3D Visual Comfort Assessment via Sparse Coding

Qiuping Jiang and Feng Shao(✉)

Faculty of Information Science and Engineering, Ningbo University, Ningbo, China
jqp910707@126.com, shaofeng@nbu.edu.cn

Abstract. The issue of visual discomfort has long been restricting the development of advanced stereoscopic 3D video technology. Bolstered by the requirement of highly comfortable three-dimensional (3D) content service, predicting the degree of visual comfort automatically with high accuracy has become a topic of intense study. This paper presents a novel visual comfort assessment (VCA) metric based on sparse coding strategy. The proposed VCA metric comprises three stages: feature representation, dictionary construction, sparse coding, and pooling strategy, respectively. In the feature representation stage, visual saliency labeled disparity statistics and neural activities are computed to capture the overall degree of visual comfort for a certain stereoscopic image. A set of stereoscopic images with a wide range degree of visual comfort are selected to construct dictionary for sparse coding. Given an input stereoscopic image, by representing features in the constructed dictionary via sparse coding algorithm, the corresponding visual comfort score can be estimated by weighting mean opinion scores (MOSs) using the sparse coding coefficients. In addition, we conduct a new 3D image benchmark database for performance validation. Experimental results on this database demonstrate that the proposed metric outperforms some representative VCA metrics in the regard of consisting with human subjective judgment.

Keywords: Visual comfort assessment (VCA) · Sparse coding · Stereoscopic image · Disparity statistic · Neural activity

1 Introduction

Due to the additional depth perception, stereoscopic three-dimensional (3D) content service has become particularly popular during the past few years and the market for 3D products is surging continuously. Compared with the traditional two-dimensional (2D) media, 3D media development faces new challenges, especially in terms of providing good end-user 3D quality of experience (QoE). Consumers have a very high expectation for QoE of the 3D services they receive. As one of the most important aspect of 3D QoE, visual comfort refers to the subjective fatigue sensation under 3D viewing in terms of various physiological symptoms, such as eye strain, headache, nausea, diplopia and so on [1, 2]. Therefore, 3D visual comfort assessment (VCA) is of crucial significance for the further development of 3D video and related technologies. Especially, it is challenging to develop such kind of objective VCA algorithm that can be directly used to guide the production and post-processing of 3D content for the sake of visual safety.

Y.-J. Zhang (Ed.): ICIG 2015, Part I, LNCS 9217, pp. 10–20, 2015.
DOI: 10.1007/978-3-319-21978-3_2

In the past decades, by exploiting the underlying factors that affect the end-users' experienced visual comfort, many objective VCA metrics have been developed [3–9]. Technically, most of state-of-the-art objective VCA metrics involve two stages, namely feature representation and mapping model construction. Specifically, the stage of feature representation involves extracting proper features to characterize the degree of experienced visual comfort and the stage of mapping model construction involves pooling a final visual comfort score from the extracted features. In detail, for feature representation, multiple forms of holistic disparity statistics are computed from disparity maps to capture the overall experienced visual comfort of stereoscopic images, such as the ratio of absolute disparity summation between the region near the screen and far from the screen [3], spatial complexity of depth image [4], horizontal disparity and vertical disparity magnitude [5], disparity range [6], disparity variance [7], etc. For mapping model construction, the most commonly used solution that has been demonstrated to be efficient is regression based pooling strategy by using linear regression [8], support vector regression (SVR) [9, 10], and some other regression algorithms. However, all of these traditional VCA metrics consistently encounter the following problems without exception: (1) they are based on holistic statistical analysis from disparity maps, which omits the important visual attention mechanism of human visual system (HVS); (2) they require a large number of stereoscopic samples with a wide range degree of visual comfort and corresponding subjective scores to train a robust visual comfort predictor, let alone the problem that it is difficult even impossible to obtain sufficient number of training samples in many practical situations; (3) they are usually sensitive to different databases, a trained visual comfort predictor on one database may lead to rather poor performance on another database. The main reason is that the trained predictor always depends on some database-specific parameters which may not be suitable for another database. That means when new samples comes the predictor need to be retained, which is time-consuming and unacceptable in many real-time applications.

