

Saliency-aware Image Quality Assessment

Zhanghui Liu¹, Yize Huang¹, Yuzhen Niu^{1, 2, a}, Lening Lin¹

¹College of Mathematics and Computer Science, Fuzhou University, Fuzhou, 350116, China

²Fujian Province Key Laboratory of Network Computing and Intelligent Information Process, Fuzhou 350116, China

^aemail:yuzhenniu@gmail.com

Keywords: Visual Saliency; Image Quality Assessment; Saliency Detection;

Abstract. Currently, visual saliency detection has been discovered and widely applied in the field of computer vision and image processing. Recent works have shown that considering human visual system features in image quality assessment will improve the consistence between objective assessment results and subjective visual perception. So we propose saliency-aware image quality assessment through a universal method to combine saliency detection with image quality assessment metrics. Since there is no image dataset used for image quality assessment and saliency detection researches simultaneously, we create a new image database, called TID-2013S, which is inheriting from image quality assessment database TID-2013 and contains 25 manually labeled ground truth images. In this paper, we select 8 representative full reference image quality assessment metrics and 7 state-of-the-art saliency detection algorithms to validate the performance of our proposed method. Experiment results show that when using the appropriate parameter of mapping function, our saliency-aware approach achieves the best performance improvement by 5.41%. And for most of image quality assessment metrics, the MCA saliency detection algorithm works best.

Introduction

With the development of social informatization and computer technology, digital image quickly applied to every aspect of human life, such as face recognition, image restoration, video monitoring, and medical image analysis, etc. In practical applications, however, it will produce different types and degrees of distortion during image transmission, compression, saving and other image processing. In response to the rapid development of image and video applications demand, domestic and foreign scholars devote a great deal of time and energy on related researches. Image quality assessment (IQA) is a basic and challenging issue of image processing which aims to establish an intelligent mathematical model to evaluate the distorted image, making the results consistent with subjective visual perception.

In academic, IQA can be divided into two categories: subjective assessment and objective assessment. Subjective assessment evaluate image quality directly by human visual system (HVS), usually recorded as mean opinion score (MOS), so IQA algorithms regard subjective evaluation as standard. Objective assessment depends on low-level image features to design mathematical model to evaluate images. It costs much time, manpower and material resources while doing subjective assessment, so that objective way is recognized and concerned by many scholars and experts because of its convenience, time-saving and high-efficiency. According to the level of usage of reference image, there are full-reference, no-reference and reduced-reference IQAs. In this paper, we mainly focus on the full-reference IQA (FR-IQA), and we choose 8 representative FR-IQAs: MSSIM [1], SSIM [2], PSNR [3], PSNRc [3], MAD [4], FSIM [5], ESSIM [6], TBC [7]. Wang et al. [1] proposed a Multi-Scale Structural Similarity (i.e., MSSIM) method which outperforms previous single-scale methods in 2003 and a Structure Similarity Metric (i.e., SSIM) which introduces the degradation of structural information in 2004 [2]. Then Egiazarian et al. [3] proposed a metric by computing the Peak Signal to Noise Ratio (i.e., PSNR) while considering the HVS in 2006 and then they incorporated the image chromatic feature to extend PSNR to PSNRc. In 2010,

Larson et al. [4] proposed a Most Apparent Distortion (i.e., MAD) metric which attempts to explicitly model two separate strategies of a detection-based strategy and an appearance-based strategy to determine the image quality. In 2011, Zhang et al. [5] proposed a novel Feature Similarity (i.e., FSIM) metric for IQA which mainly depends on the low-level features from image.

Recently, visual saliency has been applied in computer vision and image processing applications, such as adaptive display, object recognition, image retrieval, image compression, image matching, etc. The purpose of IQA is to obtain the quality scores that are consistent with subjective perception, so considering human visual system features will enhance perception consistency. There is a rich literature on visual saliency and IQA, but lack of the systematic analyses both of them. In this paper, we choose 7 typical saliency detection algorithms: MCA [8], BSCA [9], DRFI [10], SMD [11], SO [12], VA [13] and GS [14] for test. Jiang, H.Z. et al. [10] used the supervised machine learning approach to map the regional feature vector to a saliency score, and finally fused the saliency scores across multiple levels, yielding the saliency map (i.e., DRFI). Peng, H. et al. [11] proposed a useful salient object detection method (i.e., SMD) based on a novel structured matrix decomposition model. The saliency optimization (i.e., SO) method first characterized the spatial layout of image regions with respect to image boundaries [12]. Then it built a principled optimization framework that integrated multiple low-level cues, including the background measure obtained in the first step. Jiang et al. [13] formulated saliency detection via absorbing (i.e., VA) Markov chain on an image graph model. Wei et al. [14] proposed a geodesic saliency (i.e., GS) algorithm that used the background prior. GS used the background for geodesic saliency calculation to obtain the shortest distance between the calculating region and visible background points. Figure 1 shows ground truth images of 5 reference images in TID-2013S and saliency maps produced by 7 state-of-the-art saliency detection algorithms.

