Efficient B-spline wavelets based dictionary for depth coding and view rendering

Dorsaf Sebai
Cristal laboratory, ENSI
Tunisia

Faten Chaieb
Cristal laboratory, ENSI
Tunisia

Faouzi Ghorbel
Cristal laboratory, ENSI
Tunisia

ABSTRACT

Video representations that support view synthesis based on depth maps, such as multiview plus depth, have been widely emerged raising interest in efficient depth maps coding tools. In this paper, we propose an innovative sparse decomposition on wavelets based dictionary specially designed for the piece-wise planar nature of depth signal. We also evaluate performances of the proposed dictionary for depth maps coding while paying special attention to the impact of depth coding errors on resulting synthesized images. Obtained results prove the relevance of the proposed scheme able to considerably improve the perceived quality of synthesized images.

Keywords
Depth maps, synthesized images, compression, B-spline wavelets based dictionary.

1. INTRODUCTION

Multiview Video plus Depth (MVD) includes sequences of texture images and their corresponding depth maps. The latter are bi-dimensional gray level images representing the distance of each pixel to capture camera. Recent efforts point toward an efficient coding that preserves depth maps particularities, namely their piece-wise planar conception and the critical impact of pixels near contours on perceptual quality of synthesized views [1].

In this context, many coding research work aim at faithfully reconstruct depth map specific piece-wise planar conception. Morvan et al [2] exploit the linear piece-wise nature of platelet and wedgelet functions to approximate depth planar surfaces separated by shaped edges. The wedgelet representation is retained for 3D High Efficiency Video Coding (3D-HEVC) standard [3]. Maitre et al. [4]. propose a codec that relies on a lifting implementation of Shape-Adaptive Discrete Wavelet Transform (SA-DWT). SA-DWT independently treats surfaces separated by edges which, and unlike classical wavelet transforms, provides much sparser decomposition with small coefficients along depth discontinuities. Furthermore, Shen et al. [5] present a new set of Edge-Adaptive Transform (EAT) as an alternative to the classical Discrete Cosine Transform (DCT). EAT avoids filtering across depth discontinuities and so avoids creating large coefficients. However, transform domains used in [4] [5] need an encoded representation of major edge locations to be shared between both encoder and decoder sides.

Since depth images are used for view synthesis and are not themselves displayed, later efforts aim at reducing depth maps coding artifacts that cause severe distortion of synthesized views. Cheung et al. [6] define ”Don’t Care Regions” (DCRs), for each pixel, where a depth value outside the DCR will lead to a synthesis distortion larger than a threshold value. Then, they perform sparsification of the depth map in an orthogonal basis, optimally trading off its representation sparsity and its adverse effect on synthesized view distortion. More recently, this idea is reused by Cheung et al. [7] replacing DCRs by penalty function. For each pixel, a quadratic penalty function is defined based on sensitivity of interpolated images to pixel depth values during rendering process. Transform domains used in [6] [7] are classical orthogonal basis that represent dictionaries of minimum size, concentrating the signal energy over a set of few vectors. However, vectors sets larger than basis, particularly redundant dictionaries, are needed to build sparse representations of complex signals. In the last few years, the emerging attention is to enlarge common orthogonal bases through the design of suitable redundant dictionaries positioned as an interesting alternative. The latter can be a mixture of orthogonal bases and/or dictionaries. Such merging approach aims to design domains where each sub-dictionary is suitable for representing one of the signal components. The approaches for learning dictionaries from large input data sets have also been envisioned in order to enhance the correlation of dictionary atoms to signals. However, learned dictionaries are further sensitive to image variations of practical scenarios. Furthermore, if the learning process of the dictionary cannot be repeated in the decoder side, the dictionary transmission is necessary. Increases in terms of storage expense and codec complexity are also noticed due to the feature-dependent nature of learned transform domains.

In this paper, we are interested in studying a predefined mixed dictionary adapted to depth maps sparse representation. In fact, many efforts were carried out to study the most appropriate dictionary for a given class of images such as astronomical images and cartoon-images. This is not the case for the particular class of depth maps. Being redundant, the proposed dictionary, unlike orthogonal basis used in [6] [7], promotes sparsity and avoids high coefficients mainly near contours. Being predefined, the proposed dictionary, unlike the non-fixed EAT and SA-DWT, does not imply a coding overhead for the transform reconstruction in the decoder side. The proposed dictionary is then exploited for depth maps compression to evaluate its relevance for synthesis quality.

