacts on depth data can dramatically influence the quality of the synthesized view. Particularly, the sharp edges of the depth maps are prone to synthesis errors even when depth maps are uncompressed. As pointed out in a recent study¹⁵, the synthesis process, with DIBR methods, induces specific artifacts located around the edges of objects. These errors are notably due to depth map inaccuracy, numerical rounding, hole filling method in DIBR, or both. Consequently, errors occurring in these specific critical areas of the depth maps are enhanced by coarse compression. The impacts of depth compression on visual quality of synthesized views can be explained by the fact that 2D codecs are optimized for human visual perception of color images. Thus, artifacts, that may be imperceptible when visualizing the depth map, produce distortions because during the synthesis, the warping process relies on wrong depth values. The impacts of depth compression was observed in different studies^{7,18}.

Consequently, efforts have been directed in order to propose depth compression methods more adapted to the special features of depth maps. Morvan *et al.*¹² proposed to represent the depth map thanks to platelets (piecewise linear functions). The depth map is first divided through quad-tree decomposition and each block is approximated by a platelet. The platelet-based compression outperformed JPEG2000 in the study. An additional interesting comparison would be that against H.264/AVC. Moreover, in this study, the gain is evaluated with respect to the depth distortion (in PSNR). This protocol of validation is questionable because since the artifacts in the two compared methods are different, their impact on the synthesis may also be different. Yet, the quality of the synthesized views generated from the decoded depth maps is not presented. Graziosi *et al.*¹³ also proposed a block partitioning method associated to a least-square prediction for depth map compression. In this method, the validation is also achieved by comparing the depth map

distortion from different compression scheme (JPEG2000 and H.264 intra). The method includes the use of a dictionary, containing concatenations of scaled versions of previously encoded image blocks. Sarkis et al.¹⁴ proposed a depth compression method based on a subsampling in the frequency domain followed by a reconstruction using a non-linear conjugate gradient minimization scheme. The method also meant to preserve the particular features of the depth map. The method outperformed JPEG and JPEG2000.

As Morvan *et al.* and Graziosi *et al.*, our method relies on a block partitioning. Contrary to the aforementioned methods, we choose to evaluate the performances of our method against H.264, whose artifacts in the depth map induce less annoying distortions than JPEG2000 (Gibbs effect in the depth map induce disastrous distortions in the synthesized view). The quantization used in our method is also different since we choose to modify the block partitioning of the depth map according to the target bit-rate. The next section will present our proposed method that is meant to preserve object edges to maximize the synthesized view quality.

3. PROPOSED METHOD

Since color and depth sequences can be encoded through two different schemes, the proposed method enables the use of any compression method for the color sequences. In our case, we use H.264 for color sequence encoding since it has proved its efficiency for conventional color media compression. Moreover, encoding color with a standard codec enables backward compatibility with classical 2D video. This section presents the depth map encoding strategy.

3.1 Depth map encoding method

To address the first constraint regarding the preservation of the depth map edges, a content-based representation and encoding is required. We choose to base our method on the LAR method because the quad-tree representation of this method matches the characteristics of the depth maps. The LAR method is based on the assumption that an image can be considered as the combination of two components: the global information and the details, which are respectively the flat image and the local texture. The flat image and the local texture both rely on the same quad-tree representation. Each pixel in the flat image is assigned the mean value of the pixels of the block it belongs to. Each pixel in the local texture is then assigned the compensated error. Figure 1 depicts this principle.

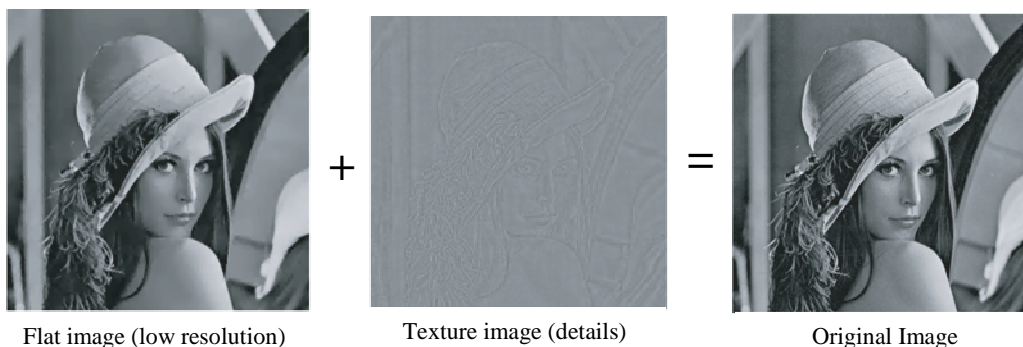


Figure 1: Assumption of LAR method.