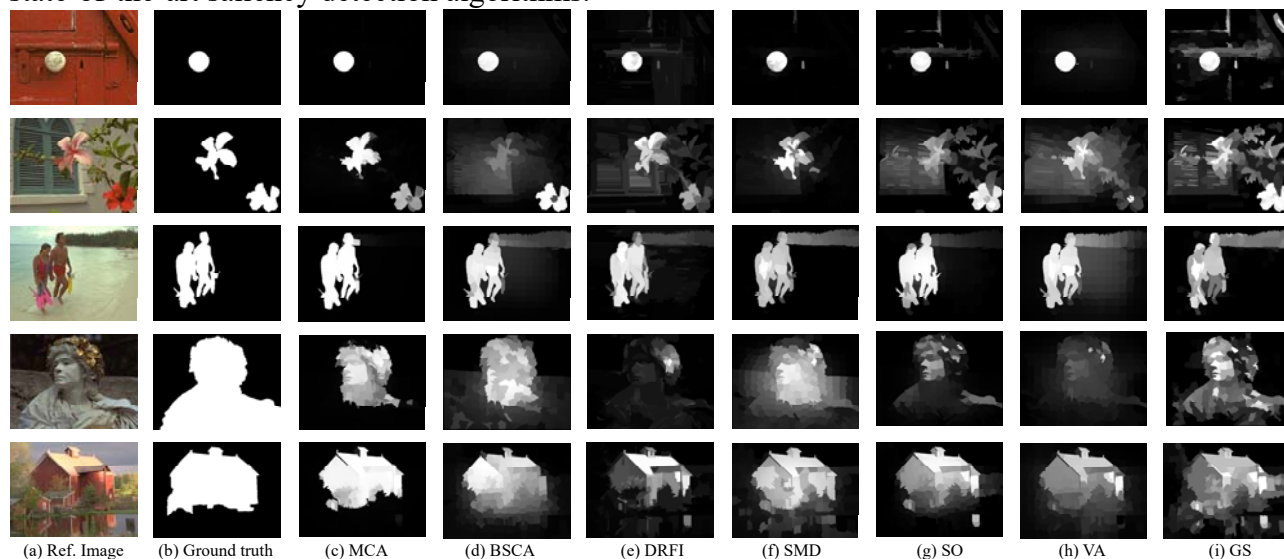


Fig.1. Ground truth images of 5 reference images in TID-2013S and saliency maps produced by 7 state-of-the-art saliency detection algorithms

Saliency Detection for Image Quality Assessment

TID2013S Database

As we know, TID-2008 [15], CSIQ [4], TID-2013 [16] and LIVE-2005 [17] are proposed for image quality assessment researches, and MSRA-1000 [18], THUS-10000 [19] and PASCAL-1500 [20] are used for salient object detection applications. However, there is no database established for these two issues simultaneously since there are only few works on saliency detection applied to image quality assessment. For this specific research, we create TID-2013S database extending from the image quality assessment database TID-2013.

New database includes 25 reference images, 3000 distorted images, and 25 ground truth images which are produced by following steps:

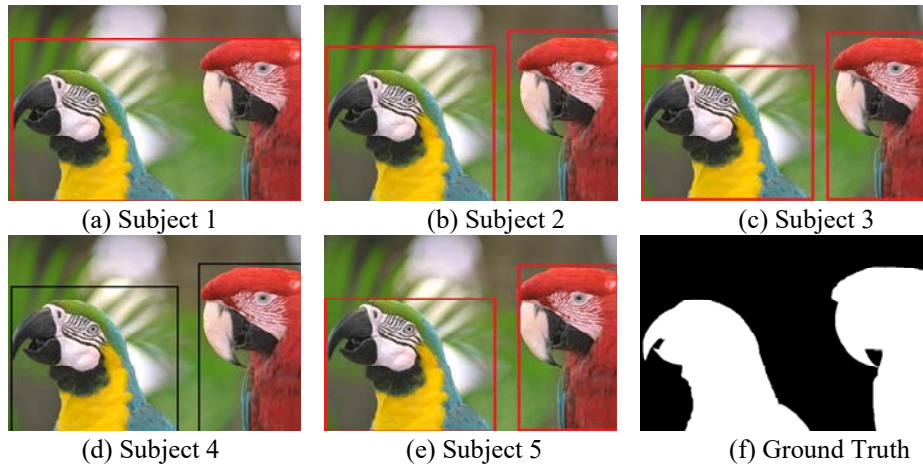


Fig.2. Rectangle masks of salient object and ground truth according to 5 subjects' opinion

- (1) Collect subjects' cognition of salient objects from the reference image. In this step, 5 subjects are told to mark the salient objects using rectangles. The cognitions of salient objects in the same image are different from different subjects. As it has happened, we follow the principle that the minority is subordinate to the majority. Figure 2 (a ~ e) show the rectangle masks made by 5 subjects.
- (2) According to rectangle marks, we manually label the salient object along its boundary to get the ground truth images. We use professional image processing software (Adobe Photoshop CS6) to label the salient objects. Finally, we mark the salient object regions with white color and the other regions with black color. In the end, ground truth image is a black-white binary image. Figure 2 (f) shows the ground truth image.

Proposed Algorithm

Most proposed FR-IQA methods use the information from low-level image characteristic, such as color, edge, luminous and saturation, to assess the image quality. These methods generate a quality map between reference image and distorted image and then simply calculate the average value of quality map as image quality score. In the paper, we introduce saliency detection into IQA method by two steps. First, we use traditional IQA metrics to calculate the quality map from distorted image. As shown in Figure 3, image (d) is the quality map of reference image (a) and distorted image (b), which is calculated by FSIM metric. Second, we use saliency detection algorithm to generate a saliency map of distorted image and obtain a weighing map after mapping process. The weighing map is applied to quality map to calculate weighted average value, i.e., the final image quality score. Figure 3 show the flow of traditional IQA and our proposed methods.

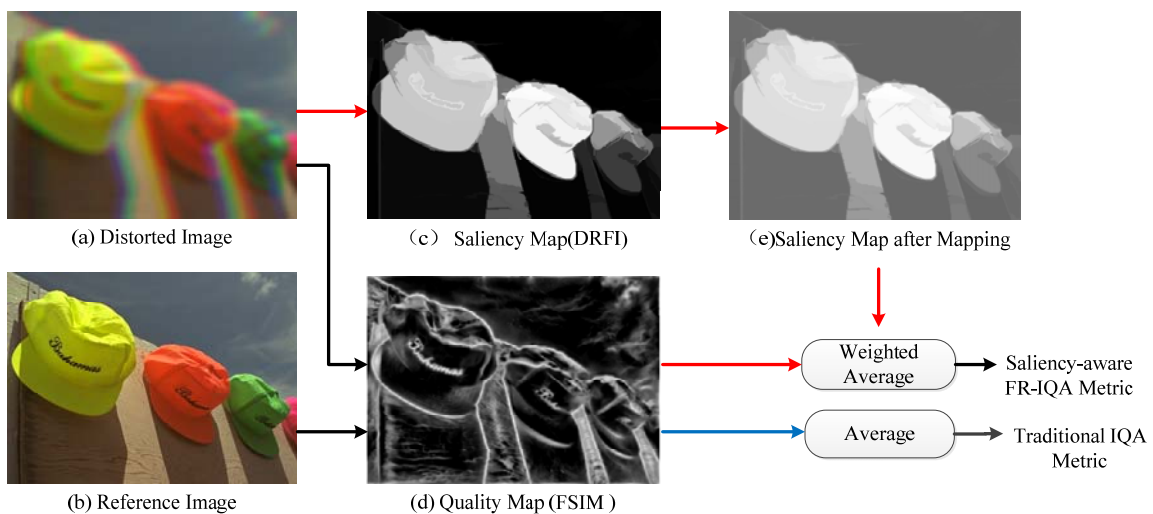


Fig.3. The flow chart of traditional IQA metric and our saliency-aware FR-IQA metric (Black arrows indicate the same process, blue arrow indicates the traditional IQA metric and red arrows indicate saliency-aware FR-IQA metric)

In the second step, we obtain the saliency map from distorted image by saliency detection algorithms. As shown in Figure 3, the image (c) is the saliency map of image (a) generated by DRFI algorithm. Since the saliency value in foreground is higher than the background, we can increase the influence of salient objects in assessment by taking saliency map as the weighing map. However, the information of background also influences the image quality, so we use Equation (1) to map the value of saliency map, which reserves the considerable influence of salient objects and enhance the influence of background to a certain extent simultaneously.

