Digital Watermarking for 3D Images

Prof. S R Zadokar¹, Prof. Suryawanshi V.G.²
¹Imperial College of Engineering & Research Wagholi, Pune, Maharashtra, India. (PDEA’s College of Engg. Manjari, Pune, Maharashtra, India)
²PDEA’s College of Engg. Manjari, Pune, Maharashtra, India
¹sanjuzadokar@gmail.com
²vijaysuryawanshi@gmail.com

Abstract—The content protection for image-based 3D data is getting more importance with the advance of low cost 3D display devices. The depth-image-based rendering (DIBR) 3D image is one of the image-based 3D data which consists of the center image and the depth image generated by the content provider. The left-eye image and the right-eye image are rendered from the center image and the depth image at the content consumer side. In this paper, we propose a robust watermarking scheme for DIBR 3D images by quantization on dual-tree complex wavelet transform (DT-CWT) coefficients. To make the proposed scheme robust to DIBR process, we select certain coefficient sub-blocks and group the coefficient rows based on the properties of DIBR. The simulation results show that the embedded watermark is stably extracted from the center view and the synthesized left and right views. In addition, even if the synthesized left and right views are distorted by general attacks, the watermark is success-fully extracted.

Index Terms—Depth-image-based rendering (DIBR), dual-tree complex wavelet transform (DT-CWT), watermarking, 3D watermarking.

I.INTRODUCTION

As the 3D display market has grown rapidly, demand for 3D content has also increased naturally. In fact, numerous recent movies have been produced in 3D form, and audiences have shown great interest in immersive experiences through the 3D content. Those 3D contents are represented by two major methods. One, which is called stereo image recording (SIR), is to store the left and right views, which are simultaneously captured for the same scene from different camera positions. This method offers a high quality immersive viewing environment to viewers because two capture devices act like human eyes. Despite such an advantage, the method has several drawbacks. It is very difficult and costly to set two cameras with the same contrast, brightness, and height. Also, the initial depth condition is fixed. That is, viewers are not able to adjust the depth condition they prefer. Large storage or transmission bandwidth due to the double color images for one scene is a critical issue as well. The other 3D representation method is depth-image-based rendering (DIBR) [1], [2]. Both the left and right views are generated by a monoscopic center view and associated per-pixel depth information using the DIBR technique. Contrary to the SIR, it provides several advantages. Firstly, viewers can easily adjust the depth degree by modifying the depth image in the DIBR system. In other words, the DIBR system provides individualized depth configuration. Secondly, it has a lower transmission bit rate per scene [3] because the depth image consists of gray level pixels and many smooth areas that lead to a higher compression ratio. Lastly, the DIBR system has backwards-compatibility to 2D digital TVs because the center image itself can be utilized as 2D content.

Due to the numerous advantages of the DIBR system, many studies have been conducted on DIBR and its applications. In particular, research on watermarking for copyright protection of DIBR content has recently received attention due to the rapid growth of the 3D content market. As shown in Fig. 1, a copyright protection technique should consider situations not only in which both virtual left and right views are illegally distributed as 3D content but also in which each single view, including the original center view, is leaked to the public without permission. So far, numerous digital watermarking schemes for 2D and SIR images have been presented [4]–[7]. In DIBR 3D images, however, pixels of a watermarked center image are partially moved horizontally to different distances due to the DIBR process. Furthermore, the corresponding depth images can be modified for a certain purpose after the watermark embedding. For those reasons, conventional watermarking schemes for 2D and stereo images cannot be directly applied to the DIBR 3D images.

In this proposed DCT domain watermarking algorithm for the DIBR system by estimating the virtual left and right views from a center view and a depth image. Three watermarks are embedded in a center view for a center view and virtual left and right views. The embedded watermarks are successfully extracted with a low bit error rate.

In this paper, we develop a watermarking algorithm for 3D images in a DIBR system where the embedded watermarks are blindly extracted. Because the DIBR process can be seen as a partial translation along the horizontal plane, the DCT-CWT is exploited due to its characteristics of approximate shift invariance and good directional selectivity. In the embedding process, the watermark signals are embedded by the quantization process of the carefully selected DCT-CWT coefficients groups. Then, the embedded watermark signals are extracted using the statistical difference caused by the quantization process. The simulation results show that
proposed scheme is robust to not only the general signal and geometric distortions, but also depth map configuration and baseline distance adjusting which are special types of attacks only on DIBR 3D images.

This paper is organized as follows. Brief descriptions of the DIBR and the dual-tree wavelet transform are introduced in Sections II and III, respectively. In Section IV, the proposed watermarking scheme on the dual-tree wavelet coefficients is given. The performance of the proposed method is described in Section V. Conclusions are given in Section VI.