Since depth maps do not contain high frequency areas, the local texture (that is to say the details) is not essential and represents an avoidable additional cost of compression. Thus, only the flat image is considered and encoded in the method we propose.

3.1.1 Quad-tree decomposition

The quad-tree decomposition is dependent on the local gradient of the depth image. Given a threshold Y for the local gradient, the image is split into blocks: the higher the local activity, the more splits. This leads to small blocks around object edges and bigger ones in continuous areas.

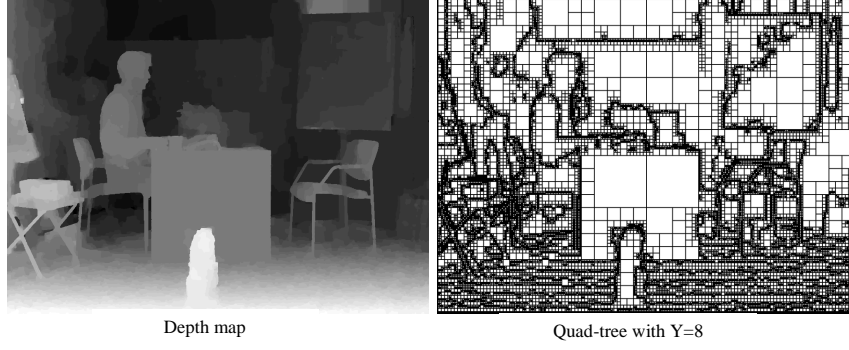


Figure 2: Quad-tree decomposition (Book Arrival).

We denote $P^{[N_{max} \dots N_{min}]}$ the quad-tree partition with N_{max} and N_{min} the maximal and minimal allowed block sizes, expressed as power of 2. Let I be an image and $I(x, y)$ a pixel of I with coordinates (x, y) . The block $b^N(i, j)$ in I is noted $I(b^N(i, j))$, expressed as:

$$b^N(i, j) = \{(x, y) \in N_x \times N_y \mid N \times i \leq x < N \times (i + 1), N \times j \leq y < N \times (j + 1), N \in [N_{min} \dots N_{max}]\} \quad (1)$$

As explained before, the quad-tree partition $P^{[N_{max} \dots N_{min}]}$ relies on the analysis of the local gradient. Then, the decomposition can be expressed as:

$$\forall I(b^N(i, j)), N = \begin{cases} N' \in [N_{max} \dots N_{min}] \text{ if } \max(I(b^{N'}(i, j))) - \min(I(b^{N'}(i, j))) \leq Y, \text{ and if } \exists(k, m) \in \{0, 1\}^2 \\ \quad \left| \max\left(I\left(b^{\frac{N'}{2}}(i+k, j+m)\right)\right) - \min\left(I\left(b^{\frac{N'}{2}}(i+k, j+m)\right)\right) > Y \right. \\ N_{min} \text{ otherwise} \end{cases} \quad (2)$$

The value of the threshold Y strongly influences the final representation of the image. Figure 2 gives an example of quad-tree decomposition for the first frame of *Book Arrival* sequence.

