GSW: A High Performance IQA Index Based on Global and Saliency Window Similarity

Yibo Fan and Zihao Meng

The State Key Lab of ASIC and System, Fudan University, Shanghai, China
Email: 18212020028@fudan.edu.cn

Abstract. Perceptual image quality assessment (IQA) attracts significant attention in recent years. It is proved that both global score and an image’s visual saliency (VS) are consistent with subjective evaluation. The global quality score reflects the consistency of the overall structure between two images, and VS map contributes complementarily in evaluating perceptual quality. This paper presents an effective IQA index based on global and saliency window similarity, namely GSW. It chooses VS map as an image feature and uses a special strategy to draw the saliency window with largest significancy. Meanwhile, the background area is taken into account to guarantee the robustness of the quality score. Experimental results on four most widely used databases verify that, compared with state-of-the-art IQA methods, GSW performs consistently well can provide more accurate quality prediction with a low computational complexity.

Keywords. Image quality assessment; full reference; visual saliency; human visual system.

1. Introduction

As one of the essential steps in image processing, Perceptual image quality assessment (IQA) plays a big role in vision applications. There are two kinds of IQA methods: subjective assessment and objective. As the assessment depending on the eyes of people is too expensive and time-consuming, objective assessment is more attractive to researchers. Because of its features, objective assessment finds its way in real-time evaluation and image quality guidance. The goal of objective assessment is to find a computational model which can accurately predict the perceptual visual quality of human effectively.

Peak signal-to-noise ratio (PSNR) is a widely used conventional fidelity metric. It has clear physical meaning and is easy for calculation. But sadly, quality scores predicted by them correlate poorly with human perceptual, especially when numerous images or distortion situations are involved.

Recently many full-reference (FR) IQA models with different methods are proposed. Researchers have found out that human visual system (HVS) tends to distinguish the structural information of the image at the very beginning. Inspired by this, structural similarity index (SSIM) [1] proposes the method to take luminance, contrast, and structure similarity into account. The feature similarity index (FSIM) [2] employs the phase congruency (PC) and gradient magnitude (GM) to calculate the similarity feature of local information. The gradient similarity-based metric (GSM) [3] is put forward considering the fact that gradients contain details of a image structure. [4] proposes an IQA index using double-random local windows to simulate HVS viewing mode and combines it with global score named GDRW.

Majority of existing FR-IQA indices only focus on global image quality. Actually, several detached distorted areas may interfere the judgment of the image quality. This indicates that in HVS, both...
global and local information play important roles in the perception process of HVS. We present an effective IQA index based on global and saliency window similarity, which two are used as complementary aspects, namely GSW.

The rest of the paper is organized as follows. Section 2 introduces the related works on IQA researches. Section 3 introduces the GSW IQA metric we put forward in detail. Section 4 presents detailed and complete experimental results and analysis. Section 5 gives the conclusion of this paper.

2. Related Works

2.1. Psychological Characteristics of HVS

Visual attention can help human respond to visual information and stimulation. Distorted pixels or regions will more likely be spotted by observer when they are in interest areas. IQA models need to get more accurate prediction score with the HVS perception. Thus, combining IQA models with effective computational VS models to correlate better with human perception has been attracting tremendous attention in recent years.

2.2. Existing Investigations of Visual Saliency in IQA

The HVS is a complex nonlinear system, but most of the current IQA methods are set up based on linear models. Therefore, the accuracy of distorted image quality prediction always relies on some premises or assumptions to be guaranteed. A number of approaches have been proposed to explore the relationship between VS and IQA metrics and aim to make their prediction results more consistent with subjective perception. In [5], Hou and Zhang proposed an algorithm that uses saliency by estimating the foreground area of an image. It can be directly calculated as a mask map of the image’s discrete cosine transform coefficients. The VSI [6] uses VS as a bottom-up feature based on the fact that human tend to notice low-level features at the very beginning and changes of them can reveal the degree of distortions. [7] put forward an image quality metric which take the saliency property into account and can accurately model the direction perception and saliency perception of HVS according to the structure shape.

In most cases, a fully automatic IQA metric is required, however, most of the mentioned approaches show poor performance in actual use. Another problem is that some methods have not undergone extensive and in-depth evaluation to verify their effectiveness.

To avoid these problems, we did experiments on four widely used datasets, and compared our method with the other 11 representative quality assessment methods. Our method shows strong robustness and effectiveness in these experiments.

2.3. Databases and Evaluation Protocols

To verify the effectiveness of our method, we conducted experiments on four widely used image databases. These large-scale benchmarks are specially constructed for IQA indices evaluation, and their information is summarized in table 1.

Four performance metrics are used to evaluate the prediction accuracy and consistency. They are SROCC, KROCC), PLCC, and RMSE.