II. DEPTH-IMAGE-BASED RENDERING SYSTEM

DIBR with a center color view and gray-level depth image is presented in [2]. Virtual left and right views can be synthesized by 3D image warping and hole-filling processes. For more natural virtual view generation, pre-processing of the depth image can be exploited [11]. The overall DIBR process is shown in Fig. 2.

Pre-processing, the first step, of a depth image is usually aimed at reducing large holes to improve the quality of the virtual left and right views by smoothing the depth image [11]–[13]. Because the original depth image can be modified in the viewer’s DIBR system, watermarking must be designed to be robust against the modification of the depth image.

3D warping, the second step, is a process in which pixels in a center color image are mapped to virtual left and right views according to a corresponding depth image. Since human eyes are placed on the horizontal line, a parallel configuration approach is commonly used in the 3D warping process [14]. According to this configuration, the pixel-wise mapping function [11] is given as follows:

\[ x_l = x_c + \left( \frac{t_x}{2} \cdot \frac{f}{Z} \right), \quad x_r = x_c - \left( \frac{t_x}{2} \cdot \frac{f}{Z} \right) \]

where \( x_c, x_l, \) and \( x_r \) are \( x \)-coordinate of the pixels in the center view, the virtual left view, and the virtual right view, respectively. \( t_x \) is a baseline distance, and \( f \) is a focal length. \( Z \) is a depth value corresponding to the coordinate of the center view pixel. If two or more pixels are mapped to the same location of virtual view, the one which has the higher depth value is selected. From this equation, we focused on two characteristics to de-sign a watermarking system. First, all pixels in the center color image are moved only horizontally moved, which means watermarks must be robust to horizontal shifting. Second, the base-line distance \( t_x \) can be changed by viewers to create more or less immersive effect. Thus, the embedded watermarks should be successfully extracted regardless of \( t_x \) changing. Then, hole filling is processed as the last step of DIBR. New exposed areas, called holes, are revealed after the 3D warping process. Because there is no pixel information for the holes, newly exposed areas are filled by average filter or interpolation. Without loss of generality, we adopt linear interpolation as a hole-filling method.

III. DUAL-TREE COMPLEX WAVELET TRANSFORM (DT-CWT)

Unlike the general discrete wavelet transform (DWT), two independent DWTs are exploited for DT-CWT. As shown in Fig. 3, a signal is decomposed and constructed by two DWTs, each of which is called a tree and uses different filter banks satisfying the perfect reconstruction (PR) condition.

Two notable characteristics of the DT-CWT for the proposed watermarking design are approximate shift-invariance and good directional selectivity. Watermarks inserted in the magnitude of DT-CWT coefficients are robust to geometric distortions related to translation. Robustness against the DIBR can be achievable when the DT-CWT coefficients sub-bands are carefully chosen with consideration of the directional selectivity and characteristic of the DIBR. To show that DC-CWT is more shift-invariant in a DIBR process than other transforms, such as Fourier transform, which is well known for its shift-invariance, the PSNR test between coefficients before and after the DIBR process was conducted. As illustrated in Fig. 4, the magnitude of DT-CWT coefficients shows less change than other transforms. In particular, with a carefully selected direction of the DC-CWT coefficients, it shows much higher PSNR than other cases. The DT-CWT for 2D signal clearly shows directional selectivity. Unlike 2D DWT, the number of high frequency coefficient sub-bands is twice as high as those of the 2D DWT for each level. As shown in Fig. 5(a), each level, denoted as \( l_v \) in this paper, of 2D DT-CWT has six high frequency sub-bands \( H_{l_v, d_v} \) where \( d_r \) \( 1, 2, \ldots, 6 \) and each \( d_r \) is a direction index of six directional filters oriented at angles of \( \pm 15^\circ, \pm 45^\circ, \pm 75^\circ, \pm 105^\circ, \pm 135^\circ, \) and \( \pm 15^\circ \). Fig. 5(b) shows the magnitude of sub-band coefficients in the DT-CWT for the “ring” image. As we clearly see, each sub-band has more energy around the edges along with its filter angle. In the DIBR system, since each pixel moves horizontally, the vertical edges are more distorted than the horizontal edges. In other if
watermarks are inserted in the $H_{l,v,3}$ and $H_{l,v,4}$.

Fig. 1. Illegal distribution scenario for each captured view by a viewer for DIBR content

Fig. 2. Depth-image-based rendering process.