3.1.2 Compression scheme

The compression scheme in the LAR method is based on a pyramidal decomposition²⁰. The pyramid, built from I , consists of a set of images, noted as $\{L_l\}_{l=0}^{l=l_{max}}$, as a multi-resolution representation of the image, where l_{max} is the top of the pyramid and $l = 0$ is the lowest level, i.e. the full resolution image. At each level, the image is expressed by:

$$\begin{cases} l = 0, L_0(x, y) = I(x, y) \\ l > 0, L_l(x, y) = \left\lfloor \frac{L_{l-1}(2x, 2y) + L_{l-1}(2x+1, 2y+1)}{2} \right\rfloor \end{cases} \quad (3)$$

The LAR method allows the prediction of each level of the pyramid, from top to bottom. For each level, the associated image of errors, also relying on the quad-tree decomposition, can be transmitted to compensate the prediction errors. At the decoder side, from the top to the bottom, the image is reconstructed.

Compression cost is mainly due to the encoding of small blocks. This is why in our proposed method small blocks are not transmitted (those are blocks whose size is such as $N = N_{min}$). This is achievable thanks to the pyramidal decomposition. The encoding of small blocks is related to the image of errors corresponding to the lowest level, i.e. L_0 . The lowest level is not encoded in the method we propose, and the image will be refined at the decoder side thanks to the analysis of the values of the nearest neighbor blocks whose size is such as $N > N_{min}$: they will be predicted, depending on the values of their closest larger blocks. This allows bit-rate savings. The pseudo code of this prediction is given in Figure 3.

```

Prediction of lowest level of the pyramidal decomposition
 $\tilde{L}_l$  is the estimated representation of the image at the decoder side, for level  $l$ .
 $P^{\lfloor N_{max} \dots N_{min} \rfloor}$  is the quad-tree partition.
For  $l=l_{max} \dots l_1$ 
    Estimate  $\tilde{L}_l$  as in the LAR method
End for
For each block of  $P$  such as  $N = N_{max} \dots 2 * N_{min}$ 
    Given  $P^{\lfloor N_{max} \dots N_{min} \rfloor}$ , then  $\tilde{L}_0(b^N(i, j)) = \tilde{L}_1(\lfloor \frac{i}{2} \rfloor, \lfloor \frac{j}{2} \rfloor)$ 
End For
For each block of  $P$  such as  $N = N_{min}$ 
     $\tilde{L}_0(b^{N_{min}}(i, j)) = \text{Mean value of the closest block } b^N \text{ of } P \text{ such as } N > N_{min}$ 
End For

```

Figure 3: Pseudo code for Prediction Algorithm of lowest level of the pyramidal decomposition.

3.2 Rate control in depth map

Pasteau *et al.*¹⁶ suggested applying a quantization step depending on the block sizes, in the case of conventional images. Our experiments revealed that in the case of depth map compression, this was not an adequate strategy because the smaller the blocks, the coarser was the quantization (this allowed bit rate savings because small block are costly). Yet, small blocks correspond to strong depth discontinuities and errors occurring in these areas may have disastrous effect at the synthesis step. Figure 4 shows the impact of the quantization as suggested in Pasteau *et al.*¹⁶ (first column) at 0.06 bpp. Depth transitions are highly degraded and will result in errors in the synthesized frame (third column, crumbling artifacts around the head and around the legs of the chair). The synthesized frames obtained in Figure 4 are generated from original color data and decoded depth maps in order to visually assess only the impact of depth quantization (i.e. not the combined effect of both color and depth compression).



Decoded depth when applying quantization suggested in Pasteau *et al.* Decoded depth when applying our proposed strategy.

Frame synthesized from decoded depth when applying quantization suggested in Pasteau *et al.* Frame synthesized from decoded depth when applying our proposed strategy.

Figure 4: Comparison of two decoded depth maps at 0.06bpp, using the LAR method or the proposed method of rate control.

Thus, we propose a quantization achieved through the evolution of the quad-tree representation of the image. Small blocks are costly and a way to reduce the bit-rate is to reduce the number of small blocks. This implies that the quad-tree representation can change according to the target bit-rate. The number of small blocks is directly related to the value of the threshold Y . Thus, an increasing threshold Y decreases the bit-rate, so that the representation of the image contains larger blocks. This corresponds to a spatial quantization that affects the values of the depth. It results in assigning the same depth to object that were not formerly in the same depth plane. The dynamic range of the depth is reduced but the global structure is preserved. Figure 4 shows that the proposed method (second column) renders sharp depth transitions. The synthesized frame in Figure 4, fourth column, shows improvements compared to the previous strategy, third column. Figure 5 gives the quad-tree representations and the resulting depth maps using two different thresholds for the quad-tree decomposition. It shows that the semantic information of the image is preserved.

