Signal Processing: Image Communication 54 (2017) 56–65

Contents lists available at ScienceDirect

Signal Processing: Image Communication

journal homepage: www.elsevier.com/locate/image

Robust contourlet-based blind watermarking for depth-image-based rendering 3D images

Lei Chen, Jiying Zhao*

Multimedia Communications Research Laboratory, School of Electrical Engineering and Computer Science, University of Ottawa, 800 King Edward Avenue, Ottawa, Ontario, Canada K1N 6N5

ARTICLE INFO

Keywords:
3D watermarking
Depth-image-based rendering
Contourlet transform
Coefficient quantization
Bit error rate

ABSTRACT

The copyright protection for 3D multimedia has attracted considerable attention and depth-image-based rendering (DIBR) has been playing a critical role in 3D content representation due to its numerous advantages. In this paper, we propose a new blind watermarking scheme for DIBR 3D images. Considering the directional multiresolution image representation and convenient tree structures of contourlet transform (CT), we perform the watermark embedding and extraction in contourlet domain. The center view and the depth map are available at the content provider side. After applying contourlet transform to the center view, we embed the watermark into the selected contourlet subbands of the center view by quantization on certain contourlet coefficients. The virtual left and right views are generated from the watermarked center view and the associated depth map using DIBR technique at the receiver side. The statistical differences between quantized and unquantized contourlet coefficients are used for watermark extraction. The watermark can be detected with a low bit error rate (BER) from the center view, the left and right views even when each view is distorted and distributed separately. The simulation results demonstrate that the proposed scheme is very robust to image compression, noise addition and geometric attacks such as rotation, scaling and cropping. Moreover, the proposed scheme has good performance in terms of depth image variation and baseline distance adjustment.

1. Introduction

Due to the rapid development of the 3D display market, the protection and authentication of the intellectual property rights of 3D multimedia has become an essential concern. As a consequence, the digital watermarking for 3D content is attracting considerable attention [1,2]. A good digital watermarking scheme for 3D images should be robust to various possible attacks. Geometric transformations and image compression are two main difficult distortions to tackle in 3D image watermarking.

There are two major existing techniques for 3D content representation: stereo image recording (SIR) [3,4] and depth-image-based rendering (DIBR) [5–8]. The left view and right view are recorded simultaneously from two cameras in the SIR approach, which creates a good viewing for the audience. In [3], a real-time stereo image watermarking method based on discrete cosine transform (DCT) and disparity map is proposed. However, it is very costly and impractical to set the two cameras under the same conditions [2]. By contrast, the DIBR technique only transmits the center view and the corresponding depth information, and generates the virtual left view and right view at the content consumer side. It has advantages in terms of bandwidth cost and 2D to 3D conversion. The depth map can be compressed before transmission and the viewers can modify the depth map to adjust the depth degree. In [5], a new scheme on 3D-TV using DIBR and its actual implementation are described. The studies in ATTEST project [6] present the advantages of DIBR 3D image representation over the traditional stereoscopic displays for 3D TV broadcasting.

Different from ordinary 2D digital watermarking [9–11], there are more specific requirements for 3D watermarking, especially for DIBR 3D image watermarking. Not only the center view, but also the virtual left and right views can be illegally distributed. Therefore, the embedded watermark information should be detected from these three views separately for identification. However, quantities of 2D watermarking techniques [12,13] fail to resist 2D to 3D conversion and cannot be directly applied to 3D watermarking.

Many researches have been conducted on digital 3D watermarking [14–16] and some DIBR watermarking algorithms have been proposed [1,2,8]. Halici et al. [8] proposed a novel DIBR watermarking method for free viewpoint television systems, where a watermark pattern was warped for every different view of multi-view video. But this method is

* Corresponding author.
E-mail address: jzhao@uottawa.ca (J. Zhao).

http://dx.doi.org/10.1016/j.image.2017.02.011
Received 18 August 2016; Received in revised form 25 February 2017; Accepted 27 February 2017
Available online 06 March 2017
0923-5965/

not blind since the original images are needed for watermark detection. Lin et al. [1] proposed a digital watermarking algorithm for DIBR 3D images, where the watermark embedded in the center view was kept for the left view and right view after rendering. This approach can resist JPEG compression and noise addition but geometric attacks were not taken into account. Kim et al. [2] proposed a robust watermarking method based on dual-tree complex wavelet transform (DT-CWT). The embedded watermark can be extracted from the center, left and right views under distortions. The algorithm is robust to baseline distance adjustment and depth map configuration. However, the approach has poor performance under combined attacks. The contourlet transform (CT), introduced by Do and Vetterli [17], gives an efficient directional multiresolution image representation. Compared with other transforms, the contourlet transform has better performance in capturing the contours and edges of an image, which can be used for watermark embedding. The contourlet transform can also reconstruct the image accurately with appropriate reconstruction filters. In our previous work [18], we proposed a new 2D image watermarking scheme robust to geometric transformations and image compression. The scheme is based on contourlet transform. The watermark is embedded into the largest contourlet coefficients of the directional subbands of the cover image. The experimental results show that the scheme is robust to image processing attacks, such as rotation, scaling, translation (RST) and image compression. In [19], we reported our new blind watermarking scheme for depth-image-based rendering (DIBR) 3D images, which is also based on contourlet transform [18].