To deal with the abovementioned problems, in this paper, we propose a novel objective VCA metric based on sparse coding algorithm as an alternative. The underlying assumption is that we consider the feature space and subjective mean opinion score (MOS) space will share almost a same intrinsic manifold. That is to say, stereoscopic images with similar degree of visual comfort should have the similar features. It is reasonable and acceptable if the extracted features are sufficiently effective to capture the characteristic of overall experienced visual comfort under 3D viewing condition. In the proposed framework, inspired by the visual attention mechanism of HVS, we use both of the disparity statistics and neural activities in visual salient regions to represent a certain stereoscopic image. The concept of neural activities refers to neural responses of disparity coding in middle temporal (MT) area of visual cortex. Previous physiological studies have revealed the essential role of MT area in controlling eye convergence movements which have innumerable links with 3D visual comfort [11]. Therefore, in this regard, we believe that the visual saliency sampled disparity statistics and neural activities are expected to give a reasonable characterization of the degree of visual comfort.

The proposed VCA metric can be sketched in the following steps. Firstly, a dictionary is directly constructed by collecting a set of stereoscopic images with a wide range degree of experienced visual comfort. Then, the visual saliency labeled disparity statistics and neural activities of a certain testing stereoscopic image are extracted and encoded using the constructed dictionary via sparse coding algorithm. Finally, the sparse coding coefficients are used to linearly weight the corresponding MOS values (corresponding to each stereoscopic image in the constructed dictionary) to predict the final visual comfort score. The performance of the proposed VCA metric is validated on a newly built 3D image database which will be briefly introduced in Sect. 3, and the experimental results in this database demonstrate the proposed VCA metric is consistent with human subjective assessment compared with some representative metrics. The significant feature of this work is two-fold: (1) since the extracted features used in the proposed VCA metric simultaneously account for human visual attention mechanism and neural responses of disparity coding, they are relatively stable and certainly less sensitive to different image contents; (2) the sparse coding algorithm provides an efficient pooling strategy to estimate the final visual comfort score from the MOS values of training samples used for dictionary construction. The rest of this paper is organized as follows. Section 2 illustrates the details of the proposed VCA framework. Experiment results and discussions are given in Sect. 3. Section 4 draws conclusions of this paper.

2 Proposed VCA Metric via Sparse Coding

In this section, we present our metric mainly from four aspects, i.e., image feature representation, dictionary construction, sparse coding, and pooling strategy, respectively. Specifically, for each stereoscopic image, we at first compute disparity statistics and neural activities in specific regions of which the mean visual saliency value is larger than a fixed threshold. The computed visual saliency labeled disparity statistics and neural activities are combined for feature representation. Then, we select a set of training stereoscopic images with a wide range MOS values to construct a dictionary for sparse coding. Finally, the final visual comfort score of a testing stereoscopic image can be estimated by linearly weighting the MOS values (corresponding to each training sample) using sparse coefficients which are obtained by representing the features via sparse coding over the constructed dictionary. Figure 1 shows the high-level diagram of the proposed VCA metric.

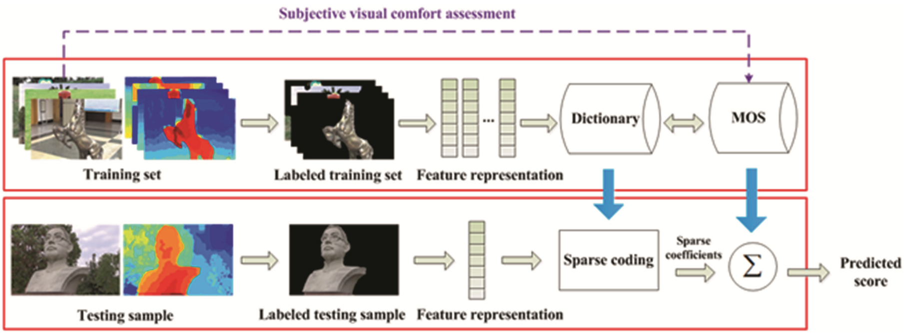

Fig. 1. High-level diagram of the proposed VCA metric.

2.1 Image Feature Representation

Unlike traditional VCA metrics, we do not extract visual comfort features by conducting global analysis of the entire image, because it is well known that visual salient regions where human pay more attention, are probably to be a primary trigger that affects overall experienced visual comfort. Features from global analysis cannot predict subjective visual comfort accurately, because the extracted features are assumed to comprehensively characterize the visual comfort, and the non-salient regions are definitely not involved with the subjective judgment of visual comfort. Inspired by this observation, in this paper, we propose to extract visual comfort related features only from those salient regions while discarding the non-salient regions.