![Image](image.png)

Fig. 3. Average PSNR between the magnitude of coefficients of the center view and the synthesized left and right views for various transforms. The baseline distance $t_x$ is 5% of the center view width. The focal length $f$ is 1.

which contain more vertical edges than other sub-bands, the embedded watermarks are easily corrupted against the 3D warping process. Fig. 6 illustrates the average peak signal-to-noise ratio (PSNR) between DT-CWT coefficient sub-blocks of the center view and its virtual views generated by the DIBR process for 12 experimental images. In the experiment, baseline distance $t_x$ and focal length $f$ are set to 5% of width of the center image and 1, respectively. For the hole filling, linear interpolation technique is employed. The result clearly shows that PSNRs in the directions of 3 and 4 are lower than other directions. That is, it is proven that the coefficient sub-blocks from directions of 1, 2, 5, and 6 are less affected by the DIBR process

IV. PROPOSED METHOD

This section describes the proposed watermarking scheme. Three-level DT-CWT is applied to the sub-blocks of the center image to embed messages by the quantization method. In the watermarking extraction process, embedded watermarks are extracted by the statistical difference between DT-CWT coefficient blocks caused by the quantization.

A. Watermark bedding

Fig. 7(a) shows a block diagram of the proposed watermark embedding process. The whole process can be divided into seven steps.

1) Step 1 (Dividing into Sub-blocks): The host image, which is a center view, is divided into $(I-w/M) * (I-h/N)$ sub-blocks $B_i$: Where $1 = 1, 2, 3, 4, ..., M, N$. $I-w$ & $I-h$ are width and height of the center image, respectively. $B_i-w = (I-w/M)$, $B_i-h = (I-h/N)$ are the width and height of the sub-block, respectively.

2) Step 2 (3-level DT-CWT): Three-level DT-CWT is applied to each sub-block as shown in Fig. 5(a). We select coefficients in level 2 and 3 to design robust watermarking against low-pass filter, noise addition, and compression such as JPEG.

3) Step3 (Pairing): The six sub-bands $H_{l,v,Dr}$ where $Dr=1,2,3,4,5,6$ are grouped into three pairs. Each of three pairs

![Image](image.png)

Fig. 5. (a) Structure of one level DT-CWT coefficient; (b) magnitude for DT-CWT coefficients of an image of a ring.
has mostly horizontal, diagonal, and vertical edges, respectively. The reason why similar sub-bands are grouped is because one of the goals of the proposed quantization embedding scheme is to make a statistical difference between two sub-bands in a pair.

4) **Step 4 (Message encoding):** The bit sequence of a meaningful signature or text, whose length is $M*N$ is shuffled to watermark sequence bits, $w_1,w_2,\ldots, w_{MN}$ with a secret key $K$ in order to enhance security.

5) **Step 5 (Quantization):** Watermark bits are embedded into the pairs by the coefficient quantization.

6) **Steps 6, 7 (Inverse DT-CWT, Merging sub-blocks):** The watermarked sub-blocks are reconstructed by inverse DT-CWT from the coefficient sub-bands. Then, each reconstructed sub-block is merged to generate a watermarked center view.

**B. Watermark Extraction**

As illustrated in Fig. 7(b), the embedded watermarks are extracted with six steps.

1) **Steps 1,2,3,4 (Dividing into sub-blocks, 3-level DT-CWT, Pairing, and Quantization):** First four steps are the same as those of the watermark embedding process.

2) **Steps 5, 6 (Quantization error comparison and counting, Watermark extraction):** Given estimated quantization errors in a sub-block, the embedded watermark bit is extracted.

![Block diagram of the proposed method.](image)

**V. EXPERIMENTAL RESULTS**

This section reports the performance of the proposed method in terms of imperceptibility and robustness to various distortions with comparison to Lin’s method. We employed two different parameter settings for Lin’s method. In Lin’s method, the block size was set to 16*16.

**A. Test Image Set and Parameter Selection**

Fig. 10 shows 12 pairs of center and depth images from Mid-dlebury Stereo Datasets [20]–[22] and Microsoft Research 3D Video Datasets [23], having various resolutions from 450*375 to 1390*1110. The depth images are gray scale images of 8-bit level.

**B. Fidelity Test**

Based on the parameter selection above, subjective and objective image quality measures were exploited for the fidelity analysis. For the subjective test, an experimental system was set with a 23-inch LG Platron full HD 3D monitor of 120 Hz re-fresh rate and an NVIDIA GeForce GTX 460 with 3D Vision active shutter glasses. Then, the subjective quality was measured by the double stimulus continuous quality scale (DSCQS) with scores from 1 (bad) to 5 (excellent) using the criteria of the mean opinion score (MOS) [25]. The result MOSs, blindly scored by 10 experts for the 12 test images watermarked by the proposed method and Lin’s methods, are listed in Table I. The result shows that both the proposed method and Lin’s methods received relatively high scores for the 3D views.

For an objective quality analysis, average PSNR and structure similarity (SSIM) [26] of the watermarked center images were measured and compared with Lin’s methods. The results are shown in Table II. Although the proposed method showed slightly lower quality measures than Lin’s methods, degradation was very hard to recognize in general.