Section 2 brings particular attention to fundamental concepts of sparse representations. In Section 3, we aim at studying an efficient dictionary in terms of depth maps sparsity-distortion tradeoff. The dictionary is then exploited, in Section 4, for compression purpose taking into account the quality of view synthesis process, the ultimate depth maps application.

2. SPARSE REPRESENTATIONS

Classical transform coding techniques make use of orthogonal basis, such as Fourier and cosine basis. In such transform domains, signal representation is unique. More recently, sparse representation concept has been developed and its exploitation in image processing is increasingly expanding. Sparse representations proved their performances for texture images compression. It is therefore interesting to explore them for depth maps compression.

Sparse representations distinguish significant components of a signal as a small number of elementary signals selected from a very large transform domain, named redundant dictionary. Sparse representations aim at finding a representation y of the original signal as a compact linear combination of a small atoms number weighted by transform coefficients : $y = D\ x$ where $y \in \mathbb{R}^M$ the representative vector of the original signal of dimension M and $D \in \mathbb{R}^{M\times n}$ a dictionary of n atoms with $n >> M$. $x \in \mathbb{R}^n$ is a sparse vector of transform coefficients. Sparsity of vector x refers to the number of zero coefficients it contains. Because of dictionary redundancy, signal representation is not unique and several combination of vector x are possible. The most appropriate combination corresponds to the sparsest one, i.e. the one with the fewest non-zero coefficients. The Orthogonal Matching Pursuit (OMP) [8] is one of the most developed decomposition algorithms devoted to search such a combination. OMP is a greedy multistage decomposition algorithm that selects, at each iteration, the most correlated atom to the original signal and then subtracts its contribution. This process is iteratively repeated for the residual signal in order to achieve an approximation tolerating an admissible reconstruction error ρ.

3. DEPTH MAPS SPARSITY-DISTORTION TRADEOFF

Efficiency of depth maps representation, both in terms of sparsity and similarity to original data, highly depends on transform domain choice. It seems useful, even required, to use atoms highly correlated to depth maps that we try to model.

3.1. Discrete Cosine/Linear Discrete B-Spline Wavelets dictionary

As introduced in Section 1, depth maps include two major components, namely smooth regions and depth discontinuities. Then, it is suitable to combine, in the same dictionary, two sets of atoms conducive to each of them. In that way, we guarantee complementarity of concatenated atoms where each type of them is capable of reconstructing some signal characteristics that the other one is unable to efficiently do. Typically, we propose the Discrete Cosine/Linear Discrete B-Spline Wavelets (DC/LDBSW) dictionary that includes two kinds of atoms :
Discrete Cosine (DC) atoms for smooth regions : DC atoms of (1) are stemmed from discrete cosine transform :

$$DC = \left\{ \cos\left(\frac{\Pi(2i-1)(k-1)}{2n}\right), i \in \{1, ..., M\}, k \in \{1, ..., n\} \right\} \tag{1}$$

where M is the signal dimension. The dictionary size n is equal to rM with $r \in \mathbb{N}^*$. If $r = 1$, DC is an orthogonal basis. Otherwise, DC is a dictionary of redundancy r. The DC atoms are indisputably adapted to smooth areas representation. This is even more valid for depth maps where smooth areas do not present texture, such as for natural images, but distances of scene objects to capture cameras.
Linear Discrete B-Spline Wavelets (LDBSW) for depth discontinuities : LDBSW atoms, defined by 2, are translated and discretized versions of linear B-spline wavelets at different resolution levels j. The discretization consists in considering the linear B-spline wavelets values at equally spaced knots on a compact interval with distance $\frac{\mathbb{Z}}{2^{j+1}}$ between two adjacent knots :

$$LDBSW = \left\{ \varphi_2(i \quad - k),\ i \ \in [1, M] \cap \mathbb{Z} \right\} \cup \left\{ 2^{\frac{j}{2}} \psi_2(2^j i - h),\ i \ \in [1, M] \cap \frac{\mathbb{Z}}{2^{j+1}} \right\}_{j \ \in \ [0, log_2(M)-1] \cap \mathbb{Z}} \quad (2)$$