In Equation (1), x is the original saliency value, and a is a parameter of mapping function. Since the saliency map is grayscale and the value of each pixels lies in the range of $[0, 1]$, the value of the mapped $f(x)$ is in the range of $[a, 1]$. We take the mapped saliency map as a weighting map and calculate the weighted average score following Equation (2) as the saliency-aware image quality score.

$$f(x) = (1-a) * x + a \quad (1)$$

$$F_s = \frac{\sum_i^p \sum_j^q [W(i, j) * Q(i, j)]}{\sum_i^p \sum_j^q W(i, j)} \quad (2)$$

where F_s is the quality score of the distorted image, p and q represent the width and height of image, $W(i, j)$ represents the mapped saliency value of pixel (i, j) , i.e., the saliency value of pixel (i, j) in Figure 3 (e), $Q(i, j)$ represents the quality map calculated by existing IQA metrics, as shown in Figure 3 (d).

Experimental Results

In order to verify the performance and generality of the proposed algorithm, we conduct experiments on TID2013S with 8 IQA metrics and 7 saliency detection algorithms described above. To judge the consistency between IQA metrics with and subjective visual perception (MOS), four commonly used performance metrics are adopted, namely, the Spearman rank correlation coefficient (SROCC), the Kendall rank correlation coefficient (KROCC), the Pearson linear correlation coefficient (PLCC), and the root mean square error (RMSE).

Experiments using ground truth map

Since the exiting saliency detection algorithms cannot generate the same saliency map as ground truth, we verify the possibility of proposed algorithm by directly using ground truth image. We select 8 typical IQA metrics and combine them with the ground truth images to obtain the quality assessment scores for each distorted image from database TID2013S, and then calculate the SROCC, KROCC, PLCC and RMSE with its corresponding MOS. The experimental results are shown in Figure 4.

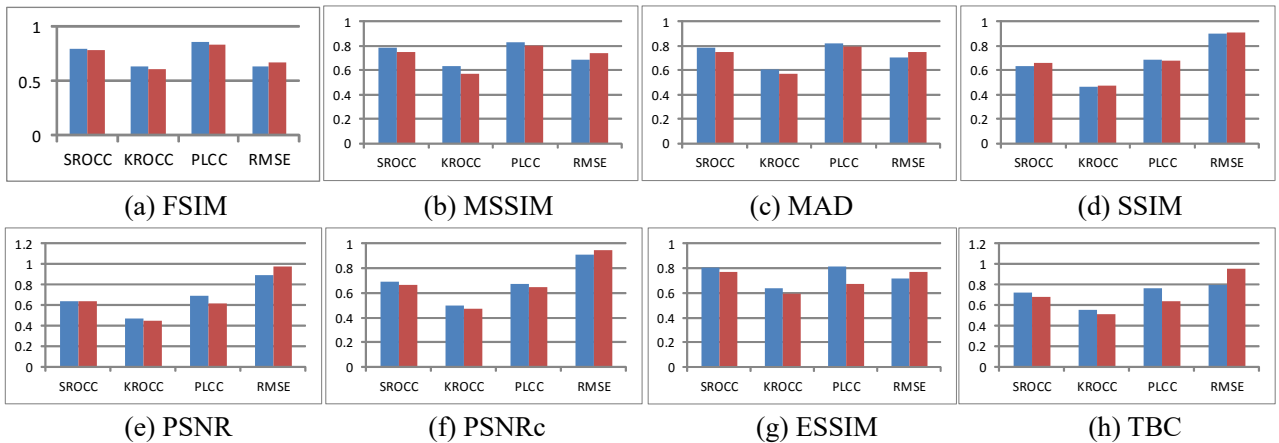


Fig.4. The correlation coefficients between 8 saliency-aware FR-IQA metrics combining with the ground truth image and their corresponding MOS (The blue chart bar indicates the original IQA metric and the red bar indicates the saliency-aware FR-IQA metric)

