IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 12, 2018, accepted February 20, 2019, date of publication February 28, 2019, date of current version March 20, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2902197

Block DCT Based Optimization for Wireless SoftCast of Depth Map

TAIHAI YANG[1], LEI LUO[1,2], (Member, IEEE), CE ZHU[2], (Fellow, IEEE), AND SHU TANG[3]
[1]School of Communications and Information Science, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
[2]School of Information and Communication, University of Electronic Science and Technology of China, Chengdu 611731, China
[3]College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Corresponding author: Lei Luo (luolei@cqupt.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61501074 and Grant 61601070, in part by the Fundamental and Advanced Technology Research Project of Chongqing under Grant cstc2015jcyjA40012 and Grant cstc2016jcyjA0455, and in part by the Science and Technology Research Project of Chongqing Municipal Education Commission under Grant KJ1500430 and Grant KJ1600411.

ABSTRACT With the recent emergence of naked-eye 3D mobile devices, the 3D video-based applications on the mobile terminal are showing great potential. However, the conventional transmission scheme based on separate source coding and channel coding may fail in the broadcast and mobile scenarios due to the drastic changes in wireless channel condition. Recently, an uncoded transmission scheme called SoftCast was introduced, which shows efficient performance for wireless 2D video transmission. However, it may not work well for 3D video wireless transmission, especially for the depth maps. Based on the piecewise smoothness of depth maps, we propose an uncoded wireless depth map transmission scheme, which incorporates three customized components of optimization into the SoftCast framework. First, compared with the frame-level discrete cosine transform (DCT) used in 2D video SoftCast, the mean removed block-based DCT is preferred as the decorrelation transform for depth maps to accommodate its piecewise smooth property, which also facilitates the following subsequent optimization. Second, view synthesis distortion-based chunk dropping is utilized for more efficient power allocation in the case of limited bandwidth, where the block DCT provides a better estimation for synthesis distortion. Third, inter-block DCT coefficients are rearranged to form chunks in terms of minimizing the intra-chunk variance, further enhancing the chunk-level power allocation efficiency. The experimental results demonstrate that the proposed scheme can provide superior performance than the state-of-the-art SoftCast methods while achieving graceful synthesis quality transition within a wide channel signal-to-noise ratio range.

INDEX TERMS 3DV, depth map, wireless transmission, SoftCast, block DCT, view synthesis, power allocation.

I. INTRODUCTION

Three dimensional video (3DV) is an exciting media with the extraordinary visual experience, which could provide arbitrary views of the real scene for audiences. It has been utilized as the visual content of the well-known 3D Television (3DTV) [1] and free viewpoint television (FTV) [2]. In order to efficiently represent the depicted 3D scene, one format called multi-view video plus depth (MVD) [3] has been extensively adopted for 3DV. The MVD format is composed of multi-view texture videos and their associated depth maps. The depth maps are used to record the geometric information of the depicted 3D scene. The most impressive advantage of the MVD format is that it could easily synthesize an arbitrary viewpoint in a region by using the depth-image-based rendering (DIBR) technology [4].

Meanwhile, with the rapid development of the wireless communication (e.g., LTE, and mobile cloud computing [5], [6]) and the recent emergence of the naked-eye 3D mobile devices (e.g., HTC EVO 3D, LG Optimus 3D, Sharp Lynx, and numerous 3D-enabled laptops), the 3D video based applications on mobile clients are showing great potential. However, the drastic channel variation of the wireless network is a great challenge for 3D video transmission, and it really hampers the development of the mobile 3D video applications.

The associate editor coordinating the review of this manuscript and approving it for publication was Dian Tjondronegoro.

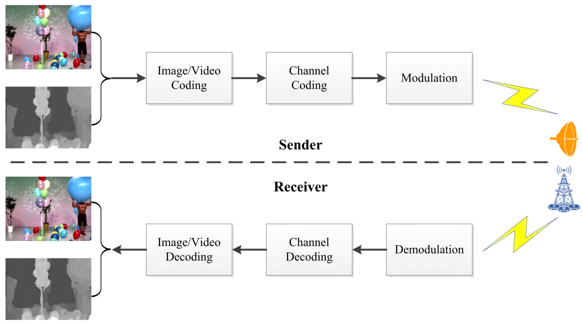

FIGURE 1. **Conventional video transmission framework.**

The conventional video transmission is built upon the *Shannon's theorem* [7], [8], where the source coding and channel coding are designed separately. As shown in FIGURE 1, the source coding (e.g., HEVC [9]) is used to eliminate the redundancy within the original video signal, while the channel coding (e.g., Turbo code, and convolutional code) is utilized to deal with the transmission errors over the wireless channel. This kind of video transmission strategy could obtain the optimal performance with the assumption that the channel condition is known before the source and channel coding to choose the appropriate coding rate [10]. Apparently, it is not suitable for the wireless scenarios where the channel condition may vary drastically. The main drawback is the *cliff effect*: if the actual channel quality falls below a specific threshold, the reconstructed quality will degrade sharply; while the actual channel quality is beyond another threshold, the reconstructed quality could not be improved any more.

Recently, a novel uncoded (or pseudo-analog) scheme called SoftCast was proposed for wireless video communication [11]–[13]. SoftCast is a clean-slate end-to-end architecture for transmitting video signal over wireless channels. The conventional video transmission applies the source and channel coding to transform the input video signal into a binary stream for modulation and transmission over wireless channels. That means the transmitted signal is the *binary* stream. Instead, the SotfCast transforms the original video signal into a series of real value coefficients from which the inverse transform could achieve the perfect reconstruction. The real value stream is directly modulated to a dense constellation for transmission, which means the transmitted signal of SoftCast is the *real value* stream. Such transmission scheme is lossy in nature and the channel noise is directly translated into the reconstruction error of the video signal. It makes the reconstruction quality commensurate with the channel signal-to-noise ratio (SNR) of the receiver. In such a way, SoftCast eliminates the cliff effect and achieves elegant quality transition within a wide range of channel SNR. Also, it could serve plenty of receivers with different channel conditions simultaneously with a single transmitted signal broadcasted by the sender.

Although the SoftCast scheme presented efficient performance for traditional 2D video wireless transmission, it may not work well for 3D video, especially for the wireless transmission of depth map. The reason is as follows. First, the depth map is just an intermediate data which is used to provide the 3D geometric information for the DIBR process. The distortion in the depth map will lead to a geometric error in the rendered virtual view. Therefore, the goal of wireless depth map transmission is to guarantee the highest possible quality of the synthesized virtual view rather than that of the depth map itself. Second, the depth map shows completely different characteristic to common 2D natural video signals. it usually contains piecewise smooth regions separated by sharp edges. An efficient depth map transmission strategy should take full advantage of this intrinsic property.

In view of the utilization and characteristics of the depth map, we propose an efficient uncoded depth map transmission scheme over wireless channels for 3D video. The contribution of this paper are the following aspects. First, due to the piecewise smooth property of the depth map, the mean removed block DCT is shown to be the preferable decorrelation transform than the frame level DCT used in 2D video SoftCast. Second, a view synthesis distortion based chunk dropping strategy is presented for more efficient power allocation in the case of limited bandwidth, where the block DCT synergizes better estimation of synthesis distortion. Third, an inter-block DCT coefficients rearrangement is developed to form chunks in terms of minimizing the intra-chunk variance, further enhancing the chunk level power allocation efficiency.

The rest of the paper is organized as follows. The related work is reviewed in Section II. In Section III, the proposed depth map transmission scheme is presented in details. In Section IV, experimental results are demonstrated to validate the proposed scheme. Finally, a conclusion is drawn in Section V.

