



Disparity-based just-noticeable-difference model for perceptual stereoscopic video coding using depth of focus blur effect ^{☆,☆☆}



Fei Xue ^a, Cheolkon Jung ^{a,*}, Joongkyu Kim ^b

^a School of Electronic Engineering, Xidian University, Xi'an 710071, China

^b College of Information and Communication Engineering, Sungkyunkwan University, Suwon 440-746, Republic of Korea

ARTICLE INFO

Article history:

Received 20 August 2015

Received in revised form 1 March 2016

Accepted 3 March 2016

Available online 3 March 2016

Keywords:

Binocular fusion limit

Depth-of-focus blur

Just-noticeable-difference model

Perceptual stereoscopic video coding

Human visual system

Visual comfort

ABSTRACT

Human 3D perception provides an important clue to the removal of redundancy in stereoscopic 3D (S3D) videos. Because objects outside the binocular fusion limit cannot be fused on retina, the human visual system (HVS) makes them blur according to the depth-of-focus (DOF) effect to increase the binocular fusion limit and suppress diplopia, i.e. double vision. Based on human depth perception, we propose a disparity-based just-noticeable-difference model (DJND) to save bit-rate and improve visual comfort in S3D videos. We combine the DOF blur effect with conventional JND models in the pixel domain into DJND. Firstly, we use disparity information to get the average disparity value of each block. Then, we integrate the DOF blur effect into luminance JND (LJND) by a selective low pass Gaussian filter to minimize the visual stimulus in S3D videos. Finally, we incorporate disparity information into the filtered JND models to obtain DJND. Experimental results demonstrate that the proposed method successfully improves both image quality and visual comfort in viewing S3D videos without increasing the bit-rate.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

With the rapid advances in the stereoscopic 3D (S3D) media technology, the demands on realistic visual experiences are also growing exponentially. S3D videos which are captured by stereo cameras provide users more vivid experiences on S3D displays. Unfortunately, compared with the real world, S3D displays often yield some distortions in 3D effects [1] which cause binocular stimulus, discomfort and fatigue for viewers. They are also usually caused by unsuitable screen disparity, large disparity change, and unnatural blur/sharpness [1–3]. Thus, for stereoscopic video coding, it is required to effectively enhance visual comfort and quality while maintaining storage space and transmission bandwidth. To achieve this, perceptual stereoscopic video coding based on human depth perception is a promising solution.

Disparity information which provides users with depth effects and vivid experiences plays an important role in many applications such as stereo vision, 3D reconstruction, and depth-image-based-

rendering (DIBR). Recently, many researchers have paid much attention to reduce visual discomfort with accurate disparity, stereoscopic distortion correction and synthetic depth of focus (DOF) blur [4,5]. Jung et al. [4] presented a selective DOF blur technique which was applied to the local regions where induced high visual discomfort but were less visually important. This technique could effectively reduce the visual discomfort based on depth maps and saliency maps. Um et al. [6] proposed a novel disparity-based asymmetrical filtering technique which used the disparity information in a stereo image to control severe blurring artifacts. It maintained image quality, but didn't consider visual comfort improvement. Leroy et al. [7] presented a visual comfort enhancement method to remove irritating high frequencies in high disparity zones. It conducted adaptive blur according to the elements such as disparity, content, and focus. Shao et al. [8] proposed a visual discomfort relaxation method to automatically adjust disparity ranges of the crossed and uncrossed disparities according to the zero disparity plane (ZDP). Jiang et al. [9] proposed a three-dimensional visual comfort assessment method using preference learning. They estimated the visual comfort score by a simple linear mapping strategy based on robust VCA and preference classification models, which achieved good performance in visual comfort enhancement. However, most methods improved the visual comfort in S3D videos without considering the bit-rates.

^{*} This paper was recommended for publication by Pen-Cheng Wang.

^{**} This work was supported by the National Natural Science Foundation of China (No. 61271298) and the International S&T Cooperation Program of China (No. 2014DFG12780).

* Corresponding author.

E-mail addresses: 2012xuefei@iiep.xidian.edu.cn (F. Xue), zhengzk@xidian.edu.cn (C. Jung), jkkim@skku.edu (J. Kim).

