Article

Robust Template-based Watermarking for DIBR 3D Images

Wook-Hyung Kim [1] iD **, Jong-Uk Hou [1], Han-Ul Jang[1], and Heung-Kyu Lee [1,*]**

[1]   School of Computing, Korea Advanced Institute of Science and Technology, Daejeon 305701, Korea;
       whkim@mmc.kaist.ac.kr (W.-H.K.); juheo@mmc.kaist.ac.kr (J.-U.H.); hanulj@mmc.kaist.ac.kr (H.-U.J);

*   Correspondence: lheungksoc@gmail.com

Abstract: Several depth image based rendering (DIBR) watermarking methods have been proposed, but they have various drawbacks, such as non-blindness, low imperceptibility, and vulnerability to signal or geometric distortion. This paper proposes a template based DIBR watermarking method that overcomes the drawbacks of previous methods. The proposed method exploits two properties to resist DIBR attacks: the pixel is only moved horizontally by DIBR, and the smaller block is not distorted by DIBR. The one dimensional (1D) discrete cosine transform (DCT) and curvelet domains are adopted to utilize these two properties. A template is inserted in the curvelet domain to restore the synchronization error caused by geometric distortion. A watermark is inserted in the 1D DCT domain to insert and detect a message from the DIBR image. Experimental results of the proposed method show high imperceptibility and robustness to various attacks, such as signal and geometric distortions. The proposed method is also robust to DIBR distortion and DIBR configuration adjustment, such as depth image preprocessing and baseline distance adjustment.

Keywords: Depth-image-based rendering (DIBR); 3D content; curvelet transform; 1D-discrete cosine transform (1D-DCT); template watermark; DIBR watermarking

1. Introduction

Three-dimensional (3D) content has been steadily increasing in popularity because of its excellent lifelike appearance. With the recent development of 3D display technology, many new 3D applications have appeared to maximize realism, such as head mounted displays, 360° virtual reality, and ultra-high definition 3D content. Consequently, interest in 3D content and the 3D market itself have both increased greatly.

Methods of representing 3D content are divided into stereo image recording (SIR) and depth-image-based rendering (DIBR). SIR stores the left and right views (as human eyes do) and provides a high-quality immersive view; however, this has many limitations, including large data size, fixed depth, high cost, and difficulty with multiple camera settings. Meanwhile, DIBR is a rendering method that creates various virtual viewpoint images using center and depth images[1–4]. The DIBR method has two main advantages: 1) The DIBR system is able to easily save and transmit 3D content,
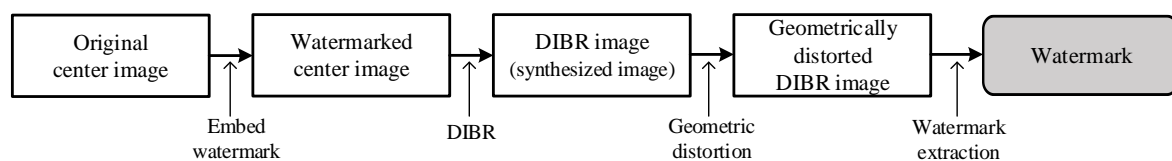

Figure 1. Watermark extraction scenario in DIBR image with geometric attack. Extraction is possible only if the watermark is survived both in the DIBR, which is a non-linear attack, and the geometric distortion, which is a linear attack.

because it requires less data compared to SIR; 2) The DIBR system can provide various viewpoints, since it allows us to adjust the 3D configuration. These advantages have led to DIBR technology being employed in 2D-to-3D conversion [5–10] and auto-stereoscopic [11,12] and multi-view stereo[13–17] that provide various viewpoints depending on the user's position

Copyright protection techniques for DIBR content has received considerable attention due to DBIRs important role and significant 3D content market growth. A typical copyright protection technique is watermarking, but many conventional two-dimensional (2D) watermarking techniques are difficult to apply to DIBR content. In the DIBR system, center image pixels are partially moved along the horizontal axis with distance depending on the depth image, using a process called non-linear geometric distortion. Watermarks inserted in the center image are strongly distorted and cannot be extracted.

Hence, several watermarking methods have been proposed as being robust to the DIBR process. Lin et al. proposed a method of embedding watermarks by predicting pixels' moving distance[18]. Protecting the center image and both the left and right images required superimposing and embedding three watermarks. This method has a low bit error rate (BER) against the DIBR process and common distortions, such as JPEG or additive noise. However, this method is vulnerable when the depth information is modified, such as in depth image preprocessing or change of baseline distance, since the moving distance for pixels during the DIBR process is predicted with unmodified depth information. Additionally, this method is vulnerable to geometric attacks due to the characteristics of the discrete cosine transform (DCT) domain.

Kim et al. suggested a watermarking method that employed quantization on dual tree complex wavelet transform (DT-CWT) domain coefficients[19]. The method used directional coefficients that were not significantly changed by the DIBR process. Row wise quantization was performed to provide robustness to horizontal pixel shifts. The method showed robustness to DIBR; JPEG compression; image scaling; and DIBR configuration adjustments, such as depth image preprocessing and baseline distance changes. However, it was vulnerable to noise addition and geometric distortions.

Wang et al.[20], Miao et al.[21], and Cui et al.[22] proposed a DIBR watermarking method that used the scale invariant feature transform (SIFT). The SIFT based-watermarking systems found similar parts between the center image and the synthesized image using a SIFT descriptor and then inserted the watermark into those parts. Because of the object matching with SIFT, these methods were robust to DIBR distortion. Additionally, the methods showed high robustness to both general signal distortions and geometric distortions. However, SIFT-based DIBR watermarking methods need SIFT descriptors during watermark extraction. These methods cannot extract watermarks blindly; therefore, SIFT-based watermarking methods are less practical than full-blind watermarking systems.

Asikuzzaman et al. proposed a DT-CWT based video watermarking method using color channels[23,24]. The method inserted a watermark into the U channel of YUV color space and inserted the same watermark rotated 180° into the V channel. They showed the method was robust to DIBR, due to DT-CWT domain characteristics; and geometric attacks, such as scaling and rotation, since the U and V channels suffer the same geometric transformation. However, if the image center was changed due to attacks such as crop or translation, the watermark could not be detected, and under geometric distortion it could only determine whether a watermark had been inserted or not, i.e., an on/off switch. Therefore, its application was somewhat limited.

Various templates have been proposed that are robust against geometric attacks[25–28]. However, these are only designed to be robust to linear distortions, such as affine transforms, and so are not robust to DIBR, which is non-linear distortion as discussed above.

The results of previous works show that blind watermarks are very difficult to survive in geometrically distorted DIBR images, because the watermark must be robust to both DIBR and geometric attacks, as shown in Fig. 1. The combination of non-linear and linear deformation severely damage most watermarking domains. Therefore, it is desirable to solve this problem by combining
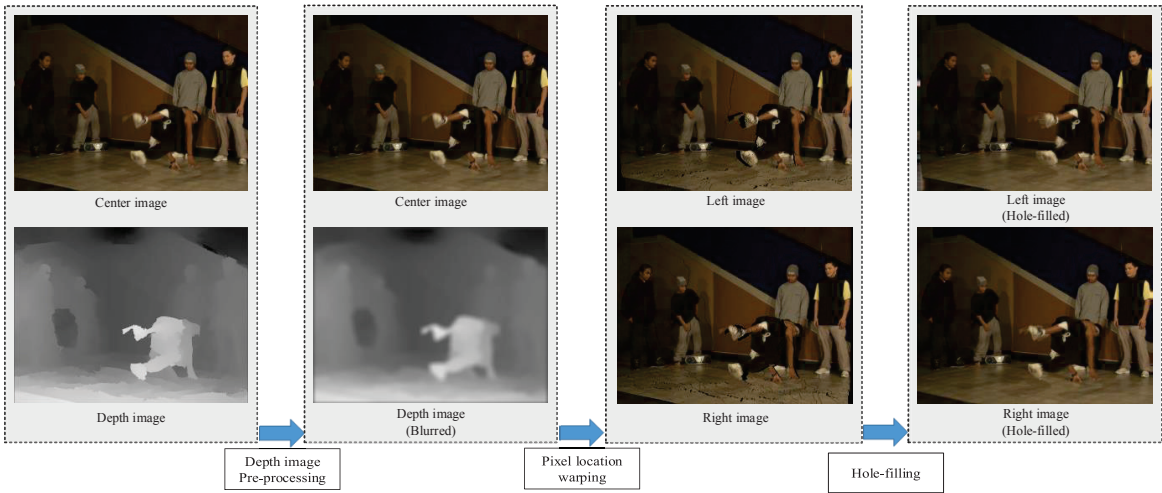
Figure 2. Overall process of DIBR system.

two watermarking methods that have different characteristics rather than using a single watermarking method.

