

Adaptively Truncating Gradient for Image Quality Assessment

Minjuan Gao

School of Electrical and Control Engineering
Shaanxi University of Science & Technology
Xi'an, China

E-mail: gaominjuan1984@163.com

Xuande Zhang

School of Electronic Information and Artificial
Intelligence

Shaanxi University of Science & Technology
Xi'an, China

E-mail: 330374142@qq.com

Hongshe Dang

School of Electrical and Control Engineering
Shaanxi University of Science & Technology
Xi'an, China

E-mail: 783940896@qq.com

Abstract—Objective image quality assessment (IQA) aims to develop computational models to predict the perceptual image quality consistent with subjective evaluations. As image information is presented by the change in intensity values in the spatial domain, the gradient, as a basic tool for measuring the change, is widely used in IQA models. However, does the change measured by the gradient actually correspond to the change perceived by the human visual system (HVS)? To explore this issue, in this paper, we analyze how the ability of the HVS to perceive changes is affected by the upper threshold, and we propose an IQA index based on an adaptively truncating gradient. Specifically, the upper threshold at each pixel in an image is adaptively determined according to the image content, and the adaptively truncating gradient is obtained by retaining the part of the gradient magnitude that is less than the upper threshold and truncating the part that is greater than the upper threshold. Then, the distorted image quality is calculated by comparing the similarity of the adaptively truncating gradient between a reference image and the distorted image. Experimental results on six benchmark databases demonstrate that the proposed index correlates well with human evaluations.

Keywords—*Image Quality Assessment; Human Visual System; Upper Threshold; Truncating Gradient*

I. INTRODUCTION

Image quality assessment deals with the quantitative evaluation of the quality of images and can be widely used in image acquisition, compression, storage, transmission and other image processing systems. Generally, human beings are the ultimate receivers of images. Subjective evaluation by humans is a reliable IQA method, but it is cumbersome and difficult to apply in real-world scenarios. An objective IQA method aims to design mathematical models to automatically measure the image quality in a way that is consistent with human evaluations. According to the availability of ground-truth images, objective IQA indices fall into three categories: full-reference (FR), reduced-reference (RR) and no-reference (NR) models [1]. In this paper, the discussion is focused on FR models.

At present, there are two popular techniques for constructing FR models: knowledge-based and learning-based techniques. The deep learning method learns the evaluation model in an end-to-end manner, and its "black-box" lacks explanation. Furthermore, this approach requires a large number of training samples, but the cost of obtaining high-quality and

convincing samples is relatively high. Currently, the commonly used method for obtaining samples is still data augmentation. In this work, we emphasize the knowledge-based approach, which uses knowledge about the HVS to heuristically construct IQA models. Investigating these models reveals that the gradient feature is widely employed. In analyzing the relationship between the gradient feature and the IQA task, the gradient has at least the following two characteristics. 1. The information contained in natural images is presented by changes in intensity value or color in the spatial domain. In extreme cases, the constant image (smoothness) and the pure noise image (variation in all directions) cannot convey any information. Thus, the feature of measuring change is widely used in IQA, with the gradient as the basic tool for measuring change. 2. The judgment of the image quality level in IQA is different from the classic discrimination task. The features for discrimination tasks, such as face recognition and fingerprint recognition, should be robust to image distortion, while the features for IQA should be sensitive to image distortion. The gradient feature is sensitive to image distortion and image content but is weak in robustness.