From Figure 4, we can intuitively find that for SROCC and KROCC, except for the SSIM in

Figure 4 (d), the other saliency-aware IQA metrics work worse than their original ones. And the PLCC of the saliency-aware IQA metrics is lower than their original metrics. The RMSE is higher than the original ones. The experimental results show that introducing the ground truth directly does not improve the original method. The possible reason is that the salient value of background region in ground truth is 0, however, the information in background is not totally ineffective for image quality assessment. Therefore, we map the saliency value of ground truth to make full advantages of both foreground and background information.

Next, we attempt to find the best mapping function by experiment with saliency-aware FSIM and MSSIM metrics. Taking the mapping parameter as 0.1, 0.2, ..., 0.9 respectively, we obtain the image quality assessment results, and then we calculate SROCC and KROCC with their corresponding MOS so as to determine the best parameter value. Since SROCC and KROCC are more important than PLCC and RMSE for IQA problem, we only calculate SROCC and KROCC to determine the best, and the results are shown in Figure 5.

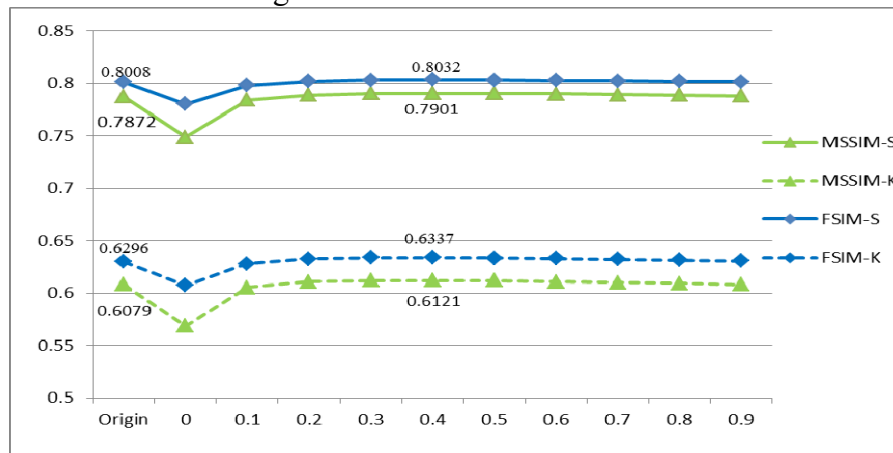


Fig.5. The SROCC and KROCC of saliency-aware FSIM and MSSIM combining with ground truth image by different values of mapping parameter

From Figure 5, we can see the following conclusions: Firstly, when gradually changing the value of mapping parameter from 0.1 to 0.9, all of the parameters for saliency-aware FSIM and 8 parameters for saliency-aware MSSIM outperform the original ones. In other words, ground truth maps after mapping make our algorithm work. Secondly, saliency-aware FSIM and MSSIM metrics achieve best performance when mapping parameter is 0.4. The SROCC and KROCC of saliency-aware FSIM is 0.8032 and 0.6337, which achieve 3.00% and 6.51% improvement against the original ones. The SROCC and KROCC of saliency-aware MSSIM is 0.7901 and 0.6121, which achieve 5.56.00% and 7.64% improvement.

So we draw a conclusion that it is useless to employ ground truth maps directly when evaluating image quality, however, mapping function makes our saliency-aware IQA strategy work, and it achieve the best performance with parameter equal to 0.4.

Experiments using saliency map

Most of time, ground truth maps are unavailable in practical applications. Alternatively, saliency maps generated by saliency detection algorithms are used in practice. In this part, we conduct experiments by using saliency maps instead of ground truth maps.

According to the experimental results and the conclusion in subsection 2.1, the mapping parameter a is set to be 0.4. Then we employ MCA [7], BSCA [8], DRFI [8], SMD [9], SO [10], VA [11], and GS [12] algorithms to get saliency maps. After getting the saliency maps, the weighting map is generated by mapping the saliency maps. Then we compute the saliency-aware IQA metrics by using the weighting map as shown in Equation 2. Finally, we calculate the KROCC and SROCC between the image quality scores calculated by our method and their corresponding MOS. We show the results in Table 1.