(1) Salient Region Segmentation

In this work, the salient region is detected by considering two aspects of factors simultaneously, namely 2D visual saliency and depth saliency. Especially, we adopt the popular graph based visual saliency (GBVS) algorithm [12] to compute the 2D visual saliency because the GBVS algorithm is one of the best-known bottom-up visual saliency models which can accurately predict human fixations for monocular images. While for stereoscopic images, the additional depth information will also affect visual attention, but it cannot be estimated by directly applying the 2D saliency models which only accounts for the monocular visual attributes. Based on the evidences reported in previous studies, it is well documented that objects with smaller depth values (i.e., larger disparity values) are more likely to capture human attention compared with the background regions in a scene. In this work, we adopt a quite simple but efficient solution to generate depth saliency map by assigning the maximum (minimum) disparity value in the disparity map to highest (lowest) depth saliency value. Thus, pixels with disparity values that closer to the maximum disparity value are assigned to larger depth saliency values while pixels with disparity values that closer to the minimum disparity value are assigned to smaller depth saliency values. After calculating the GBVS map and depth saliency map, the final salient region of a certain stereoscopic image can be obtained by directly using the well-known OSTU operation on the corresponding fused saliency map which is computed as the linear combination of the GBVS map and depth saliency map with equal weights. Figure 2 shows an example of a stereoscopic image and its corresponding GBVS map, disparity map, fused saliency map, and segmented salient region, respectively.

Fig. 2. An example of a stereoscopic image and its corresponding GBVS map, disparity map, fused saliency map, and segmented salient region, respectively.

(2) Salient Region Based Feature Extraction

(A) Disparity Statistics

In this stage, for each stereoscopic image, we conduct statistical analysis and disparity coding process of disparity map with the aid of segmented salient region to extract image-level features related with visual comfort. The used disparity statistics include disparity magnitude, disparity contrast, disparity dispersion, and disparity skewness, which have been proved to be highly correlation with visual comfort [13]. Specifically, given a stereoscopic image $I_{3D}(x, y) = \{I_L(x, y), I_R(x, y)\}$, we first detect its corresponding salient region Ω as introduced above. Then, based on the salient region Ω and the disparity map $D(x, y)$, the visual comfort related disparity statistics are given by:

(a) Mean disparity magnitude in salient region Ω, denoted as f_1:

$$f_1 = \frac{1}{d_m} \cdot \left(\sum_{i=1}^{H} \sum_{j=1}^{W} M(i,j) \cdot |D(i,j)| \right) \Bigg/ \sum_{i=1}^{H} \sum_{j=1}^{W} M(i,j) \tag{1}$$

$$M(i,j) = \begin{cases} 1, & \mathit{if}\ (i,j) \in \Omega \\ 0, & \mathit{otherwise} \end{cases} \tag{2}$$

(b) Mean disparity contrast in salient region Ω, denoted as f_2:

$$f_2 = \frac{1}{d_m} \cdot \left(\sum_{i=1}^{H} \sum_{j=1}^{W} M(i,j) \cdot |D_c(i,j)| \right) \Bigg/ \sum_{i=1}^{H} \sum_{j=1}^{W} M(i,j) \tag{3}$$

(c) Disparity dispersion in salient region Ω, denoted as f_3:

$$f_3 = \frac{1}{d_m} \cdot \sqrt{\left(\sum_{i=1}^{H} \sum_{j=1}^{W} M(i,j) \cdot D(i,j)^2 \right) \Bigg/ \sum_{i=1}^{H} \sum_{j=1}^{W} M(i,j)} \tag{4}$$

(d) Disparity skewness in salient region Ω, denoted as f_4:

$$f_4 = \sum_{i=1}^{H} \sum_{j=1}^{W} M(i,j) \cdot D(i,j) \Bigg/ \left| \sum_{i=1}^{H} \sum_{j=1}^{W} M(i,j) \cdot D(i,j) \right| \tag{5}$$

where $D_c(x, y)$ is the disparity contrast map calculated by adopting a simple center-surrounding operator, H and W are the height and width of $D(x, y)$, respectively, and d_m is the maximum disparity magnitude as a normalized factor.