Representative FR models using the gradient feature include the feature similarity index (FSIM) [2], gradient magnitude similarity deviation index (GMSD) [3], superpixel-based similarity index (SPSIM) [4] and directional anisotropic structure metric (DASM) [5]. In the FSIM and GMSD, the image gradient magnitude is employed as the fundamental feature. SPSIM is computed on the basis of three features: superpixel luminance, superpixel chrominance and pixel gradient. The DASM is obtained by incorporating the gradient magnitude, anisotropy and local directivity features. Objective IQA models are designed by simulating the behaviors of the HVS, which integrates perception, understanding and assessing functions, that is, humans evaluate the image quality in the HVS perception space. Therefore, the features for IQA should be the subjective quantity perceived by the HVS. The gradient is often directly used in IQA models as an effective feature to measure change; however, does the change measured by the gradient actually correspond to that perceived by the HVS? In fact, the change measured by the gradient belongs to the objective quantity (objective physical stimulus), while that perceived by the HVS belongs to the subjective quantity (subjective response). Thus, how can one map the objective quantity to the subjective quantity? This mapping function is nonlinear, and it is difficult to accurately describe its form. Empirically, the ability of the human perception system to sense changes has a certain upper threshold. When

the objective change exceeds the upper threshold, the subjective change increases insignificantly in situations such as the human perception of changes in salt-solution saltiness, at an outside temperature, and in the weight of objects carried.

In this paper, we discuss the ability of the HVS to perceive changes affected by the upper threshold by employing the adaptively truncating gradient to measure the change perceived by the HVS. We propose an IQA index based on the adaptively truncating gradient. Specifically, the upper threshold at each pixel in the image is adaptively determined according to the image content, and the adaptively truncating gradient is obtained by retaining the part of the gradient magnitude that is less than the upper threshold and truncating the part that is greater than the upper threshold. Experimental results on public databases show that the proposed index correlates well with the subjective judgments.

II. AN IQA INDEX BASED ON ADAPTIVELY TRUNCATING GRADIENT

A. Definition of Adaptively Truncating Gradient

The image information is presented by the change in the intensity values in the spatial domain, and this change may be destroyed by degradation of the image quality. The gradient feature can effectively measure the change and is widely used in IQA algorithms. The image gradient can be obtained by convolving the image with a gradient operator, such as Sobel, Roberts and Scharr and Prewitt. Usually, a different gradient operator for the IQA model may yield distinguished performance. This problem was discussed in [2,6], where the experiment results showed that the Scharr operator can obtain a slightly better performance than the others. Here, we adopt a 3×3 Scharr operator whose templates along the horizontal (H) and vertical (V) directions take the following form:

$$\mathbf{h}_H = \frac{1}{16} \begin{bmatrix} 3 & 0 & -3 \\ 10 & 0 & -10 \\ 3 & 0 & -3 \end{bmatrix}, \quad \mathbf{h}_V = \frac{1}{16} \begin{bmatrix} 3 & 10 & 3 \\ 0 & 0 & 0 \\ -3 & -10 & -3 \end{bmatrix}$$

Denote $\mathbf{r} = [r_1, \dots, r_i, \dots, r_N]$ for a reference image and $\mathbf{d} = [d_1, \dots, d_i, \dots, d_N]$ for a distorted image, where i is the pixel index, and N is the number of total pixels. The image gradients in the horizontal and vertical directions can be obtained by convolution of the image with \mathbf{h}_H and \mathbf{h}_V , and the gradient magnitude is

computed from their root mean square. The gradient magnitudes of \mathbf{r} and \mathbf{d} at each pixel i , denoted as $G(\mathbf{r},i)$ and $G(\mathbf{d},i)$, are calculated as

$$G(\mathbf{r},i) = \sqrt{(\mathbf{r} \otimes \mathbf{h}_H)^2(i) + (\mathbf{r} \otimes \mathbf{h}_V)^2(i)} \quad (1)$$

$$G(\mathbf{d},i) = \sqrt{(\mathbf{d} \otimes \mathbf{h}_H)^2(i) + (\mathbf{d} \otimes \mathbf{h}_V)^2(i)} \quad (2)$$

Where the symbol \otimes denotes the convolution operation.

The image gradient only reflects the objective changes in images. Since human evaluation of image quality is carried out in the HVS perception space, the image features extracted for IQA models should reflect the subjective changes perceived by the HVS. We consider that the ability of HVS to perceive changes is subject to the upper threshold. When the objective change exceeds the upper threshold, the subjective change does not obviously increase. In this study, we define the adaptively truncating gradient to measure the subjective change sensed by the HVS.