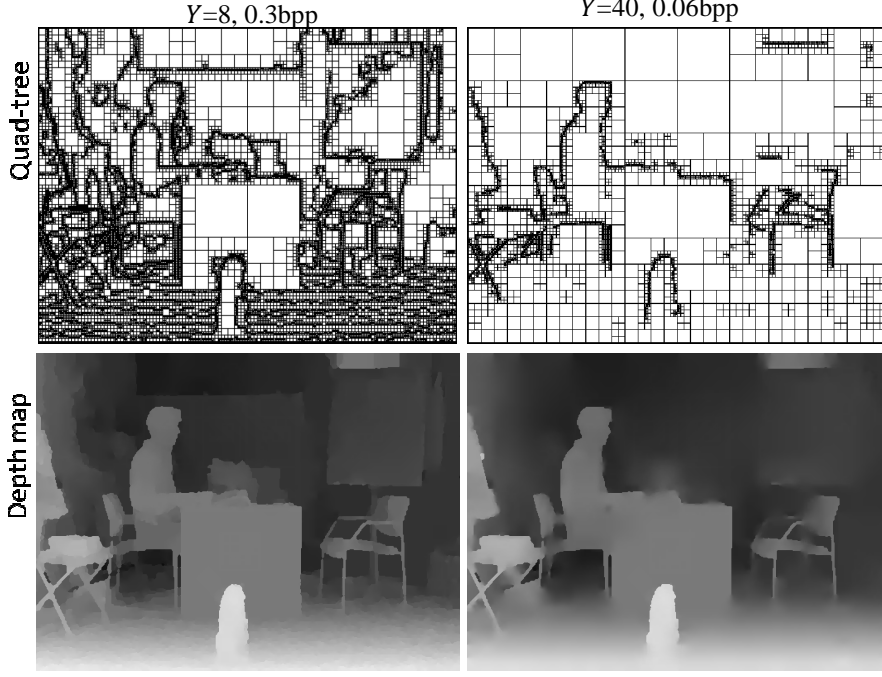


Figure 5: Quantization of the depth map.

In this study, we empirically determined a model allowing the choice of Y depending on the target bit-rate. Based on the analysis of the synthesized view quality scores obtained for various values of Y , and according to the corresponding bit-rate R of the encoded frames, we opted to an exponential model such as:

$$Y = ae^{bR} \quad (4)$$

R is the target bit-rate, a and b are two constants. Previous experiments showed that $a = 30$ and $b = -12$ gave good results for the tested sequence. Note that these values will differ from one MVD sequence to another, since the representation of the quad-tree depends on the structure of the sequence, thus Y model depends on the depth structure of the sequence.

3.3 Depth reconstruction at decoder side

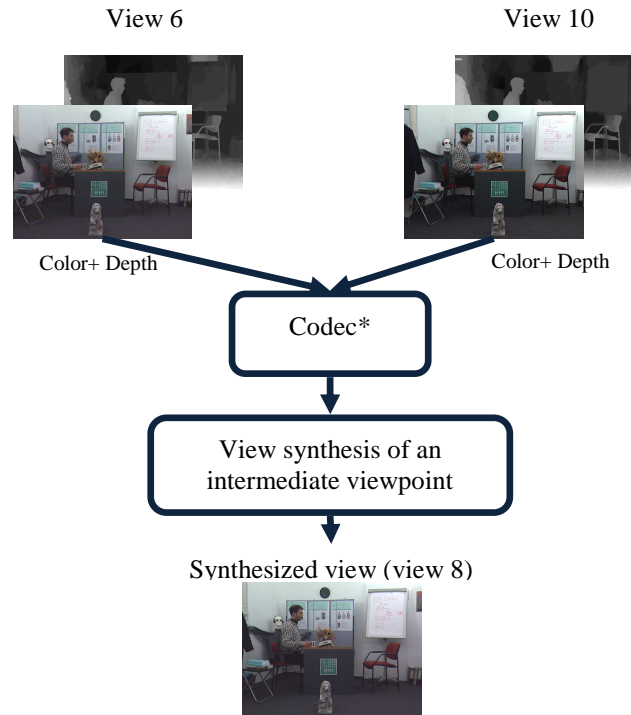
To address the second constraint regarding the consistency between color and depth edges, reconstruction step is included at the depth map decoder side, right after the first estimation of the smallest blocks, as explained in Section 3.1.2. The additional step described in this section can be considered as a second pass of the depth reconstruction. It consists of a multi-lateral filtering aided by the quad-tree representation whose principles are partially based on the description of *Lai et al.* method²¹. In our proposed method, the decoded associated color image is used to enhance only the blocks smaller or equal to 4×4 in the depth map. Small blocks are likely to be located around depth discontinuities. Thus, it is believed that improving the accuracy in these regions, according to the decoded associated color, will ensure consistency between color and depth edges. Let \tilde{C} be the decoded associated color image, and \tilde{L}_0 the lowest level image of the depth pyramidal decomposition. Let Ω be the set of pixels such as:

$$\Omega = \tilde{L}_0(x, y) | \tilde{L}_0(x, y) \in \tilde{L}_0(b^N(i, j)), \quad N \in [N_{min} \dots 4] \quad (5)$$

The reconstruction, noted $\tilde{L}_{0r}(x, y)$, of any pixel belonging to Ω is expressed as:

$$\forall \tilde{L}_0(x, y) \in \Omega, \tilde{L}_{0r}(x, y) = \tilde{L}_{0r}(p) = \frac{1}{K} \sum_{q \in \Gamma} \tilde{L}_0(p) e^{-\frac{\|p-q\|}{2\sigma_d}} e^{-\frac{\|\tilde{L}_0(p)-\tilde{L}_0(q)\|}{2\sigma_s}} e^{-\frac{\|Luma(p)-Luma(q)\|}{2\sigma_c}}, \quad (6)$$

$$K = \sum_{q \in \Gamma} e^{-\frac{\|p-q\|}{2\sigma_d}} e^{-\frac{\|\tilde{L}_0(p)-\tilde{L}_0(q)\|}{2\sigma_s}} e^{-\frac{\|Luma(p)-Luma(q)\|}{2\sigma_c}}, \quad (7)$$



*: H.264, or Proposed method

Figure 7: Overview of the experimental protocol

4.2 Results

For the performance comparisons, a pixel-based metric (PSNR) and a more perception-oriented metric VIF (Visual Information Fidelity²³) are considered. Figure 8 depicts the rate-distortion curve obtained by computing the PSNR scores, and the VIF scores of the synthesized views, with respect to the original acquired view. At high bit-rates (higher than 2bpp), the proposed method obtains better PSNR scores. However, under 2bpp, H.264 performs better.

The curve based on VIF scores shows that H.264 and the proposed method give similar results at high bit-rates (higher than 2bpp). However, contrary to the curve based on PSNR, the curve based on the perception-oriented VIF shows that the proposed method performs better at low bit-rates.

A visual appreciation is also useful to evaluate the methods. Figure 9 gives snapshots of the obtained synthesized views for 0.1bpp and 0.9bpp. Ghosting effect is perceptible with both methods behind the head of the man. However, the quad-tree based method preserves better the vertical edges: the vertical dark lines of the posters are better rendered with the data encoded with the proposed method. At low bit-rate (0.1bpp), Figure 9 gives snapshots of the synthesized views. Although, PSNR score shows lower performances for the proposed method at low bit-rate, the observation of Figure 9 shows improvements around the edges of the synthesized objects. The ghosting effect around the head of the man is less strong with the proposed method. The crumbling artifacts occurring around the leg of the chair at 0.1bpp with H.264 are no longer perceptible with the proposed method.