where $k \in [0, M[\cap\mathbb{Z}, h \in [0, 2^j M[\cap\mathbb{Z}$ and

$$\varphi_2(x) = \delta_{x,1}$$

$$\psi_2(x) = \frac{1}{12}\varphi_2(2x) - \frac{1}{2}\varphi_2(2x-1) + \frac{5}{6}\varphi_2(2x-2) - \frac{1}{2}\varphi_2(2x-3) + \frac{1}{12}\varphi_2(2x-4)$$

φ_2 and ψ_2 are scale and wavelet functions. M is the signal dimension and j the resolution level ranging from 0 to $log_2(M) - 1$. k and h are translation parameters of φ_2 and ψ_2, respectively. The cut-off approach is used for translation of φ_2 and ψ_2 at signal interval boundaries. This introduces redundancy by considering all the wavelet functions having non-trivial intersection with the interval. LDBSW atoms are piece-wise linear so they comply with the piece-wise planar definition of depth maps. Furthermore, the particular B-splines wavelets of order 2, compared to higher order B-splines wavelets, greatly limit edge smoothing and model in a better way the sharp depth details. This allows a good quality for view synthesis, the ultimate depth maps application.

To build the bi-dimensional dictionary suitable for image processing, the tensor product of the so constructed unidimensional dictionary with itself is considered.

3.2. Complementary of DC and LDBSW atoms

We consider two (8×8) blocks that respectively present areas with and without contours of *Breakdancers* [9] depth map. The original signal in figure 1 corresponds to the concatenated columns of the block into a single one-dimensional vector of depth values. The residual signal results from few OMP iterations. As shown in figure 1, the residual signal of the smooth block is already uniform and is set around zero. However, residue of the block containing discontinuities is not yet uniform and requires more iterations. We can retrieve the piece-wise linear shape of LDBSW atoms that will be useful to reduce this residual signal in next iterations of OMP.

3.3. Comparison and discussion

In this section, we aim at judging the efficiency of DC/LDBSW dictionary for depth maps sparse representation. As already mentioned, the DC atoms stemmed from discrete cosine transform are well suited for smooth areas approximation. It then remains to assess reliability of LDBSW atoms for depth discontinuities representation. For this purpose, we compare the DC/LDBSW dictionary to Discrete Cosine/Linear Discrete B-Spline (DC/LDBS) dictionary, where LDBS atoms are translated and discretized versions of linear B-spline functions of different supports. We also carry out a comparison to Discrete Cosine/Cubic Discrete B-Spline Wavelets (DC/CDBSW) and Discrete Cosine/Directional Anisotropic Atoms (DC/DAA). As well as LDBSW dictionary, atoms of CDBSW dictionary are stemmed from discrete B-spline wavelets. The difference lays in the mother wavelet order that it is no longer linear. Used B-spline wavelets in CDBSW dictionary are cubic (i.e. order 4). Being the successors of X-lets, atoms of DAA dictionary are 2D non-separable functions built by applying geometric transformations to a generating mother function [10]. The latter is a smooth low resolution function in the direction of the contour, and behaves like a wavelet in the orthogonal direction. Using LDBS, CDBSW and DAA dictionaries for comparison is not randomly made. The latter have proved among the most pertinent for signal sparse representation. Furthermore, comparison to these dictionaries would allow us to stress the relevant properties of LDBSW atoms for depth maps sparse representation.

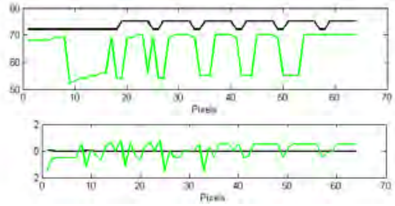


Figure 1. Original signal (top) and residual signal (bottom) issued from few OMP iterations for blocks of *Breakdancers* depth map : smooth block (black) and block with discontinuities (green).