Denote T as the upper threshold. We define a truncating function $trunc(\cdot)$. For any given variable x , it is retained when it is less than T and truncated when it is greater than T . The specific expression is

$$trunc(x) = \begin{cases} T, & \text{if } x \geq T \\ x, & \text{if } x < T \end{cases} \quad (3)$$

The truncating gradients of \mathbf{r} and \mathbf{d} at each pixel i are denoted as $G_T(\mathbf{r},i)$ and $G_T(\mathbf{d},i)$, and the upper threshold at this point is denoted as $T(i)$. Using formula (3), the calculation of $G_T(\mathbf{r},i)$ is as follows:

$$G_T(\mathbf{r},i) = trunc(G(\mathbf{r},i)) = \begin{cases} T(i), & \text{if } G(\mathbf{r},i) \geq T(i) \\ G(\mathbf{r},i), & \text{if } G(\mathbf{r},i) < T(i) \end{cases} \quad (4)$$

In Eq. (4), if the value of $G(\mathbf{r},i)$ is greater than $T(i)$, then $G(\mathbf{r},i)$ will be truncated, and the truncating gradient $G_T(\mathbf{r},i)$ is set to $T(i)$. That is, the part of the gradient magnitude that is greater than the upper threshold is masked. Otherwise, $G(\mathbf{r},i)$ is not be masked, and the truncating gradient $G_T(\mathbf{r},i)$ is set equal

to $G(\mathbf{r},i)$. That is, the part of the gradient magnitude that is less than the upper threshold can be perceived by the HVS.

Similarly, using formula (3), $G_T(\mathbf{d},i)$ is calculated as follows:

$$G_T(\mathbf{d},i) = trunc(G(\mathbf{d},i)) = \begin{cases} T(i), & \text{if } G(\mathbf{d},i) \geq T(i) \\ G(\mathbf{d},i), & \text{if } G(\mathbf{d},i) < T(i) \end{cases} \quad (5)$$

Obviously, for the calculation of the truncating gradients $G_T(\mathbf{r},i)$ and $G_T(\mathbf{d},i)$ in Eq. (4) and (5), the selection of the upper threshold $T(i)$ is very important. According to Weber's law, the ratio of the stimulus change that causes a just noticeable difference (JND) from the original stimulus intensity is a constant. In psychology, the HVS has the property of light adaptation, and the perception of luminance obeys Weber's law [7]. The just noticeable incremental luminance over the background by the HVS is related to the background luminance.

Inspired by this recognition, in contrast to Weber's law, we consider that the upper threshold for truncating the significantly perceptible stimulus change is also related to the original stimulus intensity value. Because different pixels in the image correspond to different gray values, the original stimulus intensity values will also be different. Here, we adaptively determine the upper threshold according to the background luminance of different areas of the image.

The adaptively upper threshold is defined as

$$T(i) = \frac{I(i)}{T_0} \quad (6)$$

Where T_0 is an adjustable threshold parameter. (The details of selecting T_0 will be presented in section III-A.) $I(i)$ takes the larger value of the luminance of \mathbf{r} and \mathbf{d} at point i .

$$I(i) = \max(\bar{r}(i), \bar{d}(i)) \quad (7)$$

In formula (7), the luminance values $\bar{r}(i)$ and $\bar{d}(i)$ at pixel i of \mathbf{r} and \mathbf{d} is estimated by formulas (8) and (9). For reference image \mathbf{r} , denote the square neighborhood as Ω_i^r with center of pixel i and radius

of t , and let the intensity value of any pixel in the neighborhood be $r_{i,j}$, $j \in \Omega_i^r$. Similarly, for the distorted image, denote the square neighborhood as Ω_i^d with center of pixel i and radius of t , and let the intensity value of any pixel in the neighborhood be $d_{i,j}$, $j \in \Omega_i^d$.

$$\bar{r}(i) = \frac{1}{m} \sum_{j=1}^m r_{i,j} \quad (8)$$

$$\bar{d}(i) = \frac{1}{m} \sum_{j=1}^m d_{i,j} \quad (9)$$

Where $m = (2t + 1)^2$.