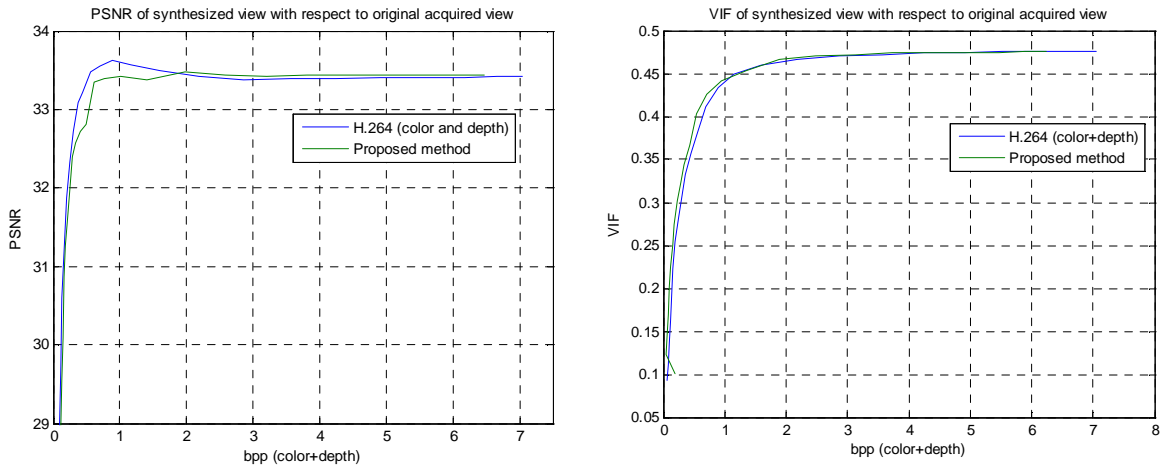


Figure 8: Performance comparisons, in terms of PSNR and VIF, between the original view and the synthesized view.



Figure 9: Snapshots of synthesized views from data encoded with H.264 and from data encoded with the proposed method.

5. CONCLUSION

We proposed a novel framework whose main purpose is to preserve consistency between color and depth edges. Depth encoding is based on a 2D still image codec, namely LAR (Locally Adapted Resolution). It consists in a quad-tree

representation of the images. The quad-tree representation contributes in the preservation of edges in depth data. The originality of the proposed method relies on the proposed quantization method and the use of the decoded color data as an anchor for the associated depth enhancement at the decoder side. The proposed method showed visual performances similar to H.264 at high bit-rates and some improvements at lower bit-rates because it preserves better the object edges. Future work should focus on the use of a more perception-oriented criterion for the quad-tree decomposition. A method to choose automatically the Y model, for any sequence, should also be investigated. Finally, the method should be extended to exploit temporal redundancies in the whole sequence.

6. ACKNOWLEDGMENTS

This work is partly supported by the French ANR-PERSEE project n° ANR-09-BLAN-0170, and ANR-CAIMAN project n° ANR-08-VERS-002.