II. RELATED WORKS

A. CONVENTIONAL 3D VIDEO TRANSMISSION

In the general video transmission framework, the source coding is followed by the channel coding. For 3D video, the most widely used source coding is 3D-HEVC [14], which provides high coding efficiency by fully exploiting data correlation in the 3D video signal. Meanwhile, to deal with the transmission error, unequal error protection (UEP) approaches are commonly used [15]. By exploiting the characteristics of the 3D video, the UEP schemes implement different level of protection for bit-streams of different level of importance. Hewage *et al.* [16] proposed to allocate unequal transmission power for texture videos and depth maps to realize different level of protection. In [17], Hewage *et al.* proposed to utilize the motion correlation between the texture video and depth map to achieve UEP. The coded motion information is allocated with high level protection to conceal the errors at the receiver side. Besides the UEP strategy, there are many sophisticated approaches to deal with the transmission tasks. Xiao *et al.* [18] proposed a scalable method to handle the bit allocation problem between texture video and depth map. Zhang *et al.* [19] proposed a packetization strategy for

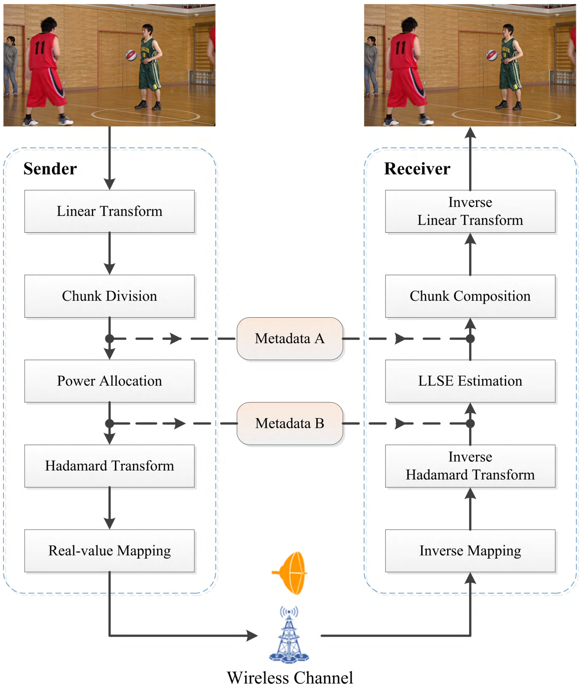

FIGURE 2. SoftCast framework.

3D video transmission, which improves the error resilience performance over wireless channels. However, these approaches all encounter the so-called cliff effect inherent in conventional source-channel coding framework.

B. SOFTCAST FOR WIRELESS VIDEO TRANSMISSION

SoftCast [11]–[13] is a simple yet effective scheme for wireless video transmission. As shown in FIGURE 2, the design of the SoftCast approach ensures that the whole system is completely linear.

On the sender side, it consists of five steps: linear transform, chunk division, power allocation, Hadamard transform, and real-value mapping. The linear transform is employed to eliminate the redundancy of the original video signals, which plays the same role as in the conventional video compression operation. The chosen transform in SoftCast is DCT. The transformed coefficients are grouped into a number of chunks for transmission via the chunk division process. A chunk is basic transmission unit in SoftCast design, which is utilized to reduce the data size of the metadata. The power allocation is performed over chunks to minimize the transmission distortion by optimally scaling the magnitude of the transformed coefficients. Hadamard transform [20] is applied to redistribute energy across chunks to form equal-importance transmission packets, thus improves the error resilience to channel error. Finally, pairs of the transformed coefficients are directly mapped to the I (in-phase) and Q (quadrature) components for OFDM transmission. The metadata A in FIGURE 2 is employed to transmit the chunk division information: mean value and variance of each chunk, while the metadata B is utilized to transmit scale factors of chunks for power allocation. As shown in [12], the total metadata is only

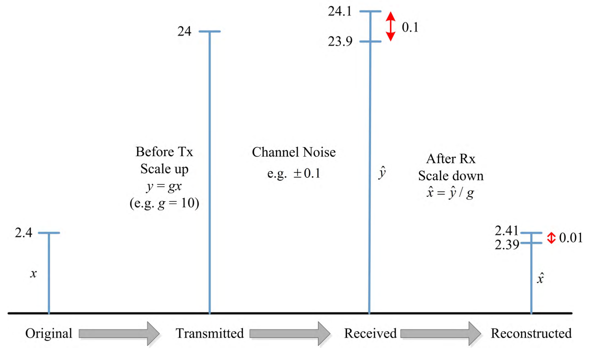

FIGURE 3. Scaling for softCast.

about 0.014 bits/pixel, which is an insignificant overhead. To protect the metadata from channel errors, it uses the lowest bitrate FEC code and BPSK modulation to guarantee its reliable reception.

On the receiver side, LLSE (Linear Least Square Estimator) [21] is employed to obtain the best estimation of the original transformed coefficients. The other steps are exactly the inverse operations of the corresponding processes on the sender side.

C. POWER ALLOCATION FOR SOFTCAST

Conventional wireless video transmission adopts the channel coding to add some parity bits against the channel noise. This approach will destroy the numerical properties of the original video signal and inevitably cause the cliff effect. Obviously, it is not suitable for SoftCast, where the real value coefficients are transmitted rather than the coded bit-stream. Thus, a novel error protection scheme is developed for SoftCast, which is based on scaling the magnitude of the transmitted Coefficients. As shown in FIGURE 3, the original signal x is scaled up to $y = gx(g > 1)$. Then, the value y is directly transmitted over the wireless channel. Suppose the channel is AWGN (Additive White Gaussian Noise), and the noise n is with zero mean and variance σ_n^2. The received signal $\hat{y}$ is scaled down to $\hat{x} = \hat{y}/g$ to get the reconstructed signal. It can be seen that the reconstruction error of the scale up/down scheme is only σ_n^2/g^2 as compared to the direct transmission error σ_n^2.

Unfortunately, the power budget of the transmitter is limited, scaling up and thus expending more power some signal samples means to assign less power on the other signal samples. Therefore, it requires to perform power allocation to find optimal scale factors to minimize the total transmission distortion of the entire signal sequence. Suppose $\mathbf{x} = (x_1, x_2, \ldots, x_N) \in \mathbf{R}^N$ is a random signal sequence to be transmitted over the wireless channel, and each element x_i represents a transformed coefficients. It has been shown in [22] and [23] that the optimal scale factor for sending x_i is

$$g_i = (E[x_i^2])^{-1/4}\left(\sqrt{\frac{P_{total}}{\sum_i \sqrt{E[x_i^2]}}}\right), \quad (1)$$

where P_{total} is total power budget, and $E[*]$ is the expectation operation. Meanwhile, the minimum total reconstruction distortion is

$$D = \sum_i \frac{\sigma_n^2}{g_i^2} = \frac{\sigma_n^2}{P_{total}}(\sum_i \sqrt{E[x_i^2]})^2. \tag{2}$$

Although assign a single scale factor for each transformed coefficient could obtain the optimal performance, it requires to transmit a large amount of metadata and hence it is infeasible. Instead, SoftCast performs power allocation over chunks to reduce the metadata amount. That means a single scale factor is assigned for each chunk.

D. SOFTCAST-LIKE UNCODED AND HYBRID VISUAL TRANSMISSION

Due to the graceful quality transition property of SoftCast, the SoftCast-like uncoded and hybrid transmission has become an attractive issue recently. An analysis for SoftCast based uncoded video transmission is made, which developes a quantitative measurement for the efficiency of decorrelation transform and the energy distribution of the signal elements [23]. In [24], two practical energy modeling schemes are proposed for power-distortion optimization of SoftCast. In order to achieve power saving within in a chunk, magnitude shift and data division are proposed to improve the efficiency of SoftCast [25], [26]. Foveation characteristic of human vision based [27] SoftCast scheme is proposed to improve the perceptual quality of the received images/videos. In [28] and [29], MCTF (motion compensate temporal filter) and DWT (discrete wavelete transform) are proposed to further remove temporal and spatial redundancy. In [30]–[33], the SoftCast are extensively incorporated into the distributed image/video coding framework to provide efficient wireless video communication. Based on SoftCast, a framework called hybrid digital-analog (HDA) coding scheme are studied in the latest years [34]–[37]. The key idea of HDA is to integrate the SoftCast-like analog method into the conventional digital video coding frameworks, thus it could provide both the advantages of digital coding and analog coding. Furthermore, the SoftCast based scheme also has been employed into satellite remote-sensing image transmission [38] and cloud image/video transmission over wireless channels [39]. A novel network slice design is proposed in [40], which consists of a rateless source compression scheme and an analog-coded SoftCast scheme. Cost-distortion optimization and resource control is conducted to use resources more efficiently for SoftCast [41]. All above mentioned are targeting on conventional natural 2D image/video, while there are only a couple of literatures on the discussion of the uncoded or pseudo-analog 3DV wireless transmission. In [42], hybrid wireless delivery scheme is presented for only stereo videos without depth information. In [43], uncoded wireless MVD transmission is considered, which simply combines the texture and depth together and extends the transform to five dimensional discrete cosine transform (5D-DCT) to exploit inter-view and texture-depth correlations for performance improvement. However, it fails to consider the unique piecewise smoothness of the depth maps, which is different from the natural texture signal.

