Signal Processing 150 (2018) 204–214

Contents lists available at ScienceDirect

Signal Processing

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

Blind stereoscopic 3D image quality assessment via analysis of naturalness, structure, and binocular asymmetry

Guanghui Yue [a,c], Chunping Hou [a], Qiuping Jiang [b,c,*], Yang Yang [a]

[a] School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China
[b] School of Information Science and Engineering, Ningbo University, Ningbo 315211, China
[c] School of Computer Science and Engineering, Nanyang Technological University, 639798, Singapore

ARTICLE INFO

Article history:
Received 30 November 2017
Revised 4 April 2018
Accepted 19 April 2018
Available online 21 April 2018

Keywords:
Stereoscopic 3D image
Quality assessment
No reference
Binocular vision
Natural scene statistic (NSS)
Asymmetric distortion

ABSTRACT

Over recent years, stereoscopic three dimensional (S3D) images have grown explosively and received increasing attention. Quality assessment, as the fundamental problem, plays an important role in promoting the prevalence of S3D images as well as the associated products. In this paper, an effective blind quality assessment method of S3D images is proposed via analysis of naturalness, structure, and binocular asymmetry. To be specific, given that natural images obey certain regular statistical properties, natural scene statistic (NSS) features of left and right views are first extracted to quantify the naturalness. Second, by considering binocular visual characteristics, statistical features are extracted from a created cyclopean map. Moreover, gray level co-occurrence matrix (GLCM) is utilized to capture quality-sensitive features from the cyclopean phase map. Third, to quantify the asymmetric distortion, a simple but effective measurement is utilized, i.e., calculating the similarity between left and right views as well as statistical features of their difference map. Finally, all extracted quality-sensitive features are combined, and trained together with the subjective ratings to form a regression model using support vector regression (SVR). Experimental results on four publicly available databases (two symmetrically distorted databases and two asymmetrically distorted databases) demonstrate that the proposed method is superior to several mainstream image quality assessment (IQA) metrics.

1. Introduction

Owing to the rapid development of multimedia and networking technologies in the past decades, the volume of digital image data has been growing explosively and played increasingly important role in daily life. However, images inevitably suffer from various distortions during acquisition, transmission, processing, and storage stages, causing unpleasant visual quality of experience. Image quality assessment (IQA) metrics, developed to evaluate or monitor image quality, show great potential on controlling and improving the execution of image processing systems, such as compression, enhancement, and transmission. To date, quantities of mature IQA metrics for two-dimensional (2D) images have been reported [1–8]. Recently, stereoscopic three-dimensional (S3D) multimedia services become popular and greatly increase the immersive experience by enabling depth perception compared to traditional 2D visual experience [9]. Undoubtedly, quality assessment of S3D images becomes more complex as more perceptual issues should be considered: visual fatigue/discomfort induced by binocular depth information [10–12] and annoying experience caused by distortions [13,14]. As the interaction between the visual fatigue/discomfort and the distortion is complex and it has not been fully understood at the current stage. This paper mainly solves the quality assessment problem of S3D images from the viewpoint of image distortions as previous works did [13–16].

Broadly speaking, quality assessment metrics can be divided into two categories: subjective method and objective method. The former case relying on subjective evaluation is more effective, reliable and is usually treated as the ground truth of the latter case. However, it is time-consuming, expensive and unsuitable for online implementation [17]. By contrast, the latter case overcomes these shortcomings and has been widely investigated. Generally, objective methods can be further categorized into full reference (FR), reduced reference (RR), and no reference (NR, also called as blind reference) according to the participation of the reference information [1]. To date, many works have been reported with rather com-

* Corresponding author at: School of Computer Science and Engineering, Nanyang Technological University, 639798, Singapore.
E-mail address: n1609861g@e.ntu.edu.sg (Q. Jiang).

https://doi.org/10.1016/j.sigpro.2018.04.019
0165-1684/

petitive performance on evaluating 2D image quality. Among them, the FR metric, assuming that the reference and distorted images' information is completely known, is the most mature and prevalent ones [2,18–20]. To deal with the problem where only part of the information is sufficient for IQA, RR metrics was proposed and get rapid development [21]. However, due to their intrinsic dependence on reference information, both FR and RR metrics can only be used in limited situations and are powerless to deal with situations, where the reference information cannot be obtained. Therefore, how to design an effective reference-free metric (i.e., NR metric) has received incremental attentions from researchers and gradually become the mainstream direction [22–27].