This paper proposes a blind template based watermarking system combining templates and message watermarks. The role of the template is to restore geometric distortion without being destroyed from DIBR attack. The template is inserted into the curvelet domain in the form of peak points, and geometric distortion is estimated using the modified iterative closest point (ICP) method. The message watermark inserts and detects messages in the DIBR image. The proposed message watermark is inserted into the 1D-DCT domain and the message is extracted from the DIBR image after geometric distortion recovery using the proposed template. The message watermark inserts the same information along the horizontal direction in the 1D-DCT domain, and the inserted message watermark is invariant to DIBR due to 1D-DCT linearity.

Experimental results showed that the proposed method has both high invisibility and robustness against various attacks. It achieved excellent scores in visual quality tests and low BER for common signal distortions, such as noise addition and JPEG compression. The watermarking system also exhibited good robustness against geometric distortions, a point of vulnerability in previous approaches, and excellent robustness against DIBR attacks and DIBR configuration adjustments.

This paper is organized as follows. Section 2 describes the DIBR system and curvelet transform, which are fundamental techniques for the proposed scheme. Section 3.1 demonstrates the main idea of the proposed watermarking method. The watermark embedding and extraction processes are presented in Sections 3.2 and 3.3, respectively, and the experimental results and conclusion are given in Sections 4 and 5, respectively.

2. Backgrounds

This section presents DIBR and the curvelet transform, fundamental techniques of the proposed watermarking system. First, the DIBR process is briefly introduced and DIBR analysis is presented, and then we provide an introduction to and analysis of the curvelet transform.

2.1. DIBR Process

The whole DIBR process is shown in Fig. 2. DIBR consists of three steps: depth image preprocessing, pixel location warping, and hole filling.

Depth image preprocessing, the first step, improves the quality of the rendered image by reducing the number of holes [29–31]. When the viewpoint is moved by the DIBR, an area where no pixel information exists is generated. These areas, referred to as holes, are the main cause of 3D image

107 quality degradation. Holes occur mainly when the depth difference between two adjacent pixels is
108 large. Hence, the image quality can be improved by reducing the number of holes through depth
109 image smoothing.

Pixel location warping, the second step, changes the position of pixels along the horizontal
direction, allowing users to feel a 3D effect. The warping equation is as follows,

$$x_L = x_C + \left(\frac{t_x}{2} \times \frac{f}{Z} \right), \quad x_R = x_C - \left(\frac{t_x}{2} \times \frac{f}{Z} \right) \tag{1}$$

110 where x_C is the pixel position on the x-axis of the center image, x_L and x_R are the pixel positions on
111 the x-axis of the left view and the right view, respectively, t_x is the baseline distance, which means the
112 distance from the center axis to the left and right, f is the focal length, and Z is the value of the depth
113 image. During warping, two or more pixels can overlap in one position. In this situation, the highest Z
114 value of a pixel has to be selected to prevent an unnatural image.

115 The last step is hole filling, which creates pixel values in holes caused by pixel location warping.
116 There are several hole filling techniques, such as interpolation and inpainting. This is a field that is
117 constantly studied in pursuit of a more natural image [4,32–34].

118 *2.2. Analysis of DIBR Attack*

119 In the DIBR process, pixels are only translated horizontally, where the translation magnitude is
120 determined by the depth . Similar to the cover model being considered as a random, the depth image
121 is also close to a random signal[35], hence pixel's moving distance can also be assumed to be random.
122 Consequently, the pixels move irregularly, unlike in common translation. Thus, the 2D transformed
123 domain coefficients are distorted by DIBR.

124 The watermark damage caused by DIBR can be confirmed by the average energy change of the
125 middle frequency at which the watermark is inserted, since the average energy is the basis of the
126 watermark embedding energy.

Table 1 shows the average energy change between center and synthesized DIBR image coefficients.
DIBR parameters used in this test were as recommended by [3], and average energy change was
defined as

$$\frac{MSE(O, S)}{\text{mean}(O^2)} \tag{2}$$

127 where O and S are the transformed domain's coefficients for the original and synthesized images,
128 respectively; and MSE is mean squared error. The various transform domain coefficients' energy is
129 severely impaired by DIBR. Therefore, the watermark energy, inserted into the transform domain's
130 coefficients, is also damaged, and watermarks damaged by more than 40% are difficult to detect.

131 However, the average energy change of the haar discrete wavelet transform (DWT) is reduced to
132 4% if the distorted coefficients are matched with the undistorted coefficients using the depth image.
133 This is because the wavelet series represent the frequency information of the small spatial block, such
134 as 2×2, 4×4, $\cdots$.

135 If the image is divided into small blocks, some of these will be undistorted, since the depth is
136 similar between adjacent pixels within an object. If all depth values are the same in the block, pixel
137 moving distances are all the same. DIBR is treated like a common translation for this case. Smaller
138 block size implies greater percentage uncorrupted blocks, as shown in Table 2 for the example of 1800
139 synthesized DIBR images. Similar to the previous average energy change test, this test also used the
140 recommended DIBR parameters from [3].

141 Since DCT and discrete Fourier transform (DFT) express a global frequency that does not include
142 spatial information, the amount of average energy change after matching is still high. In other words,
143 it can be seen that the magnitude of the coefficients is damaged in DCT and DFT, but the magnitude of
144 the coefficients is maintained in DWT.

Table 1. Average energy change between coefficients of center and synthesized images.

		2D DCT	2D DFT(mag)	2D DWT (haar)
Average	Unmatched	209%	39%	215%
energy change	Matched	207%	40%	**4%**

Table 2. Average percentage of undistorted blocks after DIBR

	Block size			
	5×5	10×10	20×20	40×40
Undistorted blocks(%)	94.32	85.68	70.67	48.35

145   In summary, the following two properties of DIBR can be identified: 1) the pixel is moved only
146 in the horizontal direction and the moving distance is determined by the depth; 2) the percentage of
147 undistorted blocks is high with a small block size. Due to the second property, the wavelet series are
148 robust to DIBR.

149 *2.3. Curvelet Transform*

     The curvelet transform is a multi-scale decomposition-like wavelet transform, and the curvelet
represents the curve shape for the various directions in the spatial domain[36–39]. The curvelet
transform is developed to improve on the limitation of wavelet-based transforms and can represent
edges more efficiently than conventional wavelet-based transforms. Moreover, curvelet bases cover
all frequencies in contrast to other directional multi-scale transforms, such as the Gabor and Ridgelet
transforms[40]. The curvelet transform is expressed as follows,

$$C(g,l,k) := \langle f, \varphi_{g,l,k} \rangle = \int_{\mathbf{R}^2} f(x)\overline{\varphi_{g,l,k}}dx$$
$$= \frac{1}{(2\pi)^2} \int \hat{f}(\omega)U_g(R_{\theta_l}\omega)e^{i\langle x_k^{(g,l)},\omega \rangle}d\omega, \qquad (3)$$

$$U_g(r,\theta) = 2^{-3g/4}W(2^{-g}r)V\left(\frac{2^{\lfloor g/2 \rfloor}\theta}{2\pi}\right), \qquad (4)$$

150 In (3), C is the curvelet coefficient, $g = 0, 1, 2, ...$ is the scale parameter, l is the rotation parameter, and
151 $k = (k_1, k_2)$ is the translation parameter. U_g is a "wedge"-shaped frequency window represented in
152 (4). R_θ is the rotation operator and $\theta_l = 2\pi \cdot 2^{-\lfloor g/2 \rfloor} \cdot l$. In (4), W and V are the radial and angular
153 windows, respectively.

154   The curvelet is illustrated in Fig. 3. Figure 3 (a) illustrates the tiling of the curvelet in the frequency
155 domain, and the curvelet shape in several directions and scales in the spatial domain are shown in Fig.
156 3 (b)–(d).

157   As shown in Eq. (3) and Fig. 3, the curvelet represents frequency information of a small spatial
158 block similar, so the curvelet is also not distorted by DIBR. In addition, energy conservation is better
159 with the curvelet transform than with conventional haar DWT when image rotation occurs[41]. For
160 example, when the image rotates 10 degrees, the energy inserted into haar DWT is reduced to 50%,
161 but that inserted into the curvelet is maintained up to 85%. In case of scaling attack, energy is well
162 maintained in most DWT.