Meanwhile, it is known that excessive binocular disparity magnitude tends to induce visual discomfort. It means that stereoscopic images with even a small amount of excessive binocular disparities may still be perceived as uncomfortable. Therefore, we are motivated to take the percentages of maximum and minimum disparity values into account for visual comfort related feature representation. The average disparity values of the maximum and minimum $p\%$ disparity values in salient region are given by:

(e) the average disparity value of the maximum $p\%$ disparity values in salient region Ω, denoted as f_5:

$$f_5 = \frac{1}{d_m} \cdot \left(\frac{1}{N(\Omega_p^+)} \sum_{(i,j) \in \Omega_p^+} D(i,j) \right) \tag{6}$$

(f) the average disparity value of the minimum $p\%$ disparity values in salient region Ω, denoted as f_6:

$$f_6 = \frac{1}{d_m} \cdot \left(\frac{1}{N(\Omega_p^-)} \sum_{(i,j) \in \Omega_p^-} D(i,j) \right) \tag{7}$$

where Ω_p^+ and Ω_p^- represent the sets of pixels whose disparity values belong to the maximum and minimum $p\%$ disparity values over the pixels in salient region Ω, respectively, $N(\Omega_p^+)$ and $N(\Omega_p^-)$ are the number of pixels in Ω_p^+ and Ω_p^-, respectively. In our experiment, the number of $p\%$ is set to 10 %, as in [9].

(B) Neural Activity

Besides disparity statistics, inspired by the recent studies in neuroscience, we also compute the neural activity of disparity coding in MT area as the biologically plausible visual comfort features. In this work, to compute the neural activity of disparity coding in MT area, we use the computational model in [14], where the disparity coding functions are approximated by a set of modified Gabor functions with different parameters

$$r_i(d) = r_0^i + A_i \cdot e^{-0.5\left(\left(d-d_0^i\right)^2 / \delta_i^2\right)} \cdot \cos\left(2\pi f_i\left(d - d_0^i\right) + \Phi_i\right) \tag{8}$$

where $r_i(d)$ is the disparity coding response of the i-th representative neuron in middle temporal area, d is the input disparity value, r_0^i is the Gabor baseline level, A_i is the Gabor amplitude, d_0^i is the Gabor center, δ_i is the Gabor width, f_i is the Gabor frequency, and Φ_i is the Gabor phase. Similar to [14], 13 representative neurons that typify the variety of responses are selected. The parameter determination of this function for different representative neurons is based on the work in [14].

Based on the disparity coding function defined by Eq. (8), same disparity value will obtain identical neural activity. Therefore, the general neural activity $\mathbf{R}$ for an input disparity map $D(x, y)$ can be represented by a neural activity matrix: $\mathbf{R} = [\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_L] \in \Re^{13 \times L}$, where $\mathbf{R}_k = p_k \cdot [r_1(k), r_2(k), \ldots, r_{13}(k)]^{\mathrm{T}} \in \Re^{13 \times 1}$ denotes the corresponding neural activity vector of the k-th ($1 \le k \le L$) bin, L is the number of bins in the disparity histogram in salient region Ω, p_k denotes the probability of the k-th bin. Afterward, we apply a max-pooling strategy on $\mathbf{R}$ to generate a final neural activity vector. Specifically, for each row in $\mathbf{R}$, denoted as $\mathbf{r}_j = [\theta_j(1), \theta_j(2), \ldots, \theta_j(L)]$ ($1 \le j \le 13$), the max-pooling strategy performed on $\mathbf{r}_j$ can be expressed as

$$\theta_j(k) = \begin{cases} \theta_j(k), & \textit{if } \theta_j(k) = \textit{max}\left(\theta_j(1), \theta_j(2), \ldots, \theta_j(L)\right) \\ 0, & \textit{otherwise} \end{cases}, k{=}1, 2, \ldots, L \tag{9}$$

then, the final neural activity vector is obtained by summing up all columns, such that:

$$[f_7, f_8, \ldots, f_{19}] = [\mathbf{R}_1 + \mathbf{R}_2 +, \ldots, +\mathbf{R}_L]^{\mathrm{T}} \tag{10}$$