As shown in Table 1, we can conclude that all the combinations of 8 FR-IQA metrics and 7 saliency detection algorithms outperform the original ones. For different FR-IQA metrics, their best performances are achieved by combining with different saliency detection algorithms. For examples,

the combination of FSIM and, that of MCA, that of SSIM and GS, and that of BSCA and MAD work best. So there is no universal saliency detection algorithm can best improve all FR-IQA metrics.

Table1. SROCC and KROCC between saliency-aware FQ-IQA metrics using different saliency detection algorithms and corresponding MOS (Red values indicate rank in the first place, Purple ones indicate rank in the second place and Green ones indicate rank in the third place)

		Origin	MCA	BSCA	DRFI	SMD	SO	VA	GS
FSIM	SROCC	0.8007	0.8044	0.8036	0.8033	0.8038	0.8035	0.8034	0.8031
	KROCC	0.6300	0.6354	0.6340	0.6340	0.6342	0.6342	0.6339	0.6340
MSSIM	SROCC	0.7872	0.7905	0.7901	0.7894	0.7903	0.7902	0.7903	0.7895
	KROCC	0.6079	0.6124	0.6114	0.6111	0.6119	0.6118	0.6118	0.6113
SSIM	SROCC	0.6370	0.6526	0.6473	0.6498	0.6494	0.6502	0.648	0.6534
	KROCC	0.4636	0.4747	0.4706	0.4726	0.4728	0.4731	0.4715	0.4760
PSNR	SROCC	0.6395	0.6681	0.6674	0.6669	0.6671	0.6688	0.6678	0.6697
	KROCC	0.4700	0.4812	0.4804	0.4807	0.4804	0.4820	0.4809	0.4827
PSNR _c	SROCC	0.6869	0.6912	0.6910	0.6896	0.6907	0.6914	0.6918	0.6927
	KROCC	0.4958	0.5010	0.5004	0.5000	0.5003	0.5014	0.5013	0.5023
MAD	SROCC	0.7522	0.7799	0.7810	0.7803	0.7805	0.7797	0.7803	0.7791
	KROCC	0.5727	0.6025	0.6037	0.6034	0.6033	0.6027	0.6031	0.6023
ESSIM	SROCC	0.8015	0.8045	0.8044	0.8034	0.8042	0.8039	0.8039	0.8026
	KROCC	0.6342	0.6387	0.6387	0.6380	0.6383	0.6381	0.6379	0.6368
TBC	SROCC	0.7278	0.7296	0.7301	0.7288	0.7304	0.7286	0.7296	0.7281
	KROCC	0.5567	0.5592	0.5594	0.5586	0.5598	0.5583	0.5592	0.5578

In order to analysis the effects of saliency-aware FR-IQA metrics with different saliency detection algorithms, we calculate the increased rate and the number of top three improvements. The MCA has the highest average improvement rate, which is increased by 1.59% (SROCC) and 1.72% (KROCC) respectively. The MCA and GS achieve 6 times best improvement performance, but the MCA reaches the highest 12 times in top three improvement performance. Overall, we suggest the MCA algorithm for saliency-aware FR-IQA metrics.

Conclusion

Existing full reference image quality assessment metrics primarily based on image low-level features and overlook the important role of human visual system. In this paper, we propose a saliency-aware full reference image quality evaluation method, and build a new image database which meets the requirements of saliency detection and image quality assessment researches simultaneously. On this new image database, we experiments our saliency-aware FR-IQA approach with the combinations of 7 representative saliency detection algorithms and 8 state-of-art full reference image quality evaluation metrics, and compare the performance of proposed metrics to original IQA metrics. According to experimental result, we got three conclusions as follows. Firstly, saliency-aware FR-IQA metrics outperform the original metrics which increase the consistency with subjective visual perception with the maximum performance increase rate by 5.41%. Secondly, when mapping parameter set to be 0.4, most of saliency-aware FR-IQA metrics reach their best improvement performance. Thirdly, by comprehensive consideration, different FR-IQA metrics with MCA saliency detection algorithm gets better performance most of time. To some extent, the performance of our proposed algorithm depends on saliency detection algorithm, so to further enhance the algorithm performance, we aims to find a more general and more effective saliency detection algorithm in our future work.