III. THE PROPOSED WIRELESS DEPTH MAP TRANSMISSION

Based on the SoftCast framework shown in FIGURE 2, to achieve best possible performance for wireless depth map transmission, we are addressing optimization of three components, namely mean removed block DCT, view synthesis distortion based chunk dropping, and inter-block DCT coefficients rearrangement.

A. MEAN REMOVED BLOCK DCT FOR DEPTH MAP

According to (2), it can be seen that the reconstruction distortion linearly increases with $E(\mathbf{x}) = \sum_i \sqrt{E[x_i^2]}$. The linear transform is employed to de-correlate the original signal. In the original SoftCast design, the frame size 2D-DCT (or 3D-DCT) is adopted to obtain the optimal decorrelation performance for conventional 2D image (or video). It should be noted that the 3D-DCT is employed to reduce the redundancy across successive frames further. However, as compared to the natural image, the depth map of 3D video presents distinct characteristics. Typical depth maps usually are composed of homogeneous regions separated by sharp edges. The correlation among pixels within a homogeneous region is very high, while the correlation across the sharp edge is negligible. Thus, the DCT transform across the sharp edge may change the relatively regular distribution of the original depth maps and hence increase $E(\mathbf{x})$ of the depth map signal. Accordingly, we will attempt to perform different size of block DCT, which does not cross the sharp edges as much as possible.

In addition, the original SoftCast approach generally consider the frame size DCT transform directly in the spatial domain, and the mean value of the DCT coefficients of each block is transmitted as metadata. Nevertheless, the block DCT scheme facilitates to transmit the mean value of each block in the pixel domain as metadata, and the block DCT is employed on the residual for each block. This strategy will take full advantage of the piecewise smooth property of depth maps and further reduce $E(\mathbf{x})$.

As shown in FIGURE 4, the decorrelation performance of different transform size of DCT is examined for *balloons* depth map, including 4×4, 8×8, 16×16, 32×32, 64×64, 128×128, 256×256, and frame size. It can be seen that the best decorrelation performance is not obtained by frame size DCT transform. On the contrary, the block DCT transform of smaller sizes obtain better performance. The performace tends to decrease with increasing DCT size. However, the smaller DCT size is, the more data amount of the metadata is required. The overall evaluation of DCT size will be discussed in Section IV-A. Therefore, the block DCT of an appropriate size is preferred as the linear transform approach rather than the frame size DCT for wireless depth map transmission.

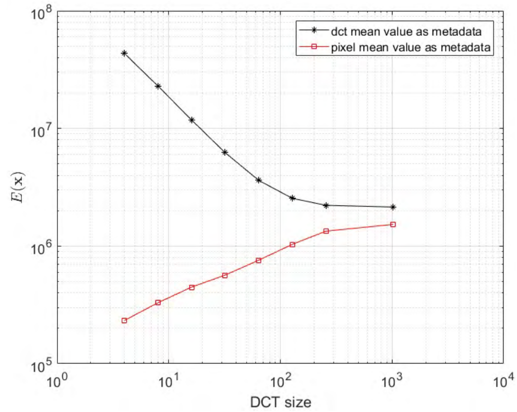

FIGURE 4. Transform performance of DCT for *balloons* depth map.

B. VIEW SYNTHESIS DISTORTION BASED CHUNK DROPPING

The depth map is employed to provide the 3D geometric information for virtual view synthesis. The distortion of depth map will lead to warping position error in the synthesized view. The depth distortion does not linearly affect the synthesis distortion. More complicatedly, the relationship is sophisticated and relevant to its co-located texture. In general, smooth regions of texture video is more tolerant to depth distortion, while the complex texture is more sensitive to depth distortion. Thus, the wireless depth map transmission needs to consider the view synthesis distortion rather than that of the depth distortion itself [44]. The adopted block DCT transform facilitates a more precise estimation of view synthesis distortion than frame size DCT. That is another important reason why the block DCT is preferred for depth map.

The distortion of the synthesized view caused by depth map error could be measured as the sum of squared error (SSE) between two versions of synthesized views as

$$\begin{aligned} SSE_V &= \sum_{(x',y')} |V_{(x',y')} - \tilde{V}_{(x',y')}|^2 \\ &= \sum_{(x,y)} |f_w(T_{x,y}, D_{x,y}) - f_w(T_{x,y}, \tilde{D}_{x,y})|^2, \end{aligned} \tag{3}$$

where $V_{(x',y')}$ denotes the synthesized view from the original texture video and the depth map, $\tilde{V}_{(x',y')}$ denotes the version generated from the distorted depth map and their corresponding texture video, T and D indicate the original texture video and depth map respectively, $\tilde{D}$ denotes the received distorted depth map, and (x', y') is the warped pixel position for the synthesized view V corresponding to (x, y) in T and D by the pre-defined 3D warping function f_w. It should be noted that, as we only discuss the wireless depth map transmission problem in this paper, the texture video is been treated unchanged; Therefore, the original texture video is employed to evaluate the depth map transmission performance.

With the 1-D parallel camera setting configuration which is commonly used in 3DTV configuration, the synthesis

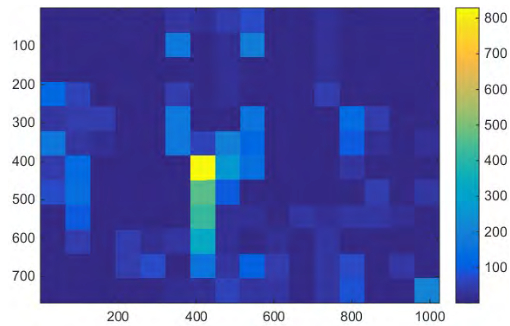

FIGURE 5. Estimated synthesis distortion of *balloons* depth map.

distortion caused by the depth error could be approximated as

$$SSE_V \approx \sum_{(x,y)} |T_{x,y} - T_{x-\Delta P(x,y),y}|^2, \tag{4}$$

where ΔP is the translational horizontal warping position error. It has been proven that it is proportional to the depth map error as [45]

$$\Delta P(x, y) = \alpha \cdot (D_{x,y} - \tilde{D}_{x,y}), \tag{5}$$

where α is a proportional coefficient determined by

$$\alpha = \frac{f \cdot L}{255}\left(\frac{1}{Z_{near}} - \frac{1}{Z_{far}}\right), \tag{6}$$

where f is the focal length, L is the baseline between the reference and the synthesized view, and Z_{near} and Z_{far} denote the nearest and farthest depth value of the 3D scene respectively. Finally, the synthesis distortion caused by depth map error could be further approximated as [46]

$$SSE_V \approx \sum_{(x,y)} \frac{|\Delta P(x, y)|}{2} \cdot G(x, y), \tag{7}$$

where $G(x, y) = |T_{x,y} - T_{x-1,y}| + |T_{x,y} - T_{x+1,y}|$, which reflects the local characteristic of the texture video. Suppose the depth map is divided into equal-size blocks. For each block B_i ($i = 1, 2, , N$), the synthesis distortion caused by dropping all DCT coefficients in the block will be

$$SSE_V^i \approx \sum_{(x,y)\in B_i} \frac{|\alpha \cdot (D_{x,y} - M_i)|}{2} \cdot G(x, y), \tag{8}$$

where M_i is the mean value of the block B_i, which is transmitted as metadata.

FIGURE 5 shows the estimated synthesis distortion caused by dropping the DCT coefficients of all blocks. It can be seen that the synthesis distortion of a large number of blocks is close to zero. In that case, we could ignore the synthesis distortion of these blocks, and just drop the DCT coefficients of these blocks for transmission. It means that, given a threshold τ, if $SSE_V^i < \tau$, the DCT coefficients of the block B_i can be discard. The power saved for these discarded blocks could be assigned to the rest ones. In addition, when the bandwidth is not enough to transmit all the blocks, to maximize the reconstruction quality on the receiver side, the blocks will be sorted in decreasing order of their estimated synthesis distortion values, and as many blocks as possible will be picked to fill the available bandwidth.