Based on Eq. (6), the value of the upper threshold at each pixel in an image can be adaptively determined according to the image content. Then, the adaptively truncating gradient is obtained by formulas (4) and (5). Figure 1 shows the gradient map and the adaptively truncating gradient map corresponding to the reference image and the distorted image. It can be seen that the maximum amplitude of the gradient map is approximately 250, while the maximum amplitude of the adaptively truncating gradient is approximately 70.

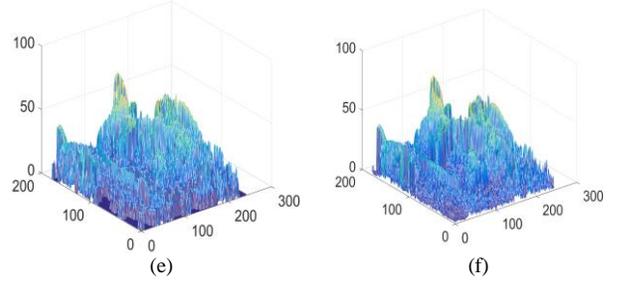
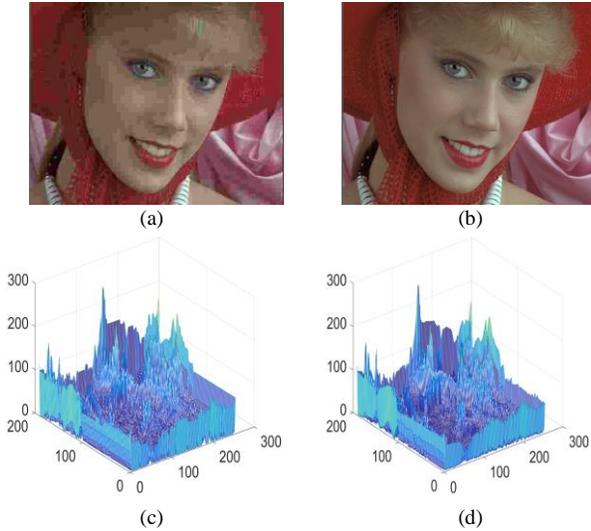


Figure 1. The gradient map and the adaptively truncating gradient map corresponding to the reference image and the distorted image. (a) the reference image. (b) the distorted image. (c) and (d) are the gradient map of (a) and (b), respectively. (e) and (f) are the adaptively truncating gradient map of (a) and (b), respectively.

B. The Proposed IQA Index

With the adaptively truncating gradient defined, the local quality of the distorted image is predicted by the similarity between the adaptively truncating gradient of \mathbf{r} and \mathbf{d} , which is defined as

$$S(i) = \frac{2G_T(\mathbf{r}, i) \cdot G_T(\mathbf{d}, i) + C}{G_T^2(\mathbf{r}, i) + G_T^2(\mathbf{d}, i) + C} \quad (10)$$

Where the parameter C is introduced to avoid the denominator becoming zero and supplies numerical stability. The range of $S(i)$ is from 0 to 1. Obviously, on the one hand, $S(i)$ is close to 0 when $G_T(\mathbf{r}, i)$ and $G_T(\mathbf{d}, i)$ are quite different. On the other hand, $S(i)$ will achieve the maximal value 1 when $G_T(\mathbf{r}, i)$ is equal to $G_T(\mathbf{d}, i)$.

The overall quality score of the distorted image is predicted by the local quality $S(i)$, which is calculated as follows :

$$score = \frac{1}{N} \sum_{i=1}^N S(i) \quad (11)$$

A higher score indicates better image quality.

