Future Generation Computer Systems 83 (2018) 85–94

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

3D visual discomfort predictor based on subjective perceived-constraint sparse representation in 3D display system

Haiyong Xu [a,b], Gangyi Jiang [b,*], Mei Yu [b], Ting Luo [a,b], Zongju Peng [b], Feng Shao [b], Hao Jiang [b]

[a] College of Science and Technology, Ningbo University, Ningbo, 315211, China
[b] Faculty of Information Science and Engineering, Ningbo University, Ningbo, 315211, China

HIGHLIGHTS

- In this paper, based on the mechanism of neural activity in V1, the feature space of visual discomfort is established by considering the significant map and spatial frequency of V1.
- Then, a subjective perceived-constraint sparse representation (SPCSR) is constructed by considering sparse coding of simple and complex cells in the receptive field and human learning mechanism.
- Finally, a 3D visual discomfort predictor (3D-VDP) with SPCSR is proposed.

ARTICLE INFO

Article history:
Received 8 March 2017
Received in revised form 24 December 2017
Accepted 9 January 2018

Keywords:
3D visual discomfort predictor
Subjective perceived-constraint sparse representation
Neural activity mechanism
Receptive field

ABSTRACT

Three-dimensional (3D) display systems have been widely adopted due to the recent increased availability of an increasing 3D contents. However, viewers may experience visual discomfort due to the limited viewing zone available of 3D display systems. Therefore, 3D visual discomfort prediction is important for optimizing 3D display systems. In this paper, we propose a 3D visual discomfort predictor (3D-VDP) that is based on the visual discomfort features of the primary visual cortex (V1) and the properties of subjective perceived-constraint sparse representation (SPCSR). Embedding subjective values of visual discomfort as a constraint into sparse representation such that the dictionary is more suitable for visual perception is the major technical contribution of this study. Specifically, the proposed 3D-VDP with SPCSR consists of two phases. In the training phase, first, the neural activity mechanism of V1 is considered, and the features of visually important disparity and spatial frequency disparity are extracted to highlight the influence of disparity on the comfort of stereoscopic images. Second, by considering the visual properties of the receptive field and learning mechanism, a perceived dictionary of visual discomfort and the corresponding perceived value of visual discomfort are obtained based on the subjective value of visual discomfort as a constraint condition applied to the construction of a supervised dictionary learning algorithm. In the testing phase, the sparse coefficient of visual discomfort of the stereoscopic image is computed according to the perceived dictionary of visual discomfort by using the sparse coding algorithm, and the final visual discomfort score of the stereoscopic image is obtained from the weighted sparse coefficients and the perceived value of visual discomfort. Experimental results obtained with the IVY LAB database and the NBU database demonstrate that, in comparison with closely related existing models, the proposed 3D-VDP with SPCSR achieves a high consistency of subjective assessment.

1. Introduction

Three-dimensional (3D) display systems are widely used nowadays, and 3D multimedia, which can provide new visual experiences such as stereoscopic viewing and viewpoint interaction, has been attracting increased amounts of attention and is regarded as being the next generation of media [1]. However, humans may experience a degree of discomfort including eye fatigue, headaches, and nausea due to the influence of this media on the human visual system (HVS), the stereoscopic image/video content, the viewing conditions, and other factors, all of which greatly reduce the quality of the experience [2]. Thus, it is of

* Corresponding author.
E-mail address: gyjiang@nbu.edu.cn (G. Jiang).

https://doi.org/10.1016/j.future.2018.01.021

great importance to be able to accurately predict the degree of discomfort that is likely to be produced by 3D image/video [3].

The human visual system relies on a large number of cues for estimating depth [4]. In general, visual depth cues can be classified into monocular and binocular cues. Our visual system utilizes monocular cues from two-dimensional natural scenes and constructs a 3D representation of the perceived scene. Apart from monocular cues, more specific depth information can be obtained from binocular cues, especially at shorter viewing distances. These cues are implemented in stereoscopic 3D display systems. In 3D display systems, stereopsis is the perception of depth that is constructed based on the difference between these two retinal images; the brain fuses the left and right image and from retinal disparity. Based on binocular cues, viewers may achieve a more natural depth perception, but sometimes at the expense of visual comfort [5]. Recent research efforts have reported that a vergence-accommodation conflict is produced when a 3D scene includes a large disparity, which may cause visual discomfort and fatigue [6]. Meanwhile, the spatial frequency also influences the binocular fusion limit and then goes on to affect the visual comfort. Schor et al. [7] examined the effect of spatial frequency using visual stimuli with Gaussian difference. Their results showed that the binocular fusion limit increased as the spatial frequency decreased, so the degree of visual discomfort also increased.