2.2 Dictionary Construction

By combining the salient region based disparity statistics and neural activities into a single vector, each stereoscopic image can be represented by a feature vector of dimension 19: $\mathbf{F} = [f_1, f_2, \ldots, f_{19}]$. In the proposed VCA framework, the dictionary construction stage is accomplished by simply combining feature vectors and corresponding subjective visual comfort scores of training samples which are selected from existing database:

$$\begin{bmatrix} \mathbf{D} \\ \mathbf{S} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_1, & \mathbf{F}_2, \ldots, \mathbf{F}_n \\ s_1, & s_2, \ldots, s_n \end{bmatrix} \tag{11}$$

where the dictionary $\mathbf{D} = [\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n]$ is a matrix of dimension $m \times n$, m is the dimension of the feature vector of each stereoscopic image (e.g., $m = 19$ in our framework) and n is the number of selected training samples, vector $\mathbf{S}$ contains the corresponding MOS values of training samples.

2.3 Sparse Coding

Motivated by the sparse coding strategy, we represent a testing stereoscopic image as a linear combination of atoms (each atom corresponds to a training sample) in the constructed dictionary $\mathbf{D}$. Denote by $\mathbf{F}_t \in \mathbb{R}^m$ the feature vector of a testing sample, its corresponding sparse representation $\mathbf{v}_t \in \mathbb{R}^{n \times 1}$ over dictionary $\mathbf{D}$ can be computed by solving the following optimization problem [15]:

$$\mathbf{v}_t^* = \arg\min_{\mathbf{v}_t} \frac{1}{2} \|\mathbf{F}_t - \mathbf{D}\mathbf{v}_t\|_2^2 + \lambda \|\mathbf{v}_t\|_1 \tag{12}$$

where the parameter λ is a positive constant balance the reconstruction error term $\|\mathbf{F}_t - \mathbf{D}\mathbf{v}_t\|_2^2$ and the sparse constraint term $\|\mathbf{v}_t\|_1$.

2.4 Pooling Strategy

By applying this sparse coding strategy to a certain testing stereoscopic image, the corresponding sparse coefficients over dictionary **D** can be easily obtained. Each element in the sparse coefficients denotes the relative importance of each atom (training sample) to reconstruct current testing sample. Based on the hypothesis that the feature space and subjective score space will share almost a same intrinsic manifold, the final visual comfort score can be pooled as:

$$Q_t = \frac{\sum_{i=1}^{n} \mathbf{v}_t^* s_i}{\sum_{i=1}^{n} \mathbf{v}_t^*} \tag{13}$$

where s_i is the MOS value of the i-th training stereoscopic image in the dictionary **D**.

3 Experimental Results and Analyses

We construct a new 3D image database named NBU 3D-VCA database for performance validation. This database contains 82 indoor and 118 outdoor stereoscopic images (a total number of 200) with a wide range degree of visual comfort. These stereoscopic images are all with a full HD resolution of 1920 × 1080 pixels. All the images are captured at the campus of Ningbo University using a 3D digital camera with dual lenses (SONY HDR-TD30E). The MOS (ranges from 1 to 5, 1 refers to extremely uncomfortable, and 5 corresponds to very comfortable) of each stereoscopic image is provided, which is obtained via a standard human subjective experiment. The subjective tests were conducted in the laboratory designed for subjective quality tests according to the recommendations of ITU-R BT.500-11 [16] and ITU-R 1438 [17]. Sixteen non-expert adult viewers (seven females and nine males) with ages range from 22 to 38 were participated in the subjective evaluation of the database. Forty selected right-view images in this database are shown in Fig. 3. We use the stereo matching algorithm presented in [18] for disparity estimation since its performance is prominent for high-quality stereoscopic images. For performance quantization, we use two commonly used criteria: Pearson linear correlation coefficient (PLCC), Spearman rank order correlation coefficient (SRCC), between the predicted visual comfort scores and MOS values to quantify the performance of a specific VCA metric. The criterion PLCC is used to measure the prediction accuracy, and SRCC is used to benchmark the prediction monotonicity. Especially, for a perfect match, we have PLCC = SRCC = 1.