In this paper, we extend our work in [19] substantially to implement a practical watermarking scheme for DIBR 3D images. The center view is transformed into contourlet domain, then the watermark is embedded into the selected contourlet subbands of the center view by quantization on certain contourlet coefficients. Considering the vertical edges and contours are more easily affected in the DIBR process, we quantize the contourlet coefficients of the contourlet subbands with more horizontal information. The left and right views are synthesized by the watermarked center view and corresponding per-pixel depth information. The normalized correlation (NC) is used to extract the watermark in the target views. The watermark can be detected blindly with a low bit error rate (BER) from these views. The experimental results show that our scheme can resist various image processing attacks and DIBR processing. The main contribution of this work is the robustness to view synthesis process. In addition, the watermark can be detected blindly with low bit error rates from the center and the virtual views.

The remainder of this paper is organized as follows. Section 2 describes the basic processes of the DIBR system. Section 3 presents the proposed watermark embedding and extraction scheme for DIBR 3D images. Section 4 gives the simulation results and analysis about the robustness of the proposed approach, and the comparison of its performance with other watermarking techniques. Finally, Section 5 concludes this paper.

2. Depth-image-based rendering system

In a DIBR system, the virtual left view and right view can be rendered by pre-processing of depth map, 3D image warping and hole-filling processes [20]. Fig. 1 illustrates the major three steps of the DIBR process [6].

2.1. Pre-processing of depth map

As the first step, pre-processing of the depth map by smoothing filter can reduce the hole occurrences in the virtual left and right views, especially in the depth regions with sharp edges and discontinuities. In this process, the gray values of an 8-bit depth image lie between Z_{near} and Z_{far}, where Z_{near} is the nearest clipping plane with gray value of 255 and Z_{far} is the farthest clipping plane with gray value of 0 [20]. All the clipping planes are parallel to the image plane. The zero-parallax setting (ZPS) is chosen to be halfway between the Z_{near} and Z_{far}. Then the depth map is further normalized so that the values lie between -0.5 and 0.5. After that, different kinds of filters can be adopted to smooth the depth image. For efficient implementation, Gaussian filter is employed in the proposed scheme because of its adjustable window size [20].

2.2. 3D image warping

Three-dimensional image warping is very important in generating the virtual left and right views. In this work, we only consider the commonly used parallel camera configuration for the DIBR system. As shown in Fig. 2 [20], one point P with the depth Z is mapped on the image plane of three cameras (C_l, C_c, and C_r) at pixels x_l, x_c and x_r, respectively. In the image warping step, the virtual left view and right view can be synthesized from the center view and its depth map by the value of baseline distance and focal length. In this case, the vertical coordinate of a pixel in the three views remains the same. From the geometry in Fig. 2, the pixel-wise mapping can be performed by the following formulas [20,21]:

$$x_l = x_c + \frac{t_x}{2}\cdot\frac{f}{Z},\ x_r = x_c - \frac{t_x}{2}\cdot\frac{f}{Z} \tag{1}$$

where x_c, x_l and x_r are the viewpoints of the center view, the virtual left and right views, respectively. f represents the focal length of the camera, t_x denotes the baseline distance and Z is the value of depth map associated with the center view. Based on the visibility property of 3D view, the pixels with the farthest depth value are warped first since the closer object can occlude the farther object [1]. The baseline distance t_x can be adjusted by the audience to create a comfortable viewing.

2.3. Hole-filling

Due to the sharp changes in the depth map and different viewpoints, there are new exposed hole areas revealed in the rendered left and right views after 3D warping process. Some pixels' information is lost during pixel-wise mapping and some areas occluded in the original image may become visible in the virtual views. The hole occurrences can be reduced by pre-processing of the associated depth map. There are many studies [20,22] on the hole-filling problem. One straightforward method to fill the holes is the interpolation of neighborhood pixel information. Better quality of the synthesized left and right views can be reconstructed with more complicated extrapolation techniques. Considering the simplicity and practicality, we use linear interpolation to fill the disocclusions.

3. Proposed watermarking scheme

As illustrated in Fig. 3, the proposed scheme applies the contourlet transform to the center image. We embed the watermark into the selected contourlet subbands of the center image by quantization [23–26] on certain contourlet coefficients. For watermark extraction, the embedded watermark is detected by the statistical difference between quantized and unquantized contourlet coefficients. In the following, we briefly introduce the contourlet transform, then we describe our proposed watermark embedding and extraction algorithms.

3.1. Contourlet transform

Contourlet transform [17] uses a double filter bank structure to obtain images with smooth contours. There are two major stages in this double filter bank: Laplacian pyramid (LP) and directional filter bank (DFB). The LP is first applied to capture the point discontinuities, and the DFB is followed to form the linear structures. The LP at each level