J. Vis. Commun. Image R. 43 (2017) 41–49

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Virtual view quality assessment based on shift compensation and visual masking effect ☆

Fen Chen *, Renzhi Jiao, Zongju Peng, Gangyi Jiang, Mei Yu

Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211, China

ARTICLE INFO

Article history:
Received 24 August 2016
Revised 17 November 2016
Accepted 17 December 2016
Available online 21 December 2016

Keywords:
Depth image based rendering
Multiview video plus depth
Visual masking effect
Virtual viewpoint quality assessment

ABSTRACT

Depth image based rendering technology is essential for free viewpoint video systems. Because of the compression of depth images and limitations of rendering algorithms, various types of distortion might occur in the virtual viewpoints and cannot effectively be evaluated by traditional two-dimensional assessment methods. Hence, this paper proposes a method for virtual viewpoint quality assessment using the visual masking effect. First, shift is compensated for the distorted virtual viewpoint and then the compensated virtual viewpoint is objectively assessed. Next, according to human visual characteristics such as texture, magnitude, and distribution masking, the corresponding visual sensitivity map and visual masks are extracted. Finally, the visual masking and all factors are pooled to create the final quality score. As verified by the experimental results, the method proposed in this paper corresponds with the characteristics of human vision and can serve as a more effective method for assessing the quality of virtual viewpoints.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Free viewpoint video (FVV) enables users to experience three-dimensional (3D) scenes from different angles [1]. Given this characteristic, FVV systems have broad potential prospects in many application fields such as medicine, games, cinema, and historic relic preservation. Multiview video plus depth (MVD) is the main representation method of a 3D scene, and is used to generate continuous virtual viewpoints. Researchers aim to obtain high quality virtual viewpoints using algorithms such as rendering optimization and depth enhancement. However, the quality of virtual views produced by these algorithms is estimated by a 2D metric that is not fully consistent with the quality of a user's perception. A perceptual-based quality assessment can be used as feedback to optimize virtual view rendering and eventually improve the perceptual quality of the virtual views.

In FVV systems, the distortion of a virtual view includes phenomena such as ghost artifacts, object geometric distortion, and temporal flickering. The distortions in virtual viewpoints are different from those in common 2D images or video and cannot be effectively evaluated by traditional 2D assessment metrics. For instance, the peak signal-to-noise ratio (PSNR) of a virtual view is inconsistent with the perceived quality of synthesized virtual views [2]. Hence, many researchers have proposed subjective and objective quality assessment algorithms for virtual views. Liu et al. built a synthesized video quality database and further proposed a full-reference quality assessment method for spatio-temporal activity distortion and flicker distortion, but it did not take into account the factors of frame rate, boundaries, and depth, which affect the perceptual quality of virtual views [3]. Tsai et al. proposed a quality assessment model for synthesized distorted images that takes the consistent object shift and ghost artifacts into consideration [4]. Chen et al. proposed a video enhancement method based on piecewise tone mapping, using discrete entropy, temporal absolute mean brightness error and histogram-intersection-based temporal error to assess the quality for enhanced video [5]. Bosc et al. detailed the inconsistency in quality between the objective metrics and subjective assessment, and proposed two approaches, an analysis of contour shifts and a mean structural similarity index (SSIM) score of unoccluded areas [6]. Chen et al. proposed a wave leader pyramids based visual information fidelity method for image quality assessment [7]. Joveluro et al. proposed the perceptual quality metric (PQM) for evaluating the quality of 3D video [8]. The method is sensitive to slight changes in image degradation and error quantification. PQM performs well when evaluating virtual views synthesized by reconstructed depth and color images. Conze et al. proposed a full-reference objective quality metric dedicated to artifact detection

☆ This paper has been recommended for acceptance by M.T. Sun.
* Corresponding author.
E-mail address: chenfen@nbu.edu.cn (F. Chen).

http://dx.doi.org/10.1016/j.jvcir.2016.12.004
1047-3203/© 2016 Elsevier Inc. All rights reserved.

in synthesized virtual views [9]. This method aims to handle the areas where disparity estimation may fail: thin objects, variations in illumination or color between the left and right views and other artifacts. Solh et al. evaluated the virtual viewpoint distortion caused by a given distorted depth map by deriving an ideal depth map [10]. Temporal outliers, temporal inconsistencies, and spatial outliers are combined to constitute a vision-based quality measure for synthesized video. Battisti et al. proposed a full reference objective quality assessment metric based on a comparison of the statistical features of wavelet subbands of the original and virtual view images [11]. The above metrics assess the quality virtual views from different aspects, and produce different results.

In order to effectively assess the quality of a virtual viewpoint, this paper presents a new virtual view quality assessment method based on the visual masking effect. In our work, we first eliminate the pixel shift distortion using a shift compensation algorithm. Then, using the results of texture masking, magnitude masking, and distribution masking, a corresponding visual mask is extracted to process the post-compensation image. Further, taking into account multiple factors such as edge, depth, and baseline distance, the preliminary evaluation results of the virtual viewpoints are weighted to derive the final evaluation results. Experimental results validate the effectiveness of the proposed virtual view quality assessment algorithm.