Table 1. Four datasets for IQA indices evaluation.

| Datasets | Reference Images No. | Distorted Images No. | Distortion Types No. | Subjects |
|----------|----------------------|----------------------|----------------------|----------|
| TID2013  | 25                   | 3000                 | 25                   | 971      |
| TID2008  | 25                   | 1700                 | 17                   | 838      |
| CSIQ     | 30                   | 866                  | 6                    | 35       |
| LIVE     | 29                   | 779                  | 5                    | 161      |
3. Our Method

We proposed a perceptual IQA index based on global and saliency window similarity, namely GSW, which is inspired by the characteristics of HVS and the purpose is to make the calculated objective quality score consistent with the subjective score of the observer.

GSW has a two-part structure and consists of global score and saliency score. The global score detects the image quality in both original size version and down-sampled version, which can give a reference score on the structure similarity and colour fidelity of pictures. The saliency score is calculated from the saliency window area and the background area and is highly consistent with HVS in saliency property.

![Figure 1. Illustration for the computational process of the saliency window.](image)

3.1. Visual Saliency

The saliency window is obtained from the image’s VS map. A significant phase congruent or visual saliency value means that this position is more attractive to HVS, and results show that both visual saliency and phase congruency will effectively reflect the high perceptual distortion which can be immediately sensed by HVS.

There are many ways and VS models can be used to extract the saliency map of an image. According to the actual experimental results, we finally use context-aware saliency detection [12] to get the saliency image in our method. It provides a compact and informative summary.

3.2. Saliency Window

HVS can immediately sense the location of the main content of the image, but the computer cannot do it directly. Therefore, we put forward the saliency window strategy. As shown in figure 1, the saliency map is evenly divided into 12x12 blocks, and the block with the largest saliency is selected as the central block. Taking it as a starting block, we extend the boundary of the saliency window in four directions, up, down, left, and right.

Here we use the left boundary extension as an example to introduce the boundary rules. If the average brightness of the left block larger than 60% of that of the current block or there is a highlighted area connected to the current block in the left block, extend the saliency window’s left boundary.
If the length/width of the drawn window is smaller than 1/3 of the length/width of the image, taking the boundary extension in the left and right directions as an example, the strategy will compare the average brightness of two neighbouring blocks, and include the brighter block into the saliency window.

3.3. Gradient Magnitude
We take the luminance component of the image to calculate the GM similarity. Inspired by FSIM, we use GSM with Prewitt operator. It can effectively assess image quality according to the change in luminance and contrast-structure, and can be used as a good compensation index in our method.

3.4. Global Score
The global score can give a reference score of the image quality based on the structure similarity and colour fidelity of the picture.

The first part of global score is \( G_D \). It is the FSIM value between the reference image and the distorted image.

Then we down sample both the reference image and distorted image to 1/4 of the original size. We calculate the GSM value \( G_D \) between the two down-sample images. \( C_D \) is a positive constant to ensure the stability of \( G_D \). \( G_O \) and \( G_D \) are combined together to get the global score:

\[
GI = G_O \cdot (G_D)^\alpha
\]  
(1)

where \( \alpha \) is a parameter used to adjust the importance of two components.

3.5. Significance Score
The image quality in the saliency window will significantly affect the subjective score. As pixels at different locations inside or outside the saliency window will contribute differently to views’ perception, we calculate weighting functions \( V_S \), the average pixel value inside the saliency window, and \( V_{SBG} \), the average pixel value outside the saliency window, and then use them to reflect the importance of different area.

We use an enhance strategy to calculate the significance score. We apply 5x5 mean blur to the reference image and the distorted image except pixels in the saliency window and get \( I_{RB} \) and \( I_{DB} \). We first calculate the GSM value \( S_W \) of two images in the saliency window \( S_W \), and then calculate the GSM value \( S_{WB} \) between \( I_{RB} \) and \( I_{DB} \). The significance index is defined as:

\[
S_S = \frac{S_W \cdot V_S + S_{WB} \cdot V_{SBG}}{V_S + V_{SBG}}
\]  
(2)

Using saliency window alone doesn’t work well on when images don’t have a specific visual center. To solve this problem, we introduce the background score.

3.6. Background Score and Saliency Score
To remedy the problem mentioned above, we introduce the background score to ensure the Robustness of the final result. Though the image content outside the saliency window has a lower impact on the HVS, adding the score of background area can help improve the robustness and comprehensiveness of the final result.

We perform 5x5 mean blur in the saliency window area of the original image and decayed image, and then calculate the image quality score between them. As the saliency window area is blurred, the calculated score will reflect the distortion in background regions between two images. Background score \( S_B \) is defined as the calculated SSIM value between the two images processed according to the above strategy.