III. EXPERIMENTAL RESULTS

A. Experimental Setup

All the experiments in this study were implemented in MATLAB R2016b and executed on a Lenovo Ideapad700 laptop with Intel Core i5-6300HQ@2.3-

GHz CPU and 4 GB RAM. Several well-known FR metrics were used when comparing performances with the proposed method, including PSNR, SSIM[1], FSIM [2], GMSD[3], DASM[5], IFC [8], VIF [9], MS-SSIM [10], and SSRM [11]. To widely evaluate the performance of these metrics, six public databases were employed for the experiments: TID2013 [12], TID2008 [13], CSIQ [14], LIVE [15], IVC [16] and A57 [17]. The TID2008 database consists of 25 reference images and a total of 1700 distorted images, each of which is distorted using 17 different types of distortions at four different levels of distortion. The TID2013 is an expanded version of TID2008, which contains 3000 distorted images with 24 distortion types. The LIVE database includes 29 reference images and 779 distorted images with five distortion types. The CSIQ database contains 30 original images and 886 distorted images degraded by six types of distortion. The IVC database consists of 10 reference images and 185 distorted images. The A57 database includes 3 reference images and 54 distorted images. Note that for the color images in these databases, only the luminance component is evaluated.

Four commonly used performance criteria are employed to evaluate the competing IQA metrics. The Spearman rank order correlation coefficient (SROCC) and Kendall rank order correlation coefficient (KROCC) are adopted for measuring the prediction monotonicity of an objective IQA metric. For compute the other two criteria, the Pearson linear correlation coefficient (PLCC) and the root mean squared error (RMSE), we need to apply a regression analysis. The PLCC measures the consistency between the objective scores after nonlinear regression and the subjective mean opinion scores (MOS). The RMSE measures the relative distance between the objective scores after nonlinear regression and MOS. For the nonlinear regression, we used the following mapping function:

$$Q_p = \beta_1 \left[\frac{1}{2} - \frac{1}{1 + e^{\beta_2(Q - \beta_3)}} \right] + \beta_4 Q + \beta_5 \quad (12)$$

where Q and Q_p are original objective scores of an IQA metric and the objective scores after regression, respectively. β_i , $i=1,2,\dots,5$ are the fixed parameters. Higher values of SROCC, KROCC, PLCC and lower RMSE values indicate a better performance of IQA metrics.

For the proposed metric, there are three parameters that need to be set to obtain the final quality score. They are T_0 , t and C . Selecting the first 8 reference images and corresponding 544 distorted images in the TID2008 database as the testing subset, we choose the parameters that can yield the highest SROCC. The result is $T_0 = 3$, $t = 51$ and $C = 1600$.

To further analyze the effect of threshold parameter T_0 , more experiments were carried out. Figure 2 shows the SROCC performance with different T_0 values on six databases. On most databases, SROCC can be best when T_0 is 3. This result indicates that the range of upper threshold T is approximately $[0, 255/3]$ for an 8-bit grayscale image according to formula (6). If the change in image intensity is above $255/3$, then it will be masked in visual perception.

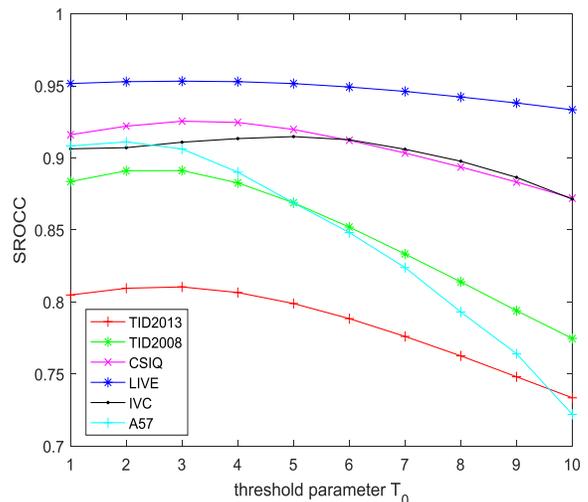


Figure 2. SROCC performance with different T_0 values on six databases

B. Performance Comparison

Table I lists the SROCC, KROCC, PLCC and RMSE results of ten metrics on six databases, and the two best results of each row are highlighted in bold. Overall, the methods which employed the gradient feature performs well across all the databases, such as FSIM, GMSD, DASM and the proposed metric. This partly demonstrates the validity of considering the degradation of gray changes in quality evaluation. Furthermore, the proposed metric performs well, outperforming SSIM and SSRM and competing with FSIM and GMSD.