The main innovations in this study are as follows: (1) the introduction of a shift compensation algorithm to eliminate the effects of different rendering algorithms on virtual viewpoint quality; (2) an analysis of human visual characteristics and extraction of visually sensitive images and distortion masks; and (3) a combination of all the factors influencing virtual viewpoint quality and integration into a single quality metric.

The rest of this paper is organized as follows. In Section 2, virtual view synthesis using the depth image-based rendering (DIBR) mechanism is presented and the distortion of virtual views is further analyzed. In Section 3, the detailed implementation of the proposed virtual view quality metric is presented. Then, in Section 4, the performance of the proposed metric is demonstrated. Finally, we draw the conclusions in Section 5.

2. Virtual viewpoint rendering process and distortion analysis

Color video and the associated depth video are used to synthesize the virtual viewpoint. The depth distortion and rendering algorithms affect the quality of virtual viewpoint. In this section, we describe the virtual viewpoint rendering process and analyze the virtual viewpoint distortion.

2.1. Virtual viewpoint rendering process

In an MVD-based FVV system, a virtual viewpoint is generated using color video and the corresponding depth video obtained from the decoder. The core idea of DIBR is to project the reference image pixels onto a target virtual viewpoint based on the depth information and camera parameters, and consists of the following main steps: First, all the pixels of the 2D reference image are re-projected onto a corresponding 3D space based on the depth information. Next, these 3D space points are projected onto a target 2D image plane. Finally, post-processing is applied to the rendered virtual viewpoint, which mainly consists of hole filling and ghost artifact elimination.

The projection process from 2D to 3D and then from 3D to 2D is called 3D warping. For a 3D video sequence captured by a parallel camera setup that has been accurately calibrated, DIBR can be represented as a disparity compensation process between viewpoints. So far, many parallel camera systems use a horizontal setup. Hence, the vertical disparity is zero. Let d be the horizontal disparity, which can be calculated using the following formula:

$$d = \frac{f \times l}{z} \tag{1}$$

where f is the camera focal length, l is the baseline distance between the reference viewpoint and virtual viewpoint, and z is the quantized depth value. Given the pixel value position (x_1, y_1) in the reference viewpoint, the position mapped to the virtual viewpoint is (x_2, y_2). The mapping process is calculated as follows:

$$x_2 = x_1 + d, \quad y_2 = y_1 \tag{2}$$

$$x_2 = x_1 - d, \quad y_2 = y_1 \tag{3}$$

Eqs. (2) and (3) respectively represent the left and right virtual viewpoints. If multiple pixels in the reference viewpoint are mapped to the same position in the virtual viewpoint, then Z-buffer technology is used for processing.

2.2. Virtual viewpoint distortion analysis

There are different types of distortion in the DIBR rendering system [12]. During virtual viewpoint rendering, depth must first be converted to disparity, which is used to determine the pixel location of the reference viewpoint in a virtual viewpoint. The depth value determines the pixel offset distance in the reference viewpoint. Depth video distortion will result in rendering displacement, and consequently deteriorate the virtual view quality. The sharp change in depth values and the global disparity result in holes at the boundary of the foreground and the border of the virtual view image. If two adjacent depth values vary drastically, a hole will appear between the two pixels, and sharper changes in depth values generate larger holes. When the depth boundary of the foreground is inaccurate or different interpolation algorithms are used to fill the holes in the foreground boundary or in the border of the virtual view image, different distortion will occur. Thus, the reasons for distortion generation in virtual viewpoint rendering can be classified into two major categories, depth distortion and the virtual viewpoint rendering algorithms.

2.2.1. Distortions caused by an inaccurate depth map

So far, the limit of depth map acquisition algorithms and encoding techniques can lead to depth distortion. The resulting virtual view will hence be degraded.

(1) Ghost artifacts and erosion: These generally occur at the boundary between the foreground and background. Inaccurate depth video can cause an inconsistency between the depth and color edges, leading to a mapping error of a foreground pixel to the background area. Consequently, the pixels at the foreground object boundary can be missing, resulting in ghost artifacts in the background region. Fig. 1(a) shows an original image and Fig. 1(b) depicts this kind of artifact.

(2) Image object offset and geometric distortion: The inaccuracy of the depth map estimation and depth compression distortion lead to an object mapping error and a rendering shift in the virtual viewpoint. As a result, holes appear in the region with sharp depth variations, and the subsequent hole-filling causes the objects in the image to be geometrically distorted. Figs. 1(c) and 1(d) respectively show the original image and distorted image with geometric distortion in the chair.

(3) Flicker distortion: When the depth value of the same image object varies greatly in two adjacent frames, an offset error of the object's position in the virtual viewpoint occurs, which often

Download English Version:

https://daneshyari.com/en/article/4969452

Download Persian Version:

https://daneshyari.com/article/4969452

Daneshyari.com