163   In summary, the curvelet transform is suitable for use as a template due to its robustness to DIBR
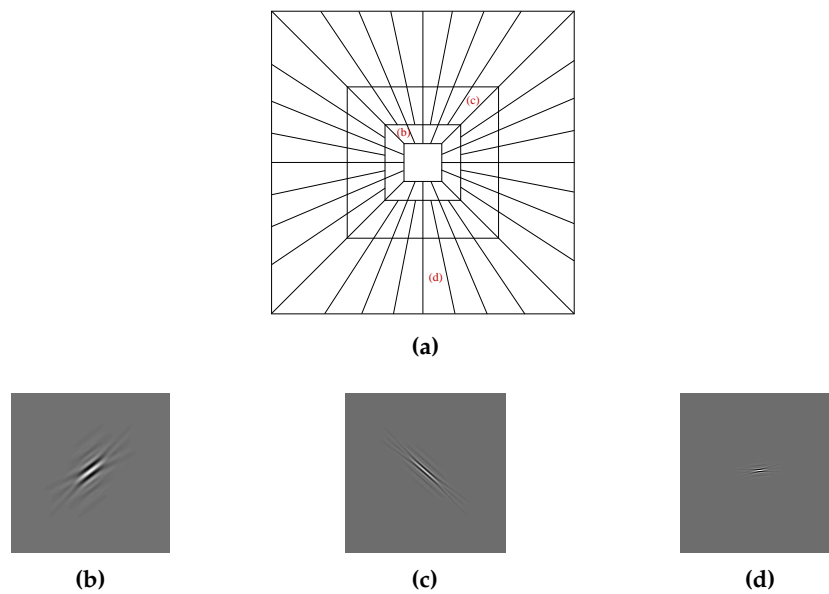164 and geometric transform.

Figure 3. Curvelet in the frequency and spatial domain. (a) Curvelet tiling of the frequency domain; (b)–(d) Curvelets for various scales and directions in the spatial domain. Curvelets are drawn on $k_1 = w/2$ and $k_2 = h/2$.
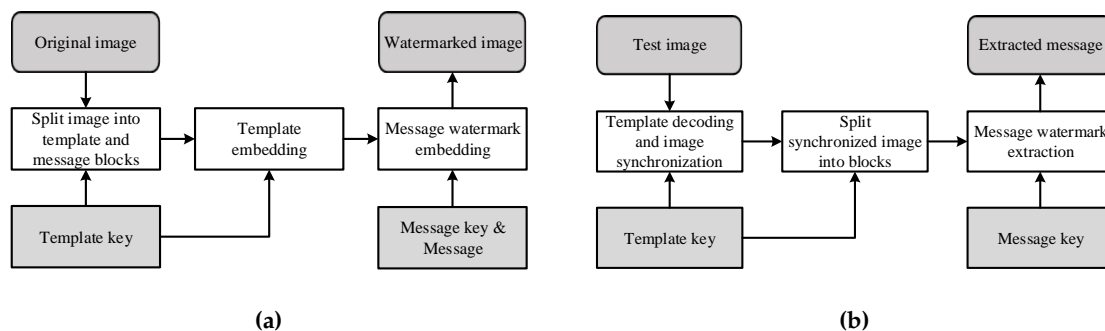

Figure 4. Watermarking process overview. (a) Embedding process; (b) Extraction process.

3. Proposed Method

3.1. Main Idea of Proposed Method

The overall flow of watermark embedding/extraction is shown in Fig. 4. As shown, the proposed watermarking system consists of a template and message watermark. This section describes the characteristics and roles of the template and message.

The DIBR watermarking system must be robust to both DIBR and geometric attacks, as shown in Fig. 1. The proposed method inserts a template into a curvelet domain robust against DIBR and geometric distortions. The inserted template enables restoring the image from geometric distortion without being destroyed by DIBR. Geometric distortion can be restored by inserting the template in peak point form and matching detected peaks with ground-truth positions, which can be obtained from the template key in the detection step.

However, although the peak matching method can recover global geometric distortion that occurs across the entire image, DIBR distortion cannot be recovered because it is treated as horizontal error. In addition, the peak point form template cannot insert messages. To address these problems, we

Table 3. The roles of template and message watermark.

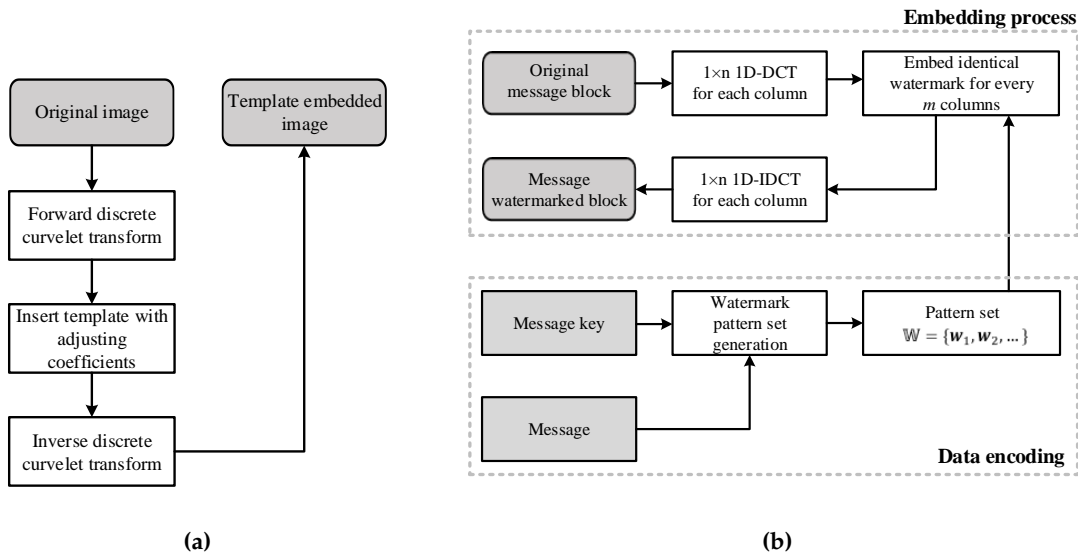|	Template	Message watermark
Domain	Curvelet	1D-DCT
Geometric distortion	Recoverable	Vulnerable
DIBR	Robust, but unrecoverable	Invariant
Payload	-	More than 1bit per block

Figure 5. Detailed embedding process. (a) Template embedding process; (b) Message watermark embedding process.

designed an invariant message watermarking technique against DIBR, where DIBR deformations that cannot be restored by the template are handled by the message watermark. Message watermarks without templates are vulnerable to geometric attacks, but combining template and message watermark advantages allows a robust DIBR watermarking system.

In summary, the template and message in the proposed watermarking system complement each other's weaknesses. The curvelet with robustness to DIBR is utilized as a template to restore images from geometric attacks. The message watermark inserted in the 1D-DCT domain resolves the unrecovered damage caused by DIBR. The message watermark itself is not robust to geometric attacks, but images that have suffered geometric attacks can be restored using templates. The roles of templates and message watermarks are summarized in Table 3.

3.2. Proposed Watermark Embedding Method

This section gives a detailed description of the proposed watermarking procedure. The whole embedding process is presented in Fig. 5.

3.2.1. Block Separation

If the template and the message watermark are inserted at the same position, the signals interfere with each other, and the robustness is degraded. To avoid overlapped insertion, the image is spatialy divided into $M \times N$ blocks, as shown in Fig. 6. The set of blocks is defined as B, and the template and message are inserted into the different blocks.
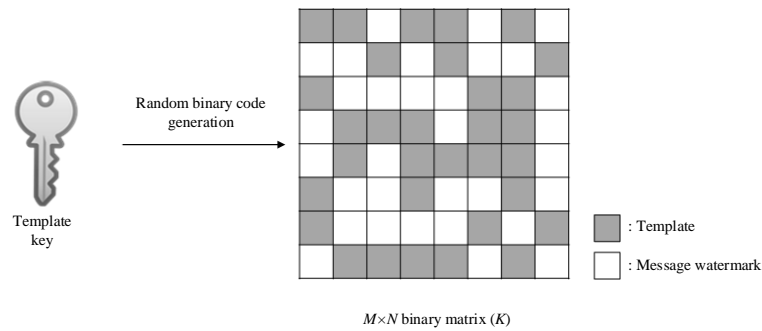
Figure 6. Example of template and message watermark block placement in spatial domain.
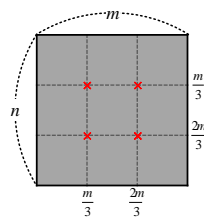

Figure 7. Insert four peak points into one template block. The points are inserted at the third division point so that the intervals are the same.

A random binary code is generated using the key, and this code is used to generate a 2D binary matrix, K, of size $M \times N$ size, which determines whether each block is a template or watermark block in an image divided into blocks. To match the number of template blocks with the number of message watermarks, average (K) should be 0.5.

The following rules distinguish the roles of the blocks.

$$B(x,y) = \begin{cases} \text{Template block, if } K(x,y) = 1 \\ \text{Watermark block, if } K(x,y) = 0 \end{cases} \tag{5}$$

where, x and y are the horizontal and vertical coordinates of B and K, $0 \leq x < M$, and $0 \leq y < N$.

3.2.2. Template Embedding

Since the curvelet coefficients contain spatial information, the template block position can be extracted from the curvelet coefficients. Therefore, the entire image is transformed into the curvelet domain without requiring a block based curvelet transform. We then insert the peak point templates into the curvelet coefficients.

As shown in Fig. 7, four peak points are inserted into one template block. If the number of peak points is too small, the robustness drops, and if there are too many points, the visual quality drops. Experimentally, four points were appropriate.