TABLE I. COMPARISON THE PERFORMANCE RESULTS OF TEN IQA METRICS ON SIX PUBLIC DATABASES.

THE FIRST TWO ARE MARKED IN BOLD

Database	criteria	PSNR	SSIM (2004)	MS-SSIM (2003)	IFC (2005)	VIF (2006)	FSIM (2011)	GMSD (2014)	DASM (2017)	SSRM (2018)	Proposed
TID2013	SROCC	0.6396	0.7417	0.7859	0.5389	0.6769	0.8015	0.8038	0.8025	0.7500	0.8105
	KROCC	0.4698	0.5588	0.6047	0.3939	0.5147	0.6289	0.6334	0.6321	0.5718	0.6387
	PLCC	0.7017	0.7895	0.8329	0.5538	0.7720	0.8589	0.8542	0.8574	0.8078	0.8601
	RMSE	0.8832	0.7608	0.6861	1.0322	0.7880	0.6349	0.6444	0.6547	0.7307	0.6324
TID2008	SROCC	0.5531	0.7749	0.8542	0.5675	0.7491	0.8805	0.8906	-	0.8332	0.8913
	KROCC	0.4027	0.5768	0.6568	0.4236	0.5860	0.6946	0.7090	-	0.6535	0.7042
	PLCC	0.5734	0.7732	0.8451	0.7340	0.8090	0.8738	0.8786	-	0.8379	0.8745
	RMSE	1.0994	0.8511	0.7173	0.9113	0.7888	0.6525	0.6408	-	0.7324	0.6458
LIVE	SROCC	0.8756	0.9479	0.9513	0.9259	0.9636	0.9634	0.9546	0.9601	0.9608	0.9531
	KROCC	0.6865	0.7963	0.8045	0.7579	0.8282	0.8337	0.8237	0.8218	0.8312	0.8211
	PLCC	0.8721	0.9449	0.9430	0.9248	0.9598	0.9597	0.9515	0.9571	0.9695	0.9379
	RMSE	13.368	8.9455	9.0956	10.392	7.6734	7.6780	7.1131	7.7716	5.6639	8.0188
CSIQ	SROCC	0.8058	0.8756	0.9133	0.7671	0.9195	0.9242	0.9571	0.9523	0.9369	0.9251
	KROCC	0.6084	0.6907	0.7393	0.5897	0.7537	0.7567	0.8122	0.8041	0.7791	0.7575
	PLCC	0.8001	0.8613	0.8998	0.8381	0.9277	0.9120	0.9543	0.9531	0.9097	0.9055
	RMSE	0.1575	0.1334	0.1145	0.1432	0.0980	0.1077	0.0791	0.0799	0.1138	0.1114
IVC	SROCC	0.6884	0.9018	0.8980	0.8993	0.8964	0.9262	0.8789	0.8966	0.9047	0.9103
	KROCC	0.5218	0.7223	0.7203	0.7202	0.7158	0.7564	0.6882	0.7179	0.7310	0.7352
	PLCC	0.7199	0.9119	0.8934	0.9080	0.9028	0.9376	0.8549	0.9190	0.9132	0.9206
	RMSE	0.8456	0.4999	0.5474	0.5105	0.5239	0.4236	0.6320	0.5220	0.4965	0.4758
A57	SROCC	0.6189	0.8066	0.8394	0.3185	0.6223	0.9181	0.9103	0.9215	0.8527	0.9062
	KROCC	0.4309	0.6058	0.6478	0.2378	0.4589	0.7639	0.7513	0.7782	0.6604	0.7289
	PLCC	0.6587	0.8017	0.8504	0.4548	0.6158	0.9252	0.9085	0.9429	0.8528	0.8530
	RMSE	0.1849	0.1469	0.1293	0.2189	0.1936	0.0933	0.1027	0.0813	0.0707	0.1283

TABLE II. COMPARISON SROCC FOR INDIVIDUAL DISTORTION OF TEN IQA METRICS ON TID2013 DATABASE.