In the detailed implementation, we first select 150 training samples from the NBU 3D-VCA database and the remainder 50 samples are used for performance test. The selected training samples cover a wide range of MOS values. For each training sample, we conduct the image feature representation to extract corresponding visual comfort related feature vector. The feature vectors of all the training samples are combined to construct dictionary **D**.

Fig. 3. Forty selected right-view images in the NBU 3D-VCA database

We compare the proposed metric with other five representative VCA metrics, i.e., Yano's metric [3], Choi's metric [4], Kim's metric [5], Sohn's metric [9], and Yong's metric [10]. For fair comparison, the same training and testing splits are conducted in the implementation of these metrics (i.e., 150 samples for training and 50 samples for testing). In addition, to better investigate the contribution of each component in the proposed VCA metric, we further design six schemes (denoted by M_1 to M_6, respectively) for comparison by considering different settings presented in Table 1. The PLCC and SRCC comparison results are listed in Table 2, where the metric with best performance has been highlighted in boldface. It is clear that our metric outperforms the five representative metrics and other modified schemes under different settings in terms of both PLCC and SRCC. The reason may lie in following aspects: first, salient region based visual comfort related feature is capable of reflecting the perceived visual comfort since it has well emphasized the important human visual attention mechanism which plays a significant role in visual perception; second, the combined use of disparity statistics and neural activities can more accurately characterize the perceived visual comfort compared with isolated disparity statistics; third, sparse coding strategy has provided an efficient way for visual comfort score pooling from the MOS values of training samples. Figure 4 shows the scatter plots between MOS values and predicted visual comfort scores obtained from the proposed metric. Obviously, the scatter plots further demonstrate the consistency of the proposed metric with respect to subjective judgment.

Table 1. Schemes under different settings of the proposed VCA framework

Saliency constraint / Adopted feature	Without salient region constraint	With salient region constraint
Disparity statistic	M_1	M_2
Neural activity	M_3	M_4
Disparity statistic + Neural activity	M_5	M_6 (Proposed)

Table 2. PLCC and SRCC comparison results of different VCA metrics

Metrics	Yano's [3]	Choi's [4]	Kim's [5]	Sohn's [9]	Yong's [10]	Proposed
PLCC	0.4569	0.6885	0.7428	0.7831	0.7762	0.8142
SRCC	0.3828	0.6046	0.6917	0.7530	0.7526	0.7629
Metrics	M_1	M_2	M_3	M_4	M_5	M_6
PLCC	0.7573	0.7960	0.7594	0.8036	0.7953	0.8142
SRCC	0.6884	0.7604	0.7108	0.7541	0.7312	0.7629

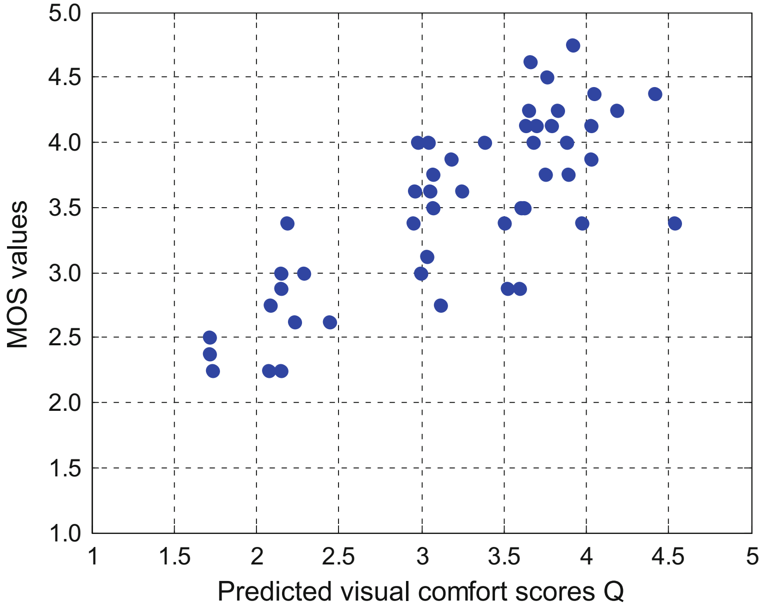

Fig. 4. Scatter plots between MOS values and predicted visual comfort scores obtained from the proposed metric.