The template embedding process is divided into three steps:

a) Forward curvelet transform: The forward curvelet transform is applied to the whole image.

b) Peak points insertion: Select two directions of a curvelet, l_1 and l_2, with scale value g. l_1 and l_2 are selected such that they differ by 90°, to increase detection rate by corner detection in the template extraction step, and g should be selected to be a middle frequency, as a compromise between invisibility and robustness. As discussed above, four peak points are inserted into one template block,

$$C_m(g, l_1, k) = C(g, l_1, k) \cdot \alpha_t + \beta_t$$
$$C_m(g, l_2, k) = C(g, l_2, k) \cdot \alpha_t + \beta_t \tag{6}$$

217   where C_m is the modified curvelet coefficient, C is the original curvelet coefficient, α_t and β_t is
218   the strength of the inserted template, and k is the location of the template point.
219   c) Inverse curvelet transform: The inverse curvelet transform is applied to the modified coefficients.

220   3.2.3. Message Watermark Embedding

The message is inserted into the 1D-DCT domain using the spread spectrum method[42]. If a watermark having the same information in the horizontal direction is inserted into the 1D-DCT, invariance can be obtained against DIBR due to the first property in Section 2.2. The 1D-DCT watermark insertion equation is as follows,

$$\mathbf{DCT}(I'_{i:}) = \mathbf{DCT}(I_{i:}) + \mathbf{w}, \tag{7}$$

221   where $i = 1, 2, ..., m$, $\mathbf{w} = [w_1, w_2, ...w_n]^T$ is the watermark signal, and m, n are the horizontal and
222   vertical sizes of the block, respectively. $I_{i:}$ denotes the i-th column of the original block, $I'_{i:}$ denotes the
223   i-th column of the watermarked block, and $\mathbf{DCT}$ denotes the 1D-DCT.

The DIBR attack is applied in the spatial domain, not in the transformed domain. Therefore, the inverse DCT (IDCT) is performed to check the change of the inserted watermark in the spatial domain. Using the IDCT, Eq. (7) can be rewritten according to linearity of the DCT as follows,

$$\begin{aligned} I'_{i:} &= \mathbf{IDCT}(\mathbf{DCT}(I'_{i:})) \\ &= \mathbf{IDCT}(\mathbf{DCT}(I_{i:}) + \mathbf{w}) = I_{i:} + \mathbf{IDCT}(\mathbf{w}) \\ &= I_{i:} + \mathbf{v}, \end{aligned} \tag{8}$$

224   where $\mathbf{v} = [v_1, v_2, ..., v_n]^T$ denotes the inverse transformed watermark.

The signal $\mathbf{v}$ is inserted as the same information in all columns of the spatial domain. This means that the embedded patterns from the j-th row have the same v_j signal. Therefore, embedded watermark $\mathbf{v}$ has the following DIBR invariance,

$$\mathbf{D}([\mathbf{v}, \mathbf{v}, ..., \mathbf{v}]) = \begin{bmatrix} v_1 & v_1 & \cdots & v_1 \\ v_2 & v_2 & \cdots & v_2 \\ \vdots & \vdots & \ddots & \vdots \\ v_n & v_n & \cdots & v_n \end{bmatrix} = [\mathbf{v}, \mathbf{v}, ..., \mathbf{v}] \tag{9}$$

225   where $\mathbf{D}(\cdot)$ is the DIBR process described in Section 2.1.

Since DIBR only translates pixels in the horizontal direction, $\mathbf{v}$ is not deformed. Therefore, the watermark can be extracted in the frequency domain as follows,

$$\begin{aligned} \mathbf{DCT}(\mathbf{D}(I'_{i:})) &= \mathbf{DCT}(\mathbf{D}(I_{i:} + \mathbf{v})) \\ &= \mathbf{DCT}(\mathbf{D}(I_{i:}) + \mathbf{D}(\mathbf{v})) \\ &= \mathbf{DCT}(\mathbf{D}(I_{i:}) + \mathbf{v}) \\ &= \mathbf{DCT}(\mathbf{D}(I_{i:})) + \mathbf{DCT}(\mathbf{v}) \\ &= \mathbf{DCT}(\mathbf{D}(I_{i:})) + \mathbf{w}. \end{aligned} \tag{10}$$

226   Hence w can be extracted in the frequency domain without being damaged by DIBR in extraction
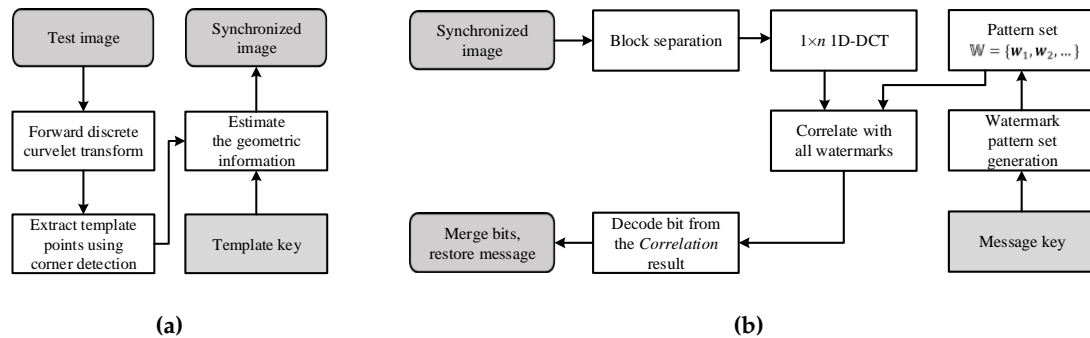227   process.

Figure 8. Proposed extraction process. (a) Template decoding and image recovery from geometric distortions, (b) Message watermark extraction.

To compensate for capacity decrease caused by template insertion, a data coding technique can be utilized to insert more than 1 bit per block [35]. The message watermark embedding process is divided into four steps as follows.

a) Column-by-column 1D-DCT: Each column of the message block is transformed by 1D-DCT.

b) Data encoding: A pseudo random watermark pattern set (i.e., a set of watermark patterns) is generated. The length of the set is determined by the user, and the capacity of the block is determined according to the length. For example, to represent 4 bits of information per block, 16 (i.e., 2^4) unique watermark patterns are generated[43].

c) Watermark embedding: The proposed method embeds the watermark based on the spread-spectrum[44]. The generated watermark pattern set $\mathbf{w}_b = \{w_{b1}, w_{b2}, \ldots, w_{bL}\}$ is embedded into the middle frequency of the DCT signal $\mathbf{s} = \{s_1, s_2, \ldots, s_L\}$ compromising both robustness and invisibility. The embedding equation is as follows,

$$s_i' = s_i + \alpha_m |s_i| w_{bi} \tag{11}$$

where i is the block column, $1 \leq i \leq L$, L is the length of the watermark signal, $\mathbf{s} = \{s_1, s_2, \ldots, s_L\}$ is the original signal, $\mathbf{s}' = \{s_1', s_2', \ldots, s_L'\}$ is the watermarked signal, α_m is the message watermark strength, w_{bL} is the watermark pattern, and b is the message inserted into the block. For example, if $b = 7$ (111 in binary form), the message inserted in the block is 111 (in this case, pseudo-random pattern set length = 2^3). The embedding step (11) is repeated for all columns in the block.

d) 1D-IDCT: The watermarked block is reconstructed by 1D-IDCT.

3.3. Proposed Watermark Extraction Method

This section describes the details of the proposed watermark extraction method. The overall process is illustrated in Fig. 8. Before extracting the messages, the image must be synchronized using a template.

3.3.1. Template Decoding

a) Forward curvelet transform: The curvelet transform is applied to the test image.

b) Extract template points using corner detection: A peak was inserted into the l_1 and l_2 pair in the embedding step, with 90° difference at each template point. Due to curvelet filter characteristics, the peak point spreads in a straight line in the corresponding direction. Therefore, the peak point is represented by an "X" shape when only coefficients of l_1 and l_2 are extracted, as shown in Fig. 9, and can be found by corner detection. This paper employed Harris corner detection, but similar approaches would also suffice.
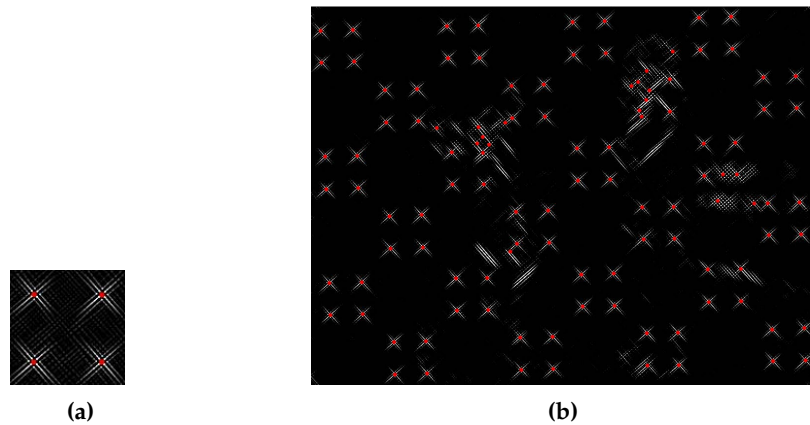
Figure 9. Template extraction results (a) for one template block; (b) for whole rotated image. Red dots are points detected by corner detection.

c) Estimate degree of geometric deformation by exploiting modified ICP: The ICP method estimates the parameters of geometric distortion when there is no matching information of two point clouds. This method assumes that the closest points between the two point clouds match each other and repeats the process of minimizing the error[45].