Some existing 3D visual discomfort assessment (3D-VDA) models are mainly based on the statistical features of global disparity, which use the mean, variance, and distribution of the disparity map to predict the visual discomfort produced by the stereoscopic image [8,9]. However, these 3D-VDA models based on global disparity are not consistent with subjective perceptions since humans are more sensitive to the visual discomfort of some significant regions when watching 3D content. Based on this, Sohn et al. [10] designed an attention-based 3D-VDA model, considering the perceptually significant regions of stereoscopic images and applying regression analysis. Furthermore, they [11] proposed object-dependent disparity features including the disparity gradient of nearby objects and the object width, and used an M5P regression tree to predict the visual discomfort. Jung et al. [12] considered the perceptually significant regions of stereoscopic images and proposed an attention-based 3D-VDA model with support vector regression (SVR). Without exception, these models used regression analysis to establish nonlinear mapping from the high-dimensional feature space to the low-dimensional subjective score space. In this case, the nonlinear mapping established by the regression models is not explicit, and so cannot reflect the visual perception of humans. Moreover, those models that are trained on a large-scale training set may also incur the risk of over-fitting, implying a rather weak generalization capability.

From a physiological point of view, the goal of the 3D-VDA model is to simulate the properties of binocular visual perception. As the most important part of HVS, the visual cortex is responsible for most visual tasks. In the visual cortex, signals are perceived through two visual pathways, namely, the dorsal stream and ventral stream. The function of stereoscopic vision is to complete the cooperation of many neurons in each region of the visual cortex between two visual pathways. However, it is not known whether the mechanisms of the neurons, which are sensitive to disparity, are present throughout the region of the visual cortex [13]. Recently, researchers have tried to simplify and simulate the neural activity responsible for stereoscopic vision to establish a 3D-VDA model. Park et al. [14] extracted the statistical features of disparity and used SVR to predict visual discomfort through an analysis of the neural coding of accommodation-vergence mismatches. Furthermore, Park et al. [15] proposed a model-based neural and statistical framework that can automatically predict the level of visual discomfort. However, relatively few studies have addressed the mechanism of the primary visual cortex (V1) and sparse coding of the stereoscopic vision.

An efficient 3D-VDA model should be capable of accurately predicting the degree of visual discomfort for stereoscopic images by simulating the related neural mechanism of human visual perception. Based on this consideration, a novel 3D-VDA model is proposed by extracting the visual disparity features of V1 and using sparse coding with a subjective perception constraint. Unlike the above mentioned models, the proposed model includes two modules: (1) the subjective perception-constraint dictionary learning module in the training phase and (2) the sparse coding module in the visual discomfort estimating phase. Specifically, in the present study, based on the mechanism of neural activity in V1, the feature space of visual discomfort is established by considering the significant map and spatial frequency of V1. Then, a subjective perceived-constraint sparse representation (SPCSR) is constructed by considering the sparse coding of simple and complex cells in the receptive field and human learning mechanism. Finally, a 3D visual discomfort predictor (3D-VDP) with SPCSR by using sparse coefficients is proposed. In summary, the main contributions of the present study are as follows:

(1) The subjective perception-constraint dictionary learning module in the training phase is proposed by adding a subjective perception constraint term to a traditional dictionary learning formulation so that the learned dictionary conforms to the human perception of visual discomfort.

(2) Once the dictionary of visual discomfort is learned, the sparse coding module in the visual discomfort estimating phase is constructed to obtain sparse coefficients. These sparse coefficients as suitable weighting coefficients for a pooling strategy are used to construct a visual discomfort predictor.

(3) Given that simple disparity features are insufficient for reflecting the degree of visual discomfort, more comprehensive features based on stereoscopic visual neural mechanism of V1 for visual discomfort are extracted.

The remainder of this paper is organized as follows. Section 2 introduces a stereoscopic visual neural mechanism of V1 for visual discomfort. Section 3 describes the proposed 3D-VDP with SPCSR. Section 4 describes the experiment and analyses. The last part is the summary.

2. Stereoscopic visual neural mechanism of V1 for visual discomfort

Stereo vision is an important part of HVS and binocular cues are implemented in 3D display systems. Stereopsis is the perception of depth that is constructed based on the difference between these two retinal images; the brain fuses the left and right image and from retinal disparity. Thus, the disparity of stereoscopic images is an important factor in determining depth perception by humans. Under normal circumstances, the accommodation and vergence processes interact with each other in the brain to deliver a 'cyclopean' image. However, a larger disparity may result in the perceived depth exceeding the Panum fusion region of the brain. This leads to conflict between the binocular accommodation and vergence, and directly causes visual discomfort [16]. From the viewpoint of neurophysiology, disparity-sensitive neurons are spread throughout the various regions of the visual cortex. Therefore, to bring visual discomfort predictions more in line with human visual perception, it is essential to simulate the neural activity resulting from any disparity. Because of the complexity of disparity processing and of the neural activity related to disparity, many mechanisms are unclear, which indirectly leads to difficulty in predicting visual discomfort. To simplify the problem, we mainly studied V1, which is closely correlated to binocular vision. The mechanism of neural activity in the other regions is not discussed in this paper.