As comparison criterion, we make use of Sparsity Ratio (SR) metric. It is defined as the number of pixels in the image divided by the number of non-zero coefficients used for its representation. A high value of SR reflects the dictionary ability to represent signals with the least number of transform coefficients. Figure 2 presents SR values obtained by sparse decomposition of *Breakdancers*, *Ballet* [9] and *Champagne* sequences on candidate dictionaries using OMP algorithm for different PSNR values.

As shown in figure 2, DC/LDBSW dictionary achieves higher SR values than DC/LDBS. This is thanks to the oscillatory behavior of LDBSW atoms that makes them visually more similar to OMP residual signals than LDBS atoms (*see* figure 1). Compared to DC/CDBSW, DC/LDBSW dictionary allows better sparsity-distortion performances. In fact, LDBSW atoms are B-spline wavelets of lower order than CDBSW ones. This allows them to strongly limit depth discontinuities smoothing, which is crucial for view synthesis.

For 1D signals, wavelets are recognized to be efficient

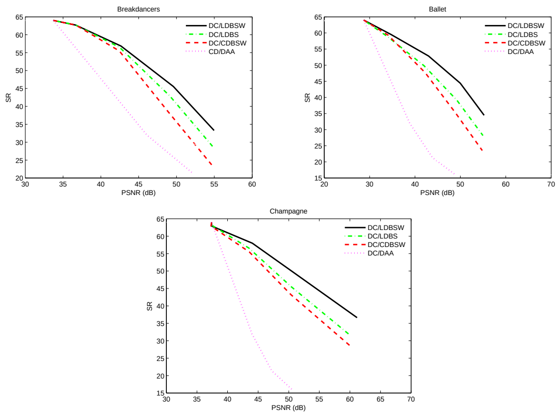


Figure 2. SR values obtained by sparse decomposition of *Breakdancers*, *Ballet* and *Champagne* sequences on DC/LDBSW, DC/LDBS, DC/CDBSW and DC/DAA dictionaries using OMP algorithm for different PSNR values.

for sparse representation of piece-wise smooth singularities. Despite their success, wavelets lose their optimality when extending them to 2D. They fail to detect regularity of contours. In order to overcome the non-optimality of 2D wavelets, it has been proposed to use geometrical-oriented atoms, i.e. the X-lets. Recently, efforts have been made towards redundant dictionaries of transformed generating function using, as DC/DAA, anisotropic geometric transformations. However, geometric atoms relevance for smooth and regular contours of natural images significantly decreases for sharp and irregular discontinuities of depth maps. In fact, DC/LDBSW dictionary achieves, as shown in figure 2, sparser depth maps representation than DC/DAA. This is particularly clear for depth maps with strong discontinuities such as *Ballet* and *Champagne*.

4. SYNTHESIZED VIEWS RATE-DISTORTION TRADEOFF

As a conclusion of the previous section, DC/LDBSW combination allows the best depth maps sparsity-distortion performances against other candidate combinations. This may presage efficient results for depth maps compression. Thus, we integrate DC/LDBSW dictionary within a depth maps compression scheme taking into account the quality of rendered views.

4.1. Compression method

As it has been observed that efficient depth maps compression is achieved by applying a down-sampling prior to encoding [11], the proposed scheme carries out a decimation by a factor of 2 of the initial depth map. One in two pixels is retained per row and per column. Then, an edge detection is applied to decimated depth map. Resulting edge image is next divided into blocks labeled as 1, if they include contours, and 0 otherwise. Sparse representation of each block is performed using the OMP algorithm on DC/LDBSW dictionary.

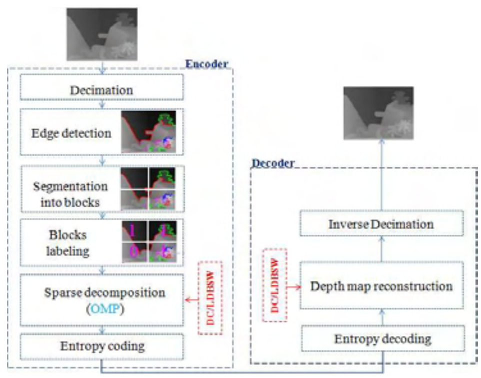


Figure 3. Flowchart of the proposed method.