THE FIRST TWO ARE MARKED IN BOLD

Database	Distortion type	PSNR	SSIM (2004)	MS-SSIM (2003)	IFC (2005)	VIF (2006)	FSIM (2011)	GMSD (2014)	DASM (2017)	SSRM (2018)	Proposed
TID2013	Awgn	0.9291	0.8671	0.8646	0.6612	0.8994	0.8973	0.9461	0.9299	0.8545	0.9293
	Awgn-color	0.8986	0.7726	0.7730	0.5352	0.8299	0.8208	0.8689	0.8612	0.7757	0.8463
	Spatial-correlated	0.9197	0.8515	0.8544	0.6601	0.8835	0.8750	0.9348	0.9301	0.8392	0.9178
	Mask-noise	0.8321	0.7767	0.8073	0.6932	0.8450	0.7944	0.7085	0.8019	0.8184	0.8068
	HF-noise	0.9140	0.8634	0.8604	0.7406	0.8972	0.8984	0.9164	0.9179	0.8754	0.9069
	Impulse-noise	0.8969	0.7503	0.7629	0.6408	0.8537	0.8072	0.7633	0.8550	0.7872	0.8336
	Quantization-noise	0.8801	0.8657	0.8706	0.6282	0.7854	0.8719	0.9057	0.9032	0.8496	0.8629
	GB	0.9155	0.9668	0.9673	0.8907	0.9650	0.9551	0.9114	0.9546	0.9674	0.9686
	Denoising	0.9481	0.9254	0.9268	0.7779	0.8911	0.9302	0.9525	0.9496	0.9288	0.9361
	JPEG	0.9189	0.9200	0.9265	0.8357	0.9192	0.9324	0.9500	0.9473	0.9287	0.9515
	JP2K	0.8840	0.9468	0.9504	0.9078	0.9516	0.9577	0.9656	0.9620	0.9562	0.9635
	JPEG-trans-error	0.7682	0.8493	0.8475	0.7425	0.8409	0.8464	0.8401	0.8534	0.8369	0.8802
	JP2K-trans-error	0.8886	0.8828	0.8889	0.7769	0.8761	0.8913	0.9135	0.8966	0.8765	0.9141
	Pattern-noise	0.6864	0.7821	0.7968	0.5737	0.7720	0.7917	0.8143	0.8138	0.7745	0.7632
	Block-distortion	0.1552	0.5720	0.4801	0.2414	0.5306	0.5489	0.6630	0.6338	0.3186	0.6635
	Mean-shift	0.7671	0.7752	0.7906	0.5522	0.6276	0.7531	0.7356	0.6127	0.6919	0.6143
	Contrast change	0.4416	0.3775	0.4634	-0.1798	0.8386	0.4686	0.3253	0.3498	0.4519	0.4889
	Saturation change	0.0944	-0.4141	-0.4099	-0.4029	-0.3099	-0.2748	-0.1907	0.0382	-0.2513	-0.2602
	Multiple-noise	0.8911	0.7803	0.7786	0.61423	0.8468	0.8469	0.8880	0.8814	0.8067	0.8698
	Comfort-noise	0.8410	0.8566	0.8528	0.81620	0.8946	0.9121	0.9298	0.9203	0.8921	0.9112
Noisy-compression	0.9144	0.9057	0.9068	0.8180	0.9204	0.9466	0.9631	0.9402	0.9164	0.9367	
Color quantization	0.9269	0.8542	0.8555	0.6006	0.8414	0.8760	0.9098	0.9177	0.8546	0.8952	
Chromatic abbr.	0.8871	0.8775	0.8784	0.8210	0.8848	0.8715	0.8517	0.8693	0.8844	0.8849	
Sparse sample	0.9044	0.9461	0.9483	0.8885	0.9353	0.9565	0.9684	0.9669	0.9541	0.9601	

Among the six databases, TID2013 has the highest number of distorted types. Table II lists the SROCC results of ten metrics about each individual distorted type of the TID2013 database. The proposed algorithm performs well in variety of distortion types. In particular, the proposed algorithm is outstanding for JPEG, JP2K and JPEG-trans-error distortion types that are sensitive to variations.

