Information Sciences 373 (2016) 251–268

Contents lists available at ScienceDirect

Information Sciences

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

Quality assessment metric of stereo images considering cyclopean integration and visual saliency

Jiachen Yang[a,*], Yafang Wang[a], Baihua Li[a], Wen Lu[b], Qinggang Meng[c], Zhihan Lv[d], Dezong Zhao[e], Zhiqun Gao[a]

[a] School of Electronic Information Engineering, Tianjin university, Tianjin 30072,, PR China
[b] School of Electronic Engineering, Xidian University, Xi'an, Shaanxi Province 710071, China
[c] Department of Computer Science, School of Science at Loughborough University, UK
[d] University College London, Dept. of Computer Science, 66-72 Gower Street, London, WC1E 6EA, United Kingdom
[e] Department of Aeronautical and Automotive Engineering, Loughborough University, Loughborough, LE11 3TU, United Kingdom

ARTICLE INFO

Article history:
Received 11 February 2016
Revised 28 August 2016
Accepted 1 September 2016
Available online 2 September 2016

Keywords:
Binocular combination
Human visual system
Saliency
Stereoscopic image quality assessment
Visual attention

ABSTRACT

In recent years, there has been great progress in the wider use of three-dimensional (3D) technologies. With increasing sources of 3D content, a useful tool is needed to evaluate the perceived quality of the 3D videos/images. This paper puts forward a framework to evaluate the quality of stereoscopic images contaminated by possible symmetric or asymmetric distortions. Human visual system (HVS) studies reveal that binocular combination models and visual saliency are the two key factors for the stereoscopic image quality assessment (SIQA) metric. Therefore inspired by such findings in HVS, this paper proposes a novel saliency map in SIQA metric for the cyclopean image called "cyclopean saliency", which avoids complex calculations and produces good results in detecting saliency regions. Moreover, experimental results show that our metric significantly outperforms conventional 2D quality metrics and yields higher correlations with human subjective judgment than the state-of-art SIQA metrics. 3D saliency performance is also compared with "cyclopean saliency" in SIQA. It is noticed that the proposed metric is applicable to both symmetric and asymmetric distortions. It can thus be concluded that the proposed SIQA metric can provide an effective evaluation tool to assess stereoscopic image quality.

1. Introduction

Visual information plays an important role in information acquisition in our daily lives. Over the past few decades, considerable stereoscopic technologies have been invented and commercialized [32,36,38]. However, new issues and challenges have also emerged with their development. During stereoscopic content creation, transmission, processing and display, various distortions causing visual fatigue and visual discomfort may be introduced. Consequently, it is necessary to develop an effective tool to measure the quality of stereoscopic images.

SIQA, like the quality assessment of monocular images (IQA), can be categorized into subjective and objective methods. Subjective assessment of stereo images represents the direct reflection of the HVS, and thus it is regarded as the most

* Corresponding author. Fax: +8613821820218.
E-mail addresses: yangjiachen@tju.edu.cn, yangjiac@163.com (J. Yang), WangYF0739@tju.edu.cn (Y. Wang), b.li@lboro.ac.uk (B. Li), luwen@mail.xidian.edu.cn (W. Lu), q.meng@lboro.ac.uk (Q. Meng), Z.Lu@cs.ucl.ac.uk (Z. Lv), D.Zhao2@lboro.ac.uk (D. Zhao), zqgao@tju.edu.cn (Z. Gao).

http://dx.doi.org/10.1016/j.ins.2016.09.004
0020-0255/

reasonable assessment method. In recent years, research using subjective experiments has achieved steady development and various factors that may affect stereoscopic perception have been investigated. However, subjective tests are time-consuming and require many duplicate experiments with a large number of participants. Therefore, objective metrics, which can be used to reliably predict the perceived quality of stereo images, has attracted significant attention from scholars and experts.

There are a number of 2D-IQA metrics with competitive performances in IQA. They can be divided into several categories: structural similarity based IQA metrics(e.g. structural similarity (SSIM) [50]; distortion distribution-based structural similarity (ADD-SSIM) and distortion distribution-based structural similarity (ADD-GSIM) [15]); deep learning based blind image quality assessment frameworks [20,23]; and natural scene statistics based metrics [11,16]. The multi-scale Geometric Analysis method [12], mimics the HVS and free-energy-based brain theory in IQA, investigating which classical HVS-inspired features could be used to evaluate the image quality [21]. Inspired by that the human visual system (HVS) exhibits obvious orientation selectivity mechanism for visual content extraction, Wu et al. proposed a reduced-reference IQA [52]. These metrics can only be used to evaluate the image quality before and after coding from one viewpoint. However, they fail to consider strong correlations with standard disparity from two adjacent viewpoints. Hence, a new objective quality assessment metric for stereoscopic images is needed.