This paper modifies the ICP to suit the problem in need of solving. Since the DIBR image has an error in the horizontal information, the weight of the horizontal distance error is set to $1/2$.

Geometric distortion parameters are estimated using the modified ICP method, comparing detected corner points with template point ground truth. The template ground truth can be generated using the key as in the embedding step.

d) Recover the test image using the estimated geometric distortion parameters.

If the degree of geometric transformation is large, the direction in which the template was inserted may have changed. For example, if the image was significantly rotated, the template inserted in direction pairs l_1 and l_2 is moved to the direction pairs $l_1 + 1$ and $l_2 + 1$. In this case, the template decoding process is repeated for all direction pairs, and the estimated geometric parameters with the lowest ICP error is used.

3.3.2. Message Watermark Extraction

The message is extracted from the image recovered by template. Message extraction consists of six steps:

a) Split synchronized image into blocks: Divide the recovered image into blocks as in the embedding step.

b) Column-by-column 1D-DCT: The same process applied in the embedding step is applied.

c) Correlation: The correlation is conducted as follows,

$$Corr(b) = \frac{\mathbf{s}' \cdot \mathbf{w}_b}{L} = \frac{1}{L} \sum_{i=1}^{L} s'_i \cdot w_{bi} \tag{12}$$

The notations are identical to those in the embedding step.

e) Bit decoding: Bits are decoded from the correlation result. For example, if correlation value with w_{3i} is the highest, then the bit decoded from the ith column is 011. The length of the bit is determined by the length of the watermark pattern set, and in this case, watermark pattern set length $= 2^3$. Different bits may be decoded for each column, and the majority voting method is used to determine the bit of the block.
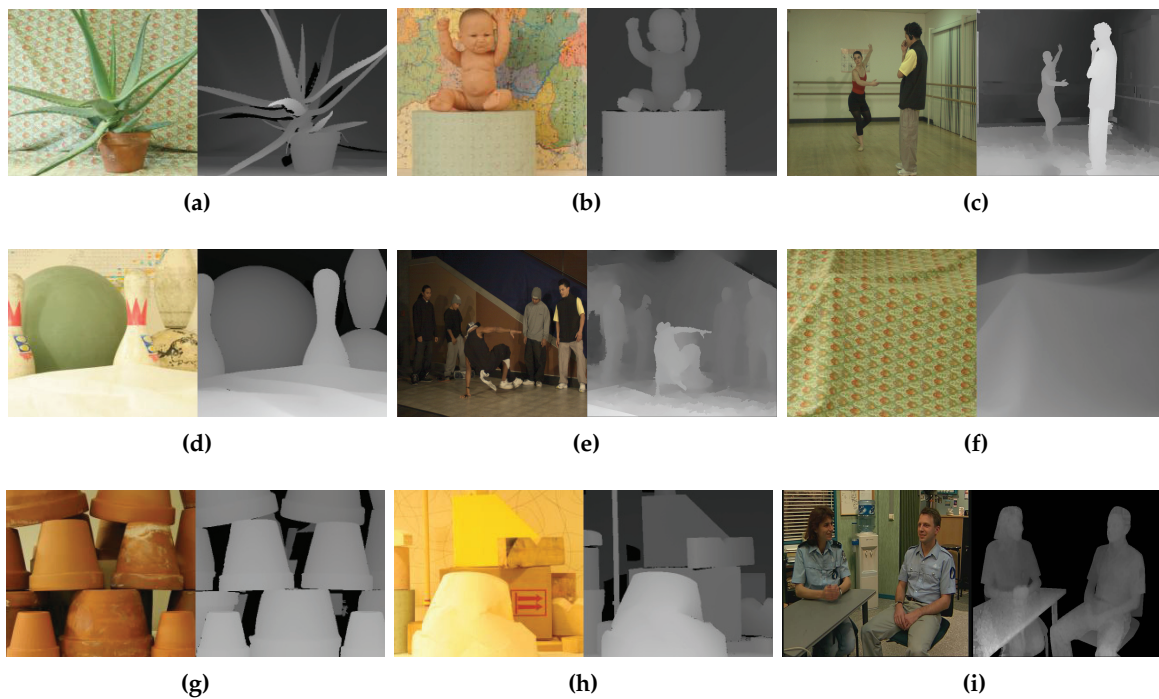
Figure 10. Test image sets. (a) aloe; (b) baby; (c) ballet; (d) bowling; (e) breakdancers; (f) cloth; (g) flowerpot; (h) lampshade; (i) interview.

f) Restore messages by merging the bits: The messages are recovered by merging the bits extracted from each message block.

4. Experimental Results

This section evaluates the proposed method performance in terms of imperceptibility and robustness to various distortions. The proposed method was compared with other blind DIBR image watermarking systems, specifically Lin's method [18] and Kim's method[19]. The bit capacity of all methods was set to 64.

4.1. Experiment Setting

The test image sets were obtained from Heinrich Hertz Institute[46], Middlebury[47–50], and Microsoft Research 3D Video Datasets[51]. Fig. 10 shows pairs of center and depth images of the test image sets. They have various resolutions from 900×720 to 1800×1500. The total number of images used in the experiment was about 1800.

The DIBR parameters are set to focal length $f = 1$ and baseline distance $t_x = 5\%$ of the image width, which are the recommended value for comfortable viewing conditions. Linear interpolation is used as hole filling for simplicity and without loss of generality.

For Lin's method, block size was set from 100×100 to 200×200 to match embedding capacity with other methods. Watermark strength was set as $\alpha = 1$, watermark pattern length = 5120, and the beginning of the embedding position is the 2560-th coefficient of the zigzag scan order in the DCT domain.

For Kim's method, block size was set to $(w/8)\times(h/8)$. The weighting factor for coefficient magnitude was set as $\mathbb{W} = 450$, maximum quantization level $maxBit = 2$, and minimum difference between paired coefficients $errMin = 8$. These values are demonstrated in [19].

In the proposed method, the size of each block is set set to $(w/8)\times(h/8)$. In the template embedding process, the template strengths α_t and β_t are set to 5 and 50, experimentally. In the message watermark
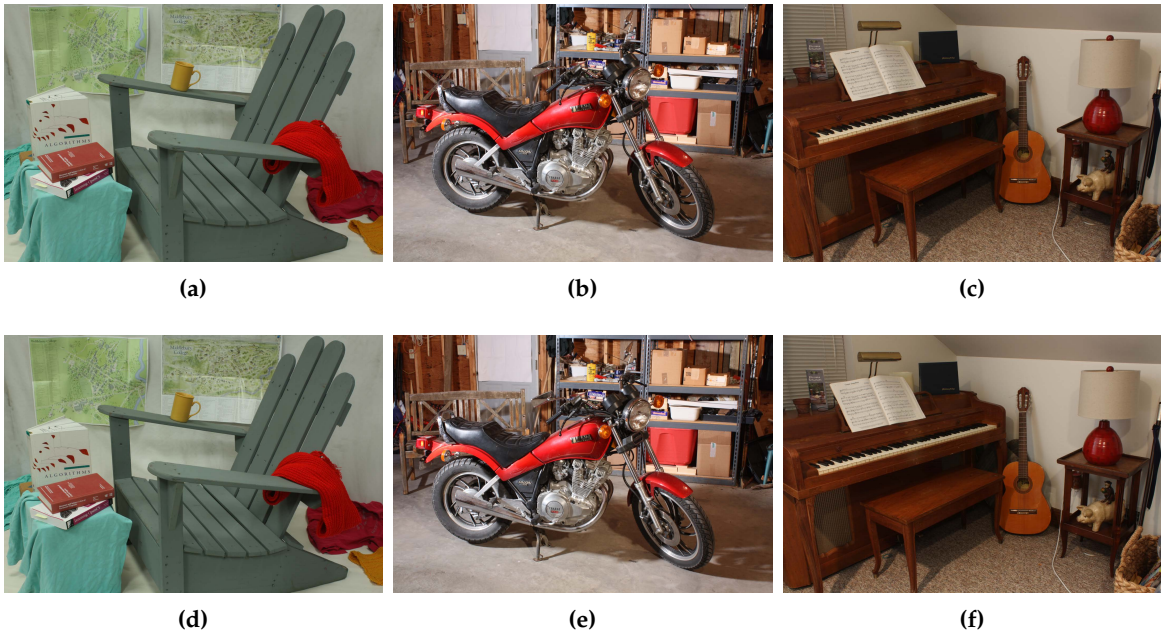
Figure 11. (a)–(c) Original 'adirondack', 'motorcycle' and 'piano'; (d)–(f) Watermarked 'adirondack', 'motorcycle' and 'piano' by the proposed method.