REFERENCES

- [1] Smolic, A., Mueller, K., Merkle, P., Fehn, C., Kauff, P., Eisert, P., and Wiegand, T., "3D Video and Free Viewpoint Video—Technologies, Applications and MPEG Standards," in Proceedings of the IEEE International Conference on Multimedia and Expo (ICME'06), 2161–2164 (2006).
- [2] Sullivan, G.J., Topiwala, P., and Luthra, A., "The H. 264/AVC advanced video coding standard: Overview and introduction to the fidelity range extensions," in Presented at the SPIE Conference on Applications of Digital Image Processing XXVII Paper No 5558, 53 (2004).
- [3] Merkle, P., Muller, K., Smolic, A., and Wiegand, T., "Efficient compression of multi-view video exploiting inter-view dependencies based on H. 264/MPEG4-AVC," in Proc. ICME, 9–12 (2006).
- [4] Muller, K., Smolic, A., Dix, K., Merkle, P., and Wiegand, T., "Coding and intermediate view synthesis of multiview video plus depth," presented at ICIP09, November 2009, 741-744.
- [5] Merkle, P., Smolic, A., Muller, K., and Wiegand, T., "Efficient prediction structures for multiview video coding," IEEE Transactions on circuits and systems for video technology 17(11), 1461–1473 (2007).
- [6] Fehn, C., "Depth-image-based rendering (DIBR), compression and transmission for a new approach on 3D-TV," in Proceedings of SPIE Stereoscopic Displays and Virtual Reality Systems XI 5291, 93–104 (2004).
- [7] Merkle, P., Morvan, Y., Smolic, A., Farin, D., Muller, K., and Wiegand, T., "The effects of multiview depth video compression on multiview rendering," Signal Processing: Image Communication 24(1-2), 73-88 (2009).
- [8] Morvan, Y., Farin, D., and de With, P.H.N., "Joint depth/texture bit-allocation for multi-view video compression," in Proceedings of Picture Coding Symposium (PCS 2007) 10, 4349 (2007).
- [9] Daribo, I., Tillier, C., and Pesquet-Popescu, B., "Motion Vector Sharing and Bit-rate Allocation for 3D Video-plus-Depth Coding," in EURASIP JASP Special Issue on 3DTV, p. 258920 (2009).
- [10] Liu, Y., Huang, Q., Ma, S., Zhao, D., and Gao, W., "Joint video/depth rate allocation for 3D video coding based on view synthesis distortion model," Signal Processing: Image Communication 24(8), 666-681 (2009).
- [11] De Silva, D., Fernando, W., Kodikaraarachchi, H., Worrall, S., and Konoz, A., "A Depth Map Post-Processing Framework for 3D-TV Systems based on Compression Artifact Analysis," IEEE Journal of Selected Topics in Signal Processing PP(99), 1-1 (0).
- [12] Morvan, Y., de With, P.H.N., and Farin, D., "Platelet-based coding of depth maps for the transmission of multiview images," in Proceedings of SPIE, Stereoscopic Displays and Applications 6055, 93–100 (2006).
- [13] Graziosi, D.B., Rodrigues, N.M.M., Pagliari, C.L., de Faria, S.M.M., da Silva, E.A.B., and De Carvalho, M.B., "Compressing depth maps using multiscale recurrent pattern image coding," Electronics letters 46(5), 340–341 (2010).
- [14] Sarkis, M., and Diepold, K., "Depth map compression via compressed sensing," in Proceedings of the 16th IEEE international conference on Image processing, 737–740 (2009).
- [15] Bosc, E., Pepion, R., Le Callet, P., Koppel, M., Ndjiki-Nya, P., Pressigout, M., and Morin, L., "Towards a new quality metric for 3D synthesized view assessment," in IEEE Journal of Selected Topics (2011).
- [16] Pasteau, F., Babel, M., Déforges, O., Strauss, C., Bédard, L., and others, "Locally Adaptive Resolution (LAR) codec," Recent Advances in Signal Processing 37-48 (2010).
- [17] Merkle, P., Smolic, A., Mueller, K., and Wiegand, T., "Multi-view video plus depth representation and coding," in Proceedings of ICIP, 201–204 (2007).
- [18] Bosc, E., Jantet, V., Pressigout, M., Morin, L., and Guillemot, C., "Bit-rate allocation for multi-view video plus

depth,” in Proc. of 3DTV Conference 2011 (2011).

[19] Vetro, A., Yea, S., and Smolic, A., “Towards a 3D video format for auto-stereoscopic displays,” Proceedings of the SPIE: Applications of Digital Image Processing XXXI, San Diego, CA, USA (2008).

[20] Babel, M., Deforges, O., and Ronsin, J., “Interleaved S+ P pyramidal decomposition with refined prediction model,” in IEEE International Conference on Image Processing, 2005. ICIP 2005 2 (2005).

[21] Lai, P.L., Tian, D., and Lopez, P., “Depth map processing with iterative joint multilateral filtering,” PCS 2010 (2010).

[22] Tanimoto, M., Fujii, T., Suzuki, K., Fukushima, N., and Mori, Y., “Reference Softwares for Depth Estimation and View Synthesis,” presented at ISO/IEC JTC1/SC29/WG11 MPEG 2008/M15377, April 2008.

[23] Sheikh, H.R., and Bovik, A.C., “Image information and visual quality,” Image Processing, IEEE Transactions on 15(2), 430–444 (2006).