After getting the significance score \( S_S \) and background score \( S_B \), the saliency score is defined as:

\[
SI = S_S \cdot (S_B)^\beta
\]  
(3)
where $\beta$ is a parameter used to adjust the weight of the saliency window and the background features.

3.7. Finally Quality Assessment
The GSW metric is defined as:

$$ GS_W = G1 \cdot (SI)^{\gamma} $$

We add $\gamma$ to adjust the influence weight of the two parts. The compute procedures of GSW are illustrated in figure 2.

![Figure 2. Illustration of the GSW computation process.](image)

4. Experimental Results and Discussion

4.1. Implementation of GSW
We use the following experimental procedure to determine the parameters in GSW. First we build a sub-dataset from TID2013 which contains 1300 images including 10 random reference images and associated 1200 different degrees of damaged images. To tune the related parameters experimentally, the variable control method is used, and the criterion for adjustment is to select parameters that can achieve higher SROCC scores. Then this set of parameters will be verified on the other three databases to confirm whether there is an improvement compared with existing methods.

According to the final experimental results, we set the parameter value as: $C_0 = 1.5$, $\alpha = 0.40$, $\beta = -0.67$, $\gamma = 0.30$. This set of fixed parameters will be used in the rest of this article.

4.2. Overall Performance Comparison
We conducted experiments to get an overall performance comparison. The prediction performance of each computing IQA index on four big-scale image datasets is listed in table 2. The IQA indices giving out the top three results of each database are highlighted in the table.

We can find that VSI and GDRW have high scores on TID2008 and TID2013, but all fail to maintain their advantage on LIVE. GSW achieves the best results in database LIVE, and ranks in the top three in all the four databases. The experimental results prove that our method is effective and stable.
In this paper, we propose an effective IQA metric which consists of global score and saliency score, namely GSW. It uses GM as a quality indicator of the global score and take VS map as a feature in saliency score. GSW uses a special strategy to draw the saliency window with largest significancy, and the background area is also taken into account to enhance the robustness. Experimental results on four most widely used databases show that, compared with state-of-the-art IQA methods, GSW performs consistently well can provide more accurate quality prediction with a low computational complexity. With the adventure of more powerful VS strategies, GSW will correlate better with human perception.

References

[1] Wang Z, Bovik A C, Sheikh H R and Simoncelli E P 2004 Image quality assessment: From error visibility to structural similarity IEEE Trans. Image Process. 13 (4) 600-612.

Table 2. Performance comparison of IQA indices.

|       | SSIM | MS-SSIM [13] | IW-SSIM [14] | IFC | VIF | MAD [10] | FSIM | FSIMc | GSM | VSI | GDRW | GSW |
|-------|------|--------------|--------------|-----|-----|----------|------|-------|-----|-----|-----|-----|
| TID2013 | SROCC 0.7417 0.7859 0.7779 0.5389 0.6769 0.7807 0.8015 0.8510 0.7946 0.8965 0.8803 0.8811 | KROCC 0.5588 0.6047 0.5977 0.3939 0.5147 0.6035 0.6289 0.6665 0.6255 0.7183 0.6978 0.7073 | PLCC 0.7895 0.8329 0.8319 0.5538 0.7720 0.8267 0.8589 0.8769 0.8464 0.9000 0.8913 0.8916 | RMSE 0.7608 0.6861 0.6880 1.0322 0.7880 0.6975 0.6349 0.5959 0.6603 0.5404 0.5621 0.5462 | SROCC 0.7749 0.8542 0.8559 0.5675 0.7491 0.8340 0.8805 0.8840 0.8504 0.8979 0.8971 0.8955 |
| TID2008 | KROCC 0.5768 0.6568 0.6636 0.4236 0.5860 0.6445 0.6946 0.6991 0.6596 0.7123 0.7125 0.7123 | PLCC 0.7732 0.8451 0.8579 0.7340 0.8084 0.8308 0.8738 0.8762 0.8422 0.8762 0.8821 0.8824 | RMSE 0.8511 0.7173 0.6995 0.9113 0.7899 0.7468 0.6525 0.6468 0.7235 0.6466 0.6322 0.6319 | SROCC 0.8756 0.9133 0.9213 0.7671 0.9195 0.9466 0.9242 0.9310 0.9108 0.9423 0.9590 0.9534 |
| CSIQ  | KROCC 0.6907 0.7393 0.7529 0.5897 0.7537 0.7970 0.7567 0.7690 0.7374 0.7857 0.8169 0.8070 | PLCC 0.8613 0.8991 0.9144 0.8348 0.9277 0.9502 0.9120 0.9192 0.8964 0.9279 0.9541 0.9486 | RMSE 0.1334 0.1149 0.1063 0.1431 0.0980 0.0818 0.1077 0.1034 0.1164 0.0979 0.0786 0.0877 | SROCC 0.9479 0.9513 0.9567 0.9295 0.9636 0.9669 0.9634 0.9645 0.9561 0.9524 0.9610 0.9781 |
| LIVE  | KROCC 0.7963 0.8045 0.8175 0.7579 0.8282 0.8421 0.8337 0.8363 0.8150 0.8058 0.8281 0.8515 | PLCC 0.9449 0.9489 0.9522 0.9268 0.9604 0.9675 0.9597 0.9613 0.9512 0.9482 0.9603 0.9680 | RMSE 8.9455 8.6188 8.3473 10.264 7.6137 6.9073 7.6780 7.5296 8.4327 8.6816 7.6247 7.1247 |