Table 4. Average PSNR and SSIM.

	PSNR	SSIM
Proposed method	50.03	0.9942
Lin's method	45.12	0.9936
Kim's method	44.02	0.9878

embedding process, α_m is set to 0.5, the length of the watermark pattern is 40, and the start of the embedding position is the 45th coefficient of 1D-DCT.

4.2. Image Quality

As shown in Fig. 11, the quality degradation of the watermarked image is not noticeable. For more accurate image quality measurements, the peak signal-to-noise ratio (PSNR) and structure similarity (SSIM)[52] were measured. The average PSNR and SSIM are shown in Table 4.

Lin's method has a lower PSNR despite it exploiting a spread spectrum-based watermarking method that is similar to the proposed method. Since the 2D-DCT is not invariant to DIBR, the Lin's method must insert three watermarks in a superimposed manner in order to protect left, right, and center image. For this reason, the inserted watermark energy is very large.

In Kim's method, the images are seriously blurred, because this method cuts the coefficient off excessively. As a result, both PSNR and SSIM values were low.

The proposed method increases imperceptibility by taking advantage of the curvelet and 1D-DCT, which are robust against DIBR. Due to this robustness, the proposed method does not require excessive insertion of the watermark. In particular, the message watermark using the 1D-DCT does not require watermark insertion in a superimposed manner, so the insertion energy can be reduced to about one-third as compared with the 2D-DCT. In addition, since the template and the message can be inserted into different blocks after dividing the block, it is possible to prevent the invisibility degradation caused by the overlapping of templates and message watermarks. For these reasons, PSNR outperforms the other methods and shows similar performance to the best methods according to SSIM.

Table 5. Average BER of DIBR images

	Center image	Left image	Right image
Proposed method	0.002	**0.002**	**0.002**
Lin's method	**0.000**	0.002	0.004
Kim's method	0.006	0.008	0.009

Table 6. Average BER of synthesized images using preprocessed depth image.

	Left image (Preprocessed depth)	Right image (Preprocessed depth)
Proposed method	**0.002**	**0.002**
Lin's method	0.003	0.007
Kim's method	0.008	0.009

4.3. Robustness against DIBR

Table 5 shows the results of a DIBR attack. All three methods have low BER against the DIBR attack. Since Lin's method inserts multiple watermarks, the BER is slightly higher on the right image. The previously inserted watermark (right watermark) was disturbed by the later inserted watermarks (left and center watermarks).

The major advantage of the DIBR process is that DIBR configurations can be adjusted to suit a user's needs. As described in Section 2.1, the user can preprocess the depth image to increase the rendered image quality. Table 6 shows robustness results for synthesized images with a preprocessed depth image. Unlike other methods, Lin's method raised the BER. This result shows that Lin's method is susceptible to the shift distance of the pixel, because the left and right watermarks were inserted by predicting the shift distance of the pixels.

Baseline distance is another DIBR configuration aspect. Various viewpoints of an image can be synthesized depending on the baseline distance change. Figure 12 shows average BER where the baseline distance is adjusted from 1 to 10%. Lin's method increases BER for the same reason as in-depth image preprocessing, whereas Kim's and the proposed method do not increase BER in the baseline distance adjustment.

4.4. Robustness against Signal Distortion

Tables 7 and 8 show the average BER for the signal distorted center, left, and right images, and Fig. 13 illustrates the average BER for the signal distorted right image.

For additive noise, Lin's method has the best performance, and the proposed method has a slightly higher BER than Lin's method. However, additive Gaussian noise with a variance of 2000 is a very
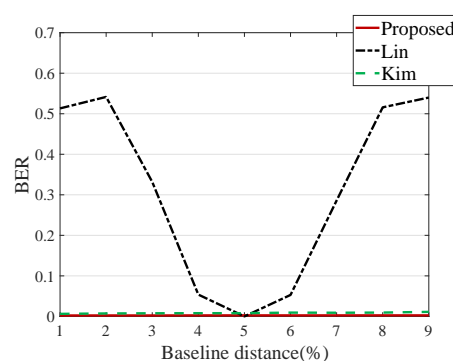

Figure 12. Average BER for baseline distance adjustment.

Table 7. Average BER for Gaussian noise attack

Noise var.	Proposed			Lin			Kim		
	Left	Center	Right	Left	Center	Right	Left	Center	Right
10	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.04	0.03	0.04
40	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.12	0.09	0.111
300	0.01	0.01	0.01	**0.00**	**0.00**	**0.00**	0.38	0.36	0.38
1250	0.04	0.04	0.05	**0.02**	**0.02**	**0.02**	0.54	0.54	0.55

Table 8. Average BER for JPEG attack

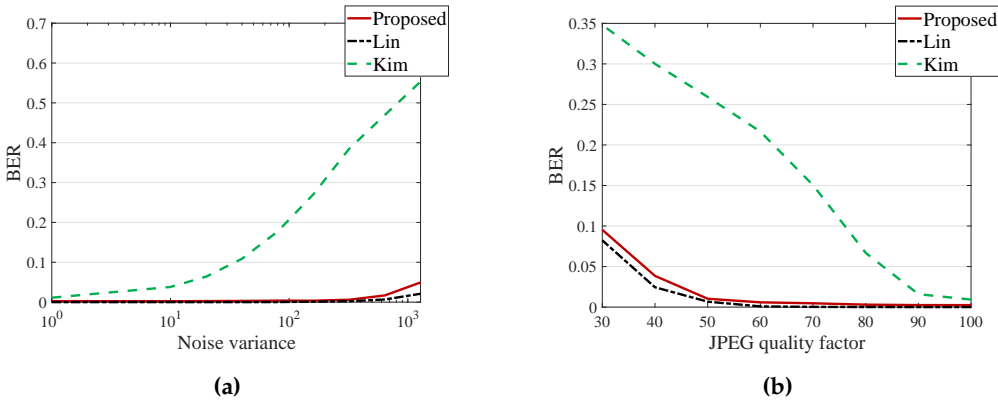| Quality factor	Proposed			Lin			Kim		
	Left	Center	Right	Left	Center	Right	Left	Center	Right
90	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.02	0.01	0.02
60	0.01	0.01	0.01	**0.00**	**0.00**	**0.00**	0.22	0.20	0.22
30	0.10	0.09	0.10	**0.09**	**0.08**	**0.10**	0.35	0.33	0.35

(a)   **(b)**

Figure 13. Average BER for signal distorted right images. (a) additive Gaussian noise; (b) JPEG compression.
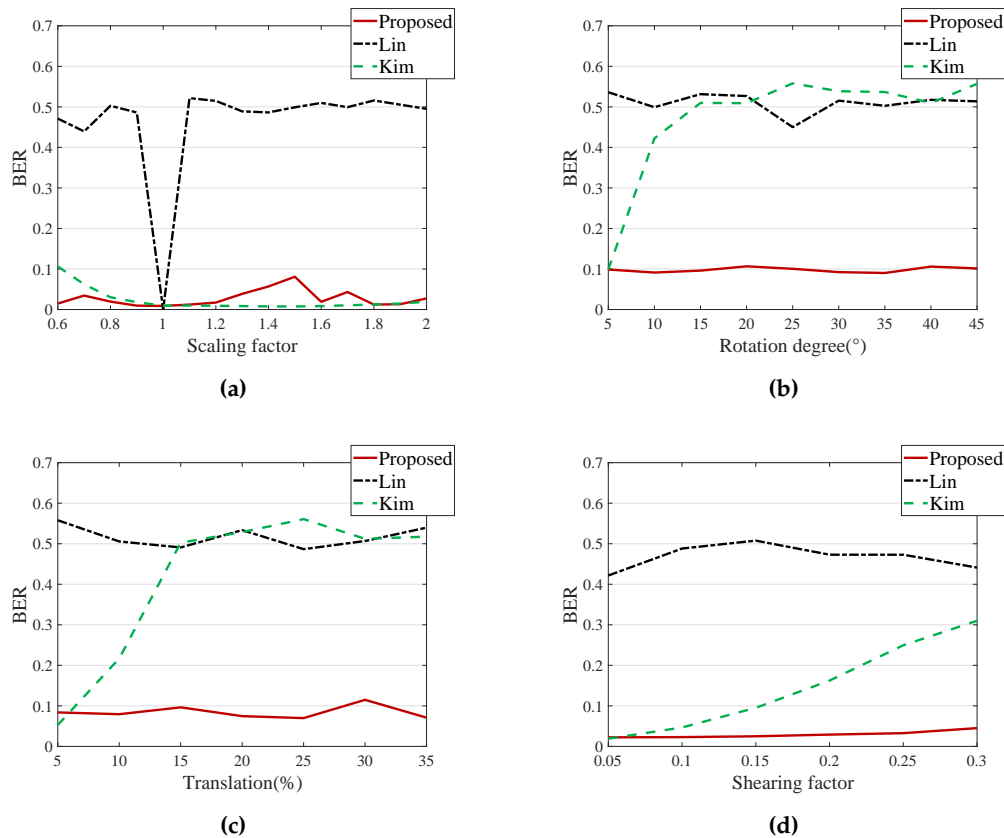
Figure 14. Average BER for geometric distortion. This experiment was conducted on the right image. (a) image size scaling; (b) image rotation; (c) image translation; (d) image shearing