As stated in Section 1, coding distortions near contours lead to harmful artifacts of synthesized views. Whereas, coding degradation in smooth surfaces has limited impact on synthesized views quality. Then, we typically adapt the stopping criterion of OMP algorithm to the nature of depth maps blocks, whether they contain contours or not. In order to favor sparsity for smooth blocks (i.e. labelled as 0), the approximation issued from the first iteration of the OMP algorithm is sufficient. In fact, smooth block decomposition on DC/LDBSW dictionary provides, thanks to DC atoms, a uniform residual signal around zero since the first OMP iteration (see figure 1).

On the contrary, blocks with contours (i.e. labelled as 1) are handled as regions of interest where distortions have to be minimized in order to achieve a good synthesis quality. To do this, OMP algorithm has to iterate until the error between the original and the approximated signals is under a fixed reconstruction error. The simple usage of depth map quadratic error can lead to suboptimal results since it only measures coding artifacts and does not reflect the real impact of the latter on the final rendering quality. Therefore, we make use of the quadratic error in the synthesized frame and not in the depth map itself. We particularly use the distortion metric of Kim et al. [1] that takes into consideration camera parameters and proves the proportional relation between the quadratic error in the synthesized view and the absolute error in the depth map.

4.2. Experimental results and analysis

Since the main use of depth maps is in view synthesis operations, experimentations are concerned with the evaluation of views that can be synthesized from already compressed depth images. The following experimentations consist in coding left and right views from *Breakdancers*, *Ballet* and *Champagne* sequences. The decoded views are then used for view synthesis using View Synthesis Reference

Software (VSRS) [12] of Nagoya University. We note that from each test data sets, the first 16 frames were used.

To evaluate the DC/LDBSW dictionary interest for compression performances, we compare results obtained by the proposed compression method with DC/LDBSW dictionary to those obtained by the same method with DC/LDBS dictionary. We particularly choose DC/LDBS dictionary for comparison since it is the most competitive one to DC/LDBSW dictionary in terms of sparsity, as shown in figure 2. Performances of the proposed scheme with DC/LDBSW dictionary are also compared to 3D-HEVC, the ongoing 3D compression standard. We do compare our method to the 3D-HEVC standard since it is the latest reference for comparison that includes latest efforts of 3D research community being approved by MPEG. We typically make use of 3D-HEVC Test Model version 4.1 (3D-HTM 4.1) [13] for which temporal and inter-view predictions are disabled because our method does not involve them.

The performances of candidate methods are compared in terms of rate-PSNR tradeoff of synthesized views. Moreover, we make use of the new human visual system based metric, Structural SIMilarity plus (SSIMplus) [14]. We also propose the visual evaluation of areas zoomed from synthesized views.

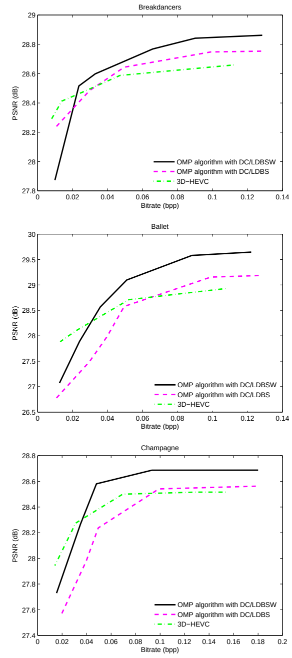


Figure 4. Rate/PSNR curves of *Breakdancers*, *Ballet* and *Champagne* synthesized views obtained from original textures and depth maps encoded using 3D-HEVC and the proposed method with DC/LDBS and DC/LDBSW dictionaries.

4.2.1 PSNR vs. Bitrate

Figure 4 depicts performances of candidate methods in terms of Bitrate-PSNR of synthesized views for *Breakdancers*, *Ballet* and *Champagne* sequences. Results of figure 2 have proved relevance of DC/LDBSW dictionary, against DC/LDBS one, in terms of sparsity. Figure 4 comes to show that DC/LDBSW dictionary is also better than DC/LDBS in terms of Bitrate-PSNR of synthesized views. Compared to 3D-HEVC, DC/LDBSW dictionary integrated within the proposed scheme provides better performances for medium and high bitrates, achieving a gain of $0.1\ dB$ at $0.1\ bpp$ for *Breakdancers*, $0.4\ dB$ at $0.08\ bpp$ for *Ballet* and $0.2\ dB$ at $0.1\ bpp$ for *Champagne*. However, 3D-HEVC allows better performances at very low bitrates since quantized values of wedgelet coefficients are restricted compared to indices of atoms dictionary that cannot be quantized.