Different from traditional 2D images, quality assessment related to S3D images is more challenging and various factors, such as 2D image quality, binocular asymmetry, depth information, etc., should also be considered together. To cope with this challenge, many works have been conducted objectively through FR and RR measures. One broadly adopted strategy is to apply existing 2D IQA metrics on both left and right views individually, and then integrate them together to infer the overall quality. Unfortunately, such simple operation usually produces moderate performance. Actually, the perceptual quality of S3D images mainly depends on the merged effects of binocular vision, such as binocular fusion, binocular suppression, and binocular rivalry. Therefore, effectively simulating the binocular visual characteristics may be particularly beneficial to design S3D IQA models. For example, Bensalma et al. proposed a FR S3D IQA metric by calculating the binocular energy difference between the reference and corrupted images [16]. Specifically, the binocular energy was calculated by simulating the properties of simple and complex cells in the visual cortex. Considering the binocular fusion effect, Chen et al. [14] first synthesized a cyclopean view from the stereopair and its associated disparity map. Then, the S3D IQA model was built by applying a traditional 2D FR IQA metric on the generated cyclopean view. According to physiological discoveries on binocular vision, Lin and Wu [15] incorporated binocular integration behaviors (i.e., binocular combination and binocular frequency integration) into existing 2D models to enhance the ability in evaluating S3D images. In [28], Wang et al. developed a 2D-to-3D pooling scheme and designed a binocular rivalry inspired multi-scale model to predict the quality of S3D images. Shao et al. [29] proposed a dictionary based FR metric by considering binocular receipted field properties. Besides, they also constructed a phase-tuned visual codebook and a phase-tuned quality lookup from the binocular energy responses [30]. In [30], the image quality was estimated by averaging the largest value of all patches' qualities, which were obtained via searching the constructed codebook and lookup. Assuming that visual primitives can effectively represent image information, Qi et al. [31] proposed a RR metric. To be specific, they first calculated the entropies of the left and right views (which were obtained based on the estimated coefficients of visual primitives via sparse coding) to represent monocular cue. Then, the mutual information of left and right views was calculated to represent binocular cue. Finally, the cues' differences between the reference and distorted images were fused by a regression function to build the final quality assessment model. Likewise, Wan et al. [32] utilized the same framework as [31] and reported a RR metric by dividing the visual primitives into three categories, i.e., DCprimary, sketch and texture.

Unlike those reference-based metrics mentioned above, NR IQA metrics for S3D images (without reference information in hand) are more difficult to design. Actually, NR IQA metrics have broader application scope compared to the reference-based ones. Although challenging, some progresses have also been achieved in this area [33]. For instance, Liu et al.[34] proposed a S3D IQA model by fully considering the impact of binocular fusion, rivalry, suppression, and a reverse saliency effect on the perception of distortion. By simulating binocular visual system, Zhou and Yu [35] extracted the local binary pattern (LBP) on binocular rivalry response and binocular energy response maps to infer image quality. Besides, they also simulated the simple and complex cells' response maps based on which the LBP descriptors were extracted as the quality-sensitive features [36]. As a stereopair has two views (i.e., left and right views), it inevitably suffers from asymmetric distortions, hereby posing a severer challenge [37]. However, the aforementioned methods ignored the measurement of distortion asymmetry. To cope with this challenge, a few works have been reported. Jiang et al. [33] addressed the quality assessment issue of S3D images using a three-column nonnegativity constrained sparse autoencoder framework, which considered both monocular view quality and the interaction of two monocular views. Shen et al. [38] designed a NR metric for asymmetrically distorted stereoscopic images by utilizing combined model. Specifically, they used the difference map between the Gabor responses of left and right views to reflect the asymmetric information. Shao et al. [39] first judged whether the degree of distortion asymmetry between the left and right views of a stereopair was subjectively noticeable, and then integrated this indicator as a guidance into a supervised dictionary framework for quality evaluation of stereopairs.

Overall, the success of the above reviewed works benefits from the simulation of binocular visual characteristics (e.g., binocular fusion, binocular rivalry and binocular suppression) and the extraction of quality-sensitive features (e.g., LBP, entropy, energy, etc.). However, naturalness, which has been proved to be closely related to image quality, is almost ignored by the aforementioned methods. Although some works applied naturalness as a quality measurement, they did not consider binocular visual characteristics [40,41]. Besides, only part of these works made specific efforts to handle the challenging binocular asymmetric distortions. Actually, evaluating asymmetric distortions is more meaningful in many practical S3D applications. In this paper, an effective blind S3D IQA method is proposed by comprehensively considering properties of binocular vision, image naturalness, and distortion asymmetry. The contributions of this work are three-fold: (1) Towards simulating binocular combination properties, two individual monocular views are synthesized to form a cyclopean image whose mean subtraction and contrast divisive normalization (MSCN) coefficients in spatial domain are fitted using general Gaussian distribution (GGD), then the associated shape and scale parameters are extracted to quantify image naturalness. Compared to competing metrics, the proposed method considers both binocular visual characteristics and image naturalness together. (2) Since phase information reflecting the critical structures in an image is more sensitive to distortion, the quality-sensitive features are captured on the phase component of the cyclopean view via gray level covariance matrix (GLCM). This is a pioneering attempt to solve S3D IQA problem by extracting GLCM on the phase map of cyclopean view. (3) Towards portraying the unique role of binocular asymmetry in affecting the quality of S3D images, the proposed method proposes to compute the similarity between gradient maps of left and right views as well as GGD parameters of their difference map to estimate the degree of distortion asymmetry. Experimental results on public 3D image databases demonstrate that the proposed method is superior to existing mainstream FR and NR IQA metrics.

The rest of the paper is organized as follows. Section 2 gives a detailed description about the proposed blind S3D IQA method. Section 3 presents the experimental results on four publicly available S3D databases and corresponding analyses. Finally, Section 4 concludes the work and gives future directions.