C. INTER-BLOCK DCT COEFFICIENTS REARRANGEMENT

As mentioned before, the optimal power allocation is to assign a single scale factor g_i for each DCT coefficient according to (1) individually. However, this optimal strategy is impractical, since the transmission overhead for the metadata is overwhelming. Chunk level power allocation is a reasonable alternative scheme, which chooses a single scale factor g_i for each chunk based on the mean energy of that chunk.

Suppose the number of divided chunks is M, and x_{ij} is the jth coefficient in chunk i, where $i = 1, 2, \ldots, M$, and $j = 1, 2, \ldots, K$ (K is the number of coefficients in each chunk), the optimal scale factor for chunk i is [12]

$$g_i = \lambda_i^{-1/4}\left(\sqrt{\frac{P_{total}}{K\sum_{i=1}^{M}\sqrt{\lambda_i}}}\right), \tag{9}$$

where λ_i is the mean energy of chunk i expressed as

$$\lambda_i = \frac{\sum_{j=1}^{K} x_{ij}^2}{K}, \tag{10}$$

and the minimal total reconstruction distortion is

$$D = \sum_{i=1}^{M} K \cdot \frac{\sigma_n^2}{g_i^2} = \frac{K^2\sigma_n^2}{P_{total}}\left(\sum_{i=1}^{M}\sqrt{\lambda_i}\right)^2. \tag{11}$$

In the original SoftCast design, nearby DCT Coefficients are directly grouped into a chunk after the frame size DCT Transform. In our case, the straightest way is to employ power allocation on every DCT transformed block. Nevertheless, this simple way is far from the optimal solution. In [22] and [23], it has been illustrated that, with the optimal power allocation strategy which is shown in (1), the more diversified the DCT coefficients are, the smaller transmission distortion is. While in order to obtain better power allocation performance for our case, a simple yet more explicit problem is discussed below.

For a given set of random numbers with length M-by-N $\mathbf{Y} = \{y_1, \cdots, y_N, y_{N+1}, \cdots, y_{2N}, \cdots, y_{MN}\}$, what is the minimum condition for the function $f = \sum_{i=1}^{M}\sqrt{\lambda_i}$, where the λ_i is the average square value of every arbitrary N numbers without overlapping in $\mathbf{Y}$?

Remark: for any $0 \leq x_1^2 \leq x_2^2 \leq x_3^2 \leq x_4^2$, $\sqrt{\frac{x_1^2+x_2^2}{2}} + \sqrt{\frac{x_3^2+x_4^2}{2}} \leq \sqrt{\frac{x_1^2+x_3^2}{2}} + \sqrt{\frac{x_2^2+x_4^2}{2}}$ and $\sqrt{\frac{x_1^2+x_3^2}{2}} + \sqrt{\frac{x_3^2+x_4^2}{2}} \leq \sqrt{\frac{x_1^2+x_4^2}{2}} + \sqrt{\frac{x_2^2+x_3^2}{2}}$. The above *Remark* can be easily proved as follows, e.g.

$$\left(\sqrt{\frac{x_1^2+x_2^2}{2}} + \sqrt{\frac{x_3^2+x_4^2}{2}}\right)^2 - \left(\sqrt{\frac{x_1^2+x_3^2}{2}} + \sqrt{\frac{x_2^2+x_4^2}{2}}\right)^2$$
$$= \sqrt{(x_1^2+x_2^2)(x_3^2+x_4^2)} - \sqrt{(x_1^2+x_3^2)(x_2^2+x_4^2)} \leq 0, \tag{12}$$

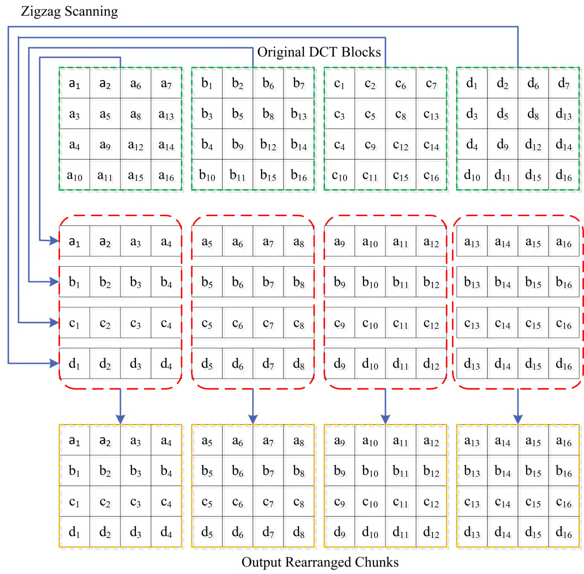

FIGURE 6. Diagram of inter-block DCT coefficients rearrangement.

because of $(x_1^2+x_2^2)(x_3^2+x_4^2) - (x_1^2+x_3^2)(x_2^2+x_4^2) = (x_1^2 - x_4^2)(x_3^2 - x_2^2) \leq 0$. The *Remark* can be extended to the general case of any M-by-N random number:

For any $0 \leq x_1^2 \leq x_2^2 \leq \cdots x_N^2 \leq \cdots x_{2N}^2 \leq \cdots x_{MN}^2$, the function $f = \sum_{i=1}^{M}\sqrt{\lambda_i}$ will achieve its minimum value only when $\lambda_i = (x_{(i-1)N+1}^2 + x_{(i-1)N+2}^2 + \cdots + x_{iN}^2)/N$ for $i = 1, 2, \cdots, M$.

The above conclusion shows that, the chunk level power allocation should sort the DCT coeffieicnts in desreasing or increasing order and group the every equal-size nearby DCT coefficients into the same chunk to achieve the optimal performance. It is intuitively easy to understand. Suppose $x_1 = 20$ and $x_2 = 200$ are two DCT coefficients in a same chunk and the scaling factor for this chunk is $g = 2$. Then, $y_1 = 40$ and $y_2 = 400$ are the transmitted values. In order to transmit these two values, the system expends different power, e.g., 1600 for the smaller one and 160000 for the larger one. The power used to transmit these two coefficient is with large difference, but their error resilience to the channel noise is the same. Therefore, there is much power wasting to transmit the large magnitude coefficients and the power allocation is inefficient at this situation. On the contrary, if the two coefficients are the same, e.g., $x_1 = x_2 = 20$, there is no power wasting and the power allocation is optimal at this time.

The DCT transform makes the signal energy compacted to a small part of coefficients, which means that a small number of coefficients located at the top-left position become large and the other coefficients become extremely small in a DCT block. Thus, the magnitude difference for the coefficients in a same DCT block is very large. We need to rearrange the DCT coefficients to make the chunk level power allocation more efficient. According to the above conclusion, the optimal way is to sort the coefficients in descending or ascending order, and group every equal-size

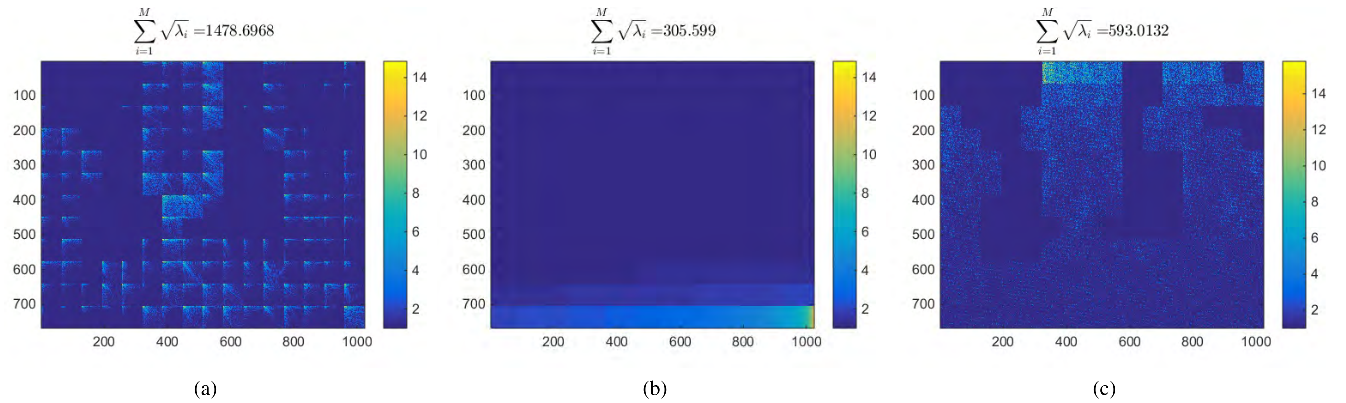

FIGURE 7. Coefficients distribution for inter-block rearrangement based chunk division, both the DCT size and chunk size are 64. (a) Original distribution, (b) ideal optimal distribution, and (c) inter-block rearrangement method.