4.2.2 SSIMplus Index

Since PSNR is a pure mathematical metric, we propose to use a new full-reference measure, SSIMplus. It provides real-time prediction of the perceptual quality of a video based on human visual system behaviors, video content characteristics, e.g. spatial and temporal complexity and video resolution, display device properties, e.g. screen size, resolution, and brightness, and viewing conditions, e.g. viewing distance and angle. Compared to most popular and widely used quality assessment measures, SSIMplus has shown a higher perceptual quality prediction accuracy and closer performances to Mean Opinion Scores [14].

Table 1 shows SSIMplus results of candidate methods obtained for test sequences at $0.01\ bpp$, $0.05\ bpp$ and $0.1\ bpp$. The evaluation is performed at these different bitrates that correspond to three critical values, namely low, medium and high bitrates. As already mentioned, the first 16 frames were used from each test data sets. It is clear from table 1 that the proposed compression scheme with DC/LDBSW dictionary produces better SSIMplus results against DC/LDBS dictionary. Confronted to the ongoing 3D-HEVC standard, the DC/LDBSW dictionary achieves a mean gain of 2 at $0.05\ bpp$ and 4 at $0.1\ bpp$. At $0.01\ bpp$, better results are performed by 3D-HEVC, achieving a mean gain of 2.

4.2.3 Zoomed areas

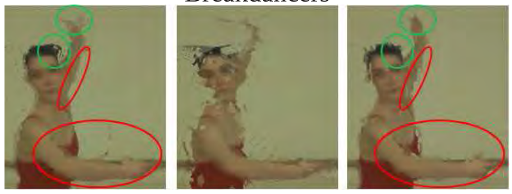
Besides PSNR and SSIMplus Human Visual System-based measure, figure 5 allows visual evaluation of areas zoomed from synthesized views of *Breakdancers*, *Ballet* and *Champagne* sequences. Since 3D-HEVC performances are better than those of our method at low bitrates, the visual evaluation is performed at $0.01bpp$. Compared to DC/LDBS dictionary, the proposed method with DC/LDBSW dictio-

Table 1. SSIMplus values of *Breakdancers*, *Ballet* and *Champagne* synthesized views obtained from original textures and depth maps encoded using 3D-HEVC and the proposed method with DC/LDBS and DC/LDBSW dictionaries at $0.01\ bpp$, $0.05\ bpp$ and $0.1\ bpp$.

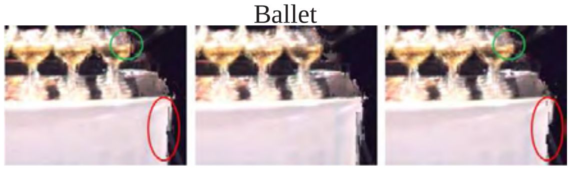
Sequence	Method	0.01 *bpp*	0.05 *bpp*	0.1 *bpp*
Breakdancers	3D-HEVC	29	34	38
	DC/LDBS	28	35	38
	DC/LDBSW	27	37	41
Ballet	3D-HEVC	32	42	43
	DC/LDBS	29	40	46
	DC/LDBSW	30	43	47
Champagne	3D-HEVC	39	46	47
	DC/LDBS	37	45	48
	DC/LDBSW	37	47	50
Mean	3D-HEVC	33	40	42
	DC/LDBS	31	40	44
	DC/LDBSW	31	42	46

nary can clearly achieve better visual synthesis quality with much less harmful distortions. Compared to 3D-HEVC, the proposed method with DC/LDBSW dictionary can achieve a competitive synthesis quality despite the outperformance of the latter at this bitrate in figure 4 and table 1. As examples, we distinguish areas circled in red where 3D-HEVC outperforms our method. The latter allows however better quality than 3D-HEVC for areas marked in green.