However, it is not easy to design an accurate SIQA metric due to the disparity and depth. The earliest study attempting to evaluate the quality of stereo images simply applied 2D-IQA to their metrics. In our previous work [53], we found that the absolute disparity map approximately reflects the disparity and depth. Therefore we proposed an objective metric for stereoscopic images by combining this discovery. As stereoscopic images are different from plain images due to additional depth information, You et al. [56] investigated the possibility of applying some common 2D quality metrics in SIQA, while also taking depth information into consideration. A similar approach was also adopted by Benoit et al. [3], in which 2D-IQA metrics were augmented with disparity information.

Other metrics in the literature simulated the perceptual route of our visual system to process input signals. Shao et al. [46] divided stereo images into monocular regions, binocular fusion regions and binocular suppression regions. Furthermore, the overall quality was given as the linear summation of the three regions. By simulating monocular and binocular visual perception and analyzing the monocular-binocular feature fidelity induced index, another SIQA algorithm was proposed in his research [44]. Realizing that the neuronal responses are directly implicated in both the control and experience of 3D perception, Park J et al. developed a model-based neuronal and statistical framework that automatically predicts the level of visual discomfort [39].

Several metrics that took the theory of binocular combination into consideration were investigated. In the subsequent work of Shao, the binocular receptive field properties in line with human visual perception were introduced into quality assessment [45]. Maalouf et al. [34] presented the definition of “cyclopean image”. They integrated left and right images into a cyclopean image to simulate brain perception, after which they used contrast sensitivity coefficients of the produced cyclopean image to derive a quality. Moreover, Chen et al. [4] proposed a cyclopean full reference metric that is able to account for binocular rivalry. Lin et al. [33] adopted three quality components in their metric to evaluate the quality of stereoscopic video, namely the cyclopean view, binocular rivalry, and the scene geometry. Wang et al. [49] set up a new subjective SIQA database and proposed a binocular rivalry inspired model to predict the quality of stereoscopic images. In addition, Gu et al. [18] introduced a metric based on the analysis of autoregressive model parameters to characterize the sharpness as an index to evaluate image blur, after which the metric was extended to assessing the sharpness of stereoscopic images. These metrics have been shown to outperform the former metrics (e.g [3,53,56].) in predicting the quality of stereoscopic images, especially in the case of asymmetric distortions.

Since HVS is the ultimate assessor of image quality that takes into account a variety of visual characteristics, it has become another important factor affecting how to choose salient features. Li et al. developed an unsupervised feature selection algorithm to select the most useful features by integrating cluster and sparse structural analysis into a joint framework, while also experimentally demonstrating the effectiveness of the proposed algorithm [31]. In [30], a Robust Structured Subspace Learning (RSSL) algorithm was presented to uncover an appropriate latent subspace for data representation. Muhammad et al. [41] proposed a feature selection algorithm avoiding the positive region to replace the conventional dependency measure. Image understanding and feature learning were first integrated into a joint learning framework. This framework was then applied to several image understanding tasks and shown to achieve good performance.

For a large number of visual characteristics, the saliency map is particularly important. This is because HVS tends to select the most relevant information from a visual scene. Zhang et al. has demonstrated that the current soundness of visual saliency modelling is sufficient for IQA to yield a statistically meaningful gain in their performance [59]. Based on this evidence, we consider visual saliency to be a factor in IQA. Meanwhile, we attempted to find an efficient 3D metric to detect saliency regions that would affect the performance of SIQA. However, evaluation metrics that consider the significance of image are based on 3D saliency maps and extraction is both difficult and tedious. In order to overcome this shortcoming, a simple but efficient saliency detection metric is set to be in line with HVS.

In this paper, inspired by previous work on cyclopean images, a full reference quality assessment metric is proposed based on binocular combination and visual attention. In this proposed metric, we take advantage of several existing binocular combination models and propose a novel method that synthesizes two saliency maps to a cyclopean image, thereby predicting the saliency of binocular combination. The main innovations of our paper are as follows: (1) A proposed saliency map for stereoscopic images based on binocular combination models, defined as “cyclopean saliency”. The definition of “cyclopean saliency” then leads to the development of a new idea for computing the saliency of stereoscopic images. (2) We