In this work, we adopt just noticeable difference (JND) models in stereoscopic video coding to reduce perceptual visual discomfort in S3D videos while maintaining the coding bit-rate. This is because the JND model can simulate the perceptual redundancy to determine the maximum distortion less than HVS. In [10], Chou and Li proposed a spatial JND model, i.e. minimally noticeable distortion (MND) profile, which considers the masking effects of background luminance and texture to quantify the perceptual redundancy. In [11,12], the foveation-based HVS JND model obtained by foveation masking effects was utilized to the video coding structure. It was useful to simulate the visual masking effect. Shao et al. [13] proposed an asymmetric stereoscopic video coding method based on the binocular vision suppression theory which the JND threshold exists between left and right views and chrominance degradation in right view would not affect the stereoscopic 3D effects. It saved bit rate significantly. Shao et al. [14] presented a depth map coding method which formulated the maximum tolerable depth distortion (MTDD) and the depth disocclusion mask to save bit rate for view synthesis. They provided a warped-SKIP mode to reduce inter-view redundancy. Thus, they achieved high view synthesis performance while improving the coding efficiency. The JND models were suitable for single-view videos, but not effective in removing the stereoscopic perceptual redundancy in stereo videos. It has been reported that HVS is more sensitive to the nearby regions and the background regions are often blurred on retina as shown in Fig. 1 [1]. In the real world (left figure in Fig. 1), focal and vergence distances are the same no matter where the viewer looks. The focused regions have higher resolution on retina while the others are blurred by HVS, called DOF blur. However, in S3D displays (the right figure in Fig. 1), focal distance is fixed at the display while vergence distance varies depending on the part of the simulated scene the viewer fixates [1]. The whole scene has higher resolution, but doesn't conform to the human depth perception. It has low visual comfort, which causes visual fatigue in S3D displays. Thus, if we apply traditional JND models in single-view videos to stereo videos, it is hard to remove the stereoscopic perceptual redundancy as well as likely to overestimate the visibility threshold of the frontal regions and underestimate that of the background regions. As shown in Fig. 2(a), viewers generally pay more attention to the frontal flowers than the background window. That is, the JND thresholds of foreground regions

should be lower than those of the background ones. However, the traditional JND map in Fig. 2(b) produces exactly the contrary effect. To address this problem, we newly propose a disparity-based JND model (DJND) to consider disparity information which can distinguish foreground regions from the background. We combine the DOF blur effect with JND model in the pixel domain based on disparity to significantly improve the image quality and visual comfort in S3D displays. The proposed method is different from [14] because the proposed method assigns fewer bits to the non-focused regions, but assigns more bits to the focused ones. Thus, the proposed method successfully improves both image quality and visual comfort in viewing S3D videos without increasing the bit-rate.

Fig. 3 illustrates the entire framework of the proposed method. It can adjust the visibility threshold of different regions into a suitable scope. First, we get the traditional JND models and use a Gaussian low pass filter on the luminance JND model to reduce the background visual stimulus and discomfort. Then, we integrate the average disparity value of each 5×5 block to update JND models and get the proposed DJND model. Finally, we use the JND thresholds to adjust quantization parameter (QP) in rate-distortion optimization of the stereo video coding structure. It can utilize the bit-rate of the perceptually unimportant regions in coding of important regions and minimize the visual stimulus of the background regions on retina. Experimental results demonstrate that the proposed method effectively improves both image quality and visual comfort in viewing S3D videos while maintaining its bit-rate.

The rest of this paper is organized as follow. In Section 2, we explain the disparity-based luminance JND model. In Section 3, we describe the disparity-based texture JND model and the proposed DJND model in detail. In Section 4, we address the use of JND thresholds on JMVC. In Section 5, we provide experimental results and draw conclusions in Section 6.

2. Disparity-based luminance JND model

The luminance JND model has been estimated based on the background luminance masking effect which is a primary factor of the human visual system [15]. The masking effect means that the human attracts higher attention in medium gray background

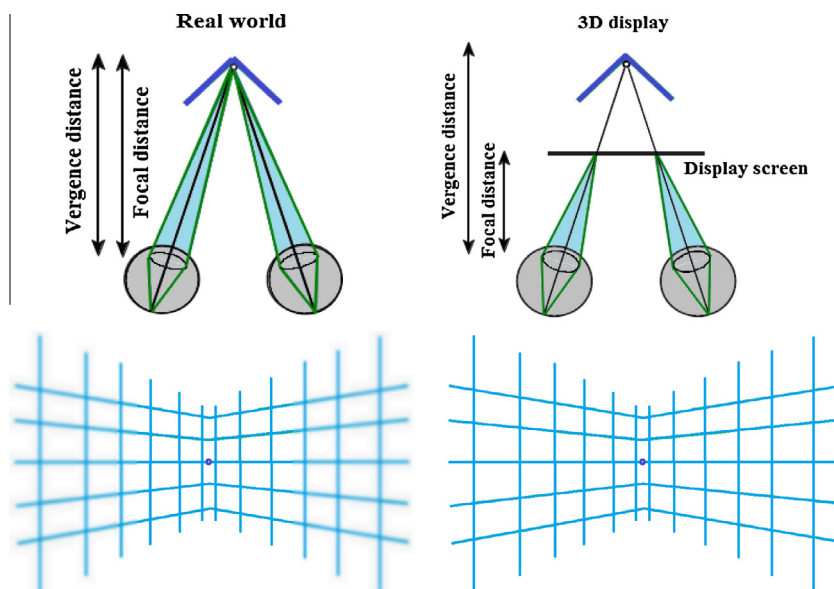


Fig. 1. Vergence and focal distances in the real world and S3D displays, redrawn from [1]. Left: real world scenes. Right: 3D displays.

Download English Version:

<https://daneshyari.com/en/article/537826>

Download Persian Version:

<https://daneshyari.com/article/537826>

[Daneshyari.com](https://daneshyari.com)