**C. Robustness Test: Signal and Geometric Distortions**

For the 12 test image pairs, a watermark was embedded into the center images. Then, the virtual left and right views were synthesized with corresponding depth images by Equation (1). As shown in Table III, the proposed method showed slightly higher BER for center images. For the virtual left and right views, however, the proposed method demonstrated robustness to the DIBR process with lower BER than Lin’s methods.

When at least one of the three views is illegally distributed, the views can be distorted by signal and geometric attacks. Accordingly, we tried to decode messages from the virtual right view, synthesized from the watermarked center view, after applying various signal and geometric distortions, such as JPEG compression, additive noise, median filtering, scaling, and rotation by the Starmark benchmark tool [27]. As illustrated in Fig. 11(a), the proposed method shows low...
BER, even after the watermarked image is compressed severely with a quality level of 15. For the additive noise, where the normalized noise level lies from 0 (same as original image) to 100 (completely random image), the proposed method shows lower BER than Lin’s method under the noise level of 2 while Lin’s method shows lower BER after the level of 3, as shown in Fig. 11(b). Practically, when the normalized noise level is greater than 3, the value of the image is seriously degraded to a PSNR of 20.72 dB, as shown in Fig. 12. This degradation barely occurs in the Internet transmission situation. From the Fig. 11(c), the robustness to the median filtering is clearly proven for the moderate median filter size.

**Table 1: Average loss of the proposed method & Lin’s method**

<table>
<thead>
<tr>
<th>Type of affine transform</th>
<th>Proposed method</th>
<th>Lin’s method</th>
<th>Lin’s method**</th>
</tr>
</thead>
<tbody>
<tr>
<td>interview</td>
<td>4.4</td>
<td>4.8</td>
<td>4.3</td>
</tr>
<tr>
<td>laundry</td>
<td>4.7</td>
<td>4.7</td>
<td>4.5</td>
</tr>
<tr>
<td>mouse</td>
<td>4.7</td>
<td>4.8</td>
<td>4.6</td>
</tr>
<tr>
<td>cups</td>
<td>4.6</td>
<td>4.7</td>
<td>4.7</td>
</tr>
<tr>
<td>cravender</td>
<td>4.4</td>
<td>4.7</td>
<td>4.7</td>
</tr>
<tr>
<td>bunny</td>
<td>4.4</td>
<td>4.7</td>
<td>4.6</td>
</tr>
</tbody>
</table>

**Fig. 13.** Comparison BER of the proposed method and Lin’s methods for various affine transforms

**Fig. 14.** Comparison BER of the proposed method and Lin’s methods for the various baseline distance ratio.

**Fig. 15.** The difference between original depth image and pre-processed depth image. (a) Original depth image; (b) asymmetrically blurred depth image.
D. Robustness Test: Baseline Distance Adjustment

One of the advantages of the DIBR system is the flexibility of depth adjustment. Because the personal preference and viewing distance vary, viewers could adjust baseline distance $t_x$ within a proper range. In the experiment, the baseline distance is ranged from 3% to 7% of the image width. As shown in Fig. 14, the performance of Lin’s methods sharply dropped when the base-line distance ratio was changed from initial $t_x$ of 5%, which was used in Lin’s embedding procedure. The proposed method, by comparison, shows lower and almost consistent BER within the range since the proposed system is designed with the characteristic of approximate shift invariance of the DT-CWT.

E. Robustness Test: Pre-Processing of Depth Image

As illustrated in Section II, the depth image can be modified to generate more natural virtual left and right views in the viewer’s DIBR system. For the experiment, asymmetric smoothing was employed [11], as shown in Fig. 15. The result BERs for the 12 test images are demonstrated in Fig. 16. The proposed method has a smaller gap between the case of original depth image and blurred depth image than Lin’s methods.

VI. CONCLUSION

The emerging 3D market inevitably produces copyright issues of the 3D content. In this work, we proposed blind watermarking scheme for DIBR 3D images with consideration of various scenarios, including a single virtual view being illegally distributed. We extended the quantization method on DWT coefficients to DT-CWT coefficients. In order to make the proposed watermarking robust to DIBR, we paired the DT-CWT sub-bands and grouped each coefficients row before quantization. Furthermore, pre-processing of depth images and baseline distance adjustment, which have not been issues in previous watermarking on DIBR 3D image, are considered to design the proposed watermarking method by exploiting the characteristics of the DT-CWT. Experimental results show low BERs for various signal processing and geometrical attacks with a good imperceptibility. As we intended, it is also robust to pre-processing of the depth image and baseline distance adjustment. Digital watermarking for 3D content is still in the early stages, and the standard of the DIBR is still being actively researched. In this situation, we will focus on extending our scheme to DIBR 3D videos and various standards of DIBR system. Also, improvement of the fidelity in terms of objective test and subjective test will be taken into consideration in future work.

REFERENCES