IV. CONCLUSION

In this paper, we discuss the problem of whether the change measured by the gradient correspond to the change perceived by the HVS. Considering that the ability of the HVS to perceive changes is affected by the upper threshold, we defined the adaptively truncating gradient and proposed a novel IQA index. Numerical experimental results showed that this index performs well on multiple databases. In addition, more studies need to be conducted to address this problem due to its complexity. In future research, we expect to using machine learning methods to further understand this issue.

ACKNOWLEDGMENT

This work was supported by the Natural Science Foundation of Shaanxi Province under Grant 2020JM-509.

REFERENCES

- [1] 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, Apr. 2004.
- [2] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
- [3] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 684–695, Feb. 2014.
- [4] W. Sun, Q. Liao, J. Xue, and F. Zhou, "SPSIM: A superpixel-based similarity index for full-reference image quality assessment," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4232–4244, Sept. 2018.
- [5] L. Ding, H. Huang, and Y. Zang, "Image quality assessment using directional anisotropy structure measurement," *IEEE Trans. Image Process.*, vol. 24, no. 4, pp. 1799–1809, Apr. 2017.
- [6] X. Zhang, X. Feng, W. Wang, and W. Xue, "Edge strength similarity for image quality assessment," *IEEE Signal Process. Lett.*, vol. 20, no. 4, pp. 319–322, Apr. 2013.
- [7] Z. Wang and A. C. Bovik, "Bottom-up approaches for full-reference image quality assessment," in *Modern image quality assessment*, Vermont, VT, USA: Morgan and Claypool, 2006, pp. 17–40.
- [8] H. R. Sheikh, A. C. Bovik, and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2117–2128, Dec. 2005.
- [9] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 430–444, Feb. 2006.
- [10] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," in *Proc. IEEE Asilomar Conf. Signals, Syst. Comput.*, Nov. 2003, pp. 1398–1402.
- [11] A. Ahar, A. Barri and P. Schelkens, "From Sparse Coding Significance to Perceptual Quality: A New Approach for Image Quality Assessment," *IEEE Trans. Image Process.*, vol. 27, no. 2, pp. 879–893, Feb. 2018.
- [12] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, and C.-C. Jay Kuo, "Color image database TID2013: Peculiarities and preliminary results," in *Proc. 4th Eur. Workshop Vis. Inf. Process.*, Jun. 2013, pp. 106–111.
- [13] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "TID2008A database for evaluation of full-reference visual quality assessment metrics," *Adv. Modern Radioelectron.*, vol. 10, pp. 30–45, May. 2009.
- [14] C. Larson and D. M. Chandler, *Categorical Image Quality (CSIQ) Database 2009* [Online]. Available: <http://vision.okstate.edu/csiq>
- [15] H. R. Sheikh, K. Seshadrinathan, A. K. Moorthy, Z. Wang, A. C. Bovik, and L. K. Cormack, *Image and Video Quality Assessment Research at LIVE 2004* [Online]. Available: <http://live.ece.utexas.edu/research/quality>
- [16] A. Ninassi, P. Le Callet, and F. Atrousseau, *Subjective Quality Assessment IVC Database 2005* [Online]. Available: <http://www2.irccyn.ecnantes.fr/ivcdb>
- [17] D. M. Chandler and S. S. Hemami, *A57 Database 2007* [Online]. Available: <http://fouard.ece.cornell.edu/dmc27/vsnr/vsnr.htm>