Hybrid Integration of Visual Attention Model into Image Quality Metric

Summary

Integrating the visual attention (VA) model into an objective image quality metric is a rapidly evolving area in modern image quality assessment (IQA) research due to the significant opportunities the VA information presents. So far, in the literature, it has been suggested to use either a task-free saliency map or a quality-task one for the integration into quality metric. A hybrid integration approach which takes the advantages of both saliency maps is presented in this paper. We compare our hybrid integration scheme with existing integration schemes using simple quality metrics. Results show that the proposed method performs better than the previous techniques in terms of prediction accuracy.

Key words: image quality metric, visual attention (VA), hybrid integration

1. Introduction

Machine evaluation of image quality plays an essential role in a wide range of applications such as image acquisition, processing, compression, transmission, reproduction, display, and so on [1]. For this reason, developing objective image quality assessment (IQA) methods has attracted much research interest over the past decade. In the literature, a large number of image quality metrics have been proposed. Depending upon the availability of reference image, they are mainly classified as full-reference (FR), reduced-reference (RR), and no-reference (NR) metrics. Each of these three types of metrics has its own usefulness and applications, and this paper gives particular attention to the problem of “FR” IQA.

Many researchers have devoted their efforts to incorporate relevant characteristics of the human visual system into image quality measures. This is because the ultimate goal of IQA research is to predict the image quality as perceived by human subjects. In recent years, integration of visual attention (VA) model into image quality metric has emerged as a promising solution [2]–[5]. Such an approach is based on the assumption that distortions occurring in visually salient areas might be more visible, and thus more annoying [3], [5].

Quantitative methods of predicting VA are usually unrealistic for real-time applications [3], [4]. On the other hand, the latter has the advantage that it enables automated deployment in various image processing systems [4], [5]. Due to this reason, numerous methods which integrate the computational VA model into quality measures continue to be reported [2], [4], [6], [7], and we focus on such a potential paradigm in this paper. The fundamental idea behind the integration is to weight image quality measurements by objective saliency map obtained from the computational VA model. Existing methods use either a “task-free” saliency map for an original image or a “quality-task” one for a test image in the design of VA-based objective metrics. Some researchers have compared the benefits of the two policies [4], [5]. In this paper, we propose a hybrid integration approach which simultaneously takes into account the two sources of VA information. The performance of the hybrid integration scheme is compared to that of conventional integration schemes using simple FR quality measures, and we find that the proposed method offers greater performance improvement than the previous approaches.

Here, before the details of our method is presented, we briefly review the above-mentioned simple and popularly used FR image quality measures. As was done in [3] and [5], we have used the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [8] in this paper to compare the performance of different VA integration policies. The PSNR is computed as follows:

$$\text{PSNR} = 20 \log_{10} \left( \frac{255}{\text{MSE}} \right).$$

(1)

where $\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (x_i - y_i)^2$.

In (1), $x_i$ and $y_i$ denote the $i$th pixel in the original image $x$ and the test image $y$, respectively, and $N$ denotes the total number of pixels in the image. The SSIM [8] basically measures the similarity between the two images based on local comparison of luminance, contrast, and structure.

2. Proposed Hybrid Integration Approach

As demonstrated in the previous studies [3], [5], both task-free and quality-task saliency maps are beneficial to the evaluation of image quality, to a greater or lesser extent. In this paper, we exploit the fact that each map is unique and informative in a different way (see examples in Fig. 1 (d) and (f)). The system diagram of the proposed method is illustrated in
Table 1 PLCC and SROCC of different VA integration approaches.

| TID2008 database [12] | PLCC  | Gain (%) | SROCC | Gain (%) |
|------------------------|-------|----------|-------|----------|
| PSNR                   | 0.5190| -        | 0.5531| -        |
| PSNR-TF [2], [3], [5]  | 0.5381| 3.7      | 0.5448| -1.5     |
| PSNR-QT [3], [5]       | 0.5314| 2.4      | 0.5428| -1.8     |
| PSNR-HYBRID            | 0.6472| 24.7     | 0.6417| 16.0     |
| PSNR                   | 0.6012| -        | 0.6251| -        |
| PSNR-TF [2], [3], [5]  | 0.6964| 15.8     | 0.6956| 11.3     |
| PSNR-QT [3], [5]       | 0.6845| 13.9     | 0.7051| 12.8     |
| PSNR-HYBRID            | 0.7354| 22.3     | 0.7206| 15.3     |

Note that, in this table, the Goferman’s VA model [9] is used.