4.3. Complexity

Further experiment is conducted to compare the efficiency of each of IQA index. The experiment is performed on Matlab R2017a and a computer equipped with Intel Core i7-7700 CPU @3.60GHz, GPU GTX 1080 Ti, 64GB RAM and 1TB SSD. The time cost of each IQA index for computing the quality is listed in table 3. We randomly picked 1500 distorted images with the resolution of 512x384 from TID2013 and calculated every IQA index’s average time cost. Results show that GSW is in the first tier and runs faster than majority of state-of-the-art metrics, which proves its efficiency.

Table 3. Running efficiency of each IQA index.

| IQA index | SSIM | MS-SSIM | IW-SSIM | IFC | VIF | MAD | FSIM | FSIMc | GSM | VSI | GDRW | GSW |
|-----------|------|---------|---------|-----|-----|-----|------|-------|-----|-----|-----|-----|
| Time cost/s | 0.0701 | 0.0978 | 0.5327 | 2.1527 | 2.2108 | 2.2746 | 0.5103 | 0.5218 | 0.0947 | 0.2631 | 0.4726 | 0.2509 |

5. Conclusions

In this paper, we propose an effective IQA metric which consists of global score and saliency score, namely GSW. It uses GM as a quality indicator of the global score and take VS map as a feature in saliency score. GSW uses a special strategy to draw the saliency window with largest significancy, and the background area is also taken into account to enhance the robustness. Experimental results on four most widely used databases show that, compared with state-of-the-art IQA methods, GSW performs consistently well can provide more accurate quality prediction with a low computational complexity. With the adventure of more powerful VS strategies, GSW will correlate better with human perception.
[2] Zhang L, Zhang D, Mou X and Zhang D 2011 FSIM: A feature similarity index for image quality assessment, IEEE Trans. Image Process. 20 (8) 2378-2386.
[3] Liu A, Lin W and Narwaria M 2012 Image quality assessment based on gradient similarity, IEEE Trans. Image Process. 21 (4) 1500-1512.
[4] Shi Z and Chen K 2017 A perceptual image quality index based on global and double-random window similarity Digital Signal Processing 60 277-286.
[5] Hou X, Harel J and Koch C 2012 Image signature: Highlighting sparse salient regions IEEE Trans. Pattern Anal. Mach. Intell. 34 (1) 194-201.
[6] Zhang L, Shen Y and Li H Y 2014 VSI: A visual saliency-induced index for perceptual image quality assessment IEEE Trans. Image Process. 23 (10) 4270-4281.
[7] Ma L, Li S N and Ngan K N 2010 Visual horizontal effect for image quality assessment IEEE Trans. Signal Process. Lett. 17 (7) 627-630.
[8] Ponomarenko N, et al. 2013, Color image database TID2013: Peculiarities and preliminary results Proc. 4th Eur. Workshop Vis. Inf. Process. pp 106-111.
[9] Ponomarenko N, et al. 2009 TID2008 - A database for evaluation of full-reference visual quality assessment metrics Adv. Modern Radioelectron. 10 (4) 30-45.
[10] Larson E C and Chandler D M 2010 Most apparent distortion: Full-reference image quality assessment and the role of strategy Journal of Electronic Imaging 19 (1) 1-21.
[11] Sheikh H R, Sabir M F and Bovik A C 2006 A statistical evaluation of recent full reference image quality assessment algorithms IEEE Trans. Image Process. 15 (11) 3440-3451.
[12] Stas G, Lihi Z M and Ayellet T 2012 Context-aware saliency detection IEEE Computer Society Conference on Computer Vision and Pattern 2011.5539929.
[13] Wang Z, Simoncelli E P and Bovik A C 2003 Multiscale structural similarity for image quality assessment Proc. 37th Asilomar Conf. Signals, Syst., Comput., 10.1109/ACSSC.2003.1292216.
[14] Wang Z and Li Q 2011 Information content weighting for perceptual image quality assessment, IEEE Trans. Image Process. 20 (5) 1185-1198.