Table 9. Average BER for scaling attack

Scaling factor	Proposed			Lin			Kim		
	Left	Center	Right	Left	Center	Right	Left	Center	Right
0.6	**0.01**	**0.01**	**0.02**	0.48	0.49	0.49	0.08	0.01	0.10
0.8	**0.02**	**0.01**	**0.02**	0.47	0.50	0.52	0.01	0.02	0.03
1.4	0.09	0.06	0.08	0.52	0.51	0.49	**0.01**	**0.01**	**0.01**
2.0	0.04	0.03	0.04	0.49	0.50	0.50	**0.02**	**0.02**	**0.02**

346 severe attack and such a large amount of noise barely occurs. Considering this, the proposed method
347 is robust enough to additive noise.
348　　　For JPEG compression, the proposed method has a slightly higher BER than Lin's method.
349 However, since the proposed method exhibits error < 0.1 for JPEG quality factor 30, which is very large
350 compression, it can be considered sufficiently strong against a JPEG attack.
351　　　Meanwhile, Kim's method is more vulnerable to signal distortion than the other methods, since
352 the quantized DT-CWT coefficients are greatly affected by the signal distortion.

4.5. Robustness against Geometric Distortion

354　　　A robustness test against geometric distortion was also conducted with center, left, and right
355 images. Tables 9–12 show the average BER for the geometrically distorted center, left, and right images,
356 and Fig. 14 illustrates the average BER for the geometrically distorted right image.
357　　　Lin's method does not show good results against geometric distortion, because this method
358 embeds the watermark in the DCT domain, which is vulnerable to geometric distortion.

Table 10. Average BER for rotation attack

Rotation degree	Proposed			Lin			Kim		
	Left	Center	Right	Left	Center	Right	Left	Center	Right
10	**0.09**	**0.05**	**0.10**	0.45	0.54	0.51	0.43	0.41	0.42
20	**0.10**	**0.07**	**0.09**	0.52	0.50	0.49	0.54	0.55	0.51
30	**0.09**	**0.06**	**0.09**	0.50	0.53	0.53	0.52	0.54	0.54
40	**0.11**	**0.08**	**0.12**	0.55	0.56	0.50	0.50	0.55	0.54

Table 11. Average BER for translation attack

Transl-ation(%)	Proposed			Lin			Kim		
	Left	Center	Right	Left	Center	Right	Left	Center	Right
10	**0.08**	**0.06**	**0.09**	0.54	0.51	0.51	0.24	0.22	0.22
20	**0.07**	**0.05**	**0.08**	0.53	0.50	0.53	0.55	0.55	0.55
30	**0.11**	**0.07**	**0.12**	0.49	0.47	0.50	0.53	0.53	0.53

Table 12. Average BER for shearing attack

Shearing factor	Proposed			Lin			Kim		
	Left	Center	Right	Left	Center	Right	Left	Center	Right
0.1	**0.03**	**0.02**	**0.02**	0.54	0.51	0.52	0.24	0.22	0.22
0.2	**0.03**	**0.03**	**0.03**	0.48	0.53	0.49	0.53	0.50	0.51
0.3	**0.05**	**0.05**	**0.05**	0.50	0.51	0.47	0.52	0.53	0.56

Table 13. Average BER of proposed method for combination geometric distortion. The experiment was conducted to the synthesized right image.

Translation(%)	Rotation°			Scaling factor	Rotation°			Scaling factor	Translation(%)		
	3	6	9		3	6	9		7	14	21
5	0.11	0.11	0.11	0.8	0.13	0.10	0.12	0.8	0.13	0.12	0.13
10	0.12	0.12	0.12	0.9	0.10	0.09	0.09	0.9	0.12	0.11	0.11
15	0.11	0.12	0.12	1.1	0.17	0.17	0.18	1.1	0.13	0.12	0.13
20	0.12	0.12	0.12	1.2	0.23	0.25	0.28	1.2	0.15	0.14	0.15

Kim's method shows good performance against scaling but is weak to rotation, translation, and shearing. This method does not lose synchronization information in scaling. However, the rotation, translation, and shearing attacks break the block synchronization. This is because the block size is specified as the ratio of the image, such as $(width/M) \times (height/N)$.

The proposed method shows good performance on this test, because it utilizes a template robust to geometric attacks. Since the images were recovered from geometric distortion using this template, the message watermark has a low error.

Experiments were also conducted on the combination of geometric distortion in the proposed method. This experiments were conducted by combining two of rotation, translation, and scaling, and the results are shown in Table 13. Rotation with translation and scaling with translation show good results. However, when rotation and scaling occur at the same time, the watermark signal is greatly weakened, so the error is higher than with other geometric combination attacks.

5. Conclusion

Due to the advent of new 3D applications, DIBR has taken on an important role in 3D content. To protect the copyright of such content, this paper proposed a template-based DIBR watermarking system. Ensuring robustness against geometric attacks in rendered images requires that watermarks should be robust to both DIBR and geometric attacks. In order to have robustness to a combination of linear and non-linear attacks, a watermarking system was designed by combining two methods: template and message watermark. Inserting a template into the curvelet domain robust to DIBR allowed this method restore the geometrically distorted image. Then, the message was extracted using a 1D-DCT message watermarking method that was invariant to the DIBR. In experimental results, the proposed method showed high image quality. In terms of robustness, it had low BER to DIBR configuration adjustment as well as standard configured DIBR. Additionally, the results showed that the proposed method is very robust against noise addition and JPEG compression. For geometric distortion, such as scaling, rotation, translation, and shearing, good performance was also demonstrated. However, this method still does not consider the robustness of video coding, which is often used in 3D video, such as high-efficiency video coding (HEVC). Therefore, future work will focus on extending this method to 3D videos.

Author Contributions: Conceptualization, Wook-Hyung Kim; Investigation, Jong-Uk Hou; Methodology, Wook-Hyung Kim and Jong-Uk Hou; Resources, Han-Ul Jang; Software, Wook-Hyung Kim; Supervision, Heung-Kyu Lee; Validation, Han-Ul Jang; Writing – original draft, Wook-Hyung Kim.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fehn, C. A 3D-TV approach using depth-image-based rendering (DIBR). Proc. of VIIP, 2003, Vol. 3.
2. Fehn, C. Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV. Electronic Imaging 2004. International Society for Optics and Photonics, 2004, pp. 93–104.
3. Zhang, L.; Tam, W.J. Stereoscopic image generation based on depth images for 3D TV. *Broadcasting, IEEE Transactions on* **2005**, *51*, 191–199.
4. Chen, W.Y.; Chang, Y.L.; Lin, S.F.; Ding, L.F.; Chen, L.G. Efficient depth image based rendering with edge dependent depth filter and interpolation. Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on. IEEE, 2005, pp. 1314–1317.
5. Ens, J.; Lawrence, P. An investigation of methods for determining depth from focus. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **1993**, *15*, 97–108.
6. Battiato, S.; Curti, S.; La Cascia, M.; Tortora, M.; Scordato, E. Depth map generation by image classification. Electronic Imaging 2004. International Society for Optics and Photonics, 2004, pp. 95–104.

7.   Moustakas, K.; Tzovaras, D.; Strintzis, M.G. Stereoscopic video generation based on efficient layered structure and motion estimation from a monoscopic image sequence. _Circuits and Systems for Video Technology, IEEE Transactions on_ **2005**, _15_, 1065–1073.

8.   Tam, W.J.; Zhang, L. 3D-TV content generation: 2D-to-3D conversion. Multimedia and Expo, 2006 IEEE International Conference on. IEEE, 2006, pp. 1869–1872.

9.   Zhang, L.; Vázquez, C.; Knorr, S. 3D-TV content creation: automatic 2D-to-3D video conversion. _Broadcasting, IEEE Transactions on_ **2011**, _57_, 372–383.

10.   Feng, Y.; Ren, J.; Jiang, J. Object-based 2D-to-3D video conversion for effective stereoscopic content generation in 3D-TV applications. _Broadcasting, IEEE Transactions on_ **2011**, _57_, 500–509.

11.   Holliman, N.; Dodgson, N.; Favalora, G.; Pockett, L. Three-Dimensional Displays: A Review and Applications Analysis. _Broadcasting, IEEE Transactions on_ **2011**, _57_, 362–371. doi:10.1109/TBC.2011.2130930.