Acknowledgement

In this paper, the research is partly sponsored by the National Natural Science Foundation of China (Project No. 61300102), the Fujian Natural Science Funds for Distinguished Young Scholar (Project No. 2015J06014), and the Natural Science Foundation of Fujian Province (Project No. 2014J01233, No 2014J01231)

References

- [1] Wang Z, Simoncelli E P, Bovik A C. Multiscale Structural Similarity for Image Quality assessment[C] //Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on IEEE, 2003, 2: 1398-1402.
- [2] Wang Z, Bovik A C, Sheikh H R, et al. Image Quality Assessment: From Error Visibility to Structural Similarity [J]. Image Processing, IEEE Transactions on, 2004, 13(4): 600-612.
- [3] Egiazarian K, Astola J, Ponomarenko N, et al. New Full-Reference Quality Metrics based on HVS[C]//CD-ROM Proceedings of the Second International Workshop on Video Processing and Quality Metrics, Scottsdale, USA, 2006, 4.
- [4] Larson E C, Chandler D M. Most Apparent Distortion: Full-Reference Image Quality Assessment and the Role of Strategy [J]. Journal of Electronic Imaging, 2010, 19(1):143-153.
- [5] Zhang L, Zhang L, Mou X, et al. FSIM: A Feature Similarity Metric for Image Quality Assessment [J]. Image Processing, IEEE Transactions on, 2011, 20(8): 2378-2386.
- [6] Zhang X, Feng X, Wang W, et al. Edge Strength Similarity for Image Quality Assessment[J]. IEEE Signal Processing Letters, 2013, 20(4):319-322.
- [7] Gu K, Liu M, Zhai G, et al. Quality Assessment Considering Viewing Distance and Image Resolution [J]. IEEE Transactions on Broadcasting, 2015, 61(3):520-531.
- [8] Wang Z, Bovik A C. A Universal Image Quality Metric [J]. IEEE Signal Processing Letters, 2002, 9(3):81 - 84.
- [9] Qin Y, Lu H, Xu Y, et al. Saliency Detection via Cellular Automata[C]//Computer Vision and Pattern Recognition. IEEE, 2015:110-119.
- [10] Jiang H, Wang J, Yuan Z, et al. Salient Object Detection: A Discriminative Regional Feature Integration Approach[C]// IEEE Conference on Computer Vision & Pattern Recognition. 2013:2083-2090.
- [11] H.Peng, B.Li, H. Ling, W. Hu, W. Xiong and J. Stephen, Salient Object Detection via Structured Matrix Decomposition, AAAI (2013).
- [12] Zhang X, Wang Z, Yan C, et al. Saliency Optimization from Robust Background Detection[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2014:2814-2821.
- [13] Jiang B, Zhang L, Lu H, et al. Saliency Detection via Absorbing Markov Chain[C]//2013 IEEE International Conference on Computer Vision (ICCV). IEEE Computer Society, 2013:1665-1672.
- [14] Wei Y, Wen F, Zhu W, et al. Geodesic Saliency Using Background Priors[C]//European Conference on Computer Vision. 2012:29-42.
- [15] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, F. Battisti, "TID2008 - A Database for Evaluation of Full-Reference Visual Quality Assessment Metrics", Advances of Modern Radioelectronics, Vol. 10, pp. 30-45, 2009.
- [16] N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, C.-C. Jay Kuo, Image database TID2013: Peculiarities, Results and Perspectives, Signal Processing: Image Communication, vol. 30, Jan. 2015, pp. 57-77.
- [17] Ghadiyaram D, Bovik A C. Massive Online Crowdsourced Study of Subjective and Objective Picture Quality [J]. IEEE Transactions on Image Processing, 2015:1-1.
- [18] Achanta R., Hemami S., Estrada F., Frequency-tuned salient region detection.[C]//In Computer Vision and Pattern Recognition, 2009. IEEE Conference: 2009 Institute of Electrical and Electronics Engineers, 2009:1597-1604.
- [19] Ming-ming Cheng, Niloy J.Mitra, Xiaolei Huang, Philip H. S. Torr, Shi-min Hu. Global Contrast based Salient Region Detection. [C]//2011 Computer Vision and Pattern Recognition, IEEE Conference: 2011 Institute of Electrical and Electronics Engineers, 2011:409-416.
- [20] Liu Tie, Sun Jian, Zheng Nan-ning, Tang Xiaou, Shum Heung-yeung. Learning to Detect a Salient Object [J]. Pattern Analysis & Machine Intelligence IEEE Transactions on, 2007, 33(2):353 - 367.