Breakdancers



Ballet



Champagne

Figure 5. Zoomed areas of views synthesized from depth maps encoded at $0.01 bpp$ using : 3D-HEVC (left), the proposed compression scheme with DC/LDBS dictionary (middle) and the proposed compression scheme with DC/LDBSW dictionary (right).

5. CONCLUSION

In this paper, we have combined depth maps compression and sparse representations that proved to be particularly relevant for compression purposes. Typically, we aimed to propose a redundant mixed dictionary adapted to depth maps sparse representation. Experimental results lead to the conclusion that it is the combination of a discrete cosine dictionary with well-localized linear B-spline wavelet atoms that yields a significant improvement in the sparsity of high-quality approximations of depth maps. Applied for depth maps compression, DC/LDBSW dictionary also shows good tradeoffs between bitrate and distortion of synthesized views. As perspective, we aim to propose an approach allowing a joint compression of the two components of MVD, namely texture and depth. In addition to the sparsity ratio, we aim to study the DC/LDBSW dictionary efficiency in terms of other comparison criteria that take into account the redundancy and the coherence of the proposed dictionary. Studies and comparisons to learned dictionaries are also in our scope.

6. REFERENCES

[1] W.-S. Kim, A. Ortega, P. Lai, D. Tian, and C. Gomila, "Depth map distortion analysis for view rendering and depth coding," in *IEEE International Conference on Image Processing*, 2009.

[2] P. Merkle, Y. Morvan, A. Smolic, D. Farin, K. Muller, P. H. N. de With, and T. Wiegand, "The effects of multiview depth video compression on multiview rendering," *Signal Processing: Image Communication Journal*, 2008.

[3] K. Muller, H. Schwarz, D. Marpe, C. Bartnik, S. Bosse, H. Brust, T. Hinz, H. Lakshman, P. Merkle, H. F. Rhee, G. Tech, M. Winken, and T. Wiegand, "3D high efficiency video coding for multi-view video and depth data," *IEEE Transactions on Image Processing, Special Issue on 3D Video Representation, Compression and Rendering*, vol. 22, pp. 3366–3378, 2013.

[4] M. Maitre and M. N. Do, "Depth and depth-color coding using shape-adaptive wavelets," *Journal of Visual Communication and Image Representation*, vol. 21, pp. 513–522, 2010.

[5] G. Shen, W.-S. Kim, S. K. Naran, A. Ortega, L. Jaejoon, and W. HoCheon, "Edge-adaptive transforms for efficient depth map coding," in *Proceedings of IEEE Picture Coding Symposium*, 2010.

[6] G. Cheung, A. Kubota, and A. Ortega, "Sparse representation of depth maps for efficient transform coding," in *Proceedings of the 28th Picture Coding Symposium*, 2010.

[7] G. Cheung, J. Ishida, A. Kubota, and A. Ortega, "Transform domain sparsification of depth maps using iterative quadratic programming," in *Proceedings of the IEEE International Conference on Image Processing*, 2011.

[8] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit recursive function approximation with

applications to wavelet decomposition," in *Proceedings of IEEE Asilomar Conference on Signals Systems and Computers*, 1993.

[9] *http://research.microsoft.com/en-us/downloads/*.

[10] R. M. F. i Ventura, P. Vandergheynst, and P. Frossard, "Low rate and flexible image coding with redundant representations," *IEEE Transactions on Image Processing*, vol. 15, pp. 726–739, 2006.

[11] K. Klimaszewski, K. Wegner, and M. Domanski, "Influence of views and depth compression onto quality of synthesized views," *ISO/IEC JTC1/SC29/WG11, M16758*, UK, 2009.

[12] M. Tanimoto, T. Fujii, K. Suzuki, N. Fukushima, and Y. Mori, "Reference softwares for depth estimation and view synthesis," *ISO/IEC JTC1/SC29/WG11 MPEG2008/M15377,* France, 2008.

[13] *https://hevc.hhi.fraunhofer.de/svn/svn 3DVCSoftware/tags/HTM-4.1*.

[14] A. Rehman, K. Zeng, and Z. Wang, "Display device-adapted video quality-of-experience assessment," in *Electronic Imaging, Human Vision and Electronic Imaging XX*, 2015.