12.   Urey, H.; Chellappan, K.; Erden, E.; Surman, P. State of the Art in Stereoscopic and Autostereoscopic Displays. _Proceedings of the IEEE_ **2011**, _99_, 540–555. doi:10.1109/JPROC.2010.2098351.

13.   Merkle, P.; Smolic, A.; Müller, K.; Wiegand, T. Multi-view video plus depth representation and coding. Image Processing, 2007. ICIP 2007. IEEE International Conference on. IEEE, 2007, Vol. 1, pp. I–201.

14.   Cheng, C.M.; Lin, S.J.; Lai, S.H. Spatio-Temporally Consistent Novel View Synthesis Algorithm From Video-Plus-Depth Sequences for Autostereoscopic Displays. _Broadcasting, IEEE Transactions on_ **2011**, _57_, 523–532. doi:10.1109/TBC.2011.2139090.

15.   Smolic, A.; Mueller, K.; Merkle, P.; Kauff, P.; Wiegand, T. An overview of available and emerging 3D video formats and depth enhanced stereo as efficient generic solution. Picture Coding Symposium, 2009. PCS 2009, 2009, pp. 1–4. doi:10.1109/PCS.2009.5167358.

16.   Shade, J.; Gortler, S.; He, L.w.; Szeliski, R. Layered depth images. Proceedings of the 25th annual conference on Computer graphics and interactive techniques. ACM, 1998, pp. 231–242.

17.   Bartczak, B.; Vandewalle, P.; Grau, O.; Briand, G.; Fournier, J.; Kerbiriou, P.; Murdoch, M.; Müller, M.; Goris, R.; Koch, R.; others. Display-independent 3D-TV production and delivery using the layered depth video format. _Broadcasting, IEEE Transactions on_ **2011**, _57_, 477–490.

18.   Lin, Y.H.; Wu, J.L. A digital blind watermarking for depth-image-based rendering 3D images. _Broadcasting, IEEE Transactions on_ **2011**, _57_, 602–611.

19.   Kim, H.D.; Lee, J.W.; Oh, T.W.; Lee, H.K. Robust DT-CWT watermarking for DIBR 3D images. _Broadcasting, IEEE Transactions on_ **2012**, _58_, 533–543.

20.   Wang, S.; Cui, C.; Niu, X. Watermarking for DIBR 3D images based on SIFT feature points. _Measurement_ **2014**, _48_, 54–62.

21.   Miao, H.; Lin, Y.H.; Wu, J.L. Image descriptor based digital semi-blind watermarking for DIBR 3D images. International Workshop on Digital Watermarking. Springer, 2014, pp. 90–104.

22.   Cui, C.; Wang, S.; Niu, X. A novel watermarking for DIBR 3D images with geometric rectification based on feature points. _Multimedia Tools and Applications_ **2015**, pp. 1–29.

23.   Asikuzzaman, M.; Alam, M.J.; Lambert, A.J.; Pickering, M.R. A blind watermarking scheme for depth-image-based rendered 3D video using the dual-tree complex wavelet transform. Image Processing (ICIP), 2014 IEEE International Conference on. IEEE, 2014, pp. 5497–5501.

24.   Asikuzzaman, M.; Alam, M.J.; Lambert, A.J.; Pickering, M.R. Robust DT CWT-Based DIBR 3D Video Watermarking Using Chrominance Embedding. _IEEE Transactions on Multimedia_ **2016**, _18_, 1733–1748.

25.   Fleet, D.J.; Heeger, D.J. Embedding invisible information in color images. Image Processing, 1997. Proceedings., International Conference on. IEEE, 1997, Vol. 1, pp. 532–535.

26.   Pereira, S.; Pun, T. Robust template matching for affine resistant image watermarks. _Image Processing, IEEE Transactions on_ **2000**, _9_, 1123–1129.

27.   Barni, M. Effectiveness of exhaustive search and template matching against watermark desynchronization. _IEEE Signal Processing Letters_ **2005**, _12_, 158–161.

28.   Zheng, D.; Liu, Y.; Zhao, J.; Saddik, A.E. A survey of RST invariant image watermarking algorithms. _ACM Computing Surveys (CSUR)_ **2007**, _39_, 5.

29.   Tam, W.J.; Alain, G.; Zhang, L.; Martin, T.; Renaud, R. Smoothing depth maps for improved steroscopic image quality. Optics East. International Society for Optics and Photonics, 2004, pp. 162–172.

30.   Fu, D.; Zhao, Y.; Yu, L. Temporal consistency enhancement on depth sequences. Picture Coding Symposium (PCS), 2010. IEEE, 2010, pp. 342–345.

31. Lee, P.J.; others. Nongeometric distortion smoothing approach for depth map preprocessing. *Multimedia, IEEE Transactions on* **2011**, *13*, 246–254.

32. Oh, K.J.; Yea, S.; Ho, Y.S. Hole filling method using depth based in-painting for view synthesis in free viewpoint television and 3-d video. Picture Coding Symposium, 2009. PCS 2009. IEEE, 2009, pp. 1–4.

33. Chen, K.Y.; Tsung, P.K.; Lin, P.C.; Yang, H.J.; Chen, L.G. Hybrid motion/depth-oriented inpainting for virtual view synthesis in multiview applications. 3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2010. IEEE, 2010, pp. 1–4.

34. Chen, W.Y.; Chang, Y.L.; Lin, S.F.; Ding, L.F.; Chen, L.G. Efficient depth image based rendering with edge dependent depth filter and interpolation. Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on. IEEE, 2005, pp. 1314–1317.

35. Cox, I.; Miller, M.; Bloom, J.; Fridrich, J.; Kalker, T. *Digital watermarking and steganography*; Morgan Kaufmann, 2007.

36. Candes, E.J.; Donoho, D.L. Curvelets: A surprisingly effective nonadaptive representation for objects with edges. Technical report, DTIC Document, 2000.

37. Candès, E.J.; Guo, F. New multiscale transforms, minimum total variation synthesis: Applications to edge-preserving image reconstruction. *Signal Processing* **2002**, *82*, 1519–1543.

38. Candès, E.J.; Donoho, D.L. New tight frames of curvelets and optimal representations of objects with piecewise C2 singularities. *Communications on pure and applied mathematics* **2004**, *57*, 219–266.

39. Candes, E.; Demanet, L.; Donoho, D.; Ying, L. Fast discrete curvelet transforms. *Multiscale Modeling & Simulation* **2006**, *5*, 861–899.

40. Sumana, I.J.; Islam, M.M.; Zhang, D.; Lu, G. Content based image retrieval using curvelet transform. Multimedia Signal Processing, 2008 IEEE 10th Workshop on. IEEE, 2008, pp. 11–16.

41. Islam, M.M.; Zhang, D.; Lu, G. Rotation invariant curvelet features for texture image retrieval. Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on. IEEE, 2009, pp. 562–565.

42. Cox, I.J.; Kilian, J.; Leighton, F.; Shamoon, T. Secure spread spectrum watermarking for multimedia. *Image Processing, IEEE Transactions on* **1997**, *6*, 1673–1687. doi:10.1109/83.650120.

43. Barni, M.; Bartolini, F.; Cappellini, V.; Piva, A. A DCT-domain system for robust image watermarking. *Signal processing* **1998**, *66*, 357–372.

44. Barni, M.; Bartolini, F.; Cappellini, V.; Piva, A. A DCT-domain system for robust image watermarking. *Signal processing* **1998**, *66*, 357–372.

45. Rusinkiewicz, S.; Levoy, M. Efficient variants of the ICP algorithm. 3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on. IEEE, 2001, pp. 145–152.

46. Fehn, C. Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV. Electronic Imaging 2004. International Society for Optics and Photonics, 2004, pp. 93–104.

47. Scharstein, D.; Szeliski, R. High-accuracy stereo depth maps using structured light. Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on. IEEE, 2003, Vol. 1, pp. I–195.

48. Scharstein, D.; Pal, C. Learning conditional random fields for stereo. Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on. IEEE, 2007, pp. 1–8.

49. Hirschmüller, H.; Scharstein, D. Evaluation of cost functions for stereo matching. Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on. IEEE, 2007, pp. 1–8.

50. Scharstein, D.; Hirschmüller, H.; Kitajima, Y.; Krathwohl, G.; Nešić, N.; Wang, X.; Westling, P. High-resolution stereo datasets with subpixel-accurate ground truth. German Conference on Pattern Recognition. Springer, 2014, pp. 31–42.

51. Zitnick, C.L.; Kang, S.B.; Uyttendaele, M.; Winder, S.; Szeliski, R. High-quality video view interpolation using a layered representation. ACM Transactions on Graphics (TOG). ACM, 2004, Vol. 23, pp. 600–608.

52. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: from error visibility to structural similarity. *Image Processing, IEEE Transactions on* **2004**, *13*, 600–612.