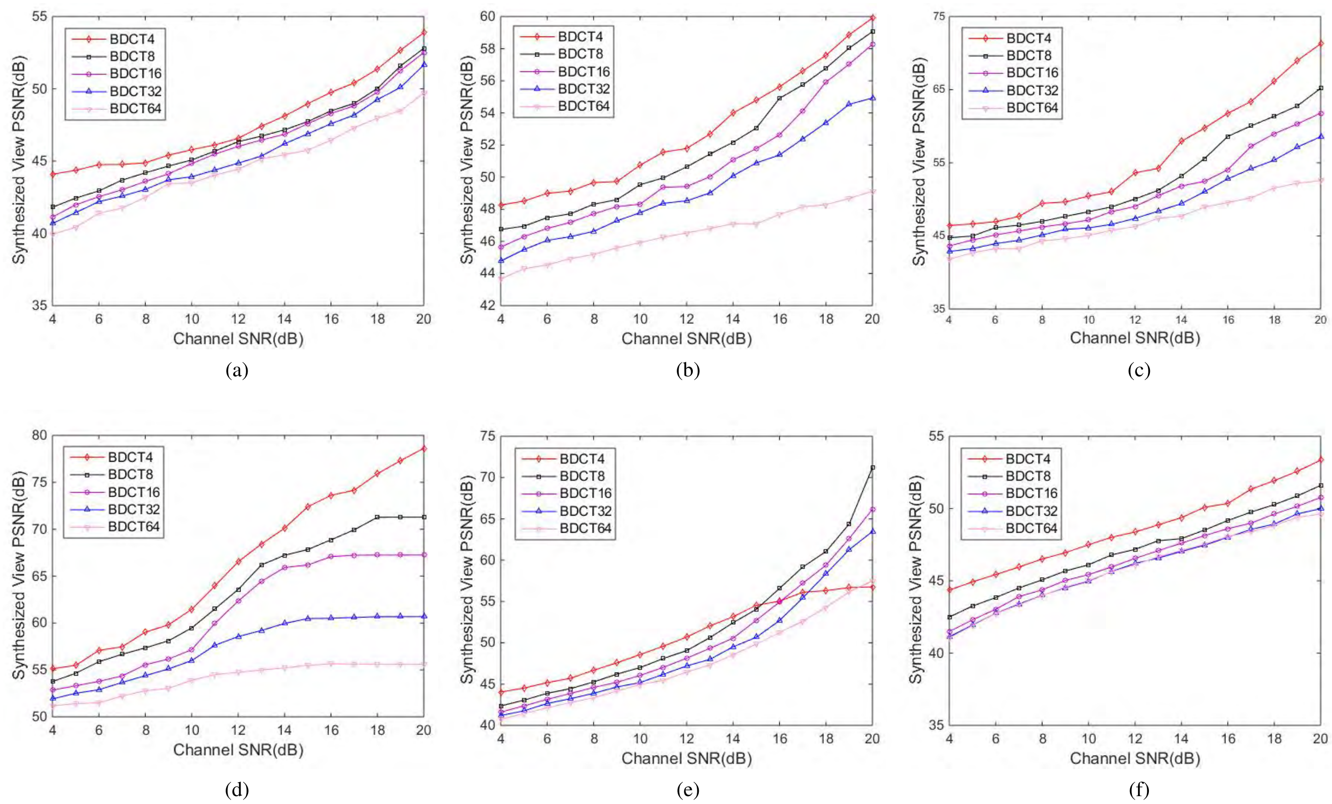

FIGURE 8. Performance of different DCT size setting. (a) balloons; (b) kendo; (c) dancer; (d) gtfly; (e) poznanstreet; (f) shark.

nearby coefficients into a chunk. However, it is impractical because the receiver side does not know the original position of coefficients while transmitting these location information is with huge overhead. Due to the distribution property of the DCT coefficients, we present an alternative way to do inter-block coefficients rearrangement. It should be noticed that the proposed scheme does not need to transmit any metadata and thus no overhead increases.

The most important property of DCT transform is that the magnitude of the DCT coefficients decays rapidly from low frequency area to high frequency region. More specifically, the coefficients approximately descends in the diagonal direction from the top left to the lower right of the DCT block. The zigzag scanning just follows this property. Therefore, as shown in FIGURE 6, inter-block coefficients rearrangement first employs the zigzag-scanning for all DCT blocks, or to say, re-arrange the coefficients in a DCT block according to their distance to the top-left point. Then, the coefficients of DCT blocks with similar frequencies are grouped into a same chunk. In such a manner, it reduces the magnitude difference in a chunk and thus improve the power allocation efficiency.

According to (11), the distortion is strictly increasing with $\sum_{i=1}^{M}\sqrt{\lambda_i}$. As shown in FIGURE 7, the original coefficients distribution, the ideal optimal distribution as mentioned, and the proposed inter-block coefficients rearrangement are com-

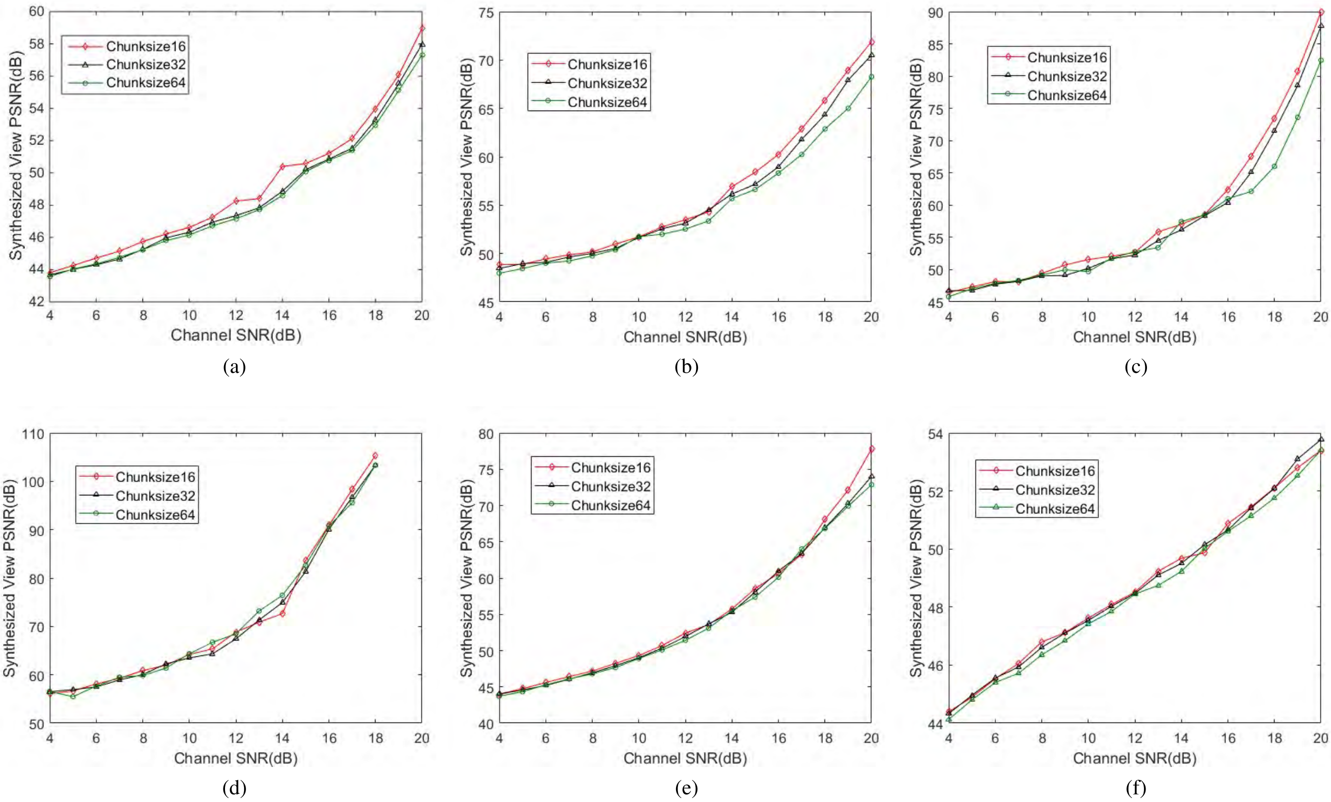

FIGURE 9. Performance of different chunk size setting. (a) balloons; (b) kendo; (c) dancer; (d) gtfly; (e) poznanstreet; (f) shark.

pared. It can be seen that the proposed scheme significantly reduces the $\sum_{i=1}^{M}\sqrt{\lambda_i}$ value and thus remarkably improves the power allocation performance. Moreover, the proposed method approaches to the optimal distribution without any transmission overhead increasing.