3. Experimental Results

In our experiments, we have used two publicly accessible subject-rated image databases: TID2008 database [12] and LIVE database [13]. We have employed two evaluation criteria: Pearson linear correlation coefficient (PLCC) and Spearman rank order correlation coefficient (SROCC). To demonstrate the effectiveness of the proposed method, we have used the cross validation scheme: we have randomly split the set of images into 25 subsets (i.e., 25-fold cross validation). In each fold, 24 subsets are used for training and the remaining subset is used for testing: in the testing phase, training examples are not included. In this paper, to show the generality of our approach, we have used two different computational VA models: Goferman’s algorithm [9] and Bruce’s algorithm [14]. Note that, as was done in [3], we have not conducted any nonlinear fitting to better visualize differences in performance: we have directly used the PLCC and SROCC between the final quality score (obtained from the algorithm) and the MOS or DMOS. We first report the performance of different VA integration approaches on the TID2008 database [12] in Tables 1 and 2. In the tables, “TF” and “QT” indicate task-free saliency map-based (in (2)) and quality-task saliency map-based (in (3)) integrations, respectively, and “HYBRID” indicates our integration approach. Note that the Goferman’s VA model [9] and Bruce’s VA model [14] are used in Tables 1 and 2, respectively. As shown in the tables, “performance gains” with the proposed
4. Conclusion

In this paper, we have presented a hybrid VA integration approach which takes advantages of task-free and quality-task saliency maps. The proposed method has simultaneously incorporated responses (i.e., VA during free viewing of unpaired images and that during image quality assessment) of the human visual system into image quality measures. The performance of our proposed scheme was compared to that of conventional schemes using simple FR quality measures. Our results show that the proposed VA integration method performs better than the existing VA integration techniques.

References

[1] Z. Wang, “Applications of objective image quality assessment methods,” IEEE Signal Process. Mag., vol.28, no.6, pp.137–142, 2011.
[2] A.K. Moorthy and A.C. Bovik, “Visual importance pooling for image quality assessment,” IEEE J. Sel. Top. Signal Process., vol.3, no.2, pp.193–201, 2009.
[3] H. Liu and I. Heynderickx, “Visual attention in objective image quality assessment: Based on eye-tracking data,” IEEE Trans. Circuits Syst. Video Technol., vol.21, no.7, pp.971–982, 2011.
[4] U. Engelke, H. Kaprykowski, H.-J. Zepernick, and P. Ndjiki-Nya, “Visual attention in quality assessment,” IEEE Signal Process. Mag., vol.28, no.6, pp.50–59, 2011.
[5] M.C.Q. Farias and W.Y.L. Akamine, “On performance of image quality metrics enhanced with visual attention computational models,” Electron. Lett., vol.48, no.11, pp.631–633, 2012.
[6] I. Gkioxalekas, G. Evangelopoulos, and P. Maragos, “Spatial Bayesian surprise for image saliency and quality assessment,” Proc. IEEE International Conference on Image Processing, pp.1081–1084, 2010.
[7] A. Guo, D. Zhao, S. Liu, X. Fan, and W. Gao, “Visual attention based image quality assessment,” Proc. IEEE International Conference on Image Processing, pp.3297–3300, 2011.
[8] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” IEEE Trans. Image Process., vol.13, no.4, pp.600–612, 2004.
[9] S. Goferman, L. Zelnik-Manor, and A. Tal, “Context-aware saliency detection,” IEEE Trans. Pattern Anal. Mach. Intell., vol.34, no.10, pp.1915–1926, 2012.
[10] M. Narwaria and W. Lin, “Objective image quality assessment based on support vector regression,” IEEE Trans. Neural Netw., vol.21, no.3, pp.515–519, 2010.
[11] M. Narwaria and W. Lin, “Scalable image quality assessment with 2D mel-cepstrum and machine learning approach,” Pattern Recognit., vol.45, no.1, pp.299–313, 2012.
[12] http://www.ponomarenko.info/tid2008.htm
[13] http://live.ece.utexas.edu/research/quality/
[14] N.D.B. Bruce and J.K. Tsotsos, “Saliency, attention, and visual search: An information theoretic approach,” J. Vision, vol.9, no.3, pp.1–24, 2009.