4 Conclusions

In this paper, we have presented an objective VCA metric for stereoscopic images based on sparse coding strategy. The main advantages of this work are two-fold: first, salient region based disparity statistics and neural activities are extracted to represent a stereoscopic image in terms of visual comfort; second, the final visual comfort score is computed by linearly weighting the MOS values of the training samples using the corresponding sparse coefficients. Experimental results on our newly built database show the promising performance at handling the VCA problem for stereoscopic images. Further work should be concentrated on exploiting more accurate visual comfort related features and investigating the influence of binocular rivalry on the perceived visual comfort of stereoscopic images.

Acknowledgments. This work was supported by the Natural Science Foundation of China (grant 612710210).

References

1. Lambooij, M., Ijsselsteijn, W., Fortuin, M., et al.: Visual discomfort and visual fatigue of stereoscopic displays: a review. J. Imaging Sci. Technol. **53**(3), 1–14 (2009)
2. Tam, W.J., Speranza, F., Yano, S., et al.: Stereoscopic 3D-TV: visual comfort. IEEE Trans. Broadcast. **57**(2), 335–346 (2011)
3. Yano, S., Ide, S., Mitsuhashi, T., et al.: A study of visual fatigue and visual comfort for 3D HDTV/HDTV images. Displays **23**(4), 191–201 (2002)
4. Choi, J., Kim, D., Choi, S., et al.: Visual fatigue modeling and analysis for stereoscopic video. Opt. Eng. **51**(1), 017206-1–017206-11 (2010)
5. Kim, D., Sohn, K.: Visual fatigue prediction for stereoscopic image. IEEE Trans. Circ. Syst. Video Technol. **21**(2), 231–236 (2011)
6. Lambooij, M., IJsselsteijn, W.A., Heynderickx, I.: Visual discomfort of 3-D TV: assessment methods and modeling. Displays **32**(4), 209–218 (2011)
7. Choi, J., Kim, D., Ham, B., et al.: Visual fatigue evaluation and enhancement for 2D-plus-depth video. In: Proceedings of IEEE International Conference on Image Processing, pp. 2981–2984 (2010)
8. Jung, Y., Lee, S., Sohn, H., et al.: Visual comfort assessment metric based on salient object motion information in stereoscopic video. J. Electron. Imaging **21**(1), 011008-1–011008-16 (2012)
9. Sohn, H., Jung, Y., Lee, S., et al.: Predicting visual discomfort using object size and disparity information in stereoscopic images. IEEE Trans. Broadcast. **59**(1), 28–37 (2013)
10. Yong, J., Sohn, H., Lee, S., et al.: Predicting visual discomfort of stereoscopic images using human attention model. IEEE Trans. Circ. Syst. Video Technol. **23**(12), 2077–2082 (2013)
11. DeAngelis, G., Cumming, B., Newsome, W.T.: Cortical area MT and the perception of stereoscopic depth. Nature **394**, 677–680 (1998)
12. Harel, J., Koch, C., Perona, P.: Graph-based visual saliency. In: Proceedings of Advances in Neural Information Processing Systems (2006)
13. Nojiri, Y., Yamanoue, H., Hanazato, A., et al.: Measurement of parallax distribution and its application to the analysis of visual comfort for stereoscopic HDTV. In: Proceedings Stereoscopic Displays Virtual Reality Syst. X, vol. 5006, pp. 195–205 (2003)
14. DeAngeliS, G.C., Uka, T.: Coding of horizontal disparity and velocity by MT neurons in the alert macaque. J. Neurophysiol. **89**(2), 1094–1111 (2003)
15. Wright, J., Yang, A., Ganesh, A., et al.: Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. **31**(2), 1–18 (2009)
16. Methodology for the subjective assessment of the quality of television pictures, ITU-R BT-500.11 (2002)
17. Subjective assessment for stereoscopic television pictures, ITU-R BT.1438 (2000)
18. Sun, D., Roth, S., Black, M.: Secrets of optical flow estimation and their principles. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, pp. 2432–2439 (2010)

http://www.springer.com/978-3-319-21977-6

Image and Graphics
8th International Conference, ICIG 2015, Tianjin, China,
August 13-16, 2015, Proceedings, Part I
Zhang, Y.-Y. (Ed.)
2015, XLI, 683 p. 362 illus., Softcover
ISBN: 978-3-319-21977-6