IV. EXPERIMENTAL RESULTS

In this section, we conduct a series of simulations to evaluate the performance of the proposed approach for wireless depth map transmission.

Test 3D video: depth maps of six 3D video sequences are employed in the experiments. Table 1 lists the used depth maps and their configurations.

Wireless Simulation Environment: the 802.11 wireless simulation link is employed as the simulation environment, which is implemented by using the Matlab Communication Toolbox. For the proposed and other SoftCast based method, the metadata is transmitted with 1/2 bit rate convolutional code and BPSK modulation, and the real value coefficients are directly transmitted over OFDM with the OFDM parameters being selected to match those of 802.11.a/g. For the conventional video transmission scheme, 3/4 bit rate convolutional code and QPSK modulation are adopted. In addition, channel SNR within the range of 4 to 20 dB is investigated in our simulations.

Metric: the PSNR metric is utilized, which is defined as a function of the mean square error (MSE) between all pixels of the reconstructed video and the original video as $PSNR = 10\log_{10}\frac{255^2}{MSE}$ (dB). For our 3D video case here, the average PSNR of all synthesized virtual views is calculated. Because only the depth map transmission issue is discussed in this paper, the original texture video is conducted to calculate the PSNR value. Therefore, the reconstructed version is synthesized [47] with the received depth maps and the original texture videos, and the original version is synthesized with the original depth maps and texture videos.

TABLE 1. 3-D video used in the experiments.

Sequence	Reference views	Synthesized virtual views	Number of Frames
Balloons	1 and 5	2, 3, 4	64
Kendo	1 and 3	1.5, 2.0, 2.5	64
Dancer	1 and 5	2, 3, 4	48
GTFly	1 and 5	2, 3, 4	48
Shark	1 and 5	2, 3, 4	48
PoznanStreet	3 and 5	3.5, 4.0, 4.5	48

A. DCT SIZE AND CHUNKS SIZE

In the proposed uncoded wireless depth map transmission, there are two important parameters: DCT size and chunk size. The influence of these two parameters on the performance is investigated first.

For the first set of simulation, the chunk size is fixed to 64, and the DCT size of 4, 8, 16, 32, and 64 is chosen. Chunk size

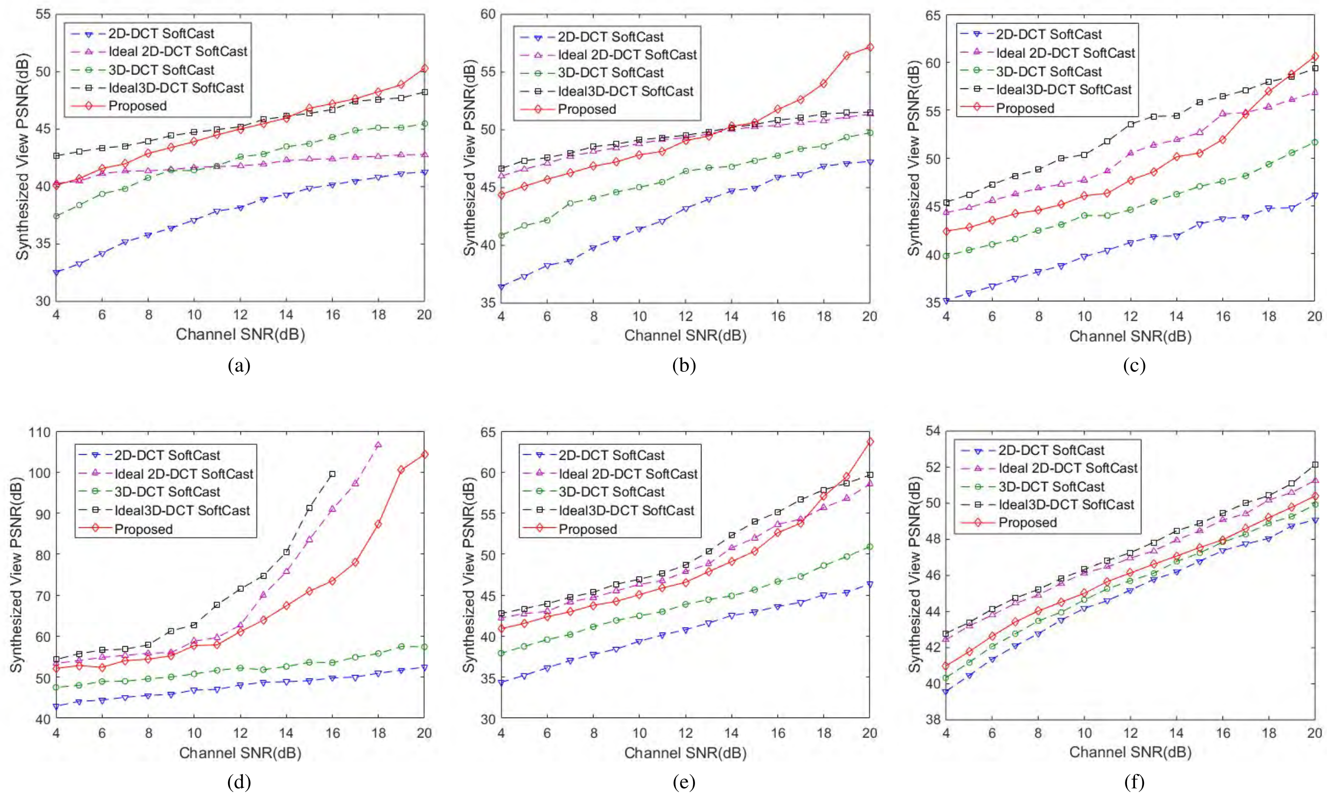

FIGURE 10. Different SoftCast based methods. (a) balloons; (b) kendo; (c) dancer; (d) gtfly; (e) poznanstreet; (f) shark.

of 64 and DCT size of 64 is set as benchmark, other settings match the same OFDM symbol number with the benchmark by dropping appropriate number of chunks. FIGURE 8 shows the comparison of different DCT size settings. It can be seen that the performance decreases with the increasing of the DCT size on the whole. It is mainly benefited from that the smaller size of mean removed DCT could obtain better decorrelation performance, which agrees with the analysis in section III-A. For the *poznanstreet* sequence with very complex depth structure, the performance of DCT size of 4 is inferior to other ones under high channel SNR. The reason is that the 4x4 size DCT does not reduce $E(\mathbf{x})$ of the depth map of the *poznanstreet* sequence significantly, yet occupies to much bit-rate for the transmission of metadata. It leads to many chunks being dropped and thus jeopardizes the performance of the high channel SNR.

For the second set of simulation, the DCT size is fixed to 4, and the chunk size of 16, 32 and 64 is chosen. As shown in FIGURE 9, it can be seen that performance of different chunk size settings are almost identical. Bigger chunk settings face slight performance degradation only when the channel SNR is very high. The chunk size mainly determines the efficiency of the power allocation. For the original SoftCast design, smaller chunk division has higher power allocation efficiency theoretically. However, inter-block DCT coefficients rearrangement of the proposed scheme has already significantly reduced the mean energy of a chunk and hence depresses the influence of the chunk size. It also proves the effectiveness of the proposed inter-block DCT coefficients rearrangement strategy.

B. SOFTCAST BASED METHODS

The original SoftCast employed 2D frame size DCT for natural images and 3D frame size DCT for natural videos. The performance of this design is investigated for depth maps. FIGURE 10 shows the performance of different SoftCast design. 2D-DCT SoftCast and 3D-DCT SoftCast represent the original chunk level power allocation design with 2D frame DCT and 3D frame DCT respectively. Ideal 2D-DCT SoftCast is the ideal optimal one for 2D-DCT case which assigns a single scale factor for each coefficient, and Ideal 3D-DCT SoftCast is the corresponding optimal one for 3D-DCT case. In the simulation, the chunk size for 2D-DCT SoftCast, 3D-DCT SoftCast, and the proposed method is 64, and the DCT size for our proposed method is also set to 64. It should be noticed that the transmit data amount of the two ideal cases is negligible, and we just want to find the ultimate performance which is unreachable. As shown in FIGURE 10, the proposed method is far superior to 2D and 3D frame DCT design ones, and close to the ultimate performance of these two case. Furthermore, the proposed one could be even better than the two ideal cases at high channel SNR. This is mainly due to the block based processing of the proposed scheme. First, the mean removed block DCT reduces $E(\mathbf{x})$ significantly as compared to the frame size DCT, which is shown in Section III-A. Second, the quality is measured by PSNR of the synthesis view rather than that of the depth map itself, which is improved by the view synthesis distortion based chunk dropping. Third, under the high SNR condition where the channel noise is small, the efficiency of the ideal power allocation is limited. The propossed inter-block coefficients rearrangement could approach to the ideal one.

V. CONCLUSION AND FURTHER WORKS

This paper proposes a block DCT based uncoded wireless depth map transmission scheme for 3D video. Based on the characteristics of the depth map, three steps of optimization are involved into the SoftCast framework, including mean removed block DCT transform, view synthesis distortion based chunk dropping, and inter-block DCT coefficients rearrangement. The experimental results show that the proposed scheme can provide superior performance to that of the state-of-the-art SoftCast methods, while producing elegant synthesis quality transition within a large range of channel SNR as compared to the source-channel coding based 3D video transmission approach.

Generally, 3D video is composed of multi-view texture video plus depth map, where this paper focuses on the optimization of uncoded wireless depth map transmission. This can be considered as one of the most important steps to realize the whole system of 3D video SoftCast, due to the uniqueness of depth map coding and transmission. In the next step, we will further integrate this scheme into 3D video SoftCast with texture video involved. Specifically, we will mainly work on the issue of joint power allocation between texture and depth to maximize the transmission efficiency of 3D video SoftCast.

REFERENCES

[1] C. Zhu, Y. Zhao, L. Yu, and M. Tanimoto, Eds., *3D-TV System With Depth-Image-Based Rendering: Architectures, Techniques and Challenges*. New York, NY, USA: Springer, 2013.

[2] M. Tanimoto, M. P. Tehrani, T. Fujii, and T. Yendo, "Free-viewpoint TV," *IEEE Signal Process. Mag.*, vol. 28, no. 1, pp. 67–76, Jan. 2011.

[3] P. Merkle, A. Smolic, K. Müller, and T. Wiegand, "Multi-view video plus depth representation and coding," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Atlanta, GA, USA, Sep. 2006, pp. 201–204.

[4] S. Zinger, L. Do, and P. H. N. de With, "Free-viewpoint depth image based rendering," *J. Vis. Commun. Image Represent.*, vol. 21, no. 5, pp. 533–541, Jan. 2010.

[5] B. Cao, S. Xia, J. Han, and Y. Li, "A distributed game methodology for crowdsensing in uncertain wireless scenario," *IEEE Trans. Mobile Comput.*, to be published.

[6] Y. Li, J. Liu, B. Cao, and C. Wang, "Joint optimization of radio and virtual machine resources with uncertain user demands in mobile cloud computing," *IEEE Trans. Multimedia*, vol. 20, no. 9, pp. 2427–2438, Sep. 2018.

[7] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul./Oct. 1948.

[8] S. Vembu, S. Verdu, and Y. Steinberg, "The source-channel separation theorem revisited," *IEEE Trans. Inf. Theory*, vol. 41, no. 1, pp. 44–54, Jan. 1995.

[9] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[10] H. Yuan, H. Fu, J. Liu, and J. Xiao, "End-to-end distortion-based multiuser bandwidth allocation for real-time video transmission over LTE network," *IEEE Trans. Broadcast.*, vol. 63, no. 2, pp. 338–349, Jun. 2017.

[11] S. Jakubczak, H. Rahul, and D. Katabi, "Softcast: One video to serve all wireless receivers," Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep. MIT-CSAIL-TR-2009-005, 2009.

[12] S. Jakubczak, H. Rahul, and D. Katabi, "One-size-fits-all wireless video," in *Proc. ACM Hotnets*, 2009, pp. 1–6.

[13] S. Jakubczak and D. Katabi, "A cross-layer design for scalable mobile video," in *Proc. Int. Conf. Mobile Comput. Netw. (MobiCom)*, 2011, pp. 289–300.

[14] G. Tech, Y. Chen, K. Müller, J.-R. Ohm, A. Vetro, and Y.-K. Wang, "Overview of the multiview and 3D extensions of high efficiency video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 35–49, Jan. 2016.

[15] L. Zhang, Y. Li, P. Cheng, and M. Li, "A layer-mixed fec scheme for scalable media transmission over mobile TV services," *IEEE Trans. Broadcast.*, vol. 63, no. 2, pp. 309–320, Jun. 2017.

[16] C. T. E. R. Hewage, Z. Ahmad, S. T. Worrall, S. Dogan, W. A. C. Fernando, and A. Kondoz, "Unequal error protection for backward compatible 3-D video transmission over WiMAX," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2009, pp. 125–129.

[17] C. T. E. R. Hewage, S. Worrall, S. Dogan, and A. M. Kondoz, "Frame concealment algorithm for stereoscopic video using motion vector sharing," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jun. 2008, pp. 485–488.

[18] J. Xiao, M. M. Hannuksela, T. Tillo, M. Gabbouj, C. Zhu, and Y. Zhao, "Scalable bit allocation between texture and depth views for 3-D video streaming over heterogeneous networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 1, pp. 139–152, Jan. 2015.

[19] X. Zhang, Y. Zhao, T. Tillo, C. Lin, J. Xiao, and A. Wang, "Packetization strategies for MVD-based 3D video transmission," in *Proc. IEEE Vis. Commun. Image Process. (VCIP)*, Nov. 2016, pp. 1–4.

[20] M. Schnell, "Hadamard codewords as orthogonal spreading sequences in synchronous DS CDMA systems for mobile radio channels," in *Proc. IEEE Int. Symp. Spread Spectr. Technol. Appl. (ISSSTA)*, Jul. 1994, pp. 505–509.

[21] C. Lawson and R. Hanson, *Solving Least Squares Problems*. Upper Saddle River, NJ, USA: Prentice-Hall, 1974.

[22] H. Cui, R. Xiong, C. Luo, Z. Song, and F. Wu, "Denoising and resource allocation in uncoded video transmission," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 1, pp. 102–112, Feb. 2015.

[23] R. Xiong, F. Wu, J. Xu, X. Fan, C. Luo, and W. Gao, "Analysis of decorrelation transform gain for uncoded wireless image and video communication," *IEEE Trans. Image Process.*, vol. 24, no. 4, pp. 1820–1833, Apr. 2016.

[24] R. Xiong, J. Zhang, F. Wu, J. Xu, and W. Gao, "Power distortion optimization for uncoded linear transformed transmission of images and videos," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 222–236, Jan. 2017.

[25] X. Lin, Y. Liu, and L. Zhang, "Scalable video softcast using magnitude shift," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2015, pp. 1996–2001.

[26] X. Lin, Y. Liu, and M. Sun, "Analog channel coding for wireless image/video softcast by data division," in *Proc. IEEE Int. Conf. Telecommun. (ICT)*, Apr. 2015, pp. 353–357.

[27] J. Shen, L. Yu, L. Li, and H. Li, "Foveation-based wireless soft image delivery," *IEEE Trans. Multimedia*, vol. 20, no. 10, pp. 2788–2800, Oct. 2018.

[28] Q. Wang, X. Lin, Y. Liu, L. Zhang, and X. Wu, "A scalable framework for mobile video broadcast using MCTF and 2D-DWT," in *Proc. IEEE Int. Symp. Wireless Pers. Multimedia Commun. (WPMC)*, 2014, pp. 118–123.

[29] M. Lv, Y. Liu, and Y. Wang, "Scalable wireless video broadcast based on unequal protection," in *Proc. IEEE Vis. Commun. Image Process. (VCIP)*, Nov. 2017, pp. 1–4.

[30] X. Fan, F. Wu, D. Zhao, and O. C. Au, "Distributed wireless visual communication with power distortion optimization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 6, pp. 1040–1053, Jun. 2013.

[31] A. Zhang, X. Fan, R. Xiong, and D. Zhao, "Distributed soft video broadcast with variable block size motion estimation," in *Proc. IEEE Vis. Commun. Image Process. (VCIP)*, Nov. 2014, pp. 1–5.

[32] M. Sun, Y. Wang, H. Yu, and Y. Liu, "Distributed cooperative video coding for wireless video broadcast system," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jun. 2015, pp. 1–6.

[33] X. Song, X. Peng, J. Xu, G. Shi, and F. Wu, "Distributed compressive sensing for cloud-based wireless image transmission," *IEEE Trans. Multimedia*, vol. 19, no. 6, pp. 1351–1364, Jun. 2017.

[34] Z. Song, R. Xiong, S. Ma, and W. Gao, "Hybridcast: A wireless image/video softcast scheme using layered representation and hybrid digital-analog modulation," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 6001–6005.
[35] L. Yu, H. Li, and W. Li, "Wireless cooperative video coding using a hybrid digital–analog scheme," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 3, pp. 436–450, Mar. 2015.
[36] N. Fan, Y. Liu, Q. Wang, and L. Zhang, "Hybrid digital-analog video multicast scheme based on H. 264/AVC and softcast," in *Proc. IEEE Int. Symp. Wireless Pers. Multimedia Commun. (WPMC)*, Sep. 2015, pp. 277–282.
[37] X. Zhao, H. Lu, C. Chen, and J. Wu, "Adaptive hybrid digital–analog video transmission in wireless fading channel," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 6, pp. 1117–1130, Jun. 2016.
[38] J. Li, X. E. Wen, H. Jia, X. Xie, and W. Gao, "AnalogCast: Full linear coding and pseudo analog transmission for satellite remote-sensing images," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Mar. 2016, pp. 1362–1366.
[39] X. Huang, J. Wu, and F. Hu, "Knowledge-enhanced mobile video broadcasting framework with cloud support," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 1, pp. 6–18, Jan. 2017.
[40] B. Tan, J. Wu, Y. Li, H. Cui, W. Yu, and C. W. Chen, "Analog coded softcast: A network slice design for multimedia broadcast/multicast," *IEEE Trans. Multimedia*, vol. 19, no. 10, pp. 2293–2306, Oct. 2017.
[41] D. Liu, J. Wu, H. Cui, D. Zhang, C. Luo, and F. Wu, "Cost-distortion optimization and resource control in pseudo-analog visual communications," *IEEE Trans. Multimedia*, vol. 20, no. 11, pp. 3097–3110, Nov. 2018.
[42] D. He, C. Luo, F. Wu, and W. Zeng, "Swift: A hybrid digital-analog scheme for low-delay transmission of mobile stereo video," in *Proc. ACM Int. Conf. Modeling, Anal., Simulation Wireless Mobile Syst.*, 2015, pp. 327–336.
[43] T. Fujibashi, T. Koike-Akino, T. Watanabe, and P. V. Orlik, "Soft video delivery for free viewpoint video," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–7.
[44] M. Yang, C. Zhu, X. Lan, and N. Zheng, "Efficient estimation of view synthesis distortion for depth coding optimization," *IEEE Trans. Multimedia*, to be published.
[45] W.-S. Kim, A. Ortega, P. Lai, D. Tian, and C. Gomila, "Depth map coding optimization using rendered view distortion for 3D video coding," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3534–3545, Nov. 2015.
[46] B. T. Oh, J. Lee, and D.-S. Park, "Depth map coding based on synthesized view distortion function," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 7, pp. 1344–1352, Nov. 2011.
[47] H. Yuan, S. Kwong, X. Wang, Y. Zhang, and F. Li, "A virtual view PSNR estimation method for 3-D videos," *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 134–140, Mar. 2016.

TAIHAI YANG received the B.Eng. degree in communication engineering from the Chongqing University of Posts and Telecommunications, Chongqing, China, in 2016, where he is currently pursuing the master's degree. His research interests include 3D video processing and video coding.

LEI LUO (M'17) received the B.S. and Ph.D. degrees from Zhejiang University, Hangzhou, China, in 2008 and 2013, respectively. He is currently an Associate Professor with the School of Communications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China, and also a Postdoctoral Fellow with the School of Information and Communication, University of Electronic Science and Technology of China, Chengdu, China. His research interests include video coding and transmission, 3D video processing, and video analysis.

CE ZHU (M'03–SM'04–F'17) received the B.S. degree from Sichuan University, Chengdu, China, in 1989, and the M.Eng. and Ph.D. degrees from Southeast University, Nanjing, China, in 1992 and 1994, respectively, all in electronic and information engineering. He did the Postdoctoral Research at The Chinese University of Hong Kong, in 1995, the City University of Hong Kong, and The University of Melbourne, Australia, from 1996 to 1998. He was with Nanyang Technological University, Singapore, from 1998 to 2012, where he was a Research Fellow, a Program Manager, an Assistant Professor, and then promoted to an Associate Professor, in 2005. He has also held visiting positions with the Queen Mary University of London and Nagoya University. He is currently a Professor with the University of Electronic Science and Technology of China, Chengdu. He has authored or co-authored over 160 publications. As a lead editor, he has edited three books and contributed four book chapters. He holds over 30 patents issued or pending. He has contributed over 40 proposals (over 20 adopted) to the international standardization bodies of JCT-3V (3D-HEVC) and MPEG Internet Video Coding as well as the China AVS standardization (mainly on encoding optimization with the proposed techniques adopted in the standard test model and reference software). His research interests include video coding and communications, video analysis and processing, 3D video, visual perception, and applications. He is a Fellow of the IET. He was a co-recipient of the two Best Paper Awards (as the first author) and the two Best Student Paper Awards at IIH-MSP 2013, the IEEE BMSB 2014, MobiMedia 2011, and the IEEE ICME 2017, respectively. He is the IEEE Broadcast Technology Society's Distinguished Lecturer. He is the Co-Chair of the Interest Group of Multimedia Processing for Communications of the Technical Committee on Multimedia Communications, the IEEE Communications Society, and the Chair of the Special Interest Group on Big Multimedia of Technical Committee on Multimedia Computing, IEEE Computer Society. He has served on technical/program committees and organizing committees, and as the Program/Track/Area/Session Chair for around 60 international conferences. He has given tutorials at international conferences (e.g., IEEE ISCAS 2015 and IEEE VCIP 2014) and has delivered keynote speeches at a few international conferences. He is an elected member of the Multimedia Signal Processing Technical Committee (MMSP-TC) of the IEEE Signal Processing Society, the Chair of the Webmaster Subcommittee of MMSP-TC, and an elected member of the Multimedia Systems and Applications Technical Committee of the IEEE Circuits and Systems Society. He has served on the Editorial Board of a few journals, as an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON BROADCASTING, and the IEEE SIGNAL PROCESSING LETTERS, an Editor for the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, and an Area Editor for *Signal Processing: Image Communication*. He has also served as a Guest Editor for a few special issues in international journals.

SHU TANG received the B.S. and M.E. degrees from the Chongqing University of Posts and Telecommunications, Chongqing, China, in 2003 and 2007, respectively, and the Ph.D. degree from Chongqing University, China, in 2013. He is currently an Associate Professor with the College of Computer Science and Technology, Chongqing University of Posts and Telecommunications. His